Computational Investigations of the Biocatalytic, Photophysical and Spectroscopic Properties of Flavins and Flavoproteins

Inaugural Dissertation

for the attainment of the title of doctor
in the Faculty of Mathematics and Natural Sciences
at the Heinrich Heine University Düsseldorf

presented by

Bora Karasulu
from Istanbul

Mülheim an der Ruhr / Düsseldorf, January 2015

From the institute for Theoretical and Computational Chemistry
at the Heinrich Heine University Düsseldorf

Published by permission of the
Faculty of Mathematics and Natural Sciences at
Heinrich Heine University Düsseldorf

Supervisor: Prof. Dr. Walter Thiel
Co-supervisor: Prof. Dr. Christel Marian

Date of the oral examination: 11.02.2015

Hiermit versichere ich, die hier vorgelegte Arbeit eigenständig und ohne unerlaubte Hilfe angefertigt zu haben. Die Dissertation wurde in der vorgelegten oder in ähnlicher Form noch bei keiner Institution eingereicht. Ich habe keine erfolglosen Promotionsversuche unternommen.

Düsseldorf, den 11.02.2015

(Bora Karasulu)

Acknowledgments

First and foremost, I would like to express my sincere and endless gratitude to my supervisor Prof. Dr. Walter Thiel for his unmatched kindness, continuous patience, support and encouragement during my PhD studies. During all these four years, I learned so much from his invaluable guidance, as a result of his remarkable knowledge and experience in many fields, his tireless endeavors for perfection and splendid meticulousness. I also gratefully acknowledge the Max-Planck-Society (Max-Planck-Gesellschaft) for the continuous funding of my PhD studies and for the financial support to attend scientific meetings, workshops and schools that contributed notably to my professional development.

I am very grateful to Prof. Dr. Christel Marian for being my co-advisor, for critically reading my dissertation, and for the pleasant and fruitful collaboration. I would also like to thank everybody who will have reviewed this thesis.

I am indebted to my collaborators Dr. Jan P. Götze and Dr. Mahendra Patil. It was a real pleasure to collaborate with them and enjoy their friendship. I truly appreciate their assistance and help throughout my PhD studies. I acknowledge Dr. Jan P. Götze for translating the abstract of this thesis into German. I also thank Jeaphianne van Rijn for her efforts in our collaborative work.

I acknowledge all my former and current colleagues from the Max-Planck-Institut für Kohlenforschung, especially Dr. Mario Barbatti, Dr. Daniele Fazzi, Dr. Pandian Sokkar, Dr. Thibaut Very, Dr.Wilmer Arbelo González, Gessenildo Pereira Rodrigues, Sumit Mittal, and Alec Owens for fruitful scientific discussions as well as for their companionship. I am also thankful to Dr. Mario Ramos da Silva for his valuable help and guidance in the initial phase of my PhD.

Last but not least, I am deeply grateful to my parents Naide and Nedim, and my brother Bahadır for their unconditional love and support throughout my life. I feel so lucky to have them. I also thank Gül for her precious support.

List of Publications

Included as part of the thesis:
Own tasks are given in italics.

(1) "Amine Oxidation Mediated by Lysine-Specific Demethylase 1: Quantum Mechanics/Molecular Mechanics Insights into Mechanism and Role of Lysine 661", Bora Karasulu, Mahendra Patil, and Walter Thiel, J. Am. Chem. Soc., **2013**, 135 (36), pp 13400–13413.
Performed all the calculations, analyzed the results, and wrote the draft of the manuscript.

(2) "Amine Oxidation Mediated by N-Methyltrytophan Oxidase: Computational Insights into the Mechanism, the Role of Active-Site Residues and Covalent Flavin Binding", Bora Karasulu and Walter Thiel, *ACS Catal.,* **2015**, 5, pp 1227–1239.
Performed all the calculations, analyzed the results, and wrote the draft of the manuscript.

(3) "Photoinduced Intra-molecular Charge Transfer in an Electronically Modified Flavin Derivative: Roseoflavin", Bora Karasulu and Walter Thiel, *J. Phys. Chem. B,* **2015,** 119, pp 928–943.
Performed all the calculations, analyzed the results, and wrote the draft of the manuscript.

(4) "Photophysics of Flavin Derivatives Absorbing in the Blue-Green Region: Thioflavins as Potential Cofactors of Photoswitches", Christel M. Marian, Setsuko Nakagawa, Vidisha Rai-Constapel, Bora Karasulu, and Walter Thiel, *J. Phys. Chem. B,* **2014**, 118 (7), pp 1743–1753.
Performed the excited-state calculations for the thioflavins (2T-LF, 4T-LF, and 2,4DT-LF) to obtain minimum geometries, excitation energies, and one-electron properties at the TD-DFT and DFT/MRCI levels of theory (in the gas phase and water).

(5) "Computing UV/Vis Spectra from the Adiabatic and Vertical Franck-Condon Schemes with the use of Cartesian and Internal Coordinates", Jan P. Götze[*], Bora Karasulu[*] and Walter Thiel, *J. Chem. Phys.*, **2013**, 139, 234108 ([*]Equal authors).
Equally contributed to the development and implementation of the main algorithm, performed part of the test calculations, and analyzed the results.

(6) "Assessment of Franck-Condon Methods for Computing Vibrationally Broadened UV/vis Absorption Spectra of Flavin Derivatives: Riboflavin, Roseoflavin and 5-thioflavin", Bora Karasulu, Jan P. Götze and Walter Thiel, *J. Chem. Theo. Comp.,* **2014**, 10 (12), pp 5549–5566.
Performed most of the calculations and analyzed the results; compiled a Python script for interfacing different software components and implemented additional features as discussed in the text; wrote the draft of the manuscript.

Other publications not submitted as part of the thesis.

(1) "Carotenoids as a Shortcut for Chlorophyll Soret-to-Q Energy Flow", Jan P. Götze, Dominik Kröner, Shiladitya Banerjee, Bora Karasulu and Walter Thiel, *ChemPhysChem,* **2014**,15, pp. 3392-3401.

(2) "QM/MM Studies of Two Strongly Coupled Chromophores in the Peridin-Chlorophyll A-Protein", Jan P. Götze, Bora Karasulu, Mahendra Patil, and Walter Thiel, *BBA–Bioenergetics*, **Submitted.**

Abstract

This dissertation addresses three different aspects of the interdisciplinary research on flavins and flavoproteins, which have been shown to be responsible for a myriad of essential biological, chemical and physical phenomena, from a computational point of view.

Amine Oxidation Mediated by Flavin-dependent Catalysts. The main goal of this part was to probe at the molecular level the catalytic mechanism of the rate-limiting oxidation step of a typical amine demethylation, a common process in diverse biochemical processes. Models were created for the active site of two amine oxidases, lysine-specific-demethylase-1 (LSD1) and N-methyltryptophan oxidase (MTOX) that utilize a covalently or noncovalently bound flavin adenine dinucleotide (FAD) as cofactor. The model systems were studied using quantum mechanics (QM) and hybrid QM/molecular mechanics (QM/MM) methods, along with classical molecular dynamics (MD) simulations, to identify the most feasible of the proposed pathways for amine oxidation and to explain the role of active-site residues in the reaction.

Photophysics and Photochemistry of Flavin Derivatives. In this part, we focus on the excited-state properties of nine different natural and artificial riboflavin (RF) analogs. We provide an in-depth computational analysis of the photophysics and photodynamics of these flavins using time-dependent density functional theory (TD-DFT) and combined DFT/multi-refence configuration interaction (MRCI) methods. On the basis of this analysis, we identify the most plausible mechanism for the intramolecular charge transfer occurring in a natural RF analog, roseoflavin (RoF). In addition, we assess the potential of eight artificial flavin derivatives to replace wild-type RF in blue-light photoreceptor flavoproteins with the aim of realizing the relevant photoinduced processes with lower energy input.

Optical Spectroscopy of Flavin Derivatives. Finally, we address the optical spectra of flavin derivatives using computational methods for simulating vibrationally broadened UV-vis spectra that are based on the Franck-Condon (FC) principle. We introduce a new technique

that combines the vertical FC scheme with curvilinear displacements and compare its performance to previous schemes with rectilinear displacements. Thereafter we present a comprehensive benchmark of the performance of most FC-based spectroscopic simulation methods using a carefully chosen set of three flavins with distinctive structural features.

Zusammenfassung

In dieser Arbeit werden drei Themenbereiche aus der interdisziplinären Forschung zu Flavinen und Flavoproteinen mit den Methoden der Theoretischen Chemie behandelt. Flavine sind verantwortlich für eine Vielzahl von essentiellen biologischen, chemischen und physikalischen Phänomenen.

Aminoxidation durch Flavin-abhängige Katalysatoren. Das Hauptziel dieses Teils war es, den katalytischen Mechanismus einer typischen Amindemethylierung in seinem ratenbestimmenden Schritt auf molekularem Niveau zu modellieren. Hierzu wurden Modelle von zwei Aminoxidasen erstellt, Lysin-spezifische Demethylase 1 (LSD1) und N-Methyltryptophan Oxidase (MTOX), die kovalent oder nicht-kovalent gebundenes Flavin-Adenin-Dinukleotid (FAD) als Cofaktor nutzen. Die Modellsysteme wurden mit Hilfe von quantenmechanischen (QM) und gekoppelten QM/Molekularmechanik (QM/MM) Methoden untersucht, begleitet von klassischen Molekulardynamik (MD) Simulationen, um den bevorzugten Mechanismus zu identifizieren und die Rolle der Aminosäuren in der aktiven Tasche der Enzyme zu verstehen.

Photophysik und Photochemie von Flavinderivaten. Dieser Teil widmet sich den Eigenschaften der elektronisch angeregten Zustände in neun Riboflavinanaloga natürlichen oder künstlichen Ursprungs. Hierzu wurde eine detaillierte computergestützte Analyse der Photophysik und Photodynamik dieser Flavine durchgeführt, mit Hilfe von zeitabhängiger Dichtefunktionaltheorie (TD-DFT) und einer Kombination aus DFT und Multireferenz-Konfigurationswechselwirkung (DFT/MRCI). Aus den Ergebnissen konnte auf den plausibelsten Mechanismus für den intramolekularen Ladungstransfer in einem natürlichen Riboflavinanalog, Roseoflavin (RoF), geschlossen werden. Zusätzlich wurde das Potential von acht künstlichen Flavinderivaten hinsichtlich ihrer Nutzung in Flavin-abhängigen Blaulichtrezeptoren evaluiert, mit dem Ziel, die gewünschten Photoprozesse mit niedrigerer Anregungsenergie induzieren zu können.

Optische Spektroskopie von Flavinderivaten. Im letzten Teil wurden die optischen Spektren von Flavinderivaten untersucht, unter Verwendung von Rechenmethoden zur Simulation von vibrationsverbreiterten UV/Vis Spektren, welche auf dem Franck-Condon (FC) Prinzip beruhen. Hierfür wurde eine neue Methode entwickelt, die das bekannte vertikale FC Schema in krummlinigen Koordinaten implementiert, gefolgt von einem Vergleich mit den üblichen FC-basierten Rechenverfahren in cartesischen Koordinaten. Mit Hilfe der entwickelten Software wurde ein umfassender Benchmark zur Leistungsfähigkeit von (fast) allen möglichen FC-basierten Simulationsmethoden zur Spektrenberechnung erstellt, anhand von drei sorgfältig ausgewählten Flavinen mit bestimmten Struktureigenschaften.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The term “flavin” refers to a group of compounds that contain a ribityl (D-Ribitol) side chain attached to a conjugated three-ring system (7,8-dimethyl-10-alkylisoalloxazine, see **Figure 1.1**). Flavins are omnipresent in nature and take part in many biochemical events as coenzymes (cofactors) and photoreceptors (chromophores). The exceptionally broad range of the biological roles of flavins includes dark-state processes (neurotransmission, cell growth and death regulation, detoxification of soil pollutants, etc.) as well as light-driven processes (photosynthesis, DNA repair via photoreduction of DNA dimers, biological clock regulation, plant phototropism, etc.), which makes them indispensable in daily life (refs. [1–4] and references therein). The first known flavin, riboflavin (RF, 7,8-dimethyl-10-(D-1′-ribityl)isoalloxazine), was the yellow-colored vitamin B_2, an essential component of the daily diet [5]. Its early discovery (in the 1800s) was followed by the continuous identification of a myriad of new flavins every year (see ref. [3] for a brief history). Given the extensive interdisciplinary research on flavins, it is an arduous task to give a complete overview on their properties, and hence only the most relevant aspects will be addressed here. Readers are referred to a recent comprehensive review for further information [6].

RF has no biological activity by itself. However, it is the precursor of naturally occurring flavins, in particular flavin mononucleotide (FMN) and flavin adenine dinucleotide (FAD) (**Figure 1.1**). Flavins are employed in biocatalytic processes as cofactors in flavoproteins due to their unusual redox properties, resulting from their extended π-conjugated system [6]. They can act as both one- and two-electron oxidative agents and may exist in three different oxidation states: oxidized (quinone), one-electron reduced (semiquinone radical), and fully reduced (hydroquinone). Each of these forms can be present as a neutral and anionic species (but not as cation considering the pKa values).

Therefore, flavins can mediate hydrogen transfers of different types (migration of a proton, a hydrogen atom or a hydride anion). They cooperate with reduction partners such as nicotinamide adenine dinucleotide (phosphate), NAD(P), or molecular oxygen to restore the oxidized form [7,8].

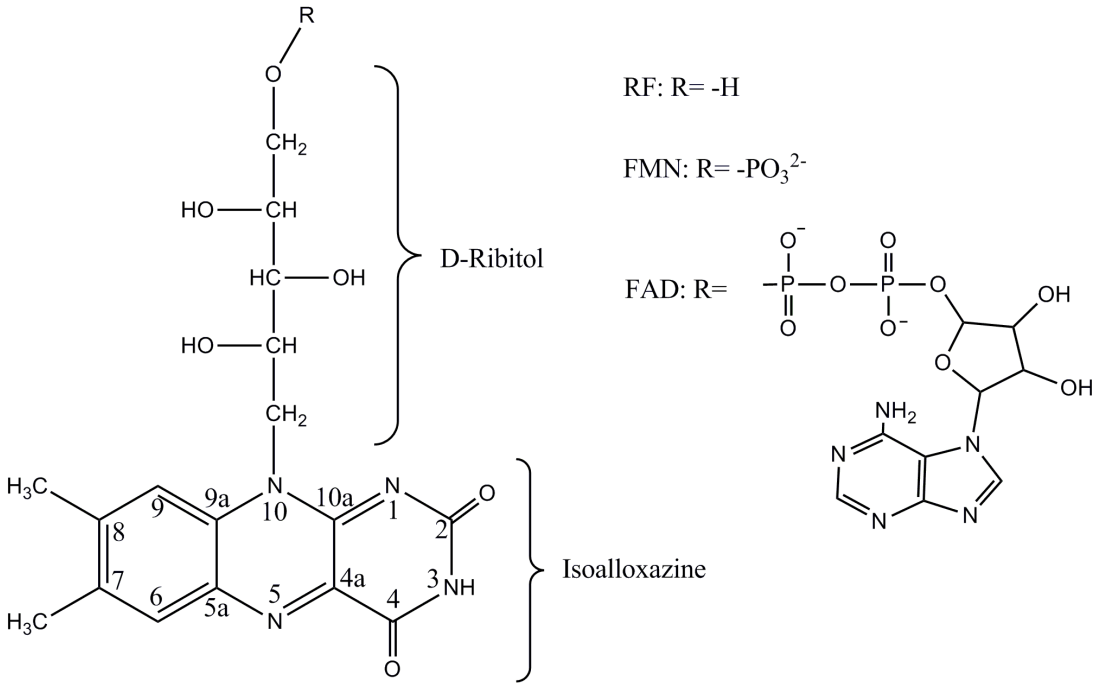


Figure 1.1. The most common natural flavins: Riboflavin (RF); flavin mononucleotide (FMN), and flavin adenine dinucleotide (FAD). Standard numbering of flavins is also given.

Thanks to their unrivaled redox properties, flavins are versatile catalysts. Likewise, flavoproteins – the biocatalysts depending on flavins as coenzymes – can mediate a miraculously large variety of chemical reaction types, although each flavoprotein has a strict catalytic specificity [4]. More than 32 families have been identified for FAD-containing flavoenzymes with regard to their sequence and structural features [9], whereas for other flavoproteins the respective families are based on their numerous common features (for instance, as laid out in reviews [3,4,10–13]). In line with their versatile and essential biochemical roles, flavoproteins have been the common target of pharmacological efforts, mainly for treatments against cancer, infectious diseases (e.g. tuberculosis, malaria, and trypanosomal infections), and neurological disorders (e.g. depression, Parkinson's disesase, Alzheimer's disease, and schizophrenia); see ref. [6] and references therein. In particular, the dysregulation of the histone methylation/demethylation balance has been associated with a

variety of severe cancer types [14], including breast [15], prostate [16], lung [17], and hepatocellular carcinoma [18]. The overexpression of lysine-specific-demethylase 1 (LSD1), the FAD-dependent amine oxidase that mediates histone demethylation [19], has been assigned as the main cause in numerous tumor types [20–22]. Consequently, various inhibitors have been designed to control the aberrant activity of LSD1 [23–27]. In this context, we will investigate the catalytic mechanism of LSD1 in Chapter 2, to provide an improved basis for the design of novel inhibitory molecules.

Due to its highly conjugated nature, the isoalloxazine moiety in flavins can also act as a chromophore. The absorption spectra of RF derivatives (lumiflavin (LF), FMN, and FAD) are generally characterized by three peaks in the ultraviolet-visible (UV/Vis) region with high molar absorptivities, which result from $\pi \rightarrow \pi^*$ excitations [6]. RF derivatives exhibit yellow-green fluorescence. The three oxidation states can be differentiated easily by their unique absorption patterns [28] and characteristic colors [4]: the dominant fully oxidized quinone form is yellow, the two-electron reduced hydroquinone form is colorless, and the single-electron reduced semiquinone radicals are red (anion) or blue (neutral). On account of these spectral properties, flavoproteins utilize FMN and FAD to serve as blue-light photoreceptors. Three common photosensor domains employing flavins have been documented in a wide range of species (archea, fungi, plants, animals, and prokaryotes): photolyase-like domains (in cryptochromes) that bind FAD, light-oxygen-voltage domains (LOV1 and LOV2) that bind FMN, and blue-light sensors using FAD (BLUF) [29]. These photoreceptors have been related to numerous light-responsive fundamental regulatory processes, such as phototropism, solar tracking, induction of flowering, leaf expansion, circadian rhythm regulation, transcriptional regulation, autophosphorylation, and stomatal opening [29–31]. Photochemical pathways for these regulatory processes have been proposed and investigated in many experimental and computational studies [31–36]. These discoveries, altogether, underline the vital photochemical functions of flavins.

Flavins are also capable of reacting with molecular oxygen in the dark (ground state) as well as upon photoexcitation (excited state). In the ground state, flavin adopts a reduced form (FlH^-, $FlH^\bullet$ or FlH_2), enabling it to reduce dioxygen (O_2) to superoxide anion ($O_2^{-\bullet}$), water or hydrogen peroxide (H_2O_2) [37]. Although thermodynamically driven [38], this conversion is not rapid for free flavins in solution, since the kinetically controlled initial single electron transfer (flavin$\rightarrow O_2$), common to all flavoproteins, is slow [38,39]. This electron transfer requires spin inversion of triplet O_2 to the reactive singlet species via

intersystem crossing (ISC) [8,37]. In a flavoprotein environment, the latter can be accelerated by 3-4 orders [39], and the resulting radical pair (semiquinone and superoxide) may follow different routes depending on the type of flavoprotein [40]. Reduction of dioxygen occurs mainly as complementary oxidative half-reaction in the oxidation of various substrates by flavoprotein oxidases (see Chapter 2), whereas the reduction of flavin makes up the reductive half-reaction [41]. On the other hand, the light-driven oxygen activation can be performed by fully oxidized flavins, which operate as photo-sensitizers that absorb visible light and in turn activate naturally inactive dioxygen molecules [29,42]. To this end, flavins require long-lived triplet species formed via spin-forbidden transitions (singlet→triplet ISC) that will react with triplet dioxygen to yield the reactive singlet dioxygen [6]. In this connection, RF and FMN have been shown to have very high quantum yields for singlet oxygen (Φ_Δ = ca. 0.5) in water, contrary to FAD (Φ_Δ = ca. 0.1) [43]. This is likely due to the stacked arrangement of the isoalloxazine ring and the adenine moiety, which enables a photo-induced intramolecular electron transfer that quenches the singlet excited state and the concomitant triplet FAD formation [44]. This converse feature of FAD facilitates its role in DNA-repair mechanisms (incorporated in DNA photolyases) and as stable photoreceptors (*vide supra*) [6].

The flavoprotein environment may significantly alter the properties of a free flavin, e.g. with regard to catalysis, optical spectroscopy, and electron paramagnetic resonance (EPR) spectroscopy [45–51]. For instance, flavoproteins are widely considered to provide a special active-site arrangement so that positive charged basic residues are placed near to the N1 and N5 positions, the main reactive sites of flavins, in order to augment the redox potential of the flavin partner as well as its reactivity towards molecular oxygen [4,9,38,45]. Likewise, bending of the isoalloxazine moiety and covalent binding of flavin by the flavoprotein have been proposed to increase the redox potential of the flavin [10,52,53]. These issues will be addressed in more detail in Chapter 2. Flavins *per se* are thermostable, but photosensitive: for instance, the isoalloxazine moiety of RF can undergo photoreduction in different solvents with the help of the ribityl side chain [54,55]. The flavoprotein environment has been shown to modulate (and improve) the photophysical and photochemical properties of flavins. As a specific example, the higher photostability of FMN and its derivatives (1-deazaFMN and 5-deazaFMN) in the LOV domain of a blue-light flavoprotein photoreceptor (YtvA) has been linked to the intervention of nearby cysteine residues [33,36].

1. Introduction

In addition to the most abundant natural RF derivatives, FMN and FAD, some others with structural modifications on their isoalloxine rings can be produced naturally or synthetically. This further expands the large diversity in the (photo)biochemical functions of flavins. Interestingly, only few natural RF analogs have been identified to date, suggesting that nature has preferred to customize the flavoprotein environment rather than to modify the structure of flavin itself, which would require novel biosynthetic pathways [6]. In contrast to the biologically inactive photolysis products of RF (i.e. LF and lumichrome, LC), other natural RF analogs (such as roseoflavin (RoF), 8-amino-riboflavin, F_0/F_{420}, schizoflavins, molybdopterin, and nekoflavin) have diverse biogenic roles and dissimilar spectral features (see ref. [56] and references therein). Among these natural analogs, we will elaborate on RoF in Chapter 3 and focus on its unusual photophysical properties. A variety of artificial RF analogs can be produced using tailored (bio)chemical synthesis techniques that enable diverse types of substitutions at specific positions of the isoalloxazine moiety (see for example refs. [3,57–59] for reviews); hence, they are more abundant than the natural ones. Flavin derivatives with reactive substituents (e.g. sulfur-containing groups) have long been used as active-site probes in flavoprotein environments [58–60]. Rather recently, an experimental in vivo method has been discovered to efficiently replace the wild-type flavin in a flavoenzyme by (natural or artificial) analogs [61].

In summary, flavin research has many facets that attract scientists from different disciplines. In this introductory overview, we have aimed to give an impression of the abundance and variety of knowledge on flavins, without attempting any complete coverage. This dissertation addresses three different aspects of flavins. It is structured as follows. Chapter 2 concentrates on the oxidative properties of flavins (for the specific case of FAD). We investigate the catalytic mechanisms of the dark-state amine oxidation process mediated by two different flavoprotein amine oxidases using classical molecular dynamics (MD) as well as quantum mechanics (QM) and hybrid QM/molecular mechanics (QM/MM) techniques. In Chapters 3 and 4, we turn our attention to photoactivated (excited-state) processes in flavins. Chapter 3 revolves around the photophysical properties of different flavin derivatives acting as photoreceptors. We analyze in detail the photodriven internal (intramolecular) charge transfer (ICT) in roseoflavin using time-dependent density functional theory (TD-DFT) and the combined DFT/multi-reference configuration interaction (MRCI) methods. Besides, we report on efforts to find riboflavin analogs, which can be used as photoswitches that absorb in the blue-green region and have properties

comparable to wild-type RF. Chapter 4 focuses on computational simulations of optical UV/Vis spectra of three flavin derivatives (riboflavin, roseoflavin, and 5-thioflavin). Here, we give a comprehensive review on simulation techniques based on Franck-Condon (FC) principle, followed by a detailed benchmark of these methods. In addition, we also present a novel technique that implements the vertical Franck-Condon approach in internal coordinates (with curvilinear displacements).

Chapter 2

Oxidative Amine Demethylation Mediated by Flavoprotein Oxidases

Oxidation is a key rate-limiting step in the synthesis of many essential chemicals. In industrial processes, oxidation often requires harsh conditions and is generally hampered by limited chemoselectivity and enantioselectivity. In nature, oxidative enzymes have evolved to perform a variety of oxidation tasks under mild conditions and with high selectivity. They can be classified as oxidases, oxygenases, and dehydrogenases [12]. These three classes differ in terms of how molecular oxygen is used as oxidant in the oxidation process. While oxygenases and hydrogenases require coenzymes as reducing equivalents (e.g., NAD^+, NAD(P)H, and quinones) in combination with molecular oxygen, oxidases only rely on O_2 for oxidation. However, oxidases may also utilize redox cofactors, which can be purely organic (flavins) or contain transition metals (with quinones) [10,12]. In this thesis, we are only interested in the flavin-dependent (or flavoprotein) oxidases, which employ FAD or FMN as cofactors (see **Figure 1.1**). By means of these cofactors, oxidases can activate organic substrates in a two-step oxidative process, where the flavin is reduced partially (to the semiquinone) or completely (to the hydroquinone) in the reductive half-reaction, and is then reoxidized in the oxidative half-reaction with formation of hydrogen peroxide, H_2O_2.

Flavoprotein oxidases are compatible with a wide range of substrates and exhibit different characteristic features [45,62]. They are composed of seven ever-growing subclasses [12]. Among these, amine oxidases have gained much interest in the literature (see [10,12] and references therein), because of the high biological and physiological importance of their target substrates, especially amino acids, the building blocks of epigenetic information. Although most amine oxidases commonly work on the C-N bond of

primary or secondary amine substrates, they can also be active in the oxidation of thioethers, nitroalkanes, and alcohols, as well as in biological processes like antibiotic nikkomycin production (NikD) and post-translational modifications (see below) [13,62]. With the constant addition of new members, flavoprotein amine oxidases are grouped into several families: D-amino acid oxidases (DAAO), monoamine oxidases (MAO), monomeric sarcosine oxidases (MSOX), and alcohol oxidases.

Another vital biological application, in which amine oxidation constitutes the key step, was revealed by the relatively recent discovery of the first histone demethylase, LSD1, in 2004 [19]. LSD1 catalyzes the demethylation of a specific lysine residue on the tail part of a histone complex that wraps the DNA strand and forms the chromatin structure [63]. Via post-translational modifications (PTM) of epigenetic information (chemical alteration of specific amino acids in the histone proteins) [64], the chromatin structure in eukaryotes can be modified without the need for unwinding the DNA strand. As a result, numerous cell-regulatory processes are controlled effectively by PTMs [65–67]. Accordingly, abnormalities in the PTM activities have been linked with many severe cancer types [14,68,69]. Methylation has long been thought to be an irreversible PTM leading to permanent epigenetic marks [70,71]; however, the discovery of LSD1 proved that methylation is indeed reversible. More importantly, it paved the route for the discovery of new flavin-dependent demethylases not only for nucleosamal substrates but also for other N-methylated amino acids and amines; hence, new families of enzymes have been identified [72,73]. Generally, demethylases are regarded as amine oxidases since the main catalyzed process is, by definition, an oxidation (see **Section 2.1**).

Flavoprotein amine demethylases can be divided into two groups on the basis of the binding mode of the flavin cofactor in the binding pocket. Most flavoproteins (ca. 90%) bind FAD or FMN tightly but noncovalently, stabilizing it through nonbonded interactions. On the other hand, some flavoproteins can attach a flavin cofactor covalently, via one specific active-site residue (Cys, Tyr, His or Thr) that is covalently bound to a predefined location at the flavin isoalloxazine moiety (–C8-αCH_2– or –C6–) or the ribityl chain (for more detailed information see the comprehensive review [10]). It has been proposed that the covalent binding of a flavin in an amine oxidase is responsible for improving protein stability by avoiding the loss of a weakly bound oxidized cofactor, and for enhancing the enzymatic activity by boosting the flavin redox potential and aiding substrate binding [10,52].

2. Oxidative Amine Demethylation Mediated by Flavoprotein Oxidases

In this thesis, we are interested in the amine demethylation pathway catalyzed by two different flavoprotein amine oxidases, LSD1 and N-methyltrypthophan oxidase (MTOX), which differ in their structural properties, the binding of their flavin cofactors, and the nature of their genuine substrates (aliphatic vs. aromatic amines).

2.1. Proposed Catalytic Mechanisms

In general, amine demethylation involves the removal of a methyl group from the N-terminus or the side group of a given amine substrate via a redox process. Shi et al. [19] proposed a three-step catalytic mechanism for histone (H3K4met) demethylation that can be generalized for amine demethylation mediated by flavoenzyme oxidases. The first step involves activation/breaking of the αC-H bond in the methyl group on the amine substrate, followed by the oxidation via transfer of a hydride equivalent. Primary deuterium isotope studies indicate that the cleavage of the chemically inert αC-H bond is the rate-limiting step in different amine oxidases [74,75]. Amine oxidation leads to two-electron reduction of flavin (FAD in MTOX), which is then (in the subsequent oxidative half-reaction) re-oxidized by molecular oxygen, with formation of a hydrogen peroxide. The re-oxidation of FAD via oxygen activation is a complementary process with very high rates [76,77] and is commonly believed to occur via a ping-pong mechanism in LSD1 [76] and MTOX [78]. In the remaining two steps of the demethylation process, the iminium intermediate is hydrolyzed, and the resulting carbinol amine spontaneously rearranges to yield formaldehyde and demethylated amine.

The rate-limiting first step of amine demethylation involves the transfer of a hydride equivalent from the amine substrate to flavin (FAD). The possibility of transferring two electrons and a proton in different order gives rise to several possible mechanisms, namely a direct hydride transfer (HT), a radical mechanism via single-electron transfer (SET) [79–81], and an adduct-forming polar nucleophilic (PN) mechanism [82]. These pathways are depicted in **Figure 2.1** for a general substrate. The simplest mechanism is the concerted transfer (HT) of the two electrons and a proton as a hydride anion from the α-carbon atom of the substrate to FAD. In the radical and adduct-forming mechanisms, in contrast, intermediate species are formed.

2.1. Proposed Catalytic Mechanisms

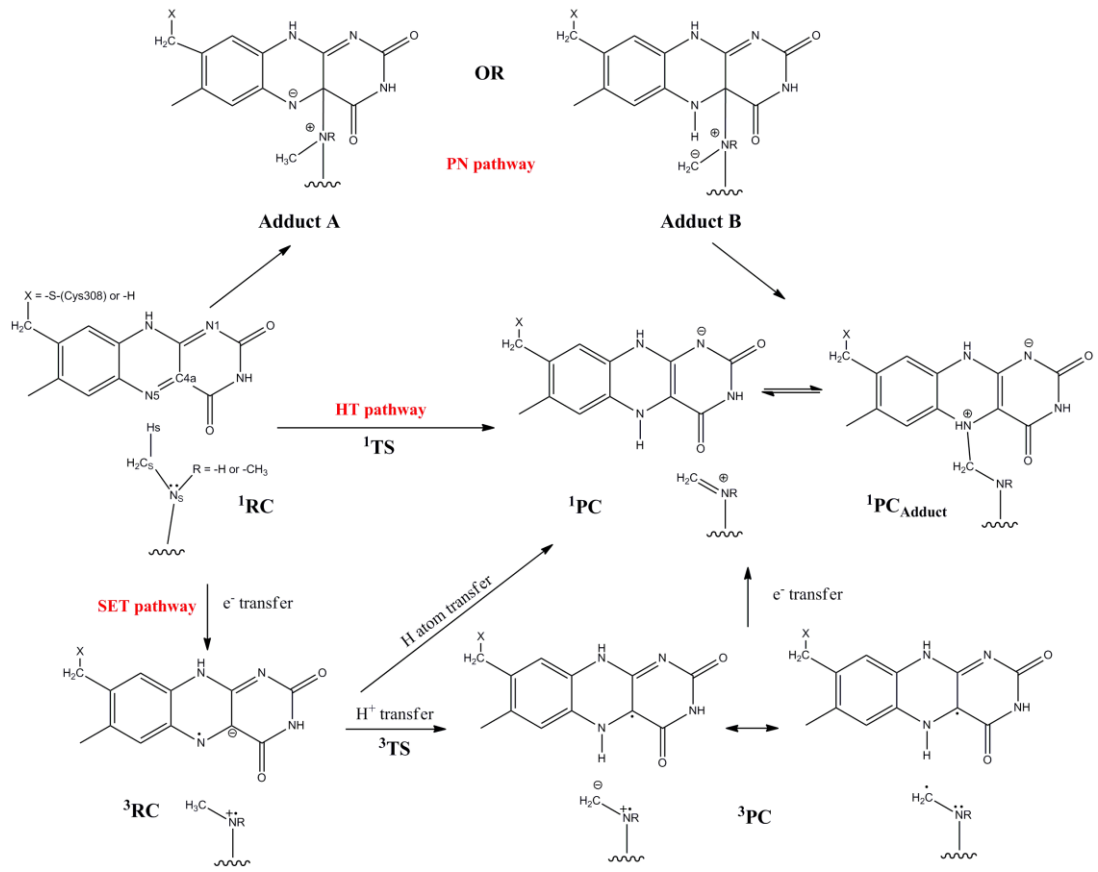


Figure 2.1. General mechanism for amine oxidation mediated by various flavoprotein oxidases. X = H in LSD1, X = S-Cys308 in MTOX ; R = H in MTOX, R = H (monomethyl lysine) or R = CH_3 (dimethyl lysine) in LSD1.

There are numerous studies in the literature that favor different mechanisms for different flavin-dependent amine oxidases (see refs [13,45,62,63,66,73] for reviews). ^{15}N kinetic isotope effect measurements and computations on MTOX did not provide clear support to either the HT or SET mechanism [74,83]. The SET mechanism was supported on the basis of cyclopropyl inhibitor studies on LSD1 [84], but this was later challenged [75] by virtue of the fact that some other flavoprotein oxidases, for which hydride transfer is considered the most likely mechanism, are also inactivated by cyclopropyl inhibitors [85,86]. Besides, the absence of any conceivable intermediate (i.e., a flavin or amine radical species) in kinetic, EPR, ENDOR and other spectroscopy studies on MTOX [87,88] as well as other amine oxidases [75,89,90] was proposed to support the HT mechanism. However, the failure to detect any intermediates was alleged to be inconclusive by others [79,83] based on the following scenario. The initial SET process could be reversible but not rate

limiting, provided that the back-transfer is much faster. This would consequently lead to low concentrations of the short-lived radical intermediates, making them undetectable by experimental techniques. Recently, a direct experimental (EPR and ENDOR) evidence was presented for the formation of a radical pair of flavin semiquinone and tyrosyl radicals in monoamine oxidase A (MAO-A) [91]. Taken together, these conflicting experimental findings on the catalytic mechanism of enzymatic amine oxidation call for a theoretical study to monitor the molecular-level details that are not directly observable in experiments.

2.2. Structural Motifs Conserved Across Flavoprotein Amine Oxidases

Flavoprotein amine oxidases show significant sequence identity, especially with regard to the domain that hosts the active-site cavity, which is located close to the isoalloxazine moiety of the flavin [92]. Likewise, the substrate is aligned in front of the *re*-face of the flavin for proper reactivity [92]. Some other specific structural motifs are also well-conserved across flavin-dependent amine oxidases [45,92]. Firstly, a conserved lysine residue (K661 in LSD1 and K259 in MTOX) forms a hydrogen bond network to the reduction site of the flavin (N5) via a conserved crystal water molecule. This water bridge (Lys-H_2O-N5) motif plays an important role in enzymatic activity, as supported by mutagenesis experiments that show a complete loss of LSD1 activity upon K661A mutation [93]. Likewise, mutation of the corresponding lysine to methionine in maize PAO (K300M) and mammalian PAO (K315M) leads to 1400-fold [94] and 1.8-fold [95] decreases in enzymatic activity, respectively. Based on these experimental findings, different roles for the conserved lysine have been proposed in different amine oxidases. These include participation in catalysis, steric positioning of the flavin ring, and acting as an active-site base by accepting a proton from a protonated substrate prior to oxidation [75,94–96]. The water-bridge motif has also been regarded as the site of oxygen activation, due to the evident loss of oxygen reactivity upon mutation of K259 in MTOX [77].

Another important conserved active site motif in amine oxidases is the aromatic cage, which consists of two or more aromatic residues, shielding the active site from the influx of external solvent molecules and contributing to its hydrophobicity [45]. In LSD1, one of the two aromatic residues is replaced by threonine (Thr810), while one conserved tyrosine (Tyr761) is still present [63]. As in other amine oxidases, the conserved tyrosine in LSD1 is

located next to the substrate binding site on the *re*-face of flavin in the plane orthogonal to the flavin ring [97,98]. In MTOX, by contrast, Tyr54, Tyr249 and His263 seem to form an aromatic cage with an orientation similar to that in LSD1 [99], with Tyr249 being the homologue of Tyr761 in LSD1. The close proximity of the conserved tyrosine residues to the substrate has been interpreted in terms of a steric role for aligning the substrate in front of the isoalloxazine ring, through H-bonding and π-cation interactions [63,100]. In addition, the cage was also linked to the increase in the nucleophilicity of the substrate amine moiety via repulsive interactions based on a mutation study on human MAO-B [101]. Apart from structural aspects, the conserved tyrosine was proposed to serve as the initial electron acceptor in MAO-A [91] as mentioned previously (see **Section 2.1**).

To summarize, there is a debate on the specific roles of the conserved structural motifs in amine oxidases. Computational studies are expected to provide detailed insight into these issues, like in the case of the catalytic mechanism.

2.3. Methods and Computational Details

We investigated the mechanism of amine oxidation in LSD1 and MTOX at the molecular level in two extensive computational studies [102,103]. In both cases, we applied the same computational procedure (with minor variations), which involved the preparation of the simulation system, classical molecular mechanics (MM) MD simulations, and pure QM (QM-only) and QM/MM calculations. The essential features of our computational approach are outlined below. For further details, readers are referred to the main text and the Supporting Information (SI) of the corresponding publications [102,103] (**Appendix A**).

2.3.1 Preparation of the Simulation System

Given that the protein environment plays an important role in the amine oxidation, like in other biochemical processes that require a biocatalyst to proceed, a realistic starting structure is crucial to model the immediate protein surrounding correctly. Therefore, crystal structures of biocatalysts determined by X-ray crystallography are generally used as the starting point in computational studies.

For LSD1, we combined two crystal structures (PDB codes 2V1D [97] and 2H94 [104]): 2V1D contains the LSD1 structure in complex with CoREST, the corepressor that enables LSD1 to bind nucleosomal substrates and enhances its activity, and with the noncovalently bound cofactor FAD. In this crystal structure, due to the vast size of the tail

part of the substrate histone-H3 protein, it was mimicked by a 21-mer, in which the fourth lysine (counterpart of the original H3K4) was mutated into a methionine. Therefore, we mutated this residue back to the original lysine and methylated it to match the reactive form. The missing crystal waters (in 2V1D) were adapted from another LSD1 crystal structure (2H94) with higher resolution (which lacks other parts except FAD). The size of the resulting model system was quite large, so that embedding it into a water droplet would lead to a system too large to be handled efficiently during the simulations. Therefore, the TOWER domain of LSD1 and CoREST were excluded from our model, since these parts are far away from the active site and have only an indirect allosteric effect on the reaction.

For MTOX, there is only a single crystal structure available so far (PDB code 2UZZ [99]), which was used our starting geometry. It contains a covalently bound FAD cofactor, but lacks crystal waters and, more importantly, the substrate N-methyltryptophan (NMT). We thus modeled the binding of this substrate via flexible docking procedures using AutoDock Vina program [105,106], which resulted in a structure with a favorable binding affinity of -7.6 kcal/mol. The predicted substrate binding is in good agreement with the model proposed in the original experimental work [99]. In addition, we manually added the water molecule that forms the Lys-H_2O-N5 motif, which is conserved across the amine oxidase family [77,107].

After preparing the initial structures of LSD1 and MTOX with the corresponding substrate and the flavin cofactor, we solvated the model systems using a water droplet with a radius of 35 Å. Oxidation states of the titratable amino acids were assigned based on the acid dissociation constants (pKa), predicted by PROPKA [108,109] and H++ webserver [110] at the optimum working conditions (in LSD1, pH=8.7 [75] and in MTOX, pH=8.0 [87]). For the active-site residues, the automated placement of hydrogens was checked by inspection to ensure that the resulting nonbonded interactions are most favorable. Thereafter, the model systems were neutralized with Na^+ and Cl^- ions and then subjected to an iterative solvation procedure ("rehydration"), in which the voids in the simulation system are filled with additional water molecules that are relaxed for a short period by some fast minimizations. The mutations considered in both studies were applied on the protein structure using Accelrys DS Visualizer software, and the setup procedure for the mutants was the same as described above for the wild-type enzymes. The overall system built was handled with the CHARMM (v37b1) program [111].

2.3.2 Classical (Molecular Mechanics) MD simulations

Considering the different oxidation states of some important active-site residues (*vide infra*), we created sets of 20-ns canonical (NVT) ensembles of the model systems using CHARMM and a time step of 1 fs. Atoms in the simulation systems were coupled to a thermal bath at constant temperature (300 K). Standard force field parameters were taken from CHARMM22 [112] for the amino acids and from TIP3P [113] for water. Nonstandard parameters were adapted from the standard parameters of related compounds (NMT from Trp) or from other publications (FAD [114], mono- and dimethylated lysines [115]).

2.3.3 QM-only and QM/MM calculations

To sample a given NVT ensemble, we picked five random snapshots along the trajectory and used these snapshots (or slightly modified variants to model transition state species) as starting points for subsequent QM-only and QM/MM calculations. The QM-only calculations were done on a truncated model system, which includes only the reactive subunits, namely the isoalloxazine moiety of FAD, the corresponding substrate, and nearby crystal water molecules (only in LSD1).

QM-only calculations were done using DFT as implemented in the Gaussian09 program suite [116]. For validation purposes (*vide infra*), we employed five different functionals (B3LYP [117] / B3LYP-D2 with Grimme-type dispersion corrections [118], M06-2X [119], LC-ωPBE [120], mPW1K [121] and ωB97xD [122]) in combination with the 6-31G(d) [123,124], TZVP or TZVPP [125] basis sets. Geometry optimizations and the following vibrational analysis were done in either the gas-phase or implicit solvent (Conductor-like Polarizable Continuum Model, CPCM [126], water), without applying any constraints in the optimizations. Intrinsic reaction coordinate (IRC) calculations were used for locating the reactant and product complexes (RC and PC) starting from the transition state (TS) connecting them.

In the combined QM/MM studies, the QM regions were extended by including several active-site LSD1 or MTOX residues that had been deemed important for amine oxidation. Other enzyme residues, solvent molecules, and ions in the model system were placed in the MM part. The QM/MM calculations were performed using the ChemShell program suite [127]. The QM part of the system was treated at the DFT level (B3LYP-D2 with Grimme-type dispersion corrections) using Gaussian09 (in the MTOX study) or TURBOMOLE [128] (in the LSD1 study). The MM calculations were handled by the DL_POLY code [129]

implemented in ChemShell using the aforementioned force field parameters. The QM/MM treatment employed an electrostatic embedding in combination with the charge-shift scheme [130] and the atoms at the QM/MM boundary were treated with the link-atom approach [127].

QM/MM geometry optimizations were carried out with the hybrid delocalized internal coordinates (HDLC) optimizer [131] implemented in ChemShell, and only residues within 15 Å of FAD were active in order to reduce the computational burden and to retain the overall protein structure. During the TS optimizations, the spatial positions of the atoms placed in the core region were optimized using the P-RFO algorithm [132,133], whereas the remaining non-frozen nuclei were treated with the L-BFGS algorithm [134,135]. For flexibility and faster convergence, we included in the core region the whole isoalloxazine moiety of FAD and the substrate under investigation rather than considering only the three atoms involved in hydride-equivalent transfer (as shown in **Figure 2.1**).

The optimized QM/MM structures were subject to numerical force constant calculations in ChemShell to determine the vibrational modes and to characterize the optimized stationary points (i.e. single negative eigenvalue of the corresponding Hessian matrix for TS, none for minima). We performed an IRC-like procedure to locate the reactant and product complexes that are connected through a given TS. This procedure involves taking small steps in both directions along the imaginary mode displacements and using the resulting geometries as starting points for subsequent unconstrained QM/MM optimizations with a small step size. Gibbs free energies (ΔG) and other thermodynamic properties were evaluated using the standard rigid-rotor harmonic-oscillator approximation. Theoretical reaction rates were approximated from the relevant Gibbs free energy barrier $\Delta G^{\ddagger}$ (the free energy of activation) in the usual manner using the Arrhenius equation.

In both QM-only and QM/MM calculations, we treated the ground-state singlet and the lowest triplet states using restricted and unrestricted Kohn-Sham (RKS and UKS) treatments, respectively. Likewise, we also employed UKS calculations for the putative open-shell singlet species to check whether they would yield an open-shell (radical-type) configuration with energy lower than that of the closed-shell configuration. Another point to note here is that we cut the residues included in the QM parts at appropriate sp3-hybridized carbon atoms whenever possible. Specifically, FAD was taken as LF that contains the isoalloxazine ring, with a cut at the C1'-N10 bond (thus excluding the side chain, see **Figure 1.1**). Various lysines considered in both studies were truncated at their C_{β}-Cγ bond (without

the backbone part), whereas Cys308 (that binds FAD covalently in MTOX) was included as a whole (including the backbone).

2.4. Histone H3K4 Oxidation in LSD1 Environment

LSD1 is a member of the monoamine oxidase family, and depends on a noncovalently bound FAD to catalyze the demethylation of methyl and dimethyl (but not trimethyl) lysines at the fourth position of the histone H3 protein tails (H3K4). LSD1 shows two to five orders of magnitude slower turnover rates as compared to other types of flavoprotein amine oxidases [76]; evidently LSD1 has evolved for substrate specificity rather than catalytic efficiency [75]. The high substrate specificity likely arises from the need for direct interaction of the LSD1 active site with at least 21 flanking residues on the tail of H3 protein [97]. Therefore, LSD1 possesses a significantly larger binding pocket than other amine oxidases [98]. LSD1 can also operate on non-histone substrates, p53 [136], DNMT [137], and MYPT1 [138], which allows LSD1 to regulate many vital processes within the cell [139]. Regarding the latter, abnormal activity of LSD1 has been related to many diseases, especially to some cancer types, heart diseases, diabetes, and neurodegenerative disorders, such as Parkinson's disease and Alzheimer's disease [64]. Therefore, the discovery of potential inhibitor molecules to regulate any abnormal demethylation/methylation balance in LSD1 is of current research interest [140]. From a drug discovery perspective, a molecular level understanding of the mechanism of the LSD1-catalyzed demethylation process is essential. To this end, we performed a computational study (ref [102] and **Appendix A**) to elaborate on the mechanistic details of the demethylation of dimethyl H3K4 using the methods described in **Section 2.3**. In the following we only summarize the highlights from this study. Interested readers are referred to the original publication for more detailed analyses.

To validate the choice of density functional, we started with a QM-only benchmark on a small yet carefully chosen set of functionals (see **Table 2.1**). Evidently, different functionals give different relative energies; more notably however, they yield the same qualitative trend in the energetics, rendering the choice of B3LYP-D2 reasonable. Likewise, we checked the effects of extending the basis set at the QM level and enlarging the QM region in the QM/MM calculations. We computed single-point QM(B3LYP-D2/TZVPP) energies at B3LYP-D2/6-31G(d) optimized geometries, and QM(B3LYP-D2/6-31G(d))/MM energies with a larger QM region that included important neighboring LSD1

residues involved in an active-site H-bonding network. As judged by the resulting relative energies (see Tables S4 and S5 in ref. [102]), the use of a larger basis set and QM region only affords a small and rather uniform stabilization (2-3 kcal/mol), leaving the qualitative picture intact. Therefore, for the rest of the study, we adopted the B3LYP-D2/6-31G(d) basis and the standard QM region that includes parts of the FAD isoalloxazine moiety, substrate lysine (sLys), and K661.

Table 2.1. QM-only relative Gibbs free energies (kcal/mol) for the stationary points (as given in **Figure 2.1**) evaluated using different DFT functionals and the 6-31G(d) basis set.[a]

	B3LYP-D2 Gas	B3LYP-D2 Water	M06-2X Gas	LC-ωPBE Gas	mPW1K Gas
Direct Hydride Transfer (HT) (singlet manifold)					
^{1}**RC**	0.00	0.00	0.00	0.00	0.00
^{1}**TS**	31.0	26.7	33.0	32.2	36.4
^{1}**PC**	15.5	12.4	7.20	9.40	4.50
Radical mechanism (SET) (triplet manifold)					
^{3}**RC**	35.1	30.2	35.6	42.5	37.5
^{3}**TS**	37.9	36.5	38.4	46.2	39.9
^{3}**PC**	29.8	26.9	27.7	36.4	25.5

[a] The present B3LYP/6-31G(d) results are in good agreement with those reported previously [83].

According to our QM-only and QM/MM results, the HT pathway is clearly favored over other proposed mechanisms (SET and PN) in the environments considered in our study. In more detail, for the HT pathway, B3LYP-D2/6-31G(d) gives activation energies of 31.0 and 26.7 kcal/mol in the gas phase and water, respectively, lower than those for the SET pathway (37.9 and 36.5 kcal/mol). In the presence of LSD1, the activation barrier for the HT pathway is reduced drastically (15-21 kcal/mol), showing the significance of the catalyst on the hydride-equivalent transfer. This barrier lowering is at least partly due to the favorable n(sLys)→ π*(C4a-N5 of FAD) orbital interactions at the rate-determining HT transition state, in which the lone pair on sLys lies almost orthogonal to the FAD isoalloxazine ring. This transition state is further stabilized by electrostatic interactions with the surrounding protein environment. In view of the well-known tendency of B3LYP to underestimate barrier heights [121,141], the predicted approximate rate (1330 min^{-1}) is in a reasonable agreement with the observed ones (8.10 ± 0.20 min^{-1} [76]).

Table 2.2. QM(B3LYP-D/6-31G(d))/MM relative energies (Gibbs free energies in parentheses) for the stationary points in upward and downward orientation: average values and standard deviations (kcal/mol) over a set of five snapshots. Results for the K661M mutant are also given.

	Deprotonated K661		Protonated K661		K661M
	Upward	**Downward**	**Upward**	**Downward**	
Direct Hydride Transfer (HT) (singlet manifold)					
^{1}RC	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
^{1}TS	23.8 ± 1.93 (20.9 ± 1.14)	18.9 ± 0.95 (15.4 ± 0.98)	25.1 ± 2.05 (21.7 ± 2.12)	23.5 ± 2.97 (20.0 ± 2.75)	22.4 ± 1.0 (19.0 ± 1.1)
^{1}PC	2.05 ± 0.95 (2.42 ± 0.94)	-12.5 ± 3.57 (-12.1 ± 3.53)	13.5 ± 1.54 (13.7 ± 1.50)	15.8 ± 1.49 (15.6 ± 1.38)	1.80 ± 1.6 (1.60 ± 2.5)
Radical mechanism (SET) (triplet manifold)					
^{3}RC	31.5 ± 2.28 (31.1 ± 2.28)	20.9 ± 1.89 (20.4 ± 1.70)	29.5 ± 2.63 (29.2 ± 2.66)	24.6 ± 1.85 (24.2 ± 1.82)	27.2 ± 0.5 (26.9 ± 0.5)
^{3}TS	48.0 ± 1.71 (44.2 ± 1.72)	37.1 ± 1.76 (33.2 ± 1.73)	49.3 ± 2.78 (45.6 ± 2.81)	44.7 ± 3.22 (41.1 ± 2.99)	48.2 ± 0.6 (44.4 ± 0.7)
^{3}PC	27.9 ± 1.38 (27.4 ± 1.37)	24.5 ± 2.27 (23.8 ± 2.36)	35.3 ± 2.88 (35.2 ± 2.85)	37.9 ± 3.14 (37.2 ± 3.37)	32.2 ± 0.6 (31.9 ± 0.7)

On the contrary, the SET rate is only slightly affected by the protein environment (TS energies of 33-44 kcal/mol) so that the SET pathway remains highly unfeasible in the enzyme compared to the HT pathway. Besides, the energetically demanding intermediate radical pair of FAD and sLys (**^{3}RC**, see **Figure 2.1**) is not stabilized by any of the LSD1 active site residues, as is evident from the relative Gibbs free energies (**Table 2.1** and **Table 2.2**) and the spin densities calculated with the larger QM region (see Figure S6 in ref. [102]). Based on the latter finding, we propose that neither Tyr 761 nor any other nearby LSD1 residue can act as the initial single electron acceptor (instead of FAD), as opposed to the stable tyrosyl-FAD radical species observed in the case of MAO-A [91]. Furthermore, the QM-only and QM/MM optimizations could not locate any stable geometries for the adducts that are formulated as intermediates in the PN pathway (see **Adducts A and B** in **Figure 2.1**). This can be attributed to the bulky methyl groups flanking the reactive center on sLys (Ns), which prevent formation of the covalent Ns-C4a bond in these adducts. Hence, our calculations rule out both the SET and the PN mechanism in LSD1.

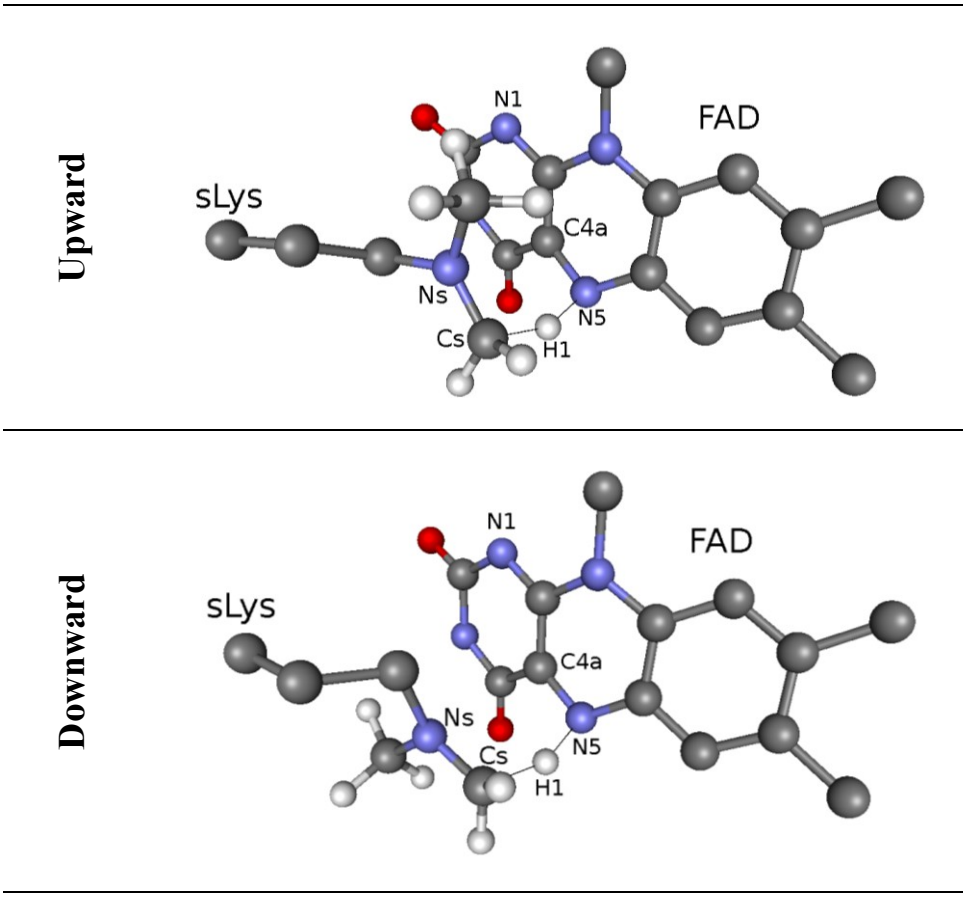


Figure 2.2. Ball-and-stick representations of the upward (top panel) and downward (bottom panel) orientations of sLys. The notation reflects the orientation of the non-reacting methyl group of sLys with respect to the N5-C4a bond of FAD. The structures show parts of the QM region from the QM/MM optimized ^{1}TS geometries, with most of the hydrogen atoms, the deprotonated K661, and the three QM water molecules removed for clarity. Also included is the numbering of the relevant atoms (used in the text). The standard flavin numbering scheme has been adopted along with special atom labels for the substrate lysine (Cs and Ns). H1 is the hydride-equivalent being transferred.

Considering the relevance of the conserved Lys-H_2O-N5 motif and the proposed roles for K661, the building block of this motif in LSD1, we considered two protonation states of K661 (and also the K661M mutant) as well as two protonation states of the substrate lysine (i.e. altogether six separate canonical ensembles). In this manner, we covered all plausible scenarios for substrate binding. The MD simulations revealed two different binding modes of sLys with respect to FAD: "upward" and "downward" (see **Figure 2.2**). The "downward"

orientation is achieved when a protonated sLys is bound in the active site, which is the more likely case as indicated by pKa analysis. In this orientation, a proton shuttle is formed between the substrate (sLys) to K661, via an extended version of the general Lys-H_2O-N5 motif with two extra water molecules (**Figure 2.3**). The "downward" orientation leads to higher HT and SET rates in comparison to the "upward" orientation (**Table 2.2**), which is likely related to the destabilization of the "downward" reactant complex (^{1}**RC**) due to reduced π-stacking interactions between the two subunits (see Figure 6 in ref. [102]). Another remarkable difference between the two orientations concerns the formation of an adduct in the product complex (^{1}**PC**). In the "downward" case, after the hydride-equivalent transfer, the two subunits ($FADH^-$ and sLys-N=CH_2^+) can form an adduct with the Cs-N5 bond (see **Figure 2.4**, left panel), whereas in the "upward" case there is only a weakly interacting complex because of the steric hindrance created by the other non-reacting methyl group (see **Figure 2.4**, right panel),.

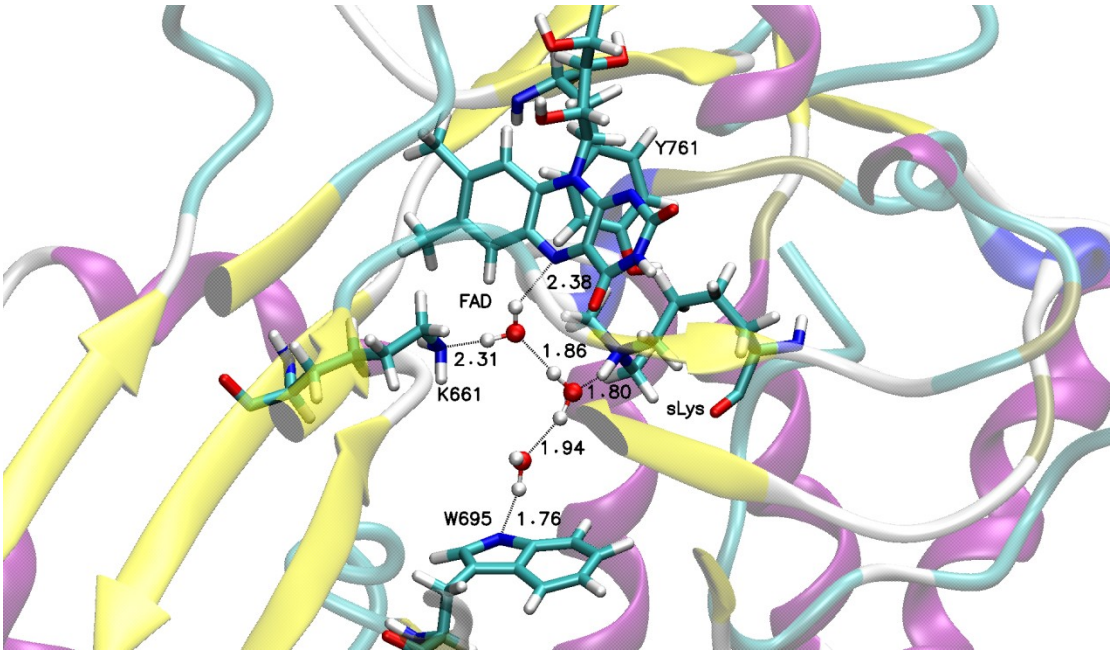


Figure 2.3. A typical snapshot showing the three-water bridge motif in the "downward" orientation. Hydrogen bonds are marked with thin black lines; the corresponding distances are given in Å.

Our study also revealed how several active-site LSD1 residues and the conserved water-bridge motif assist the catalysis in a number of ways. Tyr761 has only a steric effect

in positioning the reaction partners properly, Lys661 acts as an active-site base, and the water bridge is crucial for promoting the proton transfer, with Trp695 shielding and stabilizing this bridge through nonbonded interactions (in line with experimental point mutation studies). According to our calculations, K661 plays a crucial role in the LSD1-mediated demethylation of lysine substrates: K661 is most likely to get deprotonated as sLys (mostly in the protonated form) enters the binding pocket, and K661 is then able to act as the base that accepts a proton from sLys. Thus, K661 helps to 'liberate' the Ns(sLys) lone pair that is required for the subsequent transfer of a hydride-equivalent to FAD; moreover, deprotonated K661 is required to make the overall HT process exothermic by allowing the aforementioned adduct formation in the **^{1}PC** species (**Figure 2.4**). The studies on the K661M mutant provide further insight: the computed barriers for the HT and SET pathways are essentially unaffected by the mutation, indicating that K661 does not have a direct catalytic influence in amine (sLys) oxidation. On the contrary, the active-site base role of K661 is supported by MD simulations showing that K661M mutation disrupts the water-bridge motif and thus prevents deprotonation of sLys-NH_3^+ after binding to LSD1; this will reduce LSD1 activity (as also confirmed experimentally).

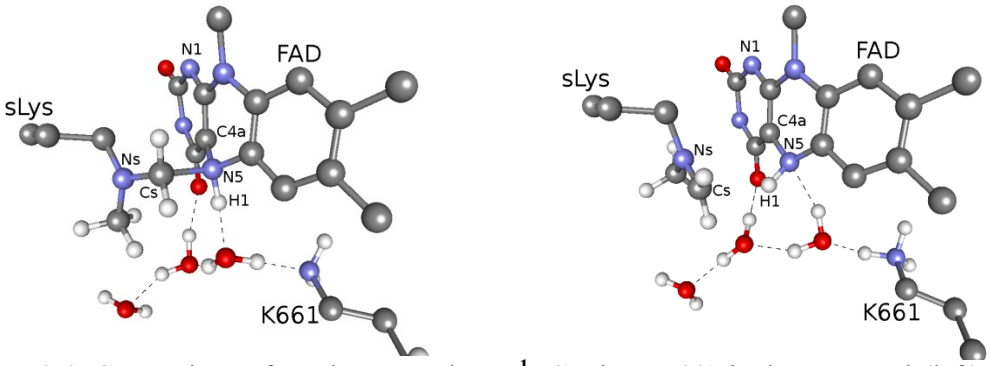


Figure 2.4. Comparison of product complexes **^{1}PC** when K661 is deprotonated (left) and protonated (right). Adduct formation following hydride transfer depends on the protonation state of K661. H-bonds are indicated by black dashed lines. Some hydrogens are not shown for clarity.

2.5. NMT Oxidation in MTOX Environment

After performing the first QM/MM studies on a flavoprotein amine oxidase (LSD1), we wanted to further augment our understanding on amine demethylation by examining an oxidase from another enzyme family that has significant structural and substrate diversity. As opposed to LSD1, MTOX catalyzes the demethylation of NMT, an aromatic substrate,

and covalently binds the flavin cofactor. Hence, we carried out another full QM/MM study (see ref. [103] or **Appendix B** for detailed results and discussions).

MTOX belongs to the family that includes MSOX, heterotetrameric sarcosine oxidase (TSOX), nikD, pipecolare oxidase (PIPOX), and fructosyl amino acid oxidase (FAOD). Although the physiological role of MTOX remains elusive up to now, other members exhibit important physiological roles including processing a soil metabolite (using sarcosine in soil as source of energy and carbon in microorganisms), neuromodulation of the GABA receptor complex, and brain metabolism (see ref [142] and references therein). Among the members of this family, MTOX shares the highest structural homology (43% sequence identity) with MSOX [143] that is, allegedly, the best characterized member of this family [52]. MTOX and MSOX possess highly analogous reactive pockets; they bind a FAD covalently through a conserved cysteine residue (MTOX/MSOX: Cys308/Cys315) [99], forming the [8α-(S-cysteinyl)FAD] complex)).

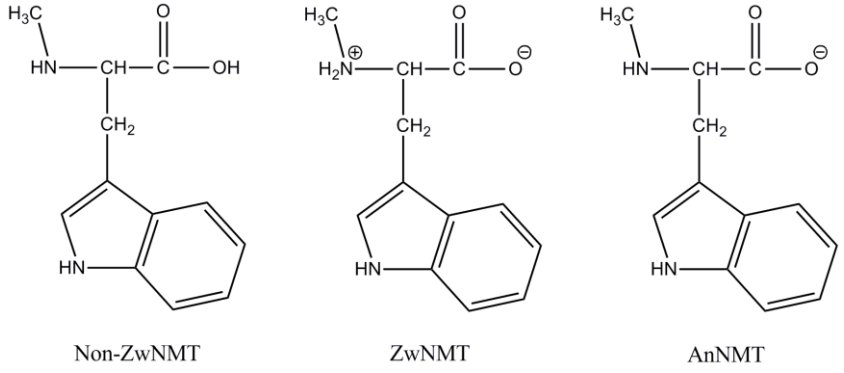


Figure 2.5. Three forms of N-methyltryptophan (NMT): Non-zwitterionic (Non-ZwNMT), zwitterionic (ZwNMT), and anionic NMT (AnNMT).

As an isolated amino acid, NMT can assume three different forms: anionic (AnNMT), neutral/non-zwitterionic (Non-ZwNMT), and neutral/zwitterionic (ZwNMT) (**Figure 2.5**). In view of the pH-dependent activity and absorption spectra of MTOX [74,77] and MSOX [144], the substrate NMT was assumed to bind in the less reactive neutral form (ZwNMT or Non-ZwNMT), which is dominant in solution at pH=8.0. Considering the optimum working conditions of MTOX (pH=8.0) [87], the substrate NMT is likely ionized within the protein environment by an active-site base to yield the more reactive anionic form (AnNMT). In this context, anionic NMT was shown to form a charge transfer (CT) complex with the oxidized flavin acting as acceptor [74,77,144], which is deemed to be stabilized by basic

residue pairs (MSOX/MTOX: R49/R48-K265/K259 and R52/R51-K348/K341 on the *si*- and *re*-faces of flavin, respectively) [52,144]. NMT is thought to be ionized after binding in the reactive pocket and before the rate-determining αC-H cleavage [74]. Within this framework, we considered the three different forms of NMT in our QM-only and QM/MM calculations on the possible pathways (see **Figure 2.6** for the model systems and **Table 2.3** and **Table 2.4** for resulting energetics). The anionic form is computed to be most reactive (having the lowest activation barriers), as previously expected [74,75,95]. The higher reactivity of the anionic form may be linked to the compactness of the two reactive centers in FAD and the substrate NMT (N5 and Cs) as seen in the related NVT ensembles (Table S2 in ref. [103]).

Table 2.3. QM-only Gibbs free relative energies (in kcal/mol) of the stationary points (**Figure 2.1**) for three different forms of NMT, computed at the B3LYP-D2/6-31G(d) level in the gas phase and water

	HT pathway (singlet manifold)			**SET pathway (singlet manifold)**			**SET pathway (triplet manifold)**		
NMT type	**^{1}RC**	**^{1}TS**	**^{1}PC[a]**	**^{3}RC**	**^{3}TS**	**^{3}PC**	**^{3}RC**	**^{3}TS**	**^{3}PC**
Gas Phase									
Non-Zwitterionic	0.0	41.3 (435*i*)	22.1	--[b]	41.0 (784*i*)	--[b]	54.0	60.7 (335*i*)	37.5
Zwitterionic[c]	0.0	--	--	--	--	--	--	--	--
Anionic	0.0	31.0 (1102*i*)	19.1	--[b]	--[b]	--[b]	25.1	36.3 (1090*i*)	36.2
Water									
Non-Zwitterionic	0.0	40.7 (693*i*)	18.7	--[b]	--[b]	--[b]	46.9	52.2 (141*i*)	39.0
Zwitterionic	0.00	58.3 (771*i*)	23.9	--[b]	54.0 (918*i*)	48.8	34.6	60.0 (1995*i*)	46.9
Anionic	0.0	26.5 (1160*i*)	14.7	--[b]	--[b]	--[b]	29.2	31.7 (924*i*)	29.2

[a] An adduct with a Cs-N5 covalent bond is formed for the non-zwitterionic and zwitterionic forms of NMT, but not for the anionic form. [b] Yields the closed-shell configuration. [c] Yields the non-zwitterionic form.

2.5. NMT Oxidation in MTOX Environment

From a mechanistic point of view, our QM-Only and QM/MM results suggest that the HT route (**Figure 2.1**) is the most feasible one for the reactive anionic form, like in the case of LSD1 (**Section 2.4**). Contrary to the LSD1 case, we could obtain stable radical pairs of FAD and NMT for the neutral NMT forms of MTOX (Non-ZwNMT and ZwNMT) with opposite spins of the unpaired electrons (singlet manifold) in the gas phase and in water (QM-only). This is probably due to the aromatic nature of NMT, which supports some unpaired electron density also on the NMT indole group (see Figure S5 in ref. [103]). Interestingly, for the neutral NMT forms, the activation barriers are lower for the SET than the HT pathway (**Table 2.3**). However, for the anionic NMT form, all singlet radical pairs collapse to the closed-shell configuration, leaving the triplet-manifold SET pathway as the only option, which, in any case, has higher barriers and lower rates than the HT route.

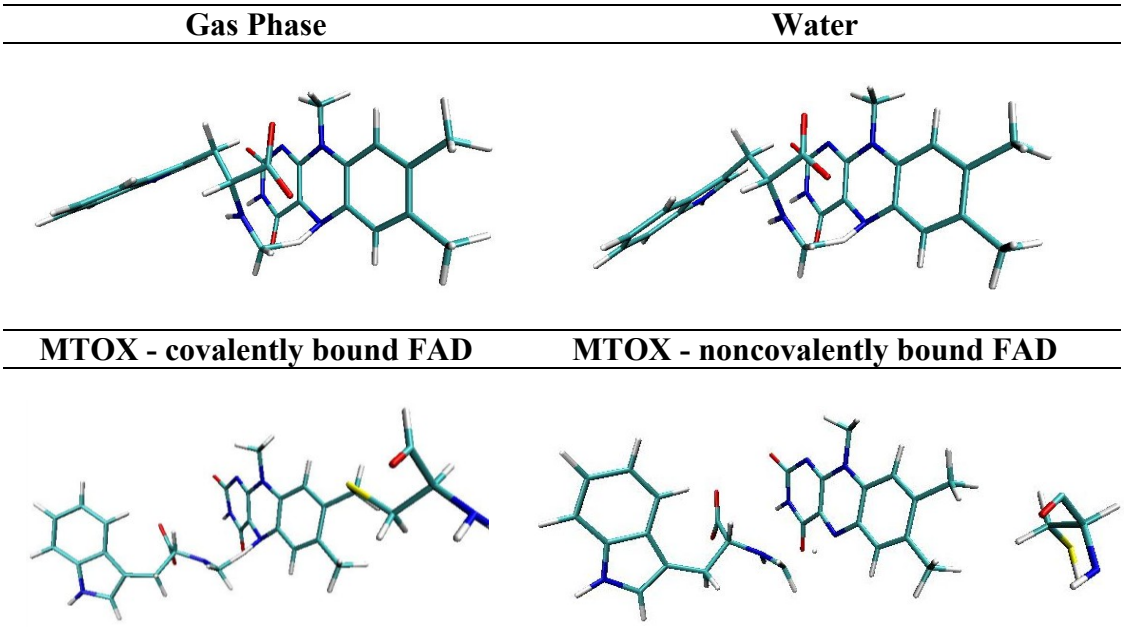


Figure 2.6. Geometries of the transition state for the HT pathway (^{1}**TS**) in different environments. Top: QM(B3LYP-D2/6-31G(d)) structures; bottom: QM/MM structures (QM region only). See **Figure 2.1** for labels and the text for the model system definitions.

Like in the LSD1 case, the MTOX environment expedites the HT rates markedly by reducing the activation barrier (21.3±2.3 kcal/mol, estimated rate of 0.087 min^{-1}), in reasonable agreement with the experimentally observed rates [87,99]. Conversely, as for LSD1, the MTOX environment does not facilitate the formation of the energetically

unfavorable radical intermediates in the triplet manifold (^{3}**RC**). On the HT pathway, the product species (^{1}**PC**) is predicted to form an adduct of NMT and FAD via a covalent Cs-N5 bond, again congruent to the LSD1 case (and also to another amine oxidase, MAO-B [145]). Likewise, the PN pathway can be ruled out due to instability of the corresponding adducts (**Adducts A and B**, **Figure 2.1**), reminiscent of the LSD1 case. Overall, in spite of their structural differences and their dissimilar substrates, LSD1 and MTOX thus show analogous mechanistic features in amine oxidation at the molecular level.

Table 2.4. QM(B3LYP-D2/6-31G(d))/MM relative energies (Gibbs relative free energies in parentheses) for the stationary points (**Figure 2.1**) of the HT and SET mechanisms for different NMT forms with noncovalently or covalently bound FAD: average values and standard deviations (kcal/mol) over a set of five snapshots. Results for the K341Q and H263N mutants are also given.

		HT pathway (singlet manifold)			**SET pathway (triplet manifold)**	
NMT type	^{1}**RC**	^{1}**TS**	^{1}**PC**	^{3}**RC**	^{3}**TS**	^{3}**PC**
covalently bound FAD						
Non-Zwitterionic	0.00 (0.00)	41.3±2.54 (38.8±2.57)	18.5±4.14 (20.8±4.24)	44.2±0.77 (41.7±0.82)	55.8±2.24 (51.0±2.15)	36.0±1.15 (34.0±1.36)
Zwitterionic	0.00 (0.00)	65.6±2.11 (63.3±2.32)	37.7±3.65 (39.2±3.84)	39.2±4.66 (37.5±4.41)	75.5±1.03 (72.4±2.15)	46.7±3.12 (43.4±3.10)
Anionic (Wild type)	0.00 (0.00)	24.1±2.46 (21.3±2.27)	2.78±2.11 (5.04±1.89)	27.7±4.56 (26.0±4.00)	37.9±2.80 (34.1±2.82)	30.8±4.57 (29.3±4.66)
Anionic (K341Q)	0.00 (0.00)	30.7±0.30 (28.4±0.33)	15.1±2.98 (16.0±2.70)	26.0±3.12 (24.2±3.10)	43.5±1.99 (40.0±2.09)	39.2±2.78 (37.8±2.84)
Anionic (H263N)	0.00 (0.00)	36.4±5.73 (33.5±5.51)	5.56±8.84 (5.42±8.10)	21.4±6.12 (19.3±6.41)	49.1±6.43 (45.1±6.09)	34.1±4.31 (32.2±3.67)
noncovalently bound FAD						
Non-Zwitterionic	0.00 (0.00)	34.8±2.41 (32.9±2.50)	6.06±1.85 (9.40±1.81)	38.4±2.47 (35.9±2.73)	46.3±4.9 (45.4±3.6)	28.3±6.98 (26.5±6.9)
Anionic	0.00 (0.00)	25.9±3.09 (23.3±3.32)	2.71±1.28 (5.56±1.46)	25.7±2.74 (23.5±2.58)	36.2±2.95 (32.0±2.92)	26.7±2.59 (24.5±2.72)

Two active-site residues in MSOX (Tyr317 and His269) were considered for the active-site base role, but were previously ruled out through mutagenesis studies [100,146]. In order to remove the extra proton out of the active pocket, a putative proton relay system was proposed for MSOX, consisting of FAD-N5, conserved Lys265 (homologous to K259 in MTOX), and several nearby water molecules [90,100], but the 60-fold loss in the rate of

the reductive half-reaction in the K259Q mutant of MTOX did not support this idea. On the contrary, the 2500-fold decrease in the reaction rate of fully-reduced flavin with oxygen observed in this mutant suggested that K259 serves as the main oxygen activation site [77]. Considering the strong basic properties and close proximity to FAD-N5 and the backbone of NMT, residues K341/K348 (in MTOX/MSOX) are promising candidates for the active-site base role, in line with the almost complete loss of amine oxidation activity in corresponding mutants of MTOX (K341Q, K341R, and K341L) [99]. To clarify this issue, we performed detailed computational analyses, including theoretical pKa analysis, various MD simulations with different protonation states of important active-site residues (K259 and K341), and subsequent QM/MM calculations with wild-type MTOX and its mutants H263N and K341Q (**Table 2.4**). Interpreting these results, we came up with the following conclusions. Tyr249 (homologous to Tyr317 in MSOX) was eliminated as a base given its location (as observed in MD simulations) and its protonation state (based on theoretical pKa analysis), in accord with the only partial loss of oxidative activity upon mutation of its homologue in MSOX (Y317F) [100]. The MD simulations and QM/MM results suggest a clear steric role for His263 in orienting the substrate in a reactive position via π-stacking interactions. His263 should play a more eminent role in MTOX than in MSOX, considering the only 2-fold lower rates upon mutation of its homologue in MSOX [146]. This may be linked to the fact that MSOX genuinely works on a smaller aliphatic substrate. By contrast, K341 likely has both electrostatic and active-site base roles, as indicated by our current computational results and the almost complete loss of MTOX activity (95% or 250-fold) upon K341Q mutation, which is apparently not due to the 10%-loss in FAD incorporation [99].

The covalent flavin linkage in MSOX has been related to the prevention of premature dissociation of an oxidized FAD from MSOX, prior to amine oxidation [52]. Along with the highly electropositive character of the immediate protein environment (due to highly basic residues), covalent binding of FAD has been proposed to lead to a sizeable enhancement of the reduction potential of FAD (estimated to ca. 120 mV) as compared to the noncovalently bound counterparts [10,52,53]. To probe the molecular basis for the catalytic role of covalent binding, we created an artificial theoretical model, in which FAD is noncovalently bound to MTOX. To this end, the Cys308-S and FAD-8α-methyl ends were capped with hydrogens, whereas the rest of the system was kept intact (see **Figure 2.6** for a visualization of the QM region). MD simulations for this setup showed that a reactive and stable binding of an anionic NMT is unlikely to occur in the case of noncovalent FAD binding. This is

expected to result from a more planar arrangement of the isoalloxazine ring with noncovalent FAD binding (42-43^{o} vs. 61-68^{o} for covalent FAD binding), as seen in the corresponding MD simulations; see Table S9 in ref. [103] for specific dihedrals used in this comparison and their values in the corresponding canonical ensembles. Besides, flavin nonplanarity was proposed to be the reason of somewhat enhanced flavin redox potentials in some amine oxidases with a covalently bound flavin [62,147]. In this regard, our QM/MM results indicate a minor increase in the activation barriers for the HT pathway (by 2.0 kcal/mol) with noncovalent FAD binding that is in the same ballpark as the observed changes in the redox potential of ca. 120 mV (≈ 2.8 kcal/mol, assuming a one-electron transfer).

2.5. NMT Oxidation in MTOX Environment

Chapter 3

Photophysics of Flavin Derivatives

After addressing the catalytic (redox) properties of some flavins and flavoproteins, we will concentrate in this chapter on the photophysical properties of selected flavin derivatives. As mentioned earlier, RF analogs can either occur naturally or be produced artificially using elaborate (bio)chemical synthesis techniques [6]. Given their different spectral, chemical, and catalytic properties compared with natural flavins, flavin derivatives have been used as active-site probes since the 1960s [3,58,59,61].

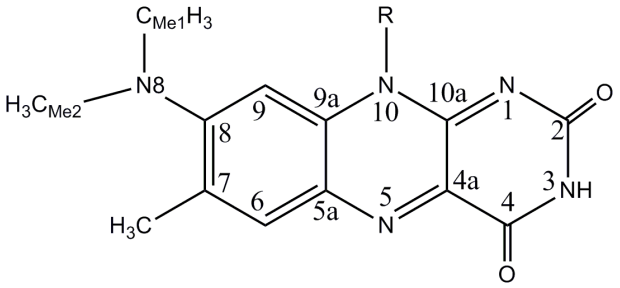


Figure 3.1. Chemical structure of roseoflavin (RoF). R is the ribityl chain and differs in RoF derivatives; e.g., for roseolumiflavin (RoLF), R = methyl; for RoFMN, R = ribityl-5'-phosphate; and for RoFAD, R = ribityl-(9-adenosyl)-pyrophosphate.

Among the few recognized natural RF analogs (*vide supra*), roseoflavin (RoF, 8-dimethylamino-8-demethyl-D-riboflavin) [148,149] has a special place due to its special biological functionalities and its intriguing photophysical properties [55,150,151]. RoF differs from RF in that a dimethylamino (DMA) group replaces the methyl group at the C8 position of the isoalloxazine ring, with the ribityl-phosphate side chain remaining intact (see **Figure 3.1**). As opposed to yellow RF, the aqueous solution of RoF is rose red, which gives RoF its name [150]. RoF is an antagonist to RF and acts as antibiotic in Gram-positive

[152,153] as well as Gram-negative bacteria (only in the presence of a flavin transporter) [154].

The optical properties of riboflavin and roseoflavin in various solvents and protein environments were subject to many theoretical [150,155–164] and experimental studies [51,54,55,61,148,150,151,164–172]. A direct comparison of the absorption and emission properties of isolated RF and RoF in a variety of solvents can be found in ref [55]. The first absorption band of RoF in water is bathochromically shifted ($\lambda_{max,RoF}$ = 500nm, $\lambda_{max,RF}$ = 445 nm) [150,168], and more intense compared to RF [55]. The second band in the adsorption spectrum of RoF is weak and hypsochromically shifted (320 nm) as compared to that of RF, a strong peak at 380 nm [150,168]. Contrary to RF and some other flavin derivatives, the second-lowest absorption band of RoF appears to be hidden under the broad first band [150]. Finally, the third adsorption band of RoF is very strong and seen in the UV region at 257 nm [150] with a bathochromic shift compared to RF [168].

In aqueous solution at pH 7, RoF has a very low fluorescence quantum yield (Φ_F = 3.2 x 10^{-4}) [150], while RF displays very strong fluorescence (Φ_F = 0.26) [168]; both maxima are located around 540 nm. The dual fluorescence character of RoF (i.e., fast and slow components in the spectrum) is deemed to arise from its ability for ICT, in contrast to RF, as also apparent from the fluorescence quenching [55]. Accordingly, the extent of charge transfer in RoF is mainly determined by the polarity of the solvent, into which RoF is placed [55,171]. Solutions of RoF are red in polar solvents indicating high levels of ICT, and yellow to orange with a strong green fluorescence in non-polar solvents that tend to suppress the ICT [55,151,171,173]. Besides, compared with RF, RoF displays interesting photodynamics as it undergoes structural changes upon photoexcitation that mainly involve torsion of its DMA group. The DMA twist in RoF has been related to the ICT phenomenon [51,55]. We will address the ICT process in RoF in various media and elaborate on its mechanism at the molecular level in **Section 3.2**.

One of the eminent aims in screening novel chemically modified flavin derivatives is to identify artificial flavin chromophores with biocatalytic, spectral, and photochemical features that are altered to achieve specific purposes. Hitherto, many artificial flavin derivatives have been designed and analyzed in depth (for instance see refs [57–59,174,175]). In particular, flavin derivatives that share the same (photo)catalytic features as their native counterpart, except for having a red-shifted absorption, are of high importance. Using these derivatives, it may be possible to perform the associated vital cellular processes

at lower energies. As a specific example, we are interested in the photochemical processes induced by blue-light photoreceptors (LOV, BLUF and cryptochromes, *vide supra*) that employ FMN or FAD as cofactor. With the aim of replacing the native cofactors to minimize the radiation damage to the biological surrounding in the LOV pocket, we investigated the photophysical properties of a carefully chosen set of eight LF derivatives with chemically modified isoalloxazine rings (**Figure 3.2**).

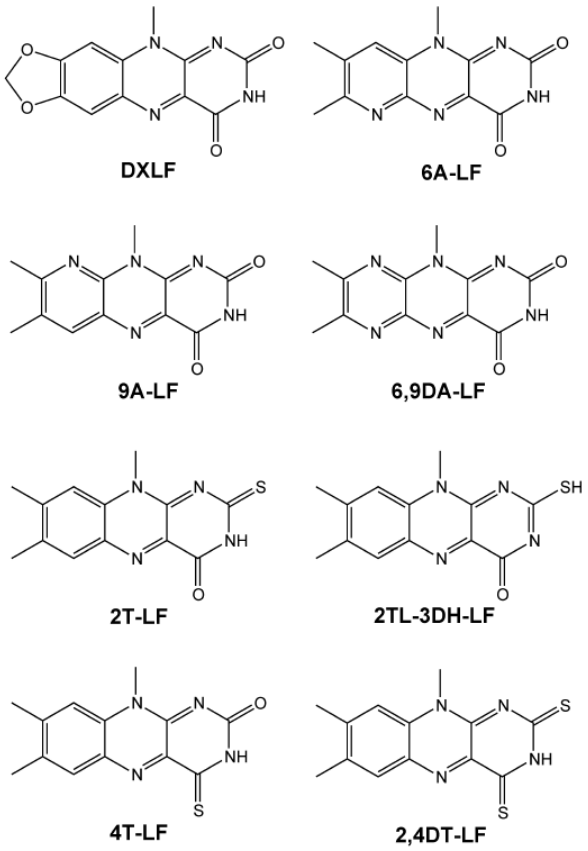


Figure 3.2. Chemical structures of the artificial LF derivatives considered in the analysis.

The modifications were designed to create π-conjugated systems by replacing original with isovalent atoms or by extending the side groups. We also accounted for the electrostatic properties and spatial limitations in the LOV pocket. To be specific, in DXLF, the methyl groups attached to C7 and C8 atoms were removed to form a dioxalane ring. In 6A-LF, 9A-LF and 6,9DA-LF, -CH groups at specific positions (as indicated in the abbreviations) were replaced with N atoms (aza substitution). In 2T-LF, 4T-LF and 2,4DT-LF, the specified carbonyl groups were substituted with thiocarbonyl groups. Finally, an

iminothiol tautomer of 2T-LF (2TL-3DH-LF) was also included in the analysis. The results of this study will be outlined in **Section 3.3**.

3.1. Methods and Computational Details

In our computational investigation of the photophysical properties of natural RoF and other selected artificial flavin derivatives (**Figure 3.1** and **Figure 3.2**), see refs. [176,177] or **Appendices C and D**, we employed very similar computational procedures, which are summarized in the following.

As mentioned earlier, the isoalloxazine moiety is the chromophore of flavin analogs and derivatives. It governs flavin photophysics and photochemistry [6,33,36,161]. While the ribityl side chain may interact with the N1 atom [178] in solution and may induce an intramolecular photoreduction of the isoalloxazine moiety under UV-vis irradiation [179], it has generally only a very minor effect on the photophysical properties of flavins, as clearly evident from the very similar experimental absorption and emission spectra of RF, FMN, and LF [180,181]. In view of this situation, and to limit the computational burden, we only considered truncated models of the flavin derivatives of interest, where the N10 is capped with a methyl group (i.e., LF analogs).

We performed geometry optimizations (and single-point calculations) with Gaussian09 [116] and Turbomole 6.3 [128] using (unrestricted) DFT for the ground and lowest triplet states, as well as linear-response time-dependent DFT (TD-DFT) [182–184] for the low-lying singlet and triplet electronically excited states. Two density functionals (B3LYP [117] and CAM-B3LYP [185]) were applied in combination with several basis sets (SVP [186], TZVP [125] and 6-31G(d) [187]). During geometry optimization, no symmetry constraints were imposed in the RoLF case not to restrain the free rotation of the DMA group, whereas the C_s point group was adopted for the set of eight artificial flavin derivatives. To check whether the optimized geometries correspond to real minima on the corresponding potential energy surface (PES), we performed analytical and numerical vibrational mode analysis using various tools implemented in Gaussian09, TURBOMOLE, or SNF [188].

We accounted for the solvent effects using an implicit continuum solvent model: CPCM [126,189] as implemented in Gaussian09 or the conductor-like screening model (COSMO) [190] in TURBOMOLE. The dielectric constants (relative permitivities) were ε_r = 78 (water), ε_r = 2.27 (benzene), ε_r = 33 (acetonitrile) and ε_r = 4 (to imitate a typical protein environment [191,192]). Implicit solvent treatments cannot properly describe solute-solvent

hydrogen bonding. To capture such specific interactions, an alternative approach is to treat nearby solvent molecules explicitly, as done in previous studies on flavin derivatives [160,161,163,193]. Therefore, we considered microsolvation (or, more precisely, microhydration) in combination with a continuum treatment of bulk solvent, with the isoalloxazine moiety being surrounded by four explicit water molecules (see **Figure 3.3** for the case of RoLF). This way, one can arrive at a more realistic description of hydrogen bonding around the hydrophilic part of a given flavin derivative in aqueous solution.

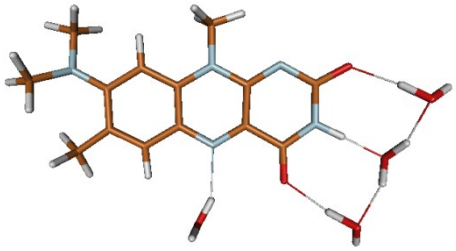


Figure 3.3. Microhydration model for RoLF with four water molecules. Hydrogen bonds are depicted with thin lines.

Single-reference DFT methods treat dynamic electron correlation efficiently [194], whereas MRCI provides a more accurate description of static correlation effects. Hence, multi-reference methods are well suited to describe different types of excited states, especially CT states and states with double-excitation character. The combination of DFT with the multi-reference CI method (DFT/MRCI) [195] benefits from the strengths of both methods. DFT/MRCI was shown to perform well in the computation of vertical excitation energies of flavins (typical errors of 0.2 eV) [163] and also of other systems, as evident from the benchmarks for a large set of valence excited states of organic molecules with ππ* and nπ* character [194]. With this in mind, we performed single-point calculations with DFT/MRCI to determine vertical excitation energies, oscillator strengths, and dipole moments of the electronically excited states at the DFT-optimized geometries. The molecular orbitals (MOs) for the subsequent DFT/MRCI calculations were generated at the DFT(BHLYP/TZVP) [196,197] level. DFT/MRCI calculations were done without symmetry constraints (C_1 point group) using the standard parameters [195] and the 20 lowest roots were computed. When required, we considered solvent effects implicitly in DFT/MRCI calculations by using COSMO while generating the KS MOs.

3.1. Methods and Computational Details

3.1.1 Calculation of the Intersystem Crossing Rates

Electronic spin-orbit matrix elements (SOME) between correlated DFT/MRCI wave functions were computed with the spin-orbit coupling kit (SPOCK) [198,199]. SPOCK uses the one-center mean-field approximation to the (time-independent) Breit-Pauli Hamiltonian [200] that has been shown to yield results with a small error margin (<5%, as compared to the full treatment) and notably faster [201,202]. The resulting SOME values were converted to intersystem crossing rates (k_{ISC}) between correlated singlet and triplet electronic states of interest using the formalism outlined in ref [203]. The k_{ISC} values were calculated based on time-dependent perturbation theory (Fermi Golden Rule), by summing over individual rates of transition from the initial vibronic level (Ψ_{S_a}, ν_{aj}) to each target vibronic level ($\Psi^{\alpha}_{T_b}, \{\nu_{bk}\}$):

$$k_{\mathrm{ISC}} = \frac{2\pi}{\hbar}\sum_a \sum_k \left|\langle \Psi_{S_a}, \nu_{aj} \left|\widehat{\mathcal{H}}_{\mathrm{SO}}\right| \Psi^{\alpha}_{T_b}, \nu_{bk}\rangle\right|^2 \delta(E_{aj} - E_{bk}) \tag{3.1}$$

Eq. 3.1 can be expanded in a Taylor series using normal coordinates as variables $\{q_k\}$ and assuming a reference point of choice ($\mathbf{q_0}$), which is taken as the equilibrium geometry of the initial electronic state within the Condon approximation. This transforms Eq. 3.1 into

$$k^{\mathrm{FC}}_{\mathrm{ISC}} \approx \frac{2\pi}{\hbar}\sum_a \left|\langle \Psi_{S_a} \left|\widehat{\mathcal{H}}_{\mathrm{SO}}\right| \Psi^{\alpha}_{T_b}\rangle\right|^2_{q_0} \sum_k \left|\langle \nu_{aj} \left|\nu_{bk}\right\rangle\right|^2 \delta(E_{aj} - E_{bk}) \tag{3.2}$$

Within the Condon approximation, the coupling between two given states can be separated into electronic and vibrational parts. The Condon approximation holds well for El-Sayed allowed transitions involving different orbital types [204]. The electronic part (i.e. the SOMEs $\left|\langle i\left|\widehat{\mathcal{H}}_{\mathrm{SO}}\right|f\rangle\right|$) are taken from SPOCK, whereas the vibrational contributions are computed using a Fourier transform of the time-dependent delta function:

$$\delta\left(E_{aj} - E_{bk}\right) = \int_{-\infty}^{\infty} \mathrm{e}^{it\left(E_{aj} - E_{bk}\right)} \mathrm{d}t \tag{3.3}$$

where the delta function is integrated numerically. Presuming the harmonic and Condon approximations, the initial equation for ISC rates (Eq. 3.1) can be reformulated:

$$k^{\mathrm{corr}}_{\mathrm{ISC}} = \left|\langle \Psi_S \left|\widehat{\mathcal{H}}_{\mathrm{SO}}\right| \Psi_T\rangle\right|^2 \int_{-\infty}^{+\infty} \mathrm{d}t G(t) \mathrm{e}^{it\left(\Delta E^0_{\mathrm{ST}} + \frac{1}{2}Tr\Omega_S\right)} \tag{3.4}$$

where G(t) is a generating function [203] adapted from Mehler's formula for the density matrix of a harmonic oscillator [205].

$$G(t) = 2^{N/2} \sqrt{\frac{\det(\mathrm{S}^{-1}\Omega_{\mathrm{S}}\Omega_{\mathrm{T}})}{\det(J^{\dagger}\Omega_{\mathrm{T}}BJ + \Omega_{\mathrm{S}})\det(J^{\dagger}\Omega_{\mathrm{T}}B^{-1}J + \Omega_{\mathrm{S}})}}$$
$$\exp(D^{\dagger}\left(\Omega_{\mathrm{T}}BJ\left(J^{\dagger}\Omega_{\mathrm{T}}BJ + \Omega_{\mathrm{S}}\right)^{-1}J^{\dagger}\Omega_{\mathrm{T}}B - \Omega_{\mathrm{T}}B\right)D) \tag{3.5}$$

Here Ω_S, Ω_T, S, and B are diagonal matrices with elements, $(\Omega_S)_{ii}=\omega_{Si}$, $(\Omega_T)_{ii}=\omega_{Ti}$, $S_{ii}=\sinh(i\omega_{Ti}t)$, $B_{ii}=\tanh(i\omega_{Ti}t/2)$, and the superscript † indicates the transpose of a given matrix; ω_{Si} and ω_{Ti} are the harmonic vibrational frequencies, and J denotes the Duschinsky transformation matrix [206], i.e., the rotation matrix that relates the mass-weighted normal modes of the initial and final states (Q_s and Q_T) as follows:

$$Q_T = JQ_S + D \tag{3.6}$$

with D being the adimensional displacement (shift) vector.

In addition, we used the multireference spin-orbit configuration interaction program (DFT/MRSOCI) [207] to compute second-order spin-forbidden properties, e.g. phosphorescence rates ($T_1 \rightarrow S_0$), which requires the four lowest roots of the MRSOCI matrix.

3.2. Photoinduced Intramolecular Charge Transfer in Roseoflavin

Donor-acceptor (D-A, or equivalently push-pull) chromophores contain an electron donor and an acceptor moiety that are interconnected by a single bond, like in aliphatic amines or aminobenzonitriles [208]. ICT is a common phenomenon following photoexcitation in D-A molecules, which involves the transfer of charge density from the donor to the acceptor, leading to the formation of a zwitterionic photoadduct and a structural rearrangement. Three different theoretical models have been proposed for the ICT mechanism (and the resulting zwitterionic form) in a typical D-A chromophore. These models differ in the relative orientation of the planes of the donor and acceptor groups: (a) planar intramolecular charge transfer (PICT) yielding the quinoid form [209], (b) twisted intramolecular charge transfer (TICT) leading to the perpendicular-twisted species [210], and (c) wagging intramolecular charge transfer (WICT) generating the zwitterion with a pyramidal donor group [211]. The planes of the donor and the acceptor moieties are (close to) parallel in PICT, and (almost) perpendicular to each other in TICT, while in WICT the plane of the donor adopts a pyramidal-like orientation by bending out of the plane of the acceptor group. So far, there has been a debate on whether (and if so, which) one of the three different mechanisms is generally able to explain the ICT process in typical D-A molecules. A broad review on D-A molecules is given in ref. [208].

3.2. Photoinduced Intramolecular Charge Transfer in Roseoflavin

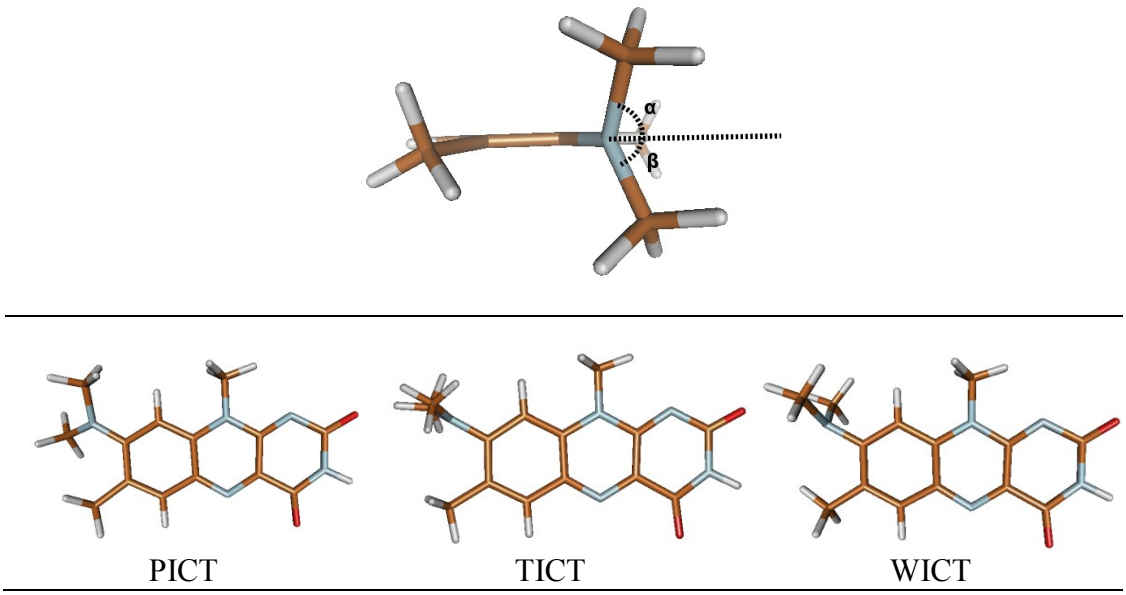


Figure 3.4. (Top) Definition of the dihedral angles (α=C9-C8-N8-C_{Me1} and β=C9-C8-N8-C_{Me2}) that describe the torsion of the DMA group with respect to the isoalloxazine plane (see text for details). (Bottom) planar (PICT), perpendicular-twisted (TICT), and wagging (WICT) forms of the zwitterionic photoproducts of RoLF.

RoF displays the characteristics of a D-A molecule, with the DMA group acting as the donor and the isoalloxazine ring as the acceptor. Hence, an ICT process in RoF may be expected to lead to one (or more) of the three zwitterionic RoF photoproducts (**Figure 3.4**). In previous experimental work on RoF, only PICT and TICT were anticipated, without any clear preference [55]. In a later low-temperature fluorescence decay study on RoF [51], the TICT was favored over the PICT mechanism. To probe the ICT mechanism in RoF and to provide insights into the photoprocesses at the molecular level, we carried out an investigation at the (TD)-DFT (B3LYP and CAM-B3LYP) and DFT/MRCI levels of theory using a truncated model (roseolumiflavin, RoLF) [176] (**Appendix C**). The key findings from this study will be covered below.

A proper description of the orientation of the donor DMA group with respect to the isoalloxazine moiety is important, as DMA torsion is the key structural feature that distinguishes the three ICT mechanisms (and the associated zwitterions). In this regard, we choose two dihedrals (α/β, see **Figure 3.4** for definitions) and use them to describe the degree of DMA torsion throughout the analysis. We find typical dihedral pairs (α/β) that characterize the optimized geometries of the PICT (ca. 21°/-115°), WICT (ca. 65°/-65°), and

TICT (ca. 87°/-87°) zwitterionic forms (with some slight variations depending on the choice of functional and environment). Vibrational mode analysis indicates that each of the optimized WICT geometries corresponds to a saddle point on the corresponding surface. Relaxed PES scan of the LE and CT states along α and β (see PES contour plots, Figure 7 in ref. [176]) confirm that the WICT geometry is generally a transition state connecting PICT and TICT species, so that the WICT zwitterion cannot be regarded as a stable photoadduct. These findings provide strong evidence against WICT being a feasible ICT mechanism in RoLF.

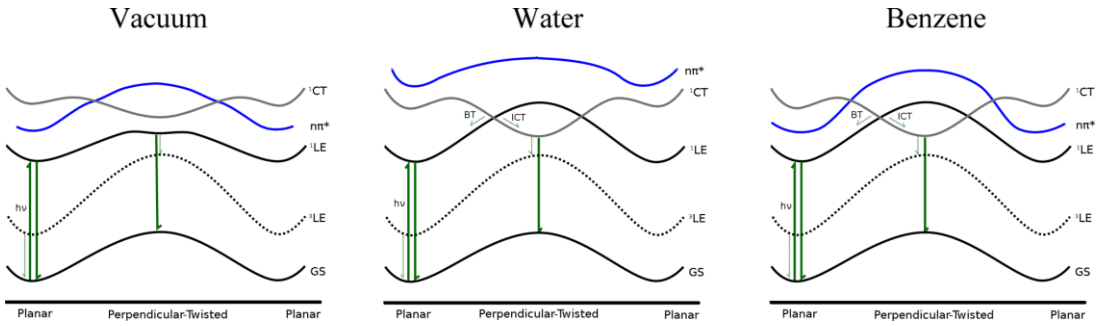


Figure 3.5. Overall photophysical behavior of roseolumiflavin in vacuum, water, and benzene based on CAM-B3LYP and DFT/MRCI calculations. BT and ICT denote back-transfer and intramolecular charge transfer.

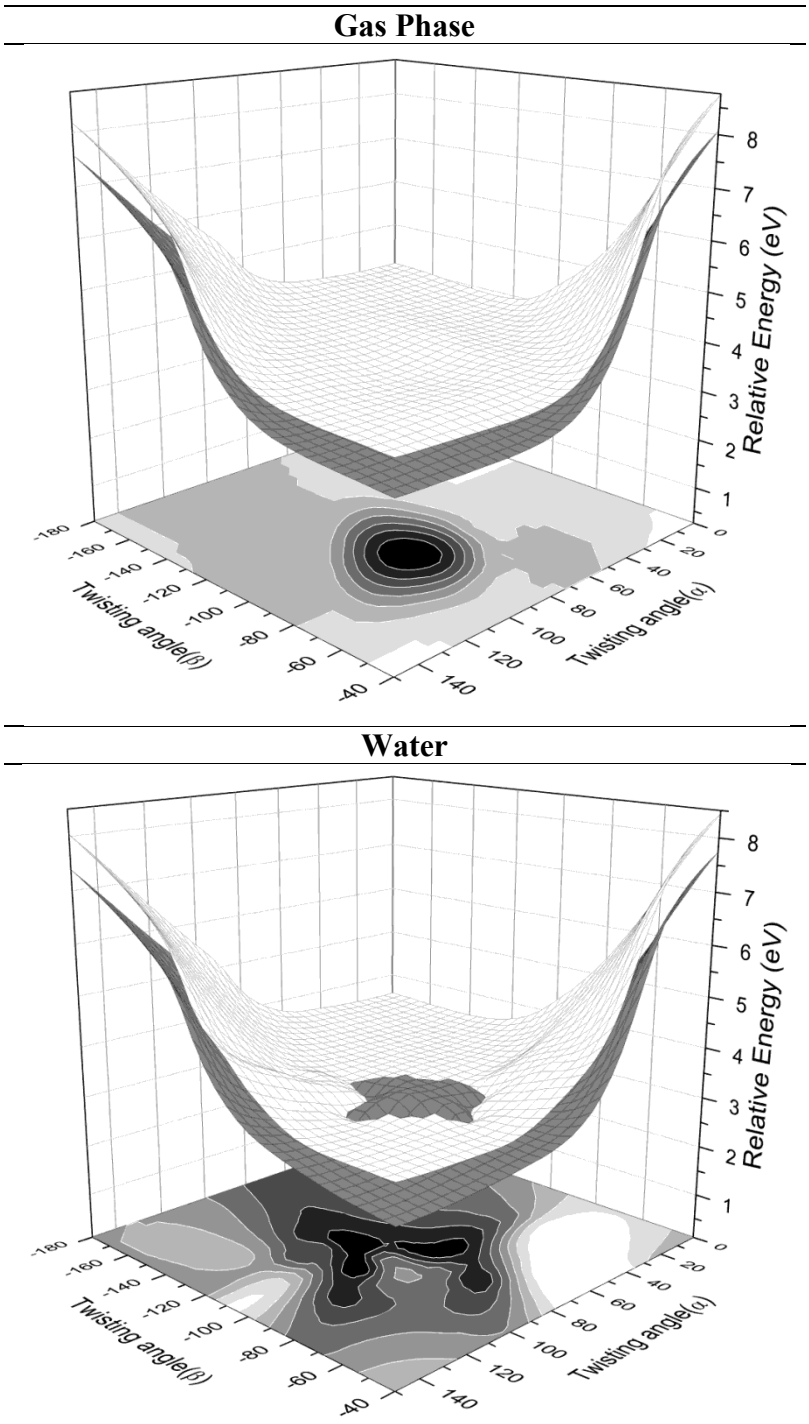


Figure 3.6. 3D representation of the relaxed energy surfaces of the LE state (gray) and CT state (white) with respect to the dihedral angles (α and β, see **Figure 3.4**) as calculated in the gas phase (top) and in water (bottom) at the CAM-B3LYP/6-31G(d) level. Energies are given in eV relative to the GS minimum at the same level. The absolute values of the energy difference between the LE and CT states at a given pair of α and β is presented as a black-

and-white contour map at the bottom, where darker color corresponds to a lower energy difference. Associated 2D contour plots can be found in Figure 7 of ref [176].

In CAM-B3LYP and DFT/MRCI calculations on RoLF, the charge transfer (CT) state has a high dipole moment, indicating a charge separation that is essential for ICT (and is apparent upon inspection of the KS MOs contributing to the excitation). However, the CT state has low oscillator strength, in contrast to the bright locally excited (LE) state with low dipole moment (see Tables 1-3 in ref. [176]). Consequently, photoexcitation will most likely populate the LE state of RoLF, and hence, a crossing of the LE and CT surfaces is required for populating the CT state. Bearing this in mind, we summarize the photophysical behavior of RoLF in the gas phase, water, and benzene in **Figure 3.5**. In the gas phase, both CAM-B3LYP and DFT/MRCI do not predict any LE/CT crossing upon DMA twisting, which precludes ICT. By contrast, in solution (water and benzene) an ICT process seems possible via an LE/CT crossing. As can be seen from the relaxed CAM-B3LYP PES scans (**Figure 3.6**), there is an LE/CT crossing seam close to the perpendicular-twisted (TICT) conformation in both water and benzene, emphasizing the need for DMA twist for ICT to occur. This clearly supports the TICT mechanism for ICT in RoLF, in line with a proposal based on low-temperature fluorescence decay measurements for RoF [51].

To elaborate further on the minimum energy crossing point (MECP), we applied a linear interpolation (LI) scheme, in which DFT/MRCI single-point calculations are used to obtain energies and one-electron properties at interpolated CAM-B3LYP geometries. In this scheme (LI of two dihedrals, LIDA), we took (at least) six intermediary values of the dihedral pairs (α/β) from the corresponding GS, PICT, and TICT geometries and performed a constrained geometry optimization keeping the two dihedrals frozen. The resulting energy profiles of different electronic states in different media are presented in **Figure 3.7**. As can be seen from these plots, the barrier that needs to be overcome to reach the LE/CT crossing (and thus to achieve ICT) is lower in water than in benzene. This can be attributed to the better stabilization of the polar CT state (compared to LE) by the more polar solvent. This finding is also in accord with the experimental observation of alleviated ICT rates of RoF in nonpolar solvents [55,61,165,169,171]. Besides, the DFT/MRCI energy barriers for the ICT and BT processes affirm diminished ICT and enhanced back-transfer (BT) rates in benzene as compared to water, again in accord with experimental findings [55].

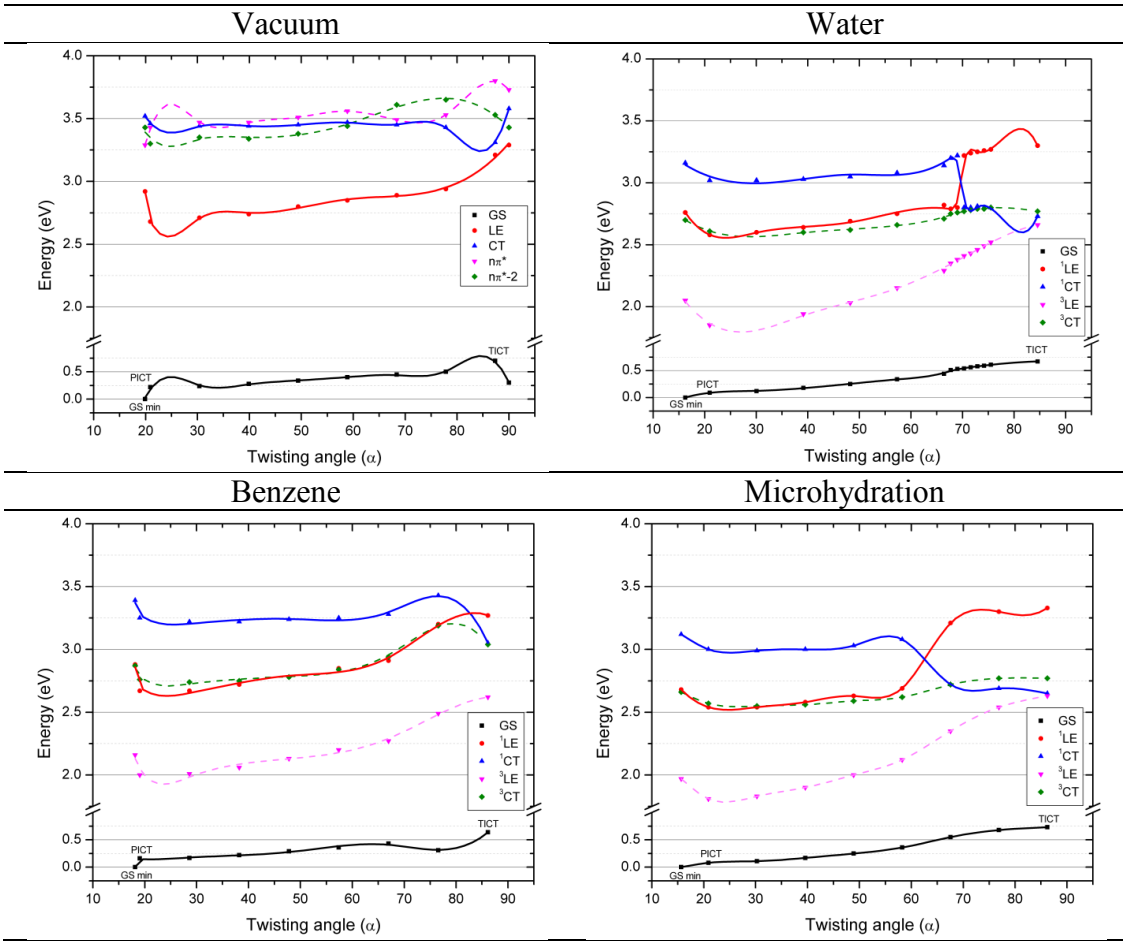


Figure 3.7. DFT/MRCI adiabatic potential energy curves of the ground and the lowest excited states as a function of twisting angle (α) obtained using the LIDA scheme (see text). Energies are given relative to the corresponding ground-state minimum. Each point was optimized with constraints (fixed values of the dihedral angles α and β) on the S_1 surface at the CAM-B3LYP level in the gas phase, water, and benzene as well as by using a microhydration model. A break was applied on the y-axes (1-1.75 eV).

3.3. Photophysics of Flavin Derivatives Absorbing in the Blue-Green Region

For a group of eight artificial LF derivatives (**Figure 3.2**) being potential candidates to replace RF in photoreceptors with a LOV domain, we present a detailed (TD)-DFT(B3LYP) and DFT/MRCI analysis of the absorption characteristics and photodynamics in ref. [177] (**Appendix D**). We evaluate these derivatives on the basis of four criteria: absorption

maximum (in the blue-green region, around 500 nm), stability, triplet formation yield (i.e. ISC rate), and lifetime of the triplet species.

Table 3.1. Vertical DFT/MRCI excitation energies [eV] of 2T-LF, 4T-LF and 2,4T-LF in the gas phase and in water.

2T-LF						
		gas phase		water solution		
state	electronic structure	ΔE [eV]	*f*(r)	ΔE [eV]	λ_{max} [a]	*f*(r)
$1^1A'$	ground state	0.00		0.00		
$1^1A''$	$n_S \rightarrow \pi_L^*$	1.80	0.000	2.55		0.002
$2^1A'$	$\pi_H \rightarrow \pi_L^*$	2.35	0.506	2.50	2.52 eV (492 nm)	0.527
$2^1A''$	$n_N \rightarrow \pi_L^*$	3.22	0.002	3.49		0.009
$3^1A'$	$\pi_{H-2} \rightarrow \pi_L^*$	3.26	0.089	3.23	3.18 eV (390 nm)	0.172
$1^3A'$	$\pi_H \rightarrow \pi_L^*$	1.68		1.91		
$1^3A''$	$n_S \rightarrow \pi_L^*$	1.74		2.49		
$2^3A'$	$\pi_{H-2} \rightarrow \pi_L^*$	2.71		2.64		
$3^3A'$	$\pi_{H-3} \rightarrow \pi_L^*$	2.92		2.95		
4T-LF						
		gas phase		water solution		
state	electronic structure	ΔE [eV]	*f*(r)	ΔE [eV]	λ_{max} [a]	*f*(r)
$1^1A'$	ground state	0.00		0.00		
$1^1A''$	$n_S \rightarrow \pi_L^*$	1.72	0.000	2.40	2.34 eV (530 nm) [b]	0.000
$2^1A'$	$\pi_H \rightarrow \pi_L^*$	2.62	0.377	2.57	2.51 eV (494 nm)	0.395
$3^1A'$	$\pi_{H-2} \rightarrow \pi_L^*$	3.07	0.061	3.11	3.14 (368 nm)	0.053
$2^1A''$	$n_N \rightarrow \pi_L^*$	3.16	0.002	3.40		0.002
$1^3A''$	$n_S \rightarrow \pi_L^*$	1.61		2.29		
$1^3A'$	$\pi_H \rightarrow \pi_L^*$	1.84		1.99		
$2^3A'$	$\pi_{H-2} \rightarrow \pi_L^*$	2.44		2.49		
$2^3A''$	$n_N \rightarrow \pi_L^*$	2.83		3.07		

* Continued on the next page.

2,4DT-LF					
		gas phase		water solution	
state	electronic structure	ΔE [eV]	*f*(r)	ΔE [eV]	*f*(r)
$1^1A'$	ground state	0.00		0.00	
$1^1A''$	$n_{S2} \rightarrow \pi_L^*$	1.57	0.000	2.22	0.000
$2^1A''$	$n_{S4} \rightarrow \pi_L^*$	1.83	0.000	2.45	0.005
$2^1A'$	$\pi_H \rightarrow \pi_L^*$	2.20	0.384	2.32	0.395
$3^1A'$	$\pi_{H-2} \rightarrow \pi_L^*$	2.52	0.252	2.73	0.174
$1^3A''$	$n_{S2} \rightarrow \pi_L^*$	1.49		2.13	
$1^3A'$	$\pi_H \rightarrow \pi_L^*$	1.62		1.77	
2^3A	$n_{S4} \rightarrow \pi_L^*$	1.77		2.40	
$2^3A'$	$\pi_{H-2} \rightarrow \pi_L^*$	1.81		2.20	

[a] Absorption maxima in 0.1 N phosphate buffer (pH 7) taken from ref. [57]. [b] Shoulder in absorption spectrum in 0.1 N phosphate buffer (pH 7), ref. [57].

Among these chromophores, only thioflavins (2T-LF, 4T-LF and 2,4DT-LF) qualified for further in-depth analyses, since the others were eliminated in the preliminary tests for the following reasons. The aza-substituted derivatives (6A-LF, 9A-LF, and 6,9DA-LF) have the maximum absorption band at 2.94-2.97 eV (417-422 nm), which almost coincides with that of wild-type LF (2.94 eV, 422 nm) [161,163], and the influence of the polar solvent environment (water) shifts the absorption maximum to only about 440 nm. Therefore, the aza-substituted derivatives were eliminated because their absorption is not sufficiently red-shifted. By contrast, DXLF passes this test: the first bright band is predicted to be at 2.84 eV (436 nm, vertical excitation), with an onset at 2.60 eV (477 nm, adiabatic excitation); in water, the absorption maximum is further shifted bathochromically to 488 nm. However, DXLF fails to provide an energetically plausible access to the $^3n\pi^*$ state, an essential component in the El-Sayed-allowed $^1\pi\pi^* \dashrightarrow {}^3n\pi^*$ ISC [204], so that triplet formation is impeded. The iminothiol tautomer (2TL-3DH-LF) has an ideal absorption band located at 2.43 eV (511 nm), but the 2-methylthiolether and the analogous 4-methylthiolether were previously shown to be unstable, due to their facile conversion back to LF via hydrolysis [57].

Hence, the thiocarbonyl-substituted compounds (2T-LF, 4T-LF, and 2,4DT-LF) are the most promising candidates for replacing the native FMN (with LF being its truncated version) in the LOV domain. In water, their first bright absorption band maxima appear in the blue-green region (495-535 nm, **Table 3.1**). More importantly, our calculations at 298 K predict that their quantum yields for triplet formation should be close to unity, as opposed to LF, for which the triplet formation ($k_{ISC} \approx 10^8\ s^{-1}$) competes with fluorescence ($k_F \approx 5 \times 10^7$

s^{-1}) [163]. In the thiocarbonyl-substituted compounds, the ISC rates are computed to be on the ps timescale, i.e. four orders of magnitude faster than in LF. This can be credited to the much larger spin-orbit coupling elements (see Tables 3, 6, and 9 in ref. [177]), which benefit from the heavy atom effect (sulfur) and lower energy of the $^1n\pi^*$and $^3n\pi^*$ states that are energetically close to the first bright $^1\pi\pi^*$ state (**Table 3.1**). Thanks to the $\pi_H \rightarrow \pi_L^*$ character of their lowest triplet states (T_1), thiolumiflavins have a long triplet lifetime, reminiscent of LF. Taken together, these features should enable these thiocarbonyl-substituted LF derivatives to operate efficiently in the photocycle of LOV domains. The triplet formation mechanisms in native LF and the three thio-LFs are illustrated in **Figure 3.8** for a direct comparison. Among the three thio-LFs considered, 2T-LF seems best suited to replace LF in the photocycle, because it has an almost identical electron density distribution over the isoalloxazine ring systems (especially at the reactive N5 and C4a centers) and should thus also have a catalytic efficiency comparable to LF.

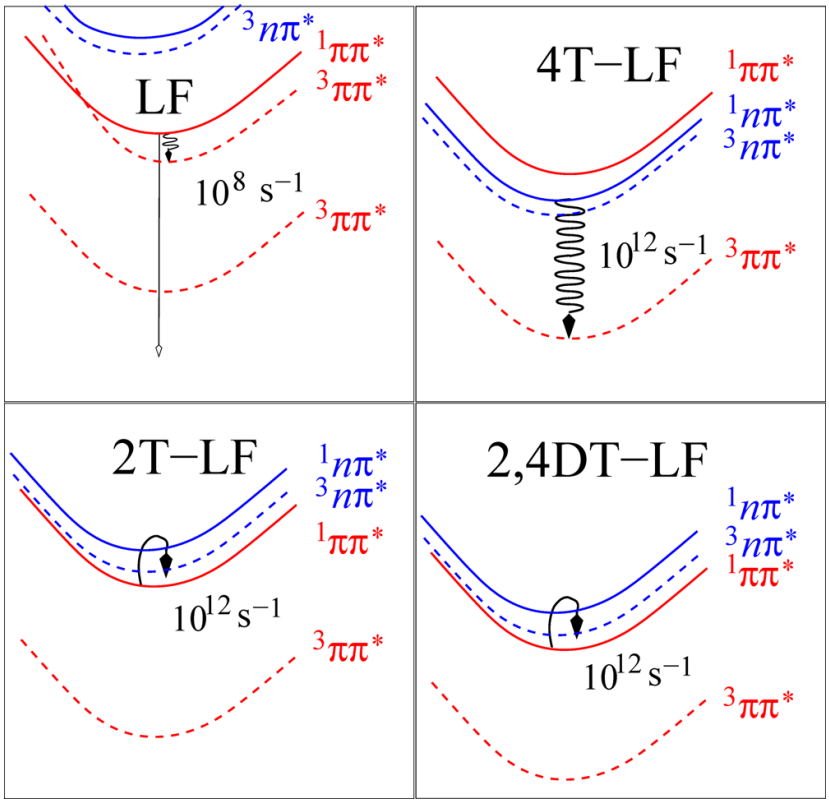


Figure 3.8. Overview of the triplet formation mechanisms and intersystem crossing rates of LF and the thiolumiflavins of interest in water. See text.

3.3. Photophysics of Flavin Derivatives Absorbing in the Blue-Green Region

Chapter 4

Computational Spectroscopy of Flavin Derivatives

After covering biocatalysis and photoreception by flavin derivatives, we will now address their spectral features and their roles as chromophores. We will start with a brief overview of Franck-Condon based methods that are frequently used to simulate vibrationally broadened optical spectra. We then introduce a new technique [212] (**Appendix E**) that complements existing FC methods, and finally present a comprehensive benchmark [213] (**Appendix F**) of most available variants of FC methods, for three flavin derivatives.

4.1. Simulation of Vibrationally Broadened Spectra using the Franck-Condon Principle

Computational spectroscopy of chromophores is a powerful tool for interpreting the experimental UV/vis spectra in terms of underlying structural and vibrational properties. Comparing the energy differences between electronic states at a fixed geometry (vertical excitations) with the measured band maxima has, to date, dominated theoretical studies of optical spectra of large molecules [214]. However, this approach is not capable of resolving the factors governing the spectral shape that can be accessed by considering the vibrational progressions in the Franck-Condon region [215,216]. Vibrational progressions in optical spectra are generally taken into consideration by methods based on the harmonic approximation and the FC principle [217–219]. In the harmonic approximation, vibrations are described with a simple harmonic oscillator, ignoring anharmonic contributions that will exist in reality; whereas the FC principle asserts that an electronic transition is most likely to occur rapidly prior to the adaptation of the nuclear positions in the molecular entity and its

environment. In this approach, the chromophore reaches the FC point through a vertical (vibronic) transition from the initial to the target electronic state, and the intensity of this transition (in absorption or emission) is proportional to the square of the overlap integral between the vibrational wave functions of these two states. The FC principle implies the Condon approximation, i.e. the electronic transition dipole moment remains unchanged within a vibrational progression.

On the basis of the FC principle, both time-dependent (TD) [220–228] and time-independent (TI) [229–237] formalisms have been devised, each having their own advantages and limitations [214]. In the TD scheme [220], the Fourier transform of the autocorrelation function for wave packet propagation on the potential energy surface of the target electronic state is used for generating the spectrum. The TD approach is rather fast, since it does not require calculation of the molecular eigenstates on the target surface [214]. The TI scheme, in contrast, involves costly FC integral calculations using recurrence relations [237], but it is more suitable for high-resolution spectra and the assignment of vibronic transitions, as individual contributions from each vibrational level are considered explicitly. The high cost of the TI approach can effectively be reduced by use of pre-screening models [231,232,235,238]. Both TI and TD treatments can account for Herzberg-Teller effects (i.e., changes in electronic transition dipole moments) [229,239] and can handle the problematic cases of anharmonicity and nonadiabaticity, for which TI needs special techniques, unlike TD [214]. At their computational limit, both frameworks are expected to yield identical spectra [214]. Apart from FC-based methods, other approaches are also available in the literature, such as normal mode sampling using the Wigner distribution [240,241] and the multiconfiguration time-dependent Hartree (MCTDH) [225] method.

In this study, we will focus on the absorption spectra of flavin derivatives, where the vibronic transition is from the initial ground state (GS) to the target electronically excited state (ES). For absorption (as well as emission), the overlap between the vibrational wave functions of the initial and the target electronic states is decisive for the resolution and quality of the simulated FC spectrum. Application of the Duschinsky transformation [206] is expected to increase the overlap because of the alignment of two normal spaces by constructing the ES vibrational modes as a linear combination of GS modes. This treatment becomes even more crucial when substantial geometrical changes occur in the chromophore

after photo-excitation. Therefore, Duschinsky mode mixing has been implemented both for the TI-FC [229,242–244] and the TD-FC approach [227,245].

If the Duschinsky treatment does not remedy the low overlap of two vibrational wave functions, an alternative would be to use the independent mode displaced harmonic oscillator model (IMDHO) [246,247] instead of the (standard) treatment with frequency alternation (IMDHO-FA) (see ref. [228] for a detailed review). Contrary to IMDHO-FA that employs two separate sets of vibrational modes from the GS and ES, IMDHO only uses the GS vibrational modes, thus maximizing the overlap. By using IMDHO, one may skip the calculation of the excited-state Hessian (which will be costly in the absence of analytical gradients). However, IMDHO introduces further approximations and disregards the differences in the shapes of the ground-state and excited-state PES, which will likely cause artifacts in the computed spectra.

4.1.1 Adiabatic and Vertical Franck-Condon Schemes

Both the TD and TI approaches require knowledge of the PES of the ground state and the electronically excited states of interest (but to different extent). FC-based methods differ in the choice of the geometry for generating the Hessian of the target state (ES for absorption). In the widely-used adiabatic FC (AFC) method, the vibrational analysis is done around the minimum geometry of the PES of the target state [248]. However, for large molecules the characterization of the excited-state PES is computationally demanding and sometimes unfeasible. Besides, following the gradients on an excited-state PES to locate minima is prone to root flipping and changes in state character due to crossing with other states. This may lead to a minimum far away from the FC region, which may no longer be a good basis for the representation of the FC region. To deal with this problem, the vertical FC (VFC) approach was introduced [228,248,249], in which an artificial extrapolated minimum (EM) generated by a Newton-Raphson step from the FC point is utilized instead of the real minimum (RM) of the excited-state PES (which would be required in AFC). In general, VFC overcomes problems related to gradient following and obeys the harmonic approximation by definition; however, it does not properly represent the vibrational character at the ES minimum. Since VFC is based on the second derivatives at the FC point, the vibrational analysis around an EM will likely lead to imaginary-frequency modes (since the EM is not a real minimum on the ES surface). In the literature, three ways have been proposed to handle such modes: (a) simply excluding them from further analysis [212,228],

(b) replacing them with a selected real frequency (or its module) [248], or (c) performing costly one-dimensional PES scans to obtain a vibrational Hamiltonian to be used in spectral calculations [249]. The latter choice appears as the most accurate solution, as it accounts for anharmonicity effects on the energies and vibrational frequencies, and it was shown to work almost perfectly in the problematic case of ethylene [249]. Nevertheless, the PES scans along each normal mode can easily become expensive for large chromophores. Given the size of the flavin chromophores considered in our studies, this option turns out to be inapplicable. An alternative would be switching to the IMDHO model, as all GS modes are guaranteed to have real eigenvalues at the GS minimum geometry; however, IMDHO has also some drawbacks (*vide supra*).

4.1.2. Choice of the Coordinate Set

Regardless of the method chosen for computing the FC spectra, the coordinates selected to represent the PES of interest and the corresponding Hessian are a crucial determinant of the quality of the resulting spectra. Cartesian coordinates (with rectilinear displacements) have, up to now, dominated the computation of the spectra [214,234,250]. Using the original framework of Wilson et al. [251], Reimers introduced a generalized internal coordinate system (with curvilinear displacements) as an alternative for generating adimensional shifts between two sets of normal modes, reorganization energies, and FC factors [244], which has been revisited by others [234]. Curvilinear displacements have been shown to outperform rectilinear ones in generating AFC spectra, in particular for non-symmetrical molecules that undergo large torsional displacements upon photoexcitation [230,234,244,250,252]. In contrast to AFC, within the VFC framework, the use of internal coordinates has not yet been considered in the literature, to the best of our knowledge. To fill this niche, we introduced a formalism for combining the VFC approach with internal coordinates (based on the methods of Wilson et al. and Reimers), which will be explained in the next section [212].

4.2. Vertical FC Scheme with Curvilinear Displacements

As the VFC approach was originally formulated for Cartesian coordinates (see refs. [249,253] for details), we will lay out only the framework and indicate the differences needed for employing internal coordinates in the normal-mode representation. We will first show that the shift vector in orthonormal coordinates along the target excited-state surface ($\boldsymbol{\Delta Q}_t$) is identical for both coordinate systems.

As already mentioned, in the VFC framework, the ground-state equilibrium geometry is used as the reference point. At this geometry, the Cartesian gradient $\boldsymbol{V}^C$ and Hessian $\boldsymbol{H}^C$ of the target excited state *t* are defined as:

$$\frac{\partial E_t}{\partial \boldsymbol{\xi}} = \boldsymbol{V}^C \tag{4.1}$$

$$\frac{\partial^2 E_t}{\partial \boldsymbol{\xi}^2} = \boldsymbol{H}^C \tag{4.2}$$

where $\boldsymbol{\xi}$ represents the set of Cartesian coordinates of all atoms included in the system. However, the ground-state minimum will generally not be a stationary point for the target state *t* and the sets of $\boldsymbol{V}^C$ and $\boldsymbol{H}^C$ of the target state may not share the same origin and orientation in *xyz* space as the ones of the ground state. Consequently, diagonalization of the Hessian of the target state $\boldsymbol{H}^C$ will mix some rotational and translational motion into the vibrational modes and thus lead to unsound eigenvalues (*i.e.,* vibrational frequencies). The resulting vibrational modes will suffer from linear dependency and will thus be somewhat coupled to each other. To alleviate the concomitant problems, some techniques have been devised that minimize the root-mean-square deviation (RMSD) between two given Cartesian coordinate sets, such as imposing the Eckart orientation [254] via a quaternion treatment [255]. Apparently, it would be better to perform the vibrational analysis in orthonormalized coordinate sets to create completely uncoupled modes and then transform back to Cartesian or internal coordinates (see below).

In the VFC method, the reference geometry of the target excited state, the extrapolated minimum (EM), is generated by performing a single Newton-Raphson step along the normal coordinates. We adopted the original VFC scheme with minor modifications: our VFC variant ignores the modes associated with a negative Hessian eigenvalue, as the resulting step would be uphill in energy and thus carry the system away from the "minimum". This is a simplification of the original VFC approach [249], in which such modes are also considered by scanning the target PES along their normal coordinates and applying the

resulting anharmonic corrections to the frequencies and energies. We discuss the merits and shortcomings of our simplification below and in more depth in ref. [212].

According to Hooke's law, the gradient at the FC point $\boldsymbol{V}_{FC}$ is related to the displacement $\boldsymbol{\Delta x}$,

$$\boldsymbol{V}_{FC} = -\boldsymbol{k}\boldsymbol{\Delta x} \tag{4.3}$$

where $\boldsymbol{x}$ can be any kind of coordinate set and $\boldsymbol{k}$ is the corresponding vector of force constants. In harmonic approximation, the displacement coordinates $\boldsymbol{\Delta Q}_t$ of the target state are computed from mass-weighted normal-mode Cartesian displacement coordinates $\boldsymbol{\Delta q}$, using

$$\boldsymbol{\Delta Q}_t^C = \boldsymbol{L}_t^{C,T}\boldsymbol{\Delta q} = \boldsymbol{L}_t^{C,T}\sqrt{\boldsymbol{m}}(\boldsymbol{\xi}_t - \boldsymbol{\xi}_0) \tag{4.4}$$

where superscript T stands for the matrix transpose. Here, the matrix $\boldsymbol{L}_t{}^C$ is composed of the orthonormal Cartesian eigenvectors of the Hessian (the normal coordinates given in the columns), and $\boldsymbol{m}$ is a diagonal matrix of the atomic masses associated with each Cartesian coordinate in $\boldsymbol{\xi}$.

The representation of the normal modes in internal coordinates ($\boldsymbol{L}^I$) is known to be advantageous for describing complex curvilinear motions [244]. A non-redundant set of internal coordinates ($\boldsymbol{S'}$) can be defined in terms of a standard (arbitrary) Z-matrix. The normal modes can be converted from Cartesian to non-redundant internal coordinates as follows [244]:

$$\boldsymbol{L}^I = \boldsymbol{B''}\boldsymbol{m}^{-1/2}\boldsymbol{L}^C \tag{4.5}$$

where the Wilson $\boldsymbol{B''}$ and $\boldsymbol{G'}$ matrices [251] are given as

$$\boldsymbol{B}^{''} = \boldsymbol{G'}^{-1/2}\boldsymbol{B'} \tag{4.6}$$

$$\boldsymbol{G'} = \boldsymbol{B'}\boldsymbol{m}^{-1}\boldsymbol{B'}^T \tag{4.7}$$

Here, the non-redundant Wilson $\boldsymbol{B'}$ matrix is defined as

$$\boldsymbol{B}^{'} = \boldsymbol{a}^{\boldsymbol{T}}\boldsymbol{B} \tag{4.8}$$

with the matrix $\boldsymbol{a}$ being composed of the eigenvectors with nonzero eigenvalues of the Wilson $\boldsymbol{G}$ matrix. The $\boldsymbol{B'}$ matrix transforms the 3N Cartesian coordinate displacements $\boldsymbol{\Delta\xi}$ to the set $\boldsymbol{\Delta S'}$ of 3N-6 (5) non-redundant internal coordinate displacements,

$$\boldsymbol{\Delta S'} = \boldsymbol{B'}\boldsymbol{\Delta\xi} \tag{4.9}$$

In this notation, primes (and double-primes) are used to label matrices associated with non-redundant (and orthonormal) internal coordinates [244]. The gradients in internal coordinates $\boldsymbol{S'}$ can be obtained from the mass-weighted Cartesian gradients:

$$\boldsymbol{V}^I = \frac{\partial E_t}{\partial \boldsymbol{S}'} = \boldsymbol{G}'^{-1/2}\boldsymbol{B}''\boldsymbol{m}^{-1}\boldsymbol{V}^C \tag{4.10}$$

The Hessian $\boldsymbol{H}^{MWI}$ in mass-weighted internal coordinates is related to the Hessian $\boldsymbol{H}^{MWC}$ in mass-weighted Cartesian coordinates by the following transformation:

$$H_{ij}^{MWI} = B_{ij}'' m_i^{-1/2} H_{ij}^{MWC} m_j^{-1/2} B_{ji}'' \tag{4.11}$$

where the Hessian $\boldsymbol{H}^{MWC}$ can be straightforwardly computed using

$$H_{ij}^{MWC} = \frac{H_{ij}^C}{\sqrt{m_i m_j}} \tag{4.12}$$

However, this transformation is not exact in general because it neglects contributions from the Wilson $\boldsymbol{B}$ gradient term; it becomes exact at stationary points with zero gradient (see Eq. 6 in ref. [256]). We have confirmed numerically for one of the currently studied flavin molecules that the approximation (4.12) is not critical, since the eigenvalues of the Cartesian and internal Hessians overall agree very well also at non-stationary points (mostly within 1%, see Table S1 in ref. [212])**.** Yet, this still causes small differences in VFC spectra depending on the choice of coordinate set.

After compiling the necessary equations for the inter-conversion between Cartesian and (non-redundant orthonormal) internal coordinates, we continue with the generation of the EM. The normal modes $\boldsymbol{L}_t^I$ of the target state are, by definition, the eigenvectors of the relevant Hessian $\boldsymbol{H}^{MWI}$ in mass-weighted internal coordinates, and can also be computed directly using Eq. 4.5. With this in mind, the Newton-Raphson displacements on the target surface in orthonormal internal coordinates are given by:

$$\boldsymbol{\Delta Q}_t^I = \boldsymbol{L}^{I,T}\boldsymbol{G}'^{-1/2}\boldsymbol{\Delta S}' \tag{4.13}$$

Both $\boldsymbol{\Delta Q}_t^C$ and $\boldsymbol{\Delta Q}_t^I$ are orthonormal displacements from the FC point on the target PES. Within the VFC scheme, we compute $\boldsymbol{\Delta Q}_t$ first by performing a Newton-Raphson step and then convert the associated displacements back to Cartesian or internal coordinates. Therefore, even though the resulting EM geometries will be different in real space, their displacements from the FC point are identical in the corresponding orthonormal spaces. The resulting real space geometry is thus of no concern. We thus conclude that:

$$\boldsymbol{\Delta Q}_t^I = \boldsymbol{\Delta Q}_t^C = \boldsymbol{\Delta Q}_t \tag{4.14}$$

Hence, Eq. 4.3 can be used for any coordinate set to construct $\boldsymbol{\Delta Q}_t$ by projecting the mass-weighted forces in the desired representation onto the corresponding normal modes:

$$\boldsymbol{L}_t^{C,T}\boldsymbol{m}^{-1/2}\boldsymbol{V}^C = \boldsymbol{L}_t^{I,T}\boldsymbol{G}'^{1/2}\boldsymbol{V}^I = \boldsymbol{V}^Q = -\boldsymbol{k}\boldsymbol{\Delta Q}_t \tag{4.15}$$

Here, each element of $\boldsymbol{\Delta Q_t}$ belongs to a normal mode n, depending on the force constants $\boldsymbol{k}$ and the gradients $\boldsymbol{V^Q}$ projected onto the normal coordinates. The force constants are then given by the eigenvalues $\boldsymbol{\lambda}$ of the mass-weighted Hessian in 3N-6 (5) mass-weighted internal coordinates ($\boldsymbol{k} = \boldsymbol{\lambda}$).

The excited-state energy at the displaced EM geometry is needed for generating FC spectra. It is given in the VFC scheme by:

$$E_{EM} = E_0 + E_{vert,t} - \frac{1}{2}\boldsymbol{\lambda \Delta Q_t^2} \tag{4.16}$$

where E_0 and $E_{vert,t}$ are the ground-state energy and the vertical excitation energy at the FC point, respectively. In our VFC version, the modes with negative λ are ignored; they are thus omitted in the last term.

For FC calculations, all or part of the items from the following list are required: the adiabatic energy shift from the GS minimum to the ES "minimum", the excited-state shift vector $\boldsymbol{\Delta Q_t}$, the vibrational frequencies $\boldsymbol{\lambda_t}$, the normal modes $\boldsymbol{L_t}$, and their ground-state counterparts along with the displacement in orthonormal space for the ground-state normal modes ($\boldsymbol{\Delta Q_0}$), and the Duschinsky matrix $\boldsymbol{J}$ [206]**.** The latter is constructed in Cartesian coordinates by

$$\boldsymbol{J^C} = \boldsymbol{L_0^{C,T} L_t^C} \tag{4.17}$$

and in internal coordinates by

$$\boldsymbol{J^I} = \boldsymbol{L_0^{I,T} L_t^I} \tag{4.18}$$

As $\boldsymbol{J}$ must be a square matrix by definition, we need to delete certain columns from $\boldsymbol{L_0}$ to allow for proper mapping, in case we discard the imaginary modes in the excited state. Thus, we remove those normal modes from the GS set that have the highest similarity to the ES imaginary modes (*i.e.,* the largest Duschinsky matrix element). We can then construct $\boldsymbol{\Delta Q_0}$ by

$$\boldsymbol{\Delta Q_0} = -\boldsymbol{J^T \Delta Q_t} \tag{4.19}$$

Since we have computed $\boldsymbol{J}$ for identical orthonormal coordinate sets in the ground and excited state, it is orthogonal before the eventual deletion of columns and rows. After this deletion, $\boldsymbol{J}$ is no longer orthogonal; this however does not affect the validity of the operation in Eq. 4.19. Since we have two different $\boldsymbol{J}$ matrices, one for each coordinate set, we may expect the corresponding displacements $\boldsymbol{\Delta Q_0}$ to differ as well (namely $\boldsymbol{\Delta Q_0^C}$ or $\boldsymbol{\Delta Q_0^I}$). Along with different $\boldsymbol{J}$ matrices, this is the main source of the diversity brought about by the use

two different coordinate sets within the VFC framework. This issue will be addressed in depth later in this Chapter.

At this point, one should also note from Eqs. 4.17-4.19 that all differences between coordinate representations would vanish when using the IMDHO approximation (i.e., neglecting the effects of Duschinsky rotation). In this regard, we use a slightly modified version of the IMDHO model for the VFC case, which differs from the original IMDHO scheme [246,247] as follows: We set $\boldsymbol{J}$ to unity after constructing $\boldsymbol{\Delta Q_0}$, and we also set $\boldsymbol{\lambda}_t$ equal to $\boldsymbol{\lambda}_0$. By contrast, in the IMDHO model for the AFC case, $\boldsymbol{\Delta Q_0}$ is set equal to $\boldsymbol{\Delta Q_t}$ as well.

4.3. Methods and Computational Details

4.3.1 The Home-Made *Python* Interface

Initially, we implemented the Newton-Raphson formalism (as given in **Section 4.2**) to enable the use of internal coordinates (with curvilinear displacements) within the vertical FC framework as a *Python* script ("*Harmonic.py*") combined with *C++* procedures for efficiency. Next, we automated the procedure for creating FC spectra starting from output files for the GS and each ES. To this end, we compiled another *Python* script ("*Curv_spec.py*") that interfaces all required software programs (see below), allows the use of both correction schemes described below, and deals with imaginary modes in case of VFC. The capabilities and the detailed workflow of this *Python* script are given in **Figure 4.1**.

The Wilson $\boldsymbol{G}$ and $\boldsymbol{B}$ matrices are generated by our code and used for computing the matrices $\boldsymbol{\Delta Q_0}$ and $\boldsymbol{J}$ that may be represented in either Cartesian or internal coordinates (as discussed in **Section 4.2**). Many of currently adopted procedures were inspired by the DUSHIN program, which is also used to produce the corresponding shift vectors and Duschinsky matrices during the computation of AFC spectra in internal coordinates; this program was kindly provided by Prof. Reimers [244]. For the computation of the TI-FC integrals, we have adapted the open-source FCClasses code [233,257] so that we can feed in our computed energies, as well as $\boldsymbol{\Delta Q}_0$, $\boldsymbol{\Delta Q_t}$ and $\boldsymbol{J}$, at the required instances. For the TD-FC spectra, we used the original version of ORCA that requires as standard input the adiabatic energy difference (0-0 transition), $\boldsymbol{\lambda}_0$, $\boldsymbol{\lambda}_t$, and $\boldsymbol{\Delta Q_t}$.

4.3. Methods and Computational Details

Our interface allows us to generate any of the variants discussed above for the TI-FC or TD-FC approaches (e.g., AFC or VFC, Cartesian rectilinear or internal curvilinear coordinates, with or without application of the Duschinsky transformation). However, the explicit use of the Duschinsky rotation in TD-FC spectra is not possible with this interface due to limitations in the underlying software (*ORCA_ASA*). As a remedy, one might consider using alternative TD-FC implementations [227,245]), which became publicly available only at a late stage of our work; they may be explored in the future through an extension of the interface.

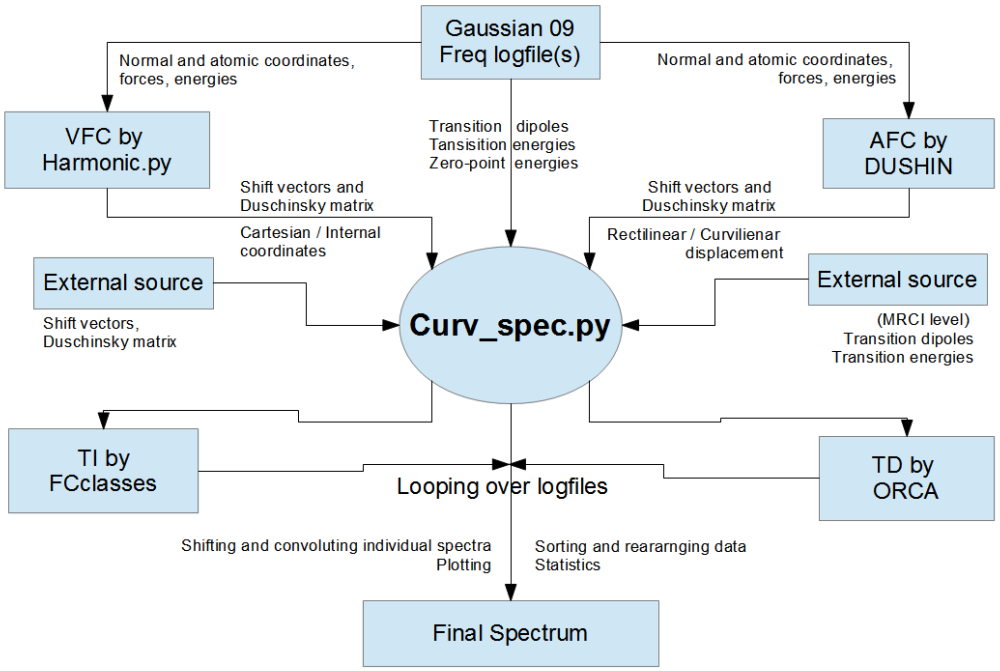


Figure 4.1. Workflow of the homemade script: Curv_spec.py

Software Capabilities:

- Generation of vibrationally broadened TI and TD spectra using AFC and VFC schemes combined with IMDHO and IMDHO-FA models (with and without Duschinsky rotation).
- Interface between various software packages: Gaussian09, DUSHIN, FCClasses, ORCA, and *Harmonic.py*.
- Compatible with Gaussian09 log files and formatted checkpoint files (may in the future be easily extended to outputs from other quantum chemistry software).
- Can work with more than one state and can convolute multiple states.

- TI-FC and TD-FC formalisms as implemented in FCClasses and ORCA.
- AFC and VFC frameworks as implemented in DUSHIN and *Harmonic.py*, compatible with the use of Cartesian and internal coordinates.
- Reading external shift vectors and Duschinsky matrix to generate spectra.
- Can apply the ZPVE and MRCI corrections (see **Section 4.3.3**) as well as customized corrections, e.g. for state-specific solvent treatments.
- Can apply the customized Duschinsky treatment for TD spectra (see below).
- Implementation of diverse methods for dealing with imaginary-frequency modes, particularly within the VFC scheme.

4.3.2. Electronic Structure Calculations

All ground- and excited-state geometry optimizations and the subsequent Hessian calculations were performed using (TD-)DFT [117,258,259] with Gaussian09 [116]. For benchmarking purposes, we combined four commonly used functionals, BP86 [260,261], B3LYP [117], CAM-B3LYP [185] and ωB97xD [122], with an Ahlrichs-type (TZVP) [125] and several Pople-type basis sets (STO-3G, 6-31G(d), 6-31G++G(d,p) and 6-311++G(d,p)) [187]. We did not apply any symmetry constraints during geometry optimizations for any of the chromophores considered here (thus allowing for side-chain torsions).

We also computed DFT/MRCI single-point energies and one-electron properties at the DFT optimized geometries. The underlying reasons and the relevant technical details were already discussed in **Section 3.1**. The resulting vertical excitation energies and transition dipole moments of the electronically excited states were used in a correction scheme, "*MRCI-corrected*" (similar to that employed by Jacquemin and coworkers [262]). In this scheme, all vibrationally resolved peak positions are shifted by the difference between the TD-DFT and DFT/MRCI vertical excitation energies, and the transition dipole moments are taken from DFT/MRCI. In other words, the difference between the DFT/MRCI and TD-DFT vertical excitation energies for the target state is added to the TD-DFT adiabatic excitation energy for that state.

Solvent effects were treated implicitly by the CPCM model [126,189] for water (ε = 78.00) and benzene (ε = 2.27). In DFT/MRCI calculations, when required, solvent effects were taken into account by generating the Kohn-Sham orbitals using the COSMO model [190] as implemented in Turbomole6.3.

4.3.3 Vibrationally Resolved Spectra Calculations

We simulated the vibrationally broadened absorption spectra at near-zero (0.01 K) and finite temperatures (298 K) by the TD-FC and TI-FC methods, using ORCA *(ORCA_ASA)* [253] and a modified version of FCClasses (as explained above) [233,263,264]. The TD approach utilized the IMDHO and IMDHO-FA models [220], which we also adapted for the TI scheme to allow for an unbiased comparison. For VFC spectra, we used a slightly modified version of the IMDHO model (see **Section 4.2**). Besides, we corrected both the VFC/IMDHO and AFC/IMDHO peak positions by the zero-point vibrational energy (ZPVE) difference between the ground state and the target excited state (similar to the procedures given in refs. [262,265]). The latter was required for a direct comparison of the spectral shapes from IMDHO to those from IMDHO-FA, for which this ZPVE correction is implicit. Computed 0-0 peak positions (and others related to it) can be further refined by accounting for solvent effects using state-specific solvent corrections [266–268]. In the absence of experimental data for the 0-0 peak position, one may introduce further theoretically motivated corrections. Here, we use the *MRCI-correction* scheme (*vide supra*).

Vibrational frequencies, normal mode displacements, and structural properties were extracted from Gaussian09 outputs and then used in the calculation of adimensional shifts and the related Duschinsky rotation. For AFC spectra, the shifts and Duschinsky rotation were obtained in either rectilinear or curvilinear coordinates with the DUSHIN program [244], whereas for VFC spectra, they were calculated using our home-made *Python* script ("*Harmonic.py*", see above). By default, ORCA does not include an implementation of the Duschinsky treatment. We thus introduced an implicit Duschinsky rotation of the shift vectors for the *ORCA* TD spectra as well, which uses the same vectors as in the TI treatment, thus allowing for more direct comparisons between the TI (FCClasses) and TD (ORCA) approaches. Likewise, we employed this implicit Duschinsky treatment in our ORCA TD+VFC spectra for addressing the problem of imaginary modes by using the Duschinsky matrix elements as a selection criterion for mode deletion (see **Section 4.2**). For generating TI spectra in FCClasses, we computed up to 10^7 FC integrals considering 60 quanta for C1 and C2 classes, unless stated otherwise. We adopted a minimum weight of 0.20 in the Boltzmann distribution for 298 K spectra.

4.3.4. Quantitative Criterion to Compare Spectra

An unbiased quantitative criterion is required for an objective comparison of two given spectra. With this in mind, we applied the following procedure to measure the overlap between a computed and an experimental spectrum quantitatively. We convoluted the stick spectra and then determined the overlap of intensities in the two convoluted spectra at each convolution point. The overall overlap (*O*) value was then computed using

$$O = 1 - \frac{\sum_i^{N_{conv}} \frac{|I_{1i} - I_{2i}|}{|I_{1i}| + |I_{2i}|}}{N_{conv}} \tag{4.20}$$

where N_{conv} is the number of convolution steps, and I_1 and I_2 are the relative intensities in spectrum 1 and 2 at a given step. By definition, the highest possible value of *O* is 1 (in case of full overlap); *O* strongly depends on the choice of half-width at half-maximum (HWHM) used for broadening the spectra (see SI, part A in ref. [213]). Hence, for a more reliable comparison, we use an average value of *O* that is computed over a set of HWHM values. When the reference (i.e. measured) spectrum lacks the stick information, a single *O* value from the HWHM value that yields the best match (in this case HWHM = 403 cm^{-1}) was used. In cases where convolution steps from two compared spectra do not coincide, the required data points were created by interpolation of the available points. When the reference 0-0 band position was not available, to minimize the effect of 0-0 peak shifts, the *O* values were further optimized by moving the computed spectrum as a whole after convolution of different bands (denoted as "optimized *O* values" in the captions of related figures).

4.4. Benchmark of FC methods

For the benchmark of the aforementioned FC methods, we compiled a test suite consisting of two small linear alkenes (ethylene and *all-trans*-1,3,5-hexatriene), two aromatic compounds (*N*,*N*,2-trimethylaniline and 6-methyl-7-dimethyl-aminoquinoxaline) and three biologically relevant flavin compounds (**Figure 4.2**). Among the myriad of known flavins, we chose riboflavin, roseoflavin, and 5-thioflavin, due to their interesting structural properties, the availability of experimental UV/vis spectra, and their biological importance (*vide supra*). For reasons discussed earlier (**Section 3.1**), we considered truncated versions of the flavins (i.e. only the isoalloxazine chromophore) in the analysis (**Figure 4.2**).

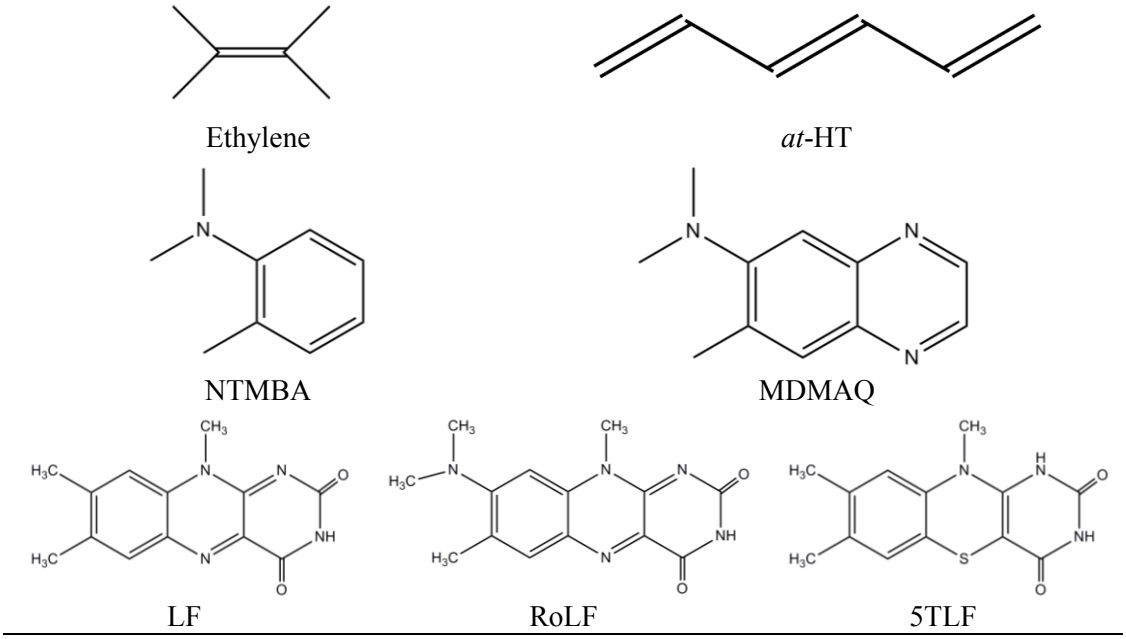


Figure 4.2. Schematic representation of the models considered here: ethylene, *all-trans*-1,3,5-hexatriene (*at*-HT), N,N,2-trimethylbenzeneamine (NTMBA, or equivalently N,N,2-trimethylaniline), 6-methyl-7-dimethyl-aminoquinoxaline (MDMAQ), and three flavin derivatives, lumiflavin (LF), roseolumiflavin (RoLF), and 5-thiolumiflavin (5TLF).

In our first publication [212], we introduced the combination of the VFC scheme and internal coordinates (with curvilinear displacements) and compared its performance to that using Cartesian coordinates (with rectilinear displacements). Towards this goal, we focused on the TI-VFC and TI-AFC spectra using rectilinear and curvilinear displacements, considering only a single excited state from each of the seven test molecules in the gas phase. Our analyses showed [212] that the choice of coordinate set does not matter for simple cases (i.e. chromophores with a plane of symmetry) for both AFC and VFC. To the contrary, for non-symmetrical chromophores, particularly those with rotatable groups (e.g. NTMBA, MDMAQ and RoLF), the curvilinear displacements outperform the Cartesian ones in terms of the accuracy of the resulting AFC spectra, in line with previous studies [230,234,244,250,252]. This is clearly evident in **Figure 4.3** (TI-AFC part): the difference of the relative intensities in curvilinear and rectilinear spectra for non-symmetrical RoLF and 5TLF can reach 60%. For the VFC spectra, however, curvilinear coordinates do not offer any significant advantages over rectilinear ones, because both descriptions are essentially equivalent after the transformation to an orthonormal space (see **Section 4.2**).

Accordingly, the choice of the coordinate set leads to only subtle differences in the computed VFC spectra (<2% in relative spectral intensity, see **Figure 4.3**, TI-VFC part), which are likely due to the approximations in the conversion of the Hessian **(Section 4.2**).

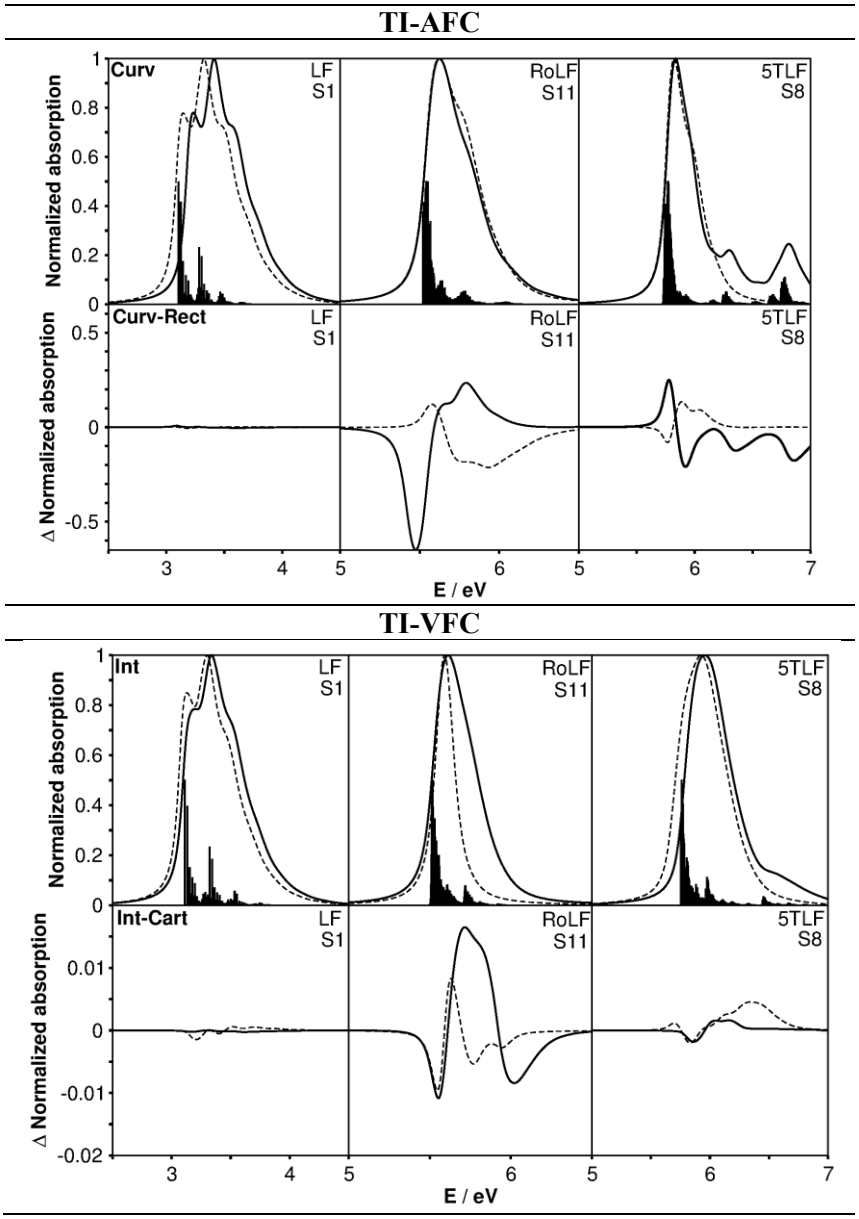


Figure 4.3. TI-AFC and TI-VFC spectra (solid lines and sticks) of LF, RoLF and 5TLF for each specified state. Dashed lines: IMDHO approximation. For TI-AFC of TI-VFC panels: (top row) Curvilinear/internal mode representation spectra; (bottom row) difference spectra (curvilinear/internal-rectilinear/Cartesian).

4.4. Benchmark of FC methods

It is well established that AFC spectra can often not be obtained, i.e., whenever a reliable minimum geometry for the target excited state is not available. This problem is frequently encountered for high-lying excited states, for example due to root flipping and changes in the excited-state character during geometry optimization. In the examples covered in our study [212], if AFC spectra were at all obtainable, they were of similar quality as the VFC spectra. Hence, to compute FC spectra, there is apparently no real need to embark on the time-consuming and often error-prone process of excited-state geometry optimization. For large systems where the excited-state minimum may be far away from the FC point, the VFC approach is actually expected to be superior to AFC, as it guarantees the inclusion of the FC region in the harmonically approximated excited-state PES.

Within the VFC framework, however, the problem of imaginary frequencies arises for the target excited state at the FC point. As already mentioned, we omit the imaginary modes of the ES (and their GS counterparts) in our analysis. We note in this context that it is a common strategy to compute FC spectra using only selected vibrational modes that affect the spectral shape most strongly [161,228]. Our approach may be regarded as a different kind of mode selection, in which we include all modes except those that are technically problematic and are not expected to influence the spectrum to a significant degree. We believe that the omission of imaginary-frequency modes will normally be justified, except for cases where such modes become crucial (*e.g.*, in a photodissociation spectrum [269]). The FC integrals connected to such modes are expected to be small, at least at lower temperatures, since the corresponding ES potential minimum is expected to be quite far away from the FC point and the associated nuclear-motion wave function is thus likely to be broad. Compliant with our expectations, the computed VFC spectra of the LF/S_1 state (with discarded imaginary-frequency modes) are found to be comparable to the available experimental data [164] (see Fig. 6 of ref. [212] for a direct comparison). Similarly, our modified (simplified) VFC method performs quite well for ethylene/$^1B_{3u}$ state as compared to the results from the original (full) VFC scheme [249]. In the considered cases, the error due to the omission of imaginary modes tends to become smaller as the system size increases, because the ratio of imaginary modes to the total number of modes decreases. This issue was further checked on a case-by-case basis in our second study [213].

In this follow-up study [213], we extended our investigation to (almost) all possible combinations of different FC methods (as discussed in previous sections). To this end, we directed our attention to the multi-state absorption bands of LF, RoLF, and 5TLF in

different media, for which experimental data is available for direct comparison [55,164,270]. Apart from their biological importance, the three flavin derivatives were chosen based on their structural properties that allowed us to perform a systematic investigation of the effect of symmetry on the resulting FC spectra. More specifically, while LF has C_s symmetry, the DMA substitution in RoLF breaks the symmetry plane and the DMA group may rotate, while the isoalloxazine moiety tends to stay planar. By contrast, 5TLF assumes a butterfly-like bent structure due to the thio substitution at the N5 position and consequently, the symmetry is completely broken. Photoexcitation causes the opening up of the "wings" in 5TLF (see **Figure 4.4**), associated with a structural rearrangement that drags the chromophore away from the FC point. This makes 5TLF a challenging case for any type of computational spectroscopy technique.

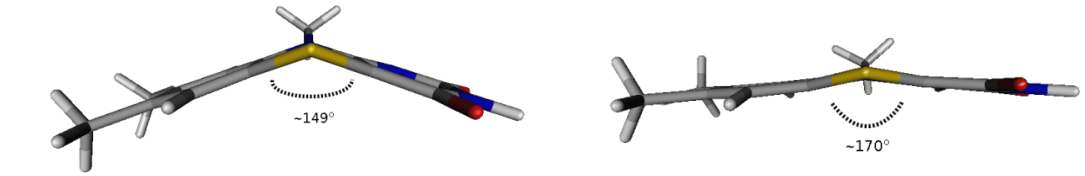


Figure 4.4. Side view of 5TLF in stick representation, neutral form: (left) S_0 and (right) S_2 geometries at the CAM-B3LYP/6-31G(d) level.

Our benchmark study [213] consisted of two main parts. In the first part, we addressed one particular state (LF/S_1 in the gas phase) covering a narrow region in the absorption spectrum. This spectral region had been measured in vacuum-like medium at almost-zero temperature (He nanodroplets at 0.3 K), with high resolution that enabled the assignment of the individual bands and progressions [164]. This spectral data is very valuable in that it makes it possible to assess the accuracy of simulated vibrationally broadened spectra free from temperature and environment effects. Therefore, we evaluated the LF/S_1spectra obtained from different density functionals, basis sets, and FC variants in terms of their similarity to the experimental reference.

The comparison of density functionals with dissimilar HF exchange (using the same basis set and FC variant) indicated that for the selected model systems the range-separated functionals (CAM-B3LYP and ωB97xD) perform equally well and yield the FC spectra with highest resemblance to the reference (with the highest *O* values; see **Figure 4.5**, top panel for visualization). Hence, we chose CAM-B3LYP for further analysis, but ωB97xD would have been an equally good choice. However, both range-separated hybrids

overestimate the 0-0 absorption band (1_0^0) due to high amount of HF exchange, and the resulting spectral envelope is thus remarkably blue shifted (by ca. 4000 cm^{-1}) with respect to the experiment. As the reference 1_0^0 band position is known in this case, we aligned the 0-0 peak from each simulated spectrum on top of it (*0-0-alignment* scheme) to be able to fairly compare the spectral topology and envelope. By contrast, to remedy 0-0 band positions in cases where no experimental reference is available, we used the *MRCI-correction* scheme, in which the 1_0^0 band position is shifted based on the DFT/MRCI vertical excitation energy. For the LF/S_1 band in the gas phase, both B3LYP and BP86 still give satisfactory spectral shapes although they perform less well than the range-separated hybrid functionals. In particular, B3LYP provides a good compromise between the 0-0 peak position (within ca. 200 cm^{-1} of the measured one) and topology; in problematic cases (e.g. 5TLF), we thus also check the results from B3LYP, being a functional of different character. The choice of the basis set affects the computed FC spectra only vaguely (see **Figure 4.5**, bottom panel) with small variations in *O* values (< 2%). This renders the use of extended basis sets unnecessary, at least for the model systems under consideration. One should, however, refrain from using too small basis sets, as evident from the artificially overestimated peak heights in the high-energy region in the STO-3G spectrum (**Figure 4.5**). The findings in our benchmarks are overall in good agreement with those for similar chromophores (anthraquinones) [265] and others [262,271–273].

The detailed analysis on LF/S_1 state also revealed another important point. To simulate AFC spectra, one needs accurate normal modes, generated at reliable geometries. Among the two available conformers of LF (Figure 4 of ref. [213]), only the eclipsed conformation could reproduce the measured spectral topology accurately, whereas the staggered one only yielded featureless spectra, as the result of low overlap between LF/S_0 and LF/S_1 states in this conformation. The VFC approach, in contrast, could overcome this problem by using the EM that is likely located in the immediate FC region, where the photon absorption occurs and the harmonic approximation holds by definition. In this regard, VFC seems superior to the AFC approach. In addition, the resulting VFC spectra are highly comparable to the experiment, reminiscent of the AFC spectra with eclipsed orientation. This finding further reinforces the confidence in VFC spectra.

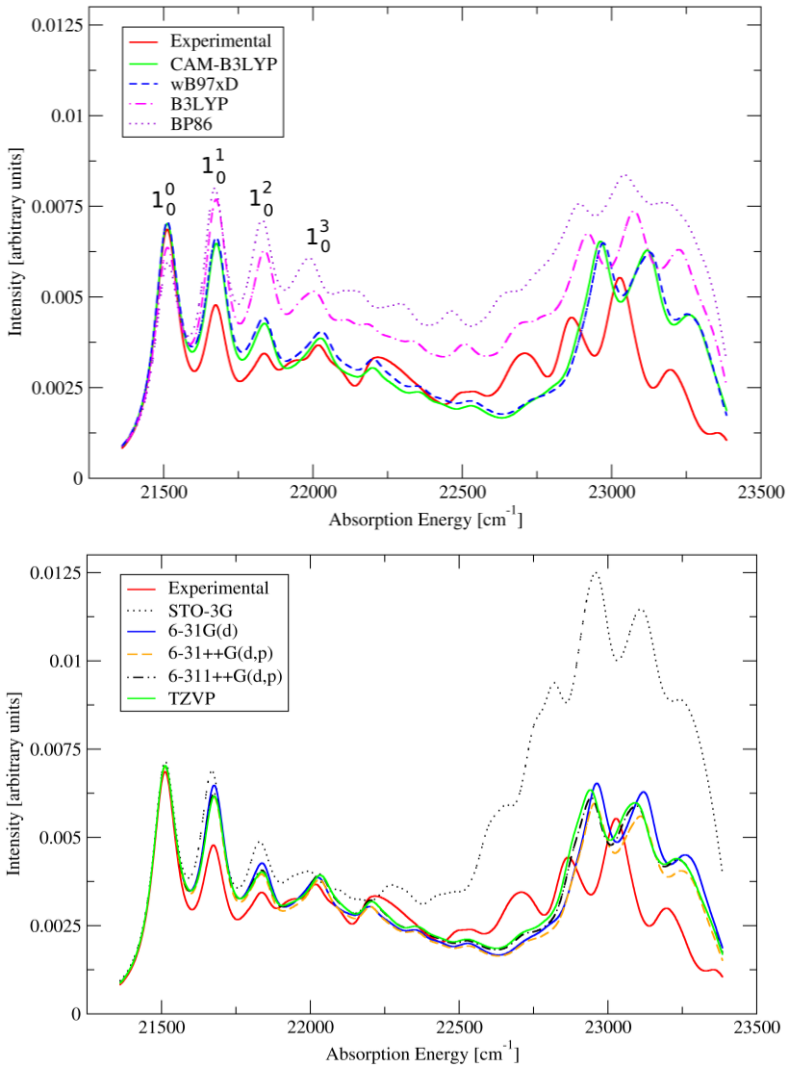


Figure 4.5. (Top) Comparison of LF/S_1 TI-AFC spectra in vacuum obtained from different density functionals with the 6-31G(d) basis set; (bottom) TI-AFC spectra from CAM-B3LYP calculations with different basis sets. TI-AFC spectra were computed at 0.01 K using Cartesian coordinates to generate the adimensional shifts and Duschinsky rotation. HWHM of 50 cm^{-1} for broadening. Spectra are *0-0-aligned* (see text). Computed spectra are compared to experiment at 0.03 K, taken from ref [164].

For the example of the LF/S_1 band in the gas phase [213], we also revisited the issue of imaginary-frequency modes within the VFC framework. As can be seen in **Figure 4.6**, the deletion of the imaginary-frequency modes (as done in our VFC approach) and of the related shift vector elements results in a spectrum with the highest similarity to the AFC spectrum that serves as reference, given the availability of a real minimum (without

imaginary modes) on the LF/S_1 energy surface. Replacing the frequency of the problematic modes with some arbitrary real value (**Figure 4.6**, dashed lines) or its modulus (black solid line) can lead to an unrealistic spectral topology. This can be explained by the high adimensional shift of an imaginary-frequency ES mode relative to its counterpart GS mode, which is caused by large unphysical gradients (due to being far away from a stationary point). By removing the contributions from imaginary-frequency modes, we can prevent introducing such unphysical effects in the computation of the broadened spectrum. These results provide further support for the our preferred practice of ignoring the imaginary-frequency modes within the VFC framework

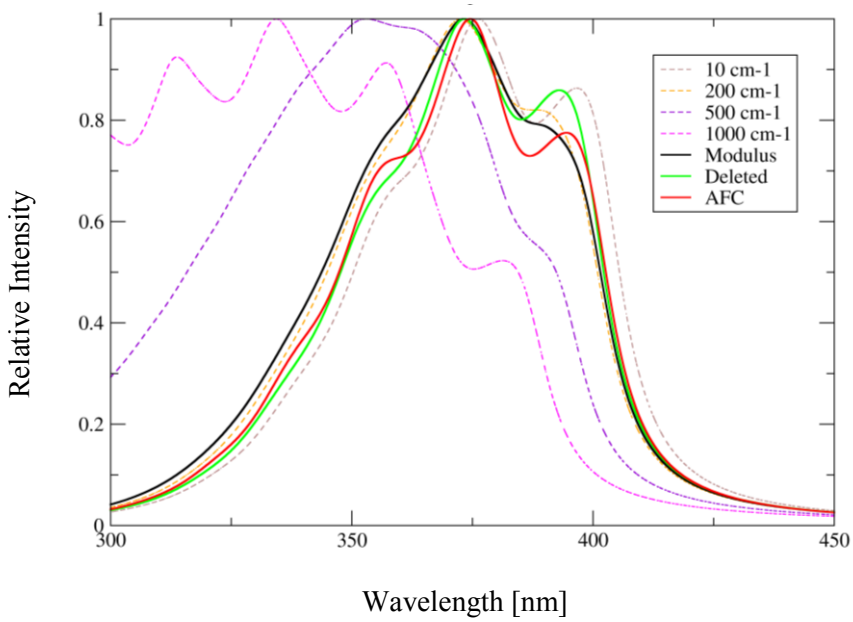
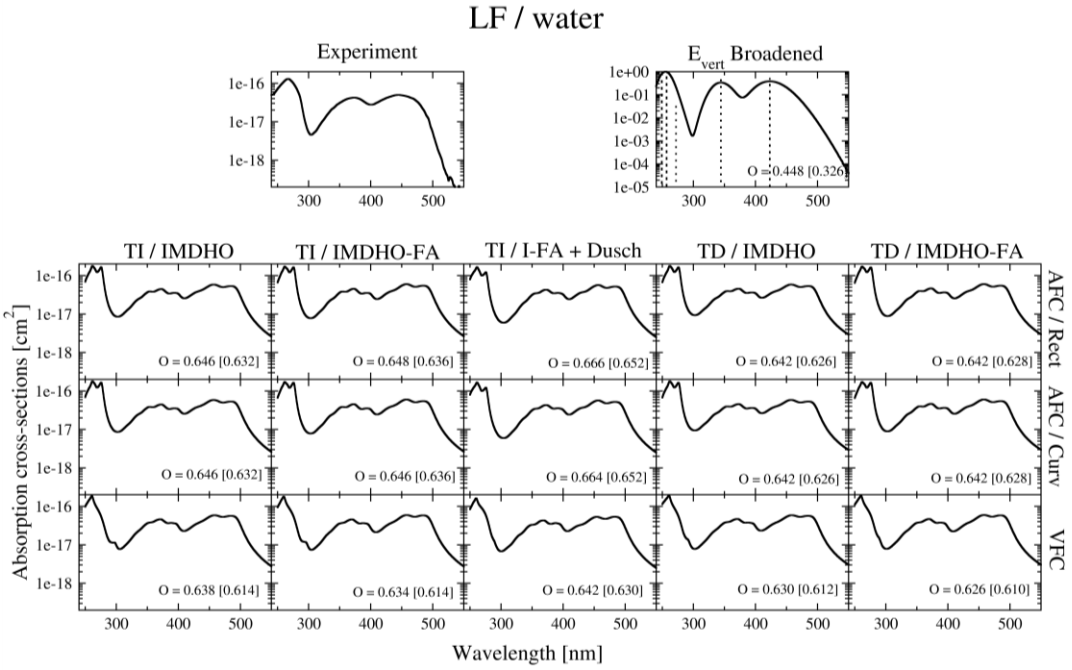


Figure 4.6. Treating the imaginary-frequency modes within the VFC scheme: CAM-B3LYP TD-VFC spectra of LF/S_1 in the gas phase, in which the imaginary frequency modes (i.e. 64 and 382*i* cm^{-1}) are deleted or replaced with an arbitrary positive number (10, 200, 500, 1000 cm^{-1}) or its modulus. The corresponding AFC spectrum is also given as a reference.

In the second part of our study [213], we extended our investigation to the full absorption spectra of LF, RoLF and 5TLF in the gas phase, water, and benzene (including bands arising from higher excitations). For brevity, we will present here only the water results at the CAM-B3LYP level (**Figure 4.7**) and summarize the key conclusions from the in-depth analyses available in the paper. Before starting the discussion of the computed spectra, the following points should be clarified. 5TLF can adopt two redox states in water: neutral (reduced N3) and anionic (oxidized N3). As it was not clear which form is the

dominant absorbing species in water at neutral pH [270], we considered both forms when simulating the spectra of 5TLF. By contrast, LF and RoLF can adopt three different oxidation states depending on the pH value: fully oxidized/neutral, reduced/ionic, and radical semiquinone. We considered them in the neutral form, which is dominant at pH=7.0 [168].

Generally speaking, we found for all chromophores that the simple (and commonly used) approach of Gaussian-broadening vertical spectra ("E_{vert} Broadened" in **Figure 4.7**) roughly reproduces the measured reference spectra but often gives only featureless bands and rather inaccurate maxima. Accordingly, for a realistic simulation of band shapes, one must account for the vibrational progressions that can be provided by FC methodology.



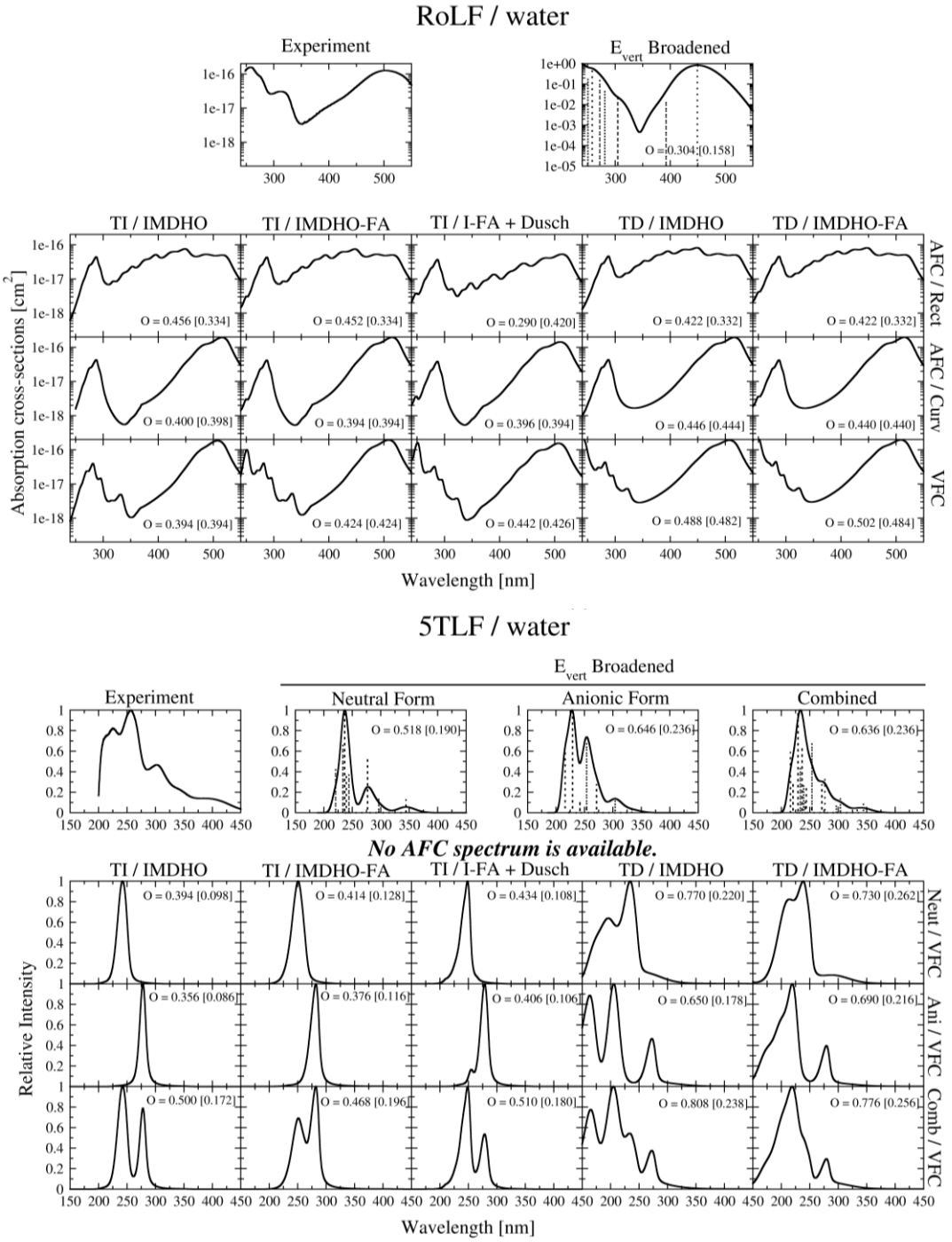


Figure 4.7. Convoluted vertical absorption spectra ("E_{vert} Broadened") along with TI-AFC, TI-VFC, TD-AFC, and TD-VFC spectra of LF, RoLF and 5TLF (neutral and anionic forms and their 1:1 mixture) computed using IMDHO or IMDHO-FA models and rectilinear or curvilinear displacements (where applicable). The FC spectra are computed at 0.01 K in

water compared to experiment at 298 K, taken from ref [55,270]. A Duschinsky treatment was applied only for TI/IMDHO-FA spectra. Optimized [original] O values are also given for computed spectra. All computed spectra are based on CAM-B3LYP geometries and Hessian and are *MRCI-corrected*. Spectra based on vertical transition energies were Gaussian-broadened with an HWHM of 1500 cm^{-1}; the FC spectra were Lorentzian-broadened with an HWHM of 403 cm^{-1}. Only VFC spectra are available for 5TLF.

As apparent from **Figure 4.7** (top panel), the computed FC spectra of LF are quite similar for different FC variants. This can be ascribed to the molecular rigidity of LF (C_s symmetry), which tends to suppress out-of-plane motions that would cause structural alterations and take the chromophore away from the FC point. However, C_s symmetry does not always guarantee the successful application of the AFC scheme, as it may not be possible to locate the RM on the target ES surface (particularly for high-lying states, as shown for the case of LF in benzene [213]). In contrast, the VFC approach can reproduce the absorption bands originating from these high-lying excited states, as it does not require the RM, and the resulting VFC spectra are overall of the same (or often higher) quality as the corresponding AFC spectra (if available).

In RoLF, the DMA group breaks the C_s symmetry and may rotate after photon absorption. As apparent from **Figure 4.7** (middle panel), the choice of the FC variant now starts to matter. Specifically, we note major changes in the AFC spectra when using different coordinate sets (Cartesian or internal coordinates) for generating the Hessian, Duschinsky mixing, and shift vectors. Even though the Duschinsky treatment helps to some extent, the Cartesian treatment fails to reproduce the bell-like shape in the measured band associated with the RoLF/S_1 state. By contrast, the AFC simulation in internal coordinates describes the spectral topology quite well since it provides a better description for the curvilinear motion during DMA twisting. On the other hand, the VFC approach does not depend on the coordinates (as shown previously) and can simulate the experimental shape very well (even better than in the AFC/internal case). Overall, the TD method performs somewhat better than the TI method, with both the VFC and AFC schemes. In benzene solution (data not shown here), only the combination of the TD method and the VFC approach was capable of reproducing the slow decrease in the tail part of the first experimental band [55].

4.4. Benchmark of FC methods

Finally, we analyzed the FC spectra for the challenging case of 5TLF, a highly asymmetric chromophore that undergoes substantial structural change upon photoexcitation (**Figure 4.4**). Consequently, we could not obtain most of the RMs on the ES surfaces and the corresponding AFC spectra. The VFC approach thus turns out to be the only option within the FC framework for 5TLF, because it avoids the error-prone excited-state geometry optimizations. Even in the VFC scheme, reliable TI spectra could not be obtained since the EM are too far away from the FC point in orthonormal space, which causes low overlap with the ground-state wave function and high shifts. The short-time approximation in the TD formalism in combination with the IMDHO-FA model appears to (partly) overcome this problem. The main features of the measured absorption spectrum can be reproduced reasonably well at this level (especially when assuming a 1:1 mixture of the neutral and anionic forms) although the agreement is less good than in the simpler cases of LF and RoLF.

On the basis of our comprehensive investigation of the three flavin derivatives considered (LF, RoLF and 5TLF), we conclude that the combination of the VFC approach with the TD formalism and the IMDHO-FA model offers the best overall performance.

Chapter 5

Concluding Remarks

The primary goal of this thesis was to address three important topics in the extensive interdisciplinary flavin research from the standpoint of computational (bio)chemistry. In our study, we applied state-of-the-art classical MD, QM, and QM/MM treatments, in particular (TD-)DFT and DFT/MRCI methods.

In the first part, our investigations focused on the catalytic properties of flavins that are employed as redox cofactors by flavin-dependent enzymes to achieve countless functions, including demethylation of amine substrates. Amine demethylation involves oxidation of the substrate amine in the rate-limiting step, in which a hydride-equivalent is transferred from the given substrate to the cofactor flavin (FMN or FAD). Amine oxidation is a fundamental reaction, being a common step in various types of (bio)chemical processes. As a specific example, the methylation of specific lysine residues on histone tails is one of the many types of post-translational modifications that regulate vital cellular processes, and aberrant levels of it have been linked with severe cancer types. In view of this, a molecular level understanding of the underlying operating mechanism of lysine demethylation, mediated by lysine-specific-demethylase 1 (LSD1), is of high importance. To this end, we studied the rate-determining oxidation step of demethylation of a dimethyl lysine (sLys) within the LSD1 binding pocket, where FAD is noncovalently bound. Our computational results and analysis show that LSD1 favors a direct hydride transfer (HT) mechanism over other proposals, including radical (SET) and adduct forming/polar nucleophilic (PN) mechanisms. We also found two different binding modes of the sLys substrate by means of MD simulations, and the QM/MM energy profiles indicate that the downward orientation is more reactive than the upward one. In addition, we scrutinized the roles of several active-

site residues and the water bridge motif (Lys-H_2O-N5), which have a direct or indirect effect on the oxidation and are mostly conserved across flavoprotein oxidases.

In a second study, we investigated the effects of covalent flavin binding on amine oxidation. We addressed the oxidation step in the demethylation of N-methyltryptophan (NMT) catalyzed by N-methyltryptophan oxidase (MTOX), which binds FAD covalently (unlike LSD1) as the redox coenzyme. Our QM/MM results confirm that the HT pathway is most facile (like in the case of LSD1), and they indicate that the covalent binding of FAD facilitates its isoalloxazine moiety of FAD to adopt a more bent geometry, which enables proper binding of the substrate NMT in its anionic form. In addition, we analyzed the roles of the active-site residues in MTOX and discussed the differences with respect to LSD1.

Flavin-dependent amine oxidases employ molecular oxygen to restore the reduced flavin cofactor and thus to prepare it for the next amine oxidation. The exact catalytic mechanism of O_2 activation by the partially or fully reduced flavin is still elusive, and several proposals for different flavoprotein types are available in the literature. A QM/MM investigation of this oxidative half-reaction would clearly be valuable. Work along these lines is currently underway in our group.

In the second part of the thesis, we explored the excited-state properties of flavin derivatives using TD-DFT and DFT/MRCI methods. We focused on the photophysical characteristics of roseoflavin, a natural riboflavin analog, and of a set of artificial RF derivatives in two separate studies. In the first one, we analyzed the intramolecular charge transfer (ICT) phenomenon in RoF, a donor-acceptor (D-A) type chromophore. RoF differs from RF by a dimethylamino (DMA) group replacing the methyl group bound to C8. The donor DMA group converts RoF into a D-A chromophore, with isoalloxazine moiety being the acceptor. Upon photoexcitation the DMA substituent may rotate, which leads to (at least) three different zwitterions: planar (PICT), perpendicular-twisted (TICT), and wagging (WICT), as in a typical D-A molecule. On the basis of our results, we assigned the mechanism involving a TICT zwitterion as the most viable one, and we also provided molecular-level explanations for experimental observations, for example why the ICT rates are higher in more polar solvents.

In the second study, we presented our efforts to find artificial, yet chemically stable, lumiflavin (LF) derivatives with absorption maxima in the blue-green region of the visible spectrum that might be used as alternative cofactors in flavin-dependent blue-light photoreceptors (e.g. with a LOV domain). The search for a replacement of native FMN in

blue-light receptors is motivated by the aim to initiate vital light-induced cell-regulatory processes at lower energies, and hence in a less destructive manner for the cell. With this in mind, we computed vertical and adiabatic excitation energies for eight LF derivatives (DXLF, 6A-LF, 9A-LF, 6,9-DA-LF, 2T-LF, 4T-LF, 2,4DT-LF, and 2TL-3DH-LF), as well as the intersystem crossing (ISC) rates for the spin-forbidden singlet→triplet transitions in the most promising thioflavins (2T-LF, 4T-LF, and 2,4DT-LF). Among these, 2T-LF appears to be the best candidate since it resembles native LF well in terms of electronic properties, is stable, has a longer-lived triplet state than LF, and absorbs in the blue-green region (ca. 500 nm). Computationally, the next step would now be to incorporate 2T-LF into a genuine LOV domain (replacing native FMN) and to perform a QM/MM study to see whether 2T-LF would properly bind and show the desired photophysical properties in the protein environment as well.

In the third part of the thesis, we turned our attention to the spectral properties of flavin derivatives and the techniques required to simulate their optical spectra accurately. In this connection, we reviewed assorted Franck-Condon (FC) based methods that have been frequently used to compute vibrationally broadened spectra of diverse chromophores. This survey of FC methods uncovered the need for a novel technique, namely the combination of vertical FC (VFC) approach and internal coordinates (with curvilinear displacements). We implemented this technique in our first study, and using a set of seven model systems, we showed that the use of internal rather than Cartesian coordinates leads to only slight differences in the computed VFC spectra, but may significantly improve the adiabatic FC (AFC) spectra. In our second study, we continued with a broader benchmark of various FC methods, including the combination of time-independent (TI) and time-dependent (TD) formalisms with the AFC or VFC scheme using the independent mode displaced harmonic oscillator model without or with frequency alteration (IMDHO/IMDHO-FA), and accounting for Duschinsky rotation when possible. We employed both Cartesian and internal coordinates for computing the Duschinsky matrix, along with the required Hessians, adimensional shifts, reorganization energies, and FC factors. For benchmarking, we chose three RF analogs (LF, roseolumiflavin RoLF, and 5-thiolumiflavin 5TLF) in view of their interesting structural properties, the availability of experimental UV/vis spectra, and their biological importance. The comparison between the results for LF (C_s symmetry, rather rigid), RoLF (essentially planar rings with rotatable DMA group), and 5TLF (asymmetric, "butterfly-wing" ring geometry) enabled us to assess the performance of FC methods in a

systematic manner. On the basis of our comprehensive computational results, we conclude that TD/IMDHO-FA within the VFC framework is overall the best combination for reproducing the measured absorption spectra of the flavin derivatives investigated. In future work, this conclusion should be tested for a larger test set of chromophores of different types. It would also be desirable to study the few missing aspects in the current benchmark (e.g. using Duschinsky mixing in the TD formalism and the Herzberg-Teller effect). We anticipate that the TD/IMDHO-FA/VFC combination will remain the preferred approach for simulating vibronic band shapes in the optical spectra of large chromophores.

Bibliography

[1] P. Macheroux, B. Kappes, S.E. Ealick, FEBS J. 278 (2011) 2625.

[2] S.O. Mansoorabadi, C.J. Thibodeaux, H. Liu, J. Org. Chem. 72 (2007) 6329.

[3] V. Massey, Biochem. Soc. Trans. 28 (2000) 283.

[4] R. Miura, Chem. Rec. 1 (2001) 183.

[5] A.W. Blyth, J. Chem. Soc. Trans. 35 (1879) 530.

[6] A. Bacher, K. Becker, W.J.H. van Berkel, R. Bittl, R. Brosi, A. Bury, T. Domratcheva, A.M. Edwards, W. Eisenreich, C. Engelhard, P. Ferreira, M. Fischer, T. Gräwert, I. Haase, S. Hay, K.J. Hellingwerf, B. Illarionov, F. Jankowitsch, E. Jortzik, H. Kandori, J.T.M. Kennis, T. Kitagawa, S. Langer, N.G.H. Leferink, J. Li, J. Ma, M. Mack, M. Martínez-Júlvez, T. Mathes, M. Medina, A.-F. Miller, F. Müller, S. Nakanishi, R.F. Pauszek, D.B. Pedrolli, C.R. Pudney, J. Schwarz, N.S. Scrutton, A.R.M. Shahi, R.J. Stanley, I.H.M. van Stokkum, A. Udvarhelyi, L. Wang, D. Yamada, Flavins and Flavoproteins: Methods and Protocols, Springer, New York, NY, 2014.

[7] J.P. Klinman, Acc. Chem. Res. 40 (2007) 325.

[8] C.A. McDonald, R.L. Fagan, F. Collard, V.M. Monnier, B.A. Palfey, J. Am. Chem. Soc. 133 (2011) 16809.

[9] O. Dym, D. Eisenberg, Protein Sci. 10 (2001) 1712.

[10] D.P.H.M. Heuts, N.S. Scrutton, W.S. McIntire, M.W. Fraaije, FEBS J. 276 (2009) 3405.

[11] M.M.E. Huijbers, S. Montersino, A.H. Westphal, D. Tischler, W.J.H. van Berkel, Arch. Biochem. Biophys. 544 (2014) 2.

[12] W.P. Dijkman, G. de Gonzalo, A. Mattevi, M.W. Fraaije, Appl. Microbiol. Biotechnol. 97 (2013) 5177.

[13] P.F. Fitzpatrick, Arch. Biochem. Biophys. 493 (2010) 13.

[14] R.A. Varier, H.T.M. Timmers, Biochim. Biophys. Acta 1815 (2011) 75.

[15] S.E. Elsheikh, A.R. Green, E.A. Rakha, D.G. Powe, R.A. Ahmed, H.M. Collins, D. Soria, J.M. Garibaldi, C.E. Paish, A.A. Ammar, M.J. Grainge, G.R. Ball, M.K. Abdelghany, L. Martinez-Pomares, D.M. Heery, I.O. Ellis, Cancer Res. 69 (2009) 3802.

[16] J. Ellinger, P. Kahl, J. von der Gathen, S. Rogenhofer, L.C. Heukamp, I. Gütgemann, B. Walter, F. Hofstädter, R. Büttner, S.C. Müller, P.J. Bastian, A. von Ruecker, Prostate 70 (2010) 61.

[17] F. Barlési, G. Giaccone, M.I. Gallegos-Ruiz, A. Loundou, S.W. Span, P. Lefesvre, F.A.E. Kruyt, J.A. Rodriguez, J. Clin. Oncol. 25 (2007) 4358.

[18] C. Magerl, J. Ellinger, T. Braunschweig, E. Kremmer, L.K. Koch, T. Höller, R. Büttner, B. Lüscher, I. Gütgemann, Hum. Pathol. 41 (2010) 181.

[19] Y.-J. Shi, F. Lan, C. Matson, P. Mulligan, J. Whetstine, P. Cole, R.A. Casero, Y. Shi, Cell 119 (2004) 941.

[20] H.-U. Schildhaus, R. Riegel, W. Hartmann, S. Steiner, E. Wardelmann, S. Merkelbach-Bruse, S. Tanaka, H. Sonobe, R. Schüle, R. Buettner, J. Kirfel, Hum. Pathol. 42 (2011) 1667.

[21] D.B. Seligson, S. Horvath, M.A. McBrian, V. Mah, H. Yu, S. Tze, Q. Wang, D. Chia, L. Goodglick, S.K. Kurdistani, Am. J. Pathol. 174 (2009) 1619.

[22] S. Lim, A. Janzer, A. Becker, A. Zimmer, R. Schüle, R. Buettner, J. Kirfel, Carcinogenesis 31 (2010) 512.

[23] Y.-C. Zheng, Y.-C. Duan, J.-L. Ma, R.-M. Xu, X. Zi, W.-L. Lv, M.-M. Wang, X.-W. Ye, S. Zhu, D. Mobley, Y.-Y. Zhu, J.-W. Wang, J.-F. Li, Z.-R. Wang, W. Zhao, H.-M. Liu, J. Med. Chem. 56 (2013) 8543.

[24] L.M. Szewczuk, J.C. Culhane, M. Yang, A. Majumdar, H. Yu, P. a Cole, Biochemistry 46 (2007) 6892.

[25] D.M. Gooden, D.M.Z. Schmidt, J.A. Pollock, A.M. Kabadi, D.G. McCafferty, Bioorg. Med. Chem. Lett. 18 (2008) 3047.

[26] J.C. Culhane, L.M. Szewczuk, X. Liu, G. Da, R. Marmorstein, P.A. Cole, J. Am. Chem. Soc. 128 (2006) 4536.

Bibliography

[27] A. Spannhoff, A.-T. Hauser, R. Heinke, W. Sippl, M. Jung, ChemMedChem 4 (2009) 1568.

[28] V. Massey, G. Palmer, Biochemistry 5 (1966) 3181.

[29] A. Losi, W. Gärtner, Photochem. Photobiol. (2011) 491.

[30] T.E. Swartz, S.B. Corchnoy, J.M. Christie, J.W. Lewis, I. Szundi, W.R. Briggs, R.A. Bogomolni, J. Biol. Chem. 276 (2001) 36493.

[31] A. Möglich, X. Yang, R.A. Ayers, K. Moffat, Plant Biol. 61 (2010) 21.

[32] A. Möglich, K. Moffat, J. Mol. Biol. 373 (2007) 112.

[33] Mario R. Silva-Junior, M. Mansurova, W. Gärtner, W. Thiel, ChemBioChem 14 (2013) 1648.

[34] J. Kennis, T. Mathes, Interface Focus 3 (2013).

[35] T.E. Swartz, S.B. Corchnoy, J.M. Christie, J.W. Lewis, I. Szundi, W.R. Briggs, R.A. Bogomolni, J. Biol. Chem. 276 (2001) 36493.

[36] S. Salzmann, M.R. Silva-Junior, W. Thiel, C.M. Marian, J. Phys. Chem. B 113 (2009) 15610.

[37] V. Massey, J. Biol. Chem. 269 (1994) 22459.

[38] G. Gadda, Biochemistry 51 (2012) 2662.

[39] A. Mattevi, Trends Biochem. Sci. 31 (2006) 276.

[40] V. Massey, Int. Congr. Ser. 1233 (2002) 3.

[41] S. Ghisla, V. Massey, Eur. J. Biochem. 181 (1989) 1.

[42] A. Losi, Photochem. Photobiol. 83 (2007) 1283.

[43] J. Baier, T. Maisch, M. Maier, E. Engel, M. Landthaler, W. Bäumler, Biophys. J. 91 (2006) 1452.

[44] S.D.M. Islam, T. Susdorf, A. Penzkofer, P. Hegemann, Chem. Phys. 295 (2003) 137.

[45] M.W. Fraaije, A. Mattevi, Trends Biochem. Sci 25 (2000) 126.

[46] N.S. Moyon, S. Mitra, J. Phys. Chem. A 115 (2011) 2456.

[47] S.-H. Song, P.L. Freddolino, A.I. Nash, E.C. Carroll, K. Schulten, K.H. Gardner, D.S. Larsen, Biochemistry 50 (2011) 2411.

[48] S. Raffelberg, M. Mansurova, W. Gärtner, A. Losi, J. Am. Chem. Soc. 133 (2011) 5346.

[49] I.A. Solov'yov, T. Domratcheva, A.R. Moughal Shahi, K. Schulten, J. Am. Chem. Soc. 134 (2012) 18046.

[50] A. Udvarhelyi, T. Domratcheva, J. Phys. Chem. B 117 (2013) 2888.

[51] P. Zirak, A. Penzkofer, T. Mathes, P. Hegemann, J. Photochem. Photobiol. B Biol. 97 (2009) 61.
[52] A. Hassan-Abdallah, G. Zhao, M.S. Jorns, Biochemistry 45 (2006) 9454.
[53] M.W. Fraaije, R.H.H. van den Heuvel, W.J.H. van Berkel, A. Mattevi, J. Biol. Chem. 274 (1999) 35514.
[54] A. Tyagi, P. Zirak, A. Penzkofer, T. Mathes, P. Hegemann, M. Mack, S. Ghisla, Chem. Phys. 364 (2009) 19.
[55] P. Zirak, A. Penzkofer, T. Mathes, P. Hegemann, Chem. Phys. 358 (2009) 111.
[56] M. Mack, S. Grill, Appl. Microbiol. Biotechnol. 71 (2006) 265.
[57] K. Dudley, A. Ehrenberg, P. Hemmerich, F. Müller, Helv. Chim. Acta 47 (1964) 1354.
[58] V. Massey, A. Claiborne, M. Biemann, S. Ghisla, J. Biol. Chem. 259 (1984) 9667.
[59] S. Ghisla, V. Massey, Biochem. J. 239 (1986) 1.
[60] J.R. Miller, D.E. Edmondson, J. Biol. Chem. 274 (1999) 23515.
[61] T. Mathes, C. Vogl, J. Stolz, P. Hegemann, J. Mol. Biol. 385 (2009) 1511.
[62] N.S. Scrutton, Nat. Prod. Rep. 21 (2004) 722.
[63] F. Forneris, E. Battaglioli, A. Mattevi, C. Binda, FEBS J. 276 (2009) 4304.
[64] B.C. Smith, J.M. Denu, Biochim. Biophys. Acta 1789 (2009) 45.
[65] C. Wood, A. Snijders, J. Williamson, C. Reynolds, J. Baldwin, M. Dickman, FEBS J. 276 (2009) 3685.
[66] J.C. Culhane, P.A. Cole, Curr. Opin. Chem. Biol. 11 (2007) 561.
[67] P. Hu, Y. Zhang, J. Am. Chem. Soc. 128 (2006) 1272.
[68] F. Lizcano, J. Garcia, Pharmaceuticals (Basel). 5 (2012) 963.
[69] N.A. Vellore, R. Baron, ChemMedChem 9 (2014) 484.
[70] J. Duerre, C. Lee, J. Neurochem. 23 (1974) 541.
[71] P. Byvoet, Arch. Biochem. Biophys. 152 (1972) 887.
[72] F. Forneris, C. Binda, E. Battaglioli, A. Mattevi, Trends Biochem. Sci. 33 (2008) 181.
[73] H. Gaweska, P.F. Fitzpatrick, BioMol Concepts 2 (2011) 365.
[74] E.C. Ralph, P.F. Fitzpatrick, Biochemistry 44 (2005) 3074.
[75] H. Gaweska, M. Henderson Pozzi, D.M.Z. Schmidt, D.G. McCafferty, P.F. Fitzpatrick, Biochemistry 48 (2009) 5440.

Bibliography

[76] F. Forneris, C. Binda, A.D. Aglio, M.W. Fraaije, E. Battaglioli, A. Mattevi, A. Dall'Aglio, J. Biol. Chem. 281 (2006) 35289.

[77] R.C. Bruckner, J. Winans, M.S. Jorns, Biochemistry 50 (2011) 4949.

[78] P. Khanna, M. Jorns, Biochemistry 40 (2001) 1441.

[79] R.B. Silverman, Acc. Chem. Res. 28 (1995) 335.

[80] R.B.R. Silverman, S.S.J. Hoffman, W.B.W. Catus, J. Am. Chem. Soc. 102 (1980) 7126.

[81] R.B. Silverman, Prog. Brain Res. 106 (1995) 23.

[82] J.R. Miller, D.E. Edmondson, Biochemistry 38 (1999) 13670.

[83] E.C. Ralph, J.S. Hirschi, M.A. Anderson, W.W. Cleland, D.A. Singleton, P.F. Fitzpatrick, Biochemistry 46 (2007) 7655.

[84] D.M.Z. Schmidt, D.G. McCafferty, Biochemistry 46 (2007) 4408.

[85] A.E. Mccann, N.S. Sampson, J. Am. Chem. Soc. 122 (2000) 35.

[86] Z. Chen, G. Zhao, S. Martinovic, Biochemistry (2005) 15444.

[87] P. Khanna, M. Schuman Jorns, Biochemistry 40 (2001) 1451.

[88] M.A. Wagner, M.S. Jorns, Biochemistry 39 (2000) 8825.

[89] F. Forneris, C. Binda, M.A. Vanoni, A. Mattevi, E. Battaglioli, FEBS Lett. 579 (2005) 2203.

[90] G. Zhao, M. Jorns, Biochemistry 45 (2006) 5985.

[91] R. V Dunn, A.W. Munro, N.J. Turner, S.E.J. Rigby, N.S. Scrutton, Chembiochem 11 (2010) 1228.

[92] A. Karytinos, F. Forneris, A. Profumo, G. Ciossani, E. Battaglioli, C. Binda, A. Mattevi, J. Biol. 284 (2009) 17775.

[93] M.G. Lee, C. Wynder, N. Cooch, R. Shiekhattar, Nature 437 (2005) 432.

[94] F. Polticelli, J. Basran, C. Faso, A. Cona, G. Minervini, R. Angelini, R. Federico, N.S. Scrutton, P. Tavladoraki, Biochemistry 44 (2005) 16108.

[95] M.H. Pozzi, V. Gawandi, P.F. Fitzpatrick, Biochemistry 48 (2009) 1508.

[96] M. Henderson Pozzi, P.F. Fitzpatrick, M.H. Pozzi, Arch. Biochem. Biophys. 498 (2010) 83.

[97] F. Forneris, C. Binda, A. Adamo, E. Battaglioli, A. Mattevi, J. Biol. Chem. 282 (2007) 20070.

[98] Y. Chen, Y. Yang, F. Wang, K. Wan, K. Yamane, Y. Zhang, M. Lei, Proc. Natl. Acad. Sci. U.S.A. 103 (2006) 13956.

[99] A. Ilari, A. Bonamore, S. Franceschini, A. Fiorillo, A. Boffi, G. Colotti, Proteins 71 (2008) 2065.

[100] G. Zhao, M.S. Jorns, Biochemistry 44 (2005) 16866.

[101] M. Li, C. Binda, A. Mattevi, D.E. Edmondson, Biochemistry 45 (2006) 4775.

[102] B. Karasulu, M. Patil, W. Thiel, J. Am. Chem. Soc. 135 (2013) 13400.

[103] B. Karasulu, W. Thiel, ACS Catalysis, Submitted.

[104] P. Stavropoulos, G. Blobel, A. Hoelz, Nat. Struct. Mol. Biol. 13 (2006) 626.

[105] G.M. Morris, R. Huey, W. Lindstrom, M.F. Sanner, R.K. Belew, D.S. Goodsell, A.J. Olson, J. Comput. Chem. 30 (2010) 2785.

[106] O. Trott, A.J. Olson, J. Comput. Chem. 31 (2010) 455.

[107] G. Zhao, R.C. Bruckner, M.S. Jorns, Biochemistry 47 (2008) 9124.

[108] M. Olsson, C. Sondergaard, M. Rostkowski, J. Jensen, J. Chem. Theory Comput. 7 (2011) 525.

[109] M. Rostkowski, M.H.M. Olsson, C.R. Søndergaard, J.H. Jensen, BMC Struct. Biol. 11 (2011) 6.

[110] J.C. Gordon, J.B. Myers, T. Folta, V. Shoja, L.S. Heath, A. Onufriev, "H++: a server for estimating pK_as and adding missing hydrogens to macromolecules", Nucleic Acids Res. 33 (2005), W368-71.

[111] B.R. Brooks, C.L. Brooks, A.D. Mackerell, L. Nilsson, R.J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caflisch, L. Caves, Q. Cui, A.R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R.W. Pastor, C.B. Post, J.Z. Pu, M. Schaefer, B. Tidor, R.M. Venable, H.L. Woodcock, X. Wu, W. Yang, D.M. York, M. Karplus, "CHARMM: the biomolecular simulation program", J. Comput. Chem. 30 (2009) 1545.

[112] A.D. MacKerell,, D. Bashford, R.L. Dunbrack,, J.D. Evanseck, M.J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F.T.K. Lau, C. Mattos, S. Michnick, T. Ngo, D.T. Nguyen, B. Prodhom, W.E. Reiher, B. Roux, M. Schlenkrich, J.C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiórkiewicz-Kuczera, D. Yin, M. Karplus, J. Phys. Chem. B 102 (1998) 3586.

[113] W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey, M.L. Klein, J. Chem. Phys. 79 (1983) 926.

[114] G. Luo, I. Andricioaei, X.S. Xie, M. Karplus, J. Phys. Chem. B 110 (2006) 9363.

Bibliography

[115] C. Grauffel, R.H. Stote, A. Dejaegere, J. Comput. Chem. 31 (2010) 2434.

[116] M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G.A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H.P. Hratchian, A.F. Izmaylov, J. Bloino, G. Zheng, J.L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J.A.M. Jr., J.E. Peralta, F. Ogliaro, M. Bearpark, J.J. Heyd, E. Brothers, K.N. Kudin, V.N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J.C. Burant, S.S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J.M. Millam, M. Klene, J.E. Knox, J.B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R.E. Stratmann, O. Yazyev, A.J. Austin, R. Cammi, C. Pomelli, J.W. Ochterski, R.L. Martin, K. Morokuma, V.G. Zakrzewski, G.A. Voth, P. Salvador, J.J. Dannenberg, S. Dapprich, A.D. Daniels, O. Farkas, J.B. Foresman, J. V. Ortiz, J. Cioslowski, D.J. Fox, Gaussian09 Rev. D.01, (2013).

[117] A. Becke, J. Chem. Phys. 98 (1993) 5648.

[118] S. Grimme, J. Comput. Chem. 27 (2006) 1787.

[119] Y. Zhao, D.G. Truhlar, Theor. Chem. Acc. 120 (2007) 215.

[120] O.A. Vydrov, G.E. Scuseria, J. Chem. Phys. 125 (2006) 234109.

[121] B.J. Lynch, P.L. Fast, M. Harris, D.G. Truhlar, J. Phys. Chem. A 104 (2000) 4811.

[122] J.-D. Chai, M. Head-Gordon, Phys. Chem. Chem. Phys. 10 (2008) 6615.

[123] R. Ditchfield, J. Chem. Phys. 54 (1971) 724.

[124] W.J. Hehre, J. Chem. Phys. 56 (1972) 2257.

[125] A. Schäfer, C. Huber, R. Ahlrichs, J. Chem. Phys. 100 (1994) 5829.

[126] M. Cossi, N. Rega, G. Scalmani, V. Barone, J. Comput. Chem. 24 (2003) 669.

[127] P. Sherwood, A.H. de Vries, M.F. Guest, G. Schreckenbach, C.R.A. Catlow, S.A. French, A.A. Sokol, S.T. Bromley, W. Thiel, A.J. Turner, S. Billeter, F. Terstegen, S. Thiel, J. Kendrick, S.C. Rogers, J. Casci, M. Watson, F. King, E. Karlsen, M. Sjøvoll, A. Fahmi, A. Schaefer, C. Lennartz, ChemShell QM/MM Interface, J. Mol. Struct. 632 (2003) 1.

[128] R. Ahlrichs, M. Bär, H. –P. Baron, S. Bauern- schmitt, S. Böcker, M. Ehrig, K. Eichkorn, S. Elliott, F. Furche, F. Haase, M. Häser, H. Horn, C. Huber, U. Huniar, M. Kattannek, C. Kölmel, M. Kollwitz, K. May, C. Ochsenfeld, H. Öhm, A. Schäfer,

U. Schneider, O. Treutler, M. von Arnim, F. Weigend, P. Weis, H. Weiss, Turbomole v.6.3 (2011).

[129] W. Smith, T.R.R. Forester, J. Mol. Graph. 14 (1996) 136.

[130] A.H. de Vries, P. Sherwood, S.J. Collins, A.M. Rigby, M. Rigutto, G.J. Kramer, J. Phys. Chem. B 103 (1999) 6133.

[131] S.R.S. Billeter, A.A.J. Turner, W. Thiel, Phys. Chem. Chem. Phys. 2 (2000) 2177.

[132] A. Banerjee, N. Adams, J. Simons, R. Shepard, J. Phys. Chem. 89 (1985) 52.

[133] J. Baker, J. Comput. Chem. 7 (1986) 385.

[134] J. Nocedal, Math. Comput. 35 (1980) 773.

[135] D.C. Liu, J. Nocedal, Math. Program. 45 (1989) 503.

[136] J. Huang, R. Sengupta, A.A.B. Espejo, M.M.G. Lee, J. Dorsey, M. Richter, S. Opravil, R. Shiekhattar, M.T. Bedford, T. Jenuwein, S.L. Berger, Nature 449 (2007) 105.

[137] J. Wang, S. Hevi, J.K. Kurash, H. Lei, F. Gay, J. Bajko, H. Su, W. Sun, H. Chang, G. Xu, F. Gaudet, E. Li, T. Chen, Nat. Genet. 41 (2009) 125.

[138] H.-S. Cho, T. Suzuki, N. Dohmae, S. Hayami, M. Unoki, M. Yoshimatsu, G. Toyokawa, M. Takawa, T. Chen, J.K. Kurash, H.I. Field, B.A.J. Ponder, Y. Nakamura, R. Hamamoto, Cancer Res. 71 (2011) 655.

[139] M. Yang, J.C. Culhane, L.M. Szewczuk, P. Jalili, H.L. Ball, M. Machius, P.A. Cole, H. Yu, Biochemistry 46 (2007) 8058.

[140] B. Lohse, J.L. Kristensen, L.H. Kristensen, K. Agger, K. Helin, M. Gajhede, R.P. Clausen, Bioorg. Med. Chem. 19 (2011) 3625.

[141] M. Lingwood, J.R. Hammond, D.A. Hrovat, J.M. Mayer, W.T. Borden, J Chem Theory Comput. 2 (2006) 740.

[142] P. Trickey, M.A. Wagner, M.S. Jorns, F.S. Mathews, Structure 7 (1999) 331.

[143] Y. Koyama, H. Ohmori, Gene 181 (1996) 179.

[144] M.A. Wagner, P. Trickey, Z.W. Chen, F.S. Mathews, M.S. Jorns, Biochemistry 39 (2000) 8813.

[145] R. Vianello, M. Repič, J. Mavri, European J. Org. Chem. 2012 (2012) 7057.

[146] G. Zhao, H. Song, Z.-W. Chen, F.S. Mathews, M.S. Jorns, Biochemistry 41 (2002) 9751.

[147] P. Trickey, J. Basran, L. Lian, Z. Chen, J.D. Barton, M.J. Sutcliffe, N.S. Scrutton, F.S. Mathews, Biochemistry 39 (2000) 7678.

Bibliography

[148] S. Kasai, R. Miura, K. Matsui, Bull. Chem. Soc. Jpn. 48 (1975) 2877.

[149] S. Otani, M. Takatsu, M. Nakano, S. Kasai, R. Miura, K. Matsu, J. Antibiot. (Tokyo). 27 (1974) 88.

[150] P.S. Song, E.B. Walker, R.D. Vierstra, K.L. Poff, Photochem. Photobiol. 32 (1980) 393.

[151] S. Shinkai, K. Kameoka, N. Honda, K. Ueda, O. Manabe, J. Lindsey, Bioorg. Chem 14 (1986) 119.

[152] S. Kasai, Y. Kubo, S. Yamanaka, T. Hirota, H. Sato, Y. Tsuzukida, K. Matusi, J. Nutr. Sci. Vitaminol. (Tokyo). 24 (1978) 339.

[153] S. Kasai, S. Yamanaka, S.-C. Wang, K. Matsui, J. Nutr. Sci. Vitaminol. (Tokyo). 25 (1978) 289.

[154] S. Grill, H. Yamaguchi, H. Wagner, L. Zwahlen, U. Kusch, M. Mack, Arch. Microbiol. 188 (2007) 377.

[155] K. Shiga, Y. Nishina, I. Ohmine, K. Horiike, S. Kasai, K. Matsui, H. Watari, T. Yamano, J. Biochem. 87 (1980) 281.

[156] Y.Y.-K. Choe, S. Nagase, K. Nishimoto, J. Comput. Chem. 28 (2007) 727.

[157] B. Klaumünzer, D. Kröner, P. Saalfrank, J. Phys. Chem. B 114 (2010) 10826.

[158] T. Merz, K. Sadeghian, M. Schütz, Phys. Chem. Chem. Phys. 13 (2011) 14775.

[159] K. Sadeghian, M. Bocola, M. Schütz, Phys. Chem. Chem. Phys. 12 (2010) 8840.

[160] K. Sadeghian, M. Schütz, J. Am. Chem. Soc. 129 (2007) 4068.

[161] S. Salzmann, V. Martinez-Junza, B. Zorn, S.E.S.E. Braslavsky, M. Mansurova, C.M.C.M. Marian, W. Gärtner, W. Gärtner, J. Phys. Chem. A 113 (2009) 9365.

[162] C. Neiss, P. Saalfrank, M. Parac, S. Grimme, J. Phys. Chem. A 107 (2003) 140.

[163] S. Salzmann, J. Tatchen, C.M. Marian, J. Photochem. Photobiol. A Chem. 198 (2008) 221.

[164] A. Vdovin, A. Slenczka, B. Dick, Chem. Phys. 422 (2013) 195.

[165] A. Tyagi, A. Penzkofer, T. Mathes, P. Hegemann, J. Photochem. Photobiol. B. 101 (2010) 76.

[166] K. Matsui, S. Kasai; F. Müller (Ed.), Chem. Biochem. Flavoenzymes, CRC Press, Boca Raton FL, USA, 1991, pp. 105–120.

[167] E. Sikorska, I. Khmelinskii, W. Prukala, S. Williams, M. Patel, D. Worrall, J. Bourdelande, J. Koput, M. Sikorski, J. Phys. Chem. A 108 (2004) 1501.

[168] P. Drössler, W. Holzer, A. Penzkofer, P. Hegemann, Chem. Phys. 282 (2002) 429.

[169] A. Tyagi, A. Penzkofer, T. Mathes, P. Hegemann, Chem. Phys. 369 (2010) 27.
[170] A. Penzkofer, Chem. Phys. 400 (2012) 142.
[171] S. Shinkai, K. Kameoka, N. Honda, K. Ueda, O. Manabe, J. Chem. Soc. Chem. Commun. (1985) 673.
[172] M.K. Otto, M. Jayaram, R.M. Hamilton, M. Delbrück, Proc. Natl. Acad. Sci. U. S. A. 78 (1981) 266.
[173] S. Shinkai, Bioorg. Chem. 15 (1987) 269.
[174] V. Massey, P. Hemmerich, Biochem. Soc. Trans. 8 (1980) 246.
[175] D. Giegel, C. Williams, V. Massey, J. Biol. Chem. 265 (1990) 6626.
[176] B. Karasulu, W. Thiel, J. Phys. Chem. B In press (2014), DOI: 10.1021/jp506101x.
[177] C.M. Marian, N. Setsuko, V. Rai-Constapel, B. Karasulu, W. Thiel, J. Phys. Chem. B 118 (2014) 1743.
[178] M.M.N. Wolf, C. Schumann, R. Gross, T. Domratcheva, R. Diller, J. Phys. Chem. B 112 (2008) 13424.
[179] W.L. Cairns, D.E. Metzler, J. Am. Chem. Soc. 225 (1971) 2772.
[180] M. Sun, T.A. Moore, P.S. Song, J. Am. Chem. Soc. 94 (1972) 1730.
[181] A. Bowd, P. Byrom, J. Hudson, J. Turnbull, Photochem. Photobiol. 8 (1968) 1.
[182] F. Furche, J. Chem. Phys. 114 (2001) 5982.
[183] F. Furche, R. Ahlrichs, J. Chem. Phys. 117 (2002) 7433.
[184] R. Bauernschmitt, R. Ahlrichs, Chem. Phys. Lett. 256 (1996) 454.
[185] T. Yanai, D.P. Tew, N.C. Handy, Chem. Phys. Lett. 393 (2004) 51.
[186] A. Schäfer, H. Horn, R. Ahlrichs, J. Chem. Phys. 97 (1992) 2571.
[187] M. Francl, W. Pietro, W. Hehre, J. Binkley, M. Gordon, D. DeFrees, J. Pople, J. Chem. Phys. 77 (1982) 3654.
[188] C. Kind, M. Reiher, J. Neugebauer, (1999).
[189] V. Barone, M. Cossi, J. Phys. Chem. A 102 (1998) 1995.
[190] A. Schäfer, A. Klamt, D. Sattel, J.C.W. Lohrenz, F. Eckert, Phys. Chem. Chem. Phys. 2 (2000) 2187.
[191] P.E.M. Siegbahn, F. Himo, Wiley Interdiscip. Rev. Comput. Mol. Sci. 1 (2011) 323.
[192] E.L. Mertz, L.I. Krishtalik, Proc. Natl. Acad. Sci. U. S. A. 97 (2000) 2081.
[193] B. Klaumünzer, D. Kröner, H. Lischka, P. Saalfrank, Phys. Chem. Chem. Phys. 14 (2012) 8693.

Bibliography

[194] M.R. Silva-Junior, M. Schreiber, S.P.A. Sauer, W. Thiel, J. Chem. Phys. 129 (2008) 104103.

[195] S. Grimme, M. Waletzke, J. Chem. Phys. 111 (1999) 5645.

[196] C. Lee, W. Yang, R.G. Parr, Phys. Rev. B 37 (1988) 785.

[197] A.D. Becke, J. Chem. Phys. 98 (1993) 1372.

[198] M. Kleinschmidt, C.M. Marian, Chem. Phys. 311 (2005) 71.

[199] M. Kleinschmidt, J. Tatchen, C.M. Marian, J. Comput. Chem. 23 (2002) 824.

[200] B.A. Heß, C.M. Marian, U. Wahlgren, O. Gropen, Chem. Phys. Lett. 251 (1996) 365.

[201] J. Tatchen, C.M. Marian, Chem. Phys. Lett. 313 (1999) 351.

[202] D. Danovich, C.M. Marian, T. Neuheuser, S.D. Peyerimhoff, S. Shaik, J. Phys. Chem. A 102 (1998) 5923.

[203] M. Etinski, J. Tatchen, C.M. Marian, J. Chem. Phys. 134 (2011) 154105.

[204] M.A. El-Sayed, J. Chem. Phys. 38 (1963) 2834.

[205] J. Markham, Rev. Mod. Phys. 31 (1959) 956.

[206] F. Duschinsky, Acta Physicochim. URSS 7 (1937) 551.

[207] M. Kleinschmidt, J. Tatchen, C.M. Marian, J. Chem. Phys. 124 (2006) 124101.

[208] Z.R. Grabowski, K. Rotkiewicz, W. Rettig, Chem. Rev. 103 (2003) 3899.

[209] K. Zachariasse, J. Photochem. Photobiol. A Chem. 105 (1997) 373.

[210] K. Rotkiewicz, K. Grellmann, Z. Grabowski, Chem. Phys. Lett. 19 (1973) 315.

[211] K.A. Zachariasse, T. von der Haar, A. Hebecker, U. Leinhos, W. Kuhnle, Pure Appl. Chem. 65 (1993) 1745.

[212] J. Götze, B. Karasulu, W. Thiel, J. Chem. Phys. 139 (2013) 234108.

[213] B. Karasulu, J.P. Götze, W. Thiel, J. Chem. Theory Comput. In print (2014), DOI: 10.1021/ct500830a.

[214] R. Improta, A. Lami, V. Barone, F. Santoro, Int. J. Quantum Chem. 110 (2010) 624.

[215] W. Domcke, L. Cederbaum, Chem. Phys. Lett. 31 (1975) 582.

[216] L.S. Cederbaum, J. Chem. Phys. 64 (1976) 603.

[217] J. Franck, Trans. Faraday Soc. 21 (1925) 536.

[218] E. Condon, Phys. Rev. 28 (1926) 1182.

[219] E. Condon, Phys. Rev. 32 (1928) 852.

[220] D. Tannor, E. Heller, J. Chem. Phys. 77 (1982) 202.

[221] J. Bloino, M. Biczysko, F. Santoro, V. Barone, J. Chem. Theory. Comput. 6 (2010) 1256.

[222] V. Barone, J. Bloino, M. Biczysko, F. Santoro, J. Chem. Theory. Comput. 5 (2009) 540.

[223] Q. Peng, Y. Niu, C. Deng, Z. Shuai, Chem. Phys. 370 (2010) 215.

[224] J. Tang, M.T. Lee, S.H. Lin, J. Chem. Phys. 119 (2003) 7188.

[225] M.H. Beck, A. Jäckle, G.A. Worth, H.-D. Meyer, Phys. Rep. 324 (2000) 1.

[226] M. de Groot, W.J. Buma, Chem. Phys. Lett. 435 (2007) 224.

[227] J. Tatchen, E. Pollak, J. Chem. Phys. 128 (2008) 164303.

[228] T. Petrenko, F. Neese, J. Chem. Phys. 127 (2007) 164319.

[229] F. Santoro, A. Lami, R. Improta, J. Bloino, V. Barone, J. Chem. Phys. 128 (2008) 224311.

[230] R. Borrelli, A. Peluso, J. Chem. Phys. 128 (2008) 044303.

[231] H.-C. Jankowiak, J.L. Stuber, R. Berger, J. Chem. Phys. 127 (2007) 234101.

[232] F. Santoro, A. Lami, R. Improta, V. Barone, J. Chem. Phys. 126 (2007) 184102.

[233] F. Santoro, R. Improta, A. Lami, J. Bloino, V. Barone, J. Chem. Phys. 126 (2007) 084509.

[234] R. Borrelli, A. Peluso, J. Chem. Phys. 125 (2006) 194308.

[235] M. Dierksen, S. Grimme, J. Chem. Phys. 122 (2005) 244101.

[236] M. Dierksen, S. Grimme, J. Chem. Phys. 120 (2004) 3544.

[237] E.V. Doktorov, J. Mol. Spectrosc. 8 (1977) 507.

[238] A. Hazra, M. Nooijen, Int. J. Quantum Chem. 95 (2003) 643.

[239] H. Ma, J. Liu, W. Liang, J. Chem. Theory Comput. 8 (2012) 4474.

[240] R. Crespo-Otero, M. Barbatti, Theor. Chem. Acc. 131 (2012) 1237.

[241] M. Nooijen, Int. J. Quantum Chem. 106 (2006) 2489.

[242] G.M. Sando, K.G. Spears, J. Phys. Chem. A 105 (2001) 5326.

[243] İ. Özkan, J. Mol. Spectrosc. 62 (1990) 147.

[244] J.R. Reimers, J. Chem. Phys. 115 (2001) 9103.

[245] S. Banerjee, D. Kröner, P. Saalfrank, J. Chem. Phys. 137 (2012) 22A534.

[246] T. Petrenko, F. Neese, J. Chem. Phys. 137 (2012) 234107.

[247] W. Domcke, L.S. Cederbaum, H. Köppel, W. von Niessen, Mol. Phys. 34 (1977) 1759.

[248] F.J. Avila Ferrer, F. Santoro, Phys. Chem. Chem. Phys. 14 (2012) 13549.

Bibliography

[249] A. Hazra, H.H. Chang, M. Nooijen, J. Chem. Phys. 121 (2004) 2125.

[250] M. Rätsep, Z.-L. Cai, J.R. Reimers, A. Freiberg, J. Chem. Phys. 134 (2011) 024506.

[251] E.B. Wilson, J.C. Decius, C.C. Cross, Molecular Vibrations: The Theory of Infrared and Raman Vibrational Spectra, McGraw-Hill, New York, 1955.

[252] A. Capobianco, R. Borrelli, C. Noce, A. Peluso, Theor. Chem. Acc. 131 (2012).

[253] F. Neese, T. Petrenko, D. Ganyushin, G. Olbrich, Coord. Chem. Rev. 251 (2007) 288.

[254] F. Jørgensen, Int. J. Quantum Chem. 14 (1978) 55.

[255] E.A. Coutsias, C. Seok, K.A. Dill, J. Comput. Chem. 25 (2004) 1849.

[256] C. Peng, P.Y. Ayala, H.B. Schlegel, M.J. Frisch, J. Comput. Chem. 17 (1996) 49.

[257] T.E. Sharp, H.M. Rosenstock, J. Chem. Phys. 41 (1964) 3453.

[258] R. Parr, W. Yang, Density Functional Theory of Atoms and Molecules, Oxford University Press, Oxford, U.K., 1994.

[259] P.J. Stephens, F.J. Devlin, C.F. Chabalowski, M.J. Frisch, J. Phys. Chem. 98 (1994) 11623.

[260] J. Perdew, Phys. Rev. B 33 (1986) 8822.

[261] A. Becke, Phys. Rev. A 38 (1988) 3098.

[262] D. Jacquemin, C. Adamo, Int. J. Quantum Chem. 112 (2012) 2135.

[263] F. Santoro, V. Barone, Int. J. Quantum Chem. 110 (2010) 476.

[264] F. Santoro, C. Cappelli, V. Barone, J. Chem. Theory Comput. 7 (2011) 1824.

[265] D. Jacquemin, E. Brémond, A. Planchat, I. Ciofini, C. Adamo, J. Chem. Theory Comput. 7 (2011) 1882.

[266] R. Improta, V. Barone, G. Scalmani, M.J. Frisch, J. Chem. Phys. 125 (2006) 054103.

[267] R. Improta, G. Scalmani, M.J. Frisch, V. Barone, J. Chem. Phys. 127 (2007) 074504.

[268] J.P. Götze, W. Thiel, Chem. Phys. 415 (2013) 247.

[269] H.F. Davis, Y.T. Lee, J. Phys. Chem. 100 (1996) 30.

[270] D. Neubert, Chemical Synthesis of 5-Thiariboflavin as a Potential Chromophore in Blue Light-Sensitive Photoreceptors, M.Sc. Thesis, Hochschule Niederrhein, University of Applied Sciences, 2011.

[271] M.J.G. Peach, P. Benfield, T. Helgaker, D.J. Tozer, J. Chem. Phys. 128 (2008) 044118.

[272] E. Stendardo, F.A. Ferrer, J Chem. Theory. Comput. 8 (2012) 4483.

[273] M. Dierksen, S. Grimme, J. Phys. Chem. A 108 (2004) 10225.

APPENDIX A

Amine Oxidation Mediated by Lysine-Specific Demethylase 1: Quantum Mechanics/Molecular Mechanics Insights into Mechanism and Role of Lysine 661

Bora Karasulu, Mahendra Patil, and Walter Thiel

J. Am. Chem. Soc., **2013**, 135 (36), pp 13400-13413.

JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Article

pubs.acs.org/JACS

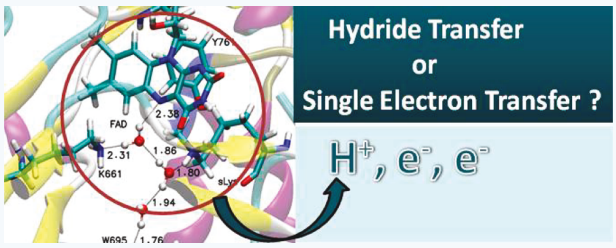
Amine Oxidation Mediated by Lysine-Specific Demethylase 1: Quantum Mechanics/Molecular Mechanics Insights into Mechanism and Role of Lysine 661

Bora Karasulu, Mahendra Patil,[†] and Walter Thiel*

Max-Planck-Institut für Kohlenforschung, Kaiser-Wilhelm-Platz 1, 45470 Mülheim, Germany

Supporting Information

ABSTRACT: We report classical molecular dynamics (MD) simulations and combined quantum mechanics/molecular mechanics (QM/MM) calculations to elucidate the catalytic mechanism of the rate-determining amine oxidation step in the lysine-specific demethylase 1 (LSD1)-catalyzed demethylation of the histone tail lysine (H3K4), with flavin adenine dinucleotide (FAD) acting as cofactor. The oxidation of substrate lysine (sLys) involves the cleavage of an α-CH bond accompanied by the transfer of a hydride ion equivalent to FAD, leading to an imine intermediate. This hydride transfer pathway is shown to be clearly favored for sLys oxidation over other proposed mechanisms, including the radical (or single-electron transfer) route as well as carbanion and polar-nucleophilic mechanisms. MD simulations on six NVT ensembles (covering different protonation states of sLys and K661 as well as the K661M mutant) identify two possible orientations of the reacting sLys and FAD subunits (called "downward" and "upward"). Calculations at the QM(B3LYP-D/6-31G*)/CHARMM22 level provide molecular-level insights into the mechanism, helping to understand how LSD1 achieves the activation of the rather inert methyl-CH bond in a metal-free environment. Factors such as proper alignment of sLys (downward orientation), transition-state stabilization (due to the protein environment and favorable orbital interactions), and product stabilization via adduct formation are found to be crucial for facilitating the oxidative α-CH bond cleavage. The current study also sheds light on the role of important active-site residues (Y761, K661, and W695) and of the conserved water-bridge motif. The steric influence of Y761 helps to position the reaction partners properly, K661 is predicted to get deprotonated prior to substrate binding and to act as an active-site base that accepts a proton from sLys to enable the subsequent amine oxidation, and the water bridge that is stabilized by K661 and W695 mediates this proton transfer.



INTRODUCTION

Chromatins, the basic structural units of genetic material, are used for DNA packaging in eukaryotes. They are mainly composed of aggregates of the nucleosome, an octamer of four different histones.[1] Eukaryotic DNA is wrapped around the histone cores, forming a bead-like structure. In order to access the genetic information, chromatin is partially unwound during gene expression, transcription, DNA repair, replication, and related processes.[2] Histone proteins regulate gene transcription by altering the chromatin structure via post-translational modifications performed on their tails.[3] These modifications are carried out at target sites, which include the N-terminus of a specific lysine or arginine residue at a specific histone tail, each site being related to important cell regulatory processes.[4] They essentially involve covalent addition/removal of a chemical group to/from the site of interest. Among these post-translational modifications, only methylation had long been thought to be irreversible, as the half-life of methylated histone tails is longer than that of the wild-type (non-methylated) histones.[5] The possibility of demethylation was raised by observations on rat kidney in 1973[6] and confirmed by the discovery in 2004 of the first human histone demethylase, lysine-specific demethylase 1 (LSD1),[7] showing that the methylation of the histone tail is dynamically controlled by the reverse demethylation process.

LSD1 is a monoamine oxidase (MAO) that depends on flavin adenine dinucleotide (FAD). It specifically catalyzes the demethylation of methyl and dimethyl (not trimethyl) lysine residues at the fourth position of the histone H3 protein tails (H3K4).[8] LSD1 shows a slight preference for dimethyl lysine over methyl lysine, with turnover rates of 8.10 ± 0.20 and 3.40 ± 0.10 min^{-1}, respectively.[8] Judging from the very low turnover rates, 2−5 orders of magnitude smaller than those of other types of flavoprotein amine oxidases, LSD1 has apparently evolved for substrate specificity rather than catalyst efficiency.[9] The high substrate specificity has been linked to the need for direct interaction of the active site of LSD1 with at least 21 neighboring residues on the tail of H3 protein and the need to orient the methylated substrate lysine in front of the *re*-face of the cofactor FAD to enable proper enzymatic activity.[10,11] Therefore, LSD1 possesses a significantly larger binding pocket

Received: April 10, 2013
Published: August 15, 2013

   dx.doi.org/10.1021/ja403582u | J. Am. Chem. Soc. 2013, 135, 13400−13413

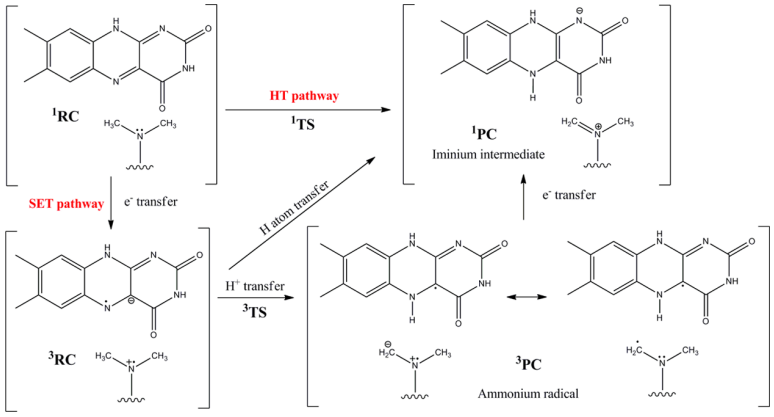


Figure 1. Direct hydride transfer (HT) and radical (SET) pathways proposed for the amine-oxidation step of lysine demethylation catalyzed by LSD1. Abbreviations: RC/PC, reactant/product complex; TS, transition state.

than other amine oxidases. LSD1 is a component of a multiprotein repressor complex, which is controlled by the repressor element 1-silencing transcription (REST). LSD1 binds a corepressor (CoREST) via its TOWER domain.[4] Its interaction with CoREST not only enables the demethylation of the histone H3K4 substrate but also stabilizes LSD1 and increases the enzymatic activity 2-fold.[1] LSD1 can also operate on nonhistone substrates, p53,[12] DNMT,[13] and MYPT1.[14] This allows LSD1 to regulate many vital cell-regulatory processes.[3] It has been shown that abnormal activity of LSD1 is related to many diseases, especially to some cancer types, heart diseases, diabetes, and neurodegenerative disorders, such as Parkinson's disease and Alzheimer's disease.[15] Therefore, the discovery of potential inhibitor molecules that can regulate any abnormal demethylation/methylation balance in LSD1 is of current research interest.[16] From a drug discovery perspective, a molecular-level understanding of the mechanism of the LSD1-catalyzed demethylation process is essential.

LSD1 shows significant sequence identity with other types of amine oxidases, especially with regard to the domain that hosts the active-site cavity. The active site of LSD1 and of other flavoprotein amine oxidases is normally located close to the isoalloxazine moiety of the flavin.[1] Some structural motifs[1,17] are well conserved across flavin-dependent amine oxidases. Most importantly, a conserved lysine residue (K661 in LSD1) forms a hydrogen-bond network to the flavin via a conserved crystal water molecule juxtaposed to the reduction site of the flavin. This water-bridge motif is commonly considered to play an important role in enzymatic activity, which is supported by mutagenesis experiments that show a loss of LSD1 activity upon K661A mutation.[18] Mutation of the corresponding lysine to methionine in maize PAO (K300M) and mammalian PAO (K315M) leads to 1400-fold[19] and 1.8-fold[20] decreases in enzymatic activity, respectively. Based on these experimental findings, different roles for the conserved lysine have been proposed in different amine oxidases. These include participation in catalysis, steric positioning of the flavin ring,[20] and acting as an active-site base by accepting a proton from protonated lysine substrate.[9,19−21] The water-bridge motif has also been considered as the site of oxygen activation during the nonenzymatic oxidative half-reaction, because of the loss of enzymatic turnover and oxygen reactivity upon mutation of K259 in *N*-methyltryptophan oxidase (MTOX).[22]

Another important conserved motif in the active site of different amine oxidases is the aromatic cage, which consists of a pair of aromatic residues shielding the active site from the influx of external solvent molecules, by contributing to its hydrophobicity.[17] In LSD1, one of the two aromatic residues is replaced by threonine (Thr810), while one conserved tyrosine (Tyr761) is still present.[1] As in other amine oxidases, the conserved tyrosine in LSD1 (Tyr761) is located[10,11] next to the substrate binding site on the *re*-face of flavin in the plane orthogonal to the flavin ring. The close proximity of the conserved tyrosine residue to the substrate has been interpreted in terms of a steric role of this residue in aligning the substrate in front of the isoalloxazine ring, through H-bonding and π-cation interactions.[1] In addition, a mutation study on the sandwich-like Tyr-Tyr aromatic cage in human MAO-B[23] suggested that the cage environment enhances the nucleophilicity of the substrate amine moiety via repulsive interactions. Apart from structural aspects, the conserved tyrosine was reported to serve as the initial electron acceptor in MAO-A, through detection of the tyrosyl radical by EPR and ENDOR spectroscopy.[24] The observation of rapid redox equilibrium between the tyrosyl radical and the flavin in MAO-A is in accord with the fact that the tyrosine is covalently bound to the flavin through a cysteine. In contrast, FAD is non-covalently attached to LSD1, which may impede the capacity of Tyr761 in LSD1 to act as the initial electron acceptor. In summary, although the steric influence of the conserved tyrosine appears to be more pronounced than its catalytic relevance, there is still a clear need to determine the precise role of Tyr761 in LSD1.

The LSD1-catalyzed demethylation process involves removal of one methyl group (or of two such groups in two consecutive steps) from the N-terminal of the amine substrate (methylated histone H3K4) via a redox process. Shi et al.[7] proposed a three-step catalytic mechanism for histone demethylation. In the first step, the α-CH bond of the methyl group is cleaved, and the amine is oxidized via transfer of a hydride equivalent from the substrate to FAD. Primary deuterium isotope studies indicate that the breaking of the chemically inert α-CH bond is the rate-limiting step.[9] Amine oxidation is accompanied by two-electron reduction of FAD, which is then (in an oxidative half-reaction) re-oxidized by molecular oxygen with formation of a hydrogen peroxide molecule. This re-oxidation of FAD is a complementary process with high rates[8] that prepares FAD for the subsequent demethylation. In the remaining two steps of the

dx.doi.org/10.1021/ja403582u | *J. Am. Chem. Soc.* 2013, 135, 13400−13413

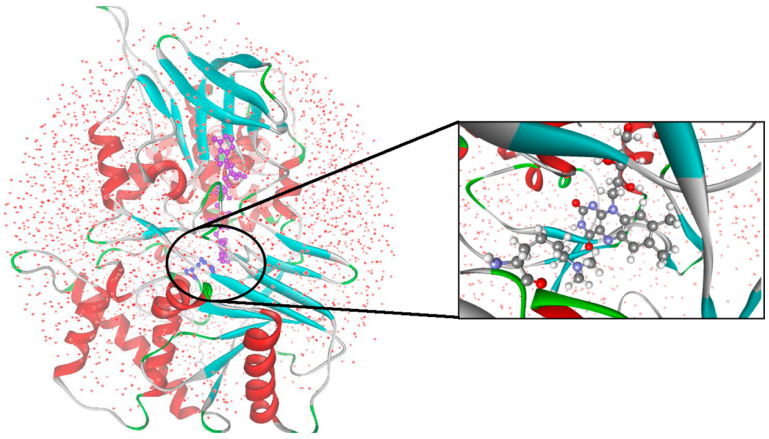


Figure 2. Representative snapshot of the simulation system: FAD and substrate H3K4 (sLys) are shown in ball-and-stick representation in blue and purple, respectively. In the enlarged active site view, the atoms are colored according to atom type (gray, carbon; white, hydrogen; blue, nitrogen; red, oxygen). The simulation system is solvated in a water droplet with radius of 35 Å.

demethylation process, the iminium intermediate is hydrolyzed and the resulting carbinol amine spontaneously rearranges to yield formaldehyde and the demethylated amine. These products have been identified by various experimental methods.[4] The proposed mechanism is supported by the binding mode of the substrate in the crystal structure of LSD1 in the presence of an inhibitor peptide[10] as well as by other experimental findings.[8] Consistent with the lack of a lone pair at the amino group in trimethyl lysines that is required in the proposed amine oxidation mechanism, LSD1 does not act on trimethyl lysines even though its spacious active site can accommodate trimethyl substrate binding.[25]

The rate-limiting first step of amine oxidation involves the transfer of a hydride equivalent from the amine substrate to FAD. The possibility of transferring two electrons and a proton in different order gives rise to several possible mechanisms, namely a direct hydride transfer (HT), a radical mechanism via single-electron transfer (SET),[26−28] and an adduct-forming mechanism.[29] The first two of these mechanisms are depicted in Figure 1 for the case of histone dimethyl lysine demethylation. Obviously, the simplest mechanism is the direct concerted transfer of the two electrons and a proton as a hydride anion from the α-carbon atom of the substrate to FAD. In the radical and adduct-forming mechanisms, there are intermediate species.

There are many studies that favor different mechanisms for different amine oxidases (see refs 1, 3, 17, 30−32 for comprehensive reviews). ^{15}N kinetic isotope effect (KIE) computations on MTOX[33] did not provide clear support to either the HT or SET mechanism. Cyclopropyl inhibitor studies on LSD1 were interpreted as being indicative of a SET mechanism,[34] but this view was challenged[9] on the grounds that some other flavoprotein oxidases, for which HT is considered the most likely mechanism, are also inactivated by cyclopropyl inhibitors.[35,36] The absence of any conceivable intermediate (i.e., a flavin or amine radical species) in kinetic, EPR, and ENDOR studies on LSD1[9,37] was taken as evidence in favor of the HT mechanism. The failure to find any intermediates by EPR or ENDOR spectroscopy was considered to be inconclusive by others,[27,33] however, on the basis of the following scenario: the initial SET process could be reversible (but not rate-limiting), with the back-transfer being much faster; this could lead to a very low concentration of the short-lived radical intermediates, which might not be visible in the EPR and ENDOR spectra.

These conflicting experimental findings on the LSD1 catalytic mechanism call for a theoretical study to monitor molecular-level details that are not directly observable experimentally. Here we report the results from molecular dynamics (MD) simulations as well as quantum mechanical (QM) and QM/molecular mechanical (QM/MM) calculations on the amine oxidation step of histone lysine demethylation. The objective is to assign the catalytic mechanism and to gain detailed molecular-level insight into the role of the active-site LSD1 residues (Lys661, Tyr761, and W695) and of the conserved water-bridge motif in the amine oxidation step.

Among the published computational studies on amine oxidases, only a few cover some aspects within the scope of the current work. A QM-only investigation on MTOX-mediated amine oxidation[33] addressed a simple model system consisting of a truncated isoalloxazine moiety of the flavin and dimethylamine $NH(CH_3)_2$ as the substrate. The computed DFT(B3LYP) energies were found to favor the direct hydride mechanism over the radical-SET mechanism for MTOX.[33] The first QM-only study on LSD1 demethylation was performed by Karasulu et al.[38] using semiempirical (PM3) and DFT (B3LYP) methods. However, the results were only preliminary and limited in scope. More recently, Kong et al.[39] reported classical MD and QM/MM (ONIOM) calculations on the mechanism of the LSD1-catalyzed amine oxidation, which support the HT mechanism. Their QM/MM study focused on the HT pathway and did not provide detailed information on other mechanisms. Moreover, they considered only one particular protonation state of sLys and the conserved K661 residue, and also only one representative snapshot of the system.[39] In the present study, we perform a more extensive sampling of the system by preparing six different NVT ensembles through 20 ns MD runs and by considering different possible protonation states and also mutation of the crucial K661 residue. This allows us to characterize two distinct orientations of sLys with respect to the isoalloxazine ring of FAD and to come up with a comprehensive mechanistic scenario of the events following substrate binding. Furthermore, the roles of important active-site LSD1 residues (K661, Y761, and W695) and of the conserved water-bridge motif are investigated in detail.

dx.doi.org/10.1021/ja403582u | J. Am. Chem. Soc. 2013, 135, 13400−13413

METHODS AND COMPUTATIONAL DETAILS

The simulations reported here are based on the crystal structure of LSD1 (PDB code 2V1D[10]) in complex with CoREST, the co-repressor that enables LSD1 to bind nucleosomal substrates and increases its activity, and with FAD, the cofactor that is non-covalently bound to LSD1 and acts as an oxidative agent in the demethylation process. The setup of the simulation system consisting of ca. 18 000 atoms (see Figure 2) is described in detail in the Supporting Information. Six different NVT ensembles were generated from six separate 20 ns MD simulation runs. Each of these six NVT ensembles corresponds to a different protonation state of K661 or the substrate dimethyl-K4, or involves K661M mutation. The classical MD simulations were performed with CHARMM[40] using a time step of 1 fs, and the atoms were coupled to a thermal bath at constant temperature (300 K). Spherical boundary conditions were imposed using the miscellaneous mean-field potential[40] to prevent the evaporation of solvent molecules. All bonds to hydrogen atoms were constrained by the SHAKE algorithm.[41] The duration (20 ns) of the productive MD simulation runs was sufficient to provide good starting points for the subsequent QM/MM calculations.

QM-only calculations were carried out using density functional theory (DFT) and the Gaussian09 program suite.[42] For this purpose, the full system was truncated to a model system that consisted of sLys, the isoalloxazine ring of FAD, and two active-site water molecules. Four different functionals (B3LYP,[43] M06-2X,[44] LC-ωPBE,[45] and mPW1K[46]) were utilized with the 6-31G* basis set in the gas-phase optimizations and the following vibrational analysis. No constraints were applied in the optimizations. Intrinsic reaction coordinate calculations were used for locating the reactant and product complexes (RC and PC) starting from the transition state (TS) connecting them.

Hybrid QM/MM studies of the full simulation system were performed with the ChemShell program suite.[47] The QM part of the system was computed at the DFT level (B3LYP/6-31G*[43]) using the TURBOMOLE 6.3 software;[48] Grimme-type dispersion corrections[49] were included in all single-point calculations and geometry optimizations. The ground-state singlet and the lowest triplet states were described using restricted and unrestricted Kohn–Sham (RKS and UKS) treatments, respectively. In addition, UKS calculations were performed for all putative open-shell singlet species to check whether they may yield an open-shell (radical-type) configuration with energy lower than that of the closed-shell configuration. The MM calculations were handled by the DL_POLY code[50] implemented in ChemShell using the CHARMM22 force-field parameters specified in the Supporting Information. The QM/MM treatment employed an electrostatic embedding in combination with the charge-shift scheme[51] and the atoms at the QM/MM boundary were treated by the link-atom approach.[47]

In the QM/MM calculations, the QM region consisted of FAD, K661, sLys, and three active-site water molecules (i.e., a total of 72–73 atoms, depending on the protonation state of K661). The included residues were truncated at appropriate sp^3-hybridized carbon atoms. To be specific, FAD was represented by the isoalloxazine ring, with a cut at the C1′–N10 bond (thus excluding the side chain), whereas the lysines were truncated at their C_β–C_γ bond (thus excluding their backbone parts). QM/MM geometry optimizations were carried out with the hybrid delocalized internal coordinates optimizer[52] implemented in ChemShell. Starting geometries for optimizations were taken from several snapshots of the canonical MD ensembles, and only residues within 15 Å of FAD were optimized in order to reduce the computational burden and also to retain the overall protein structure in the absence of CoREST. During the TS optimizations, the spatial positions of the atoms placed in the core region (i.e., those directly involved in the reaction) were optimized using the P-RFO algorithm, whereas the remaining non-frozen nuclei were treated by the L-BFGS algorithm. The optimized structures were subjected to numerical force constant calculations in ChemShell to determine the vibrational modes and to characterize the optimized stationary points (one negative eigenvalue of the corresponding Hessian matrix for TS, none for minima). Gibbs free energies (ΔG) and other thermodynamic properties were evaluated using the standard rigid-rotor harmonic oscillator approximation. Theoretical reaction rates were obtained from the relevant Gibbs free energy barrier $\Delta G^{\ddagger}$ (the free energy of activation) in the usual manner. Natural bond order analyses were performed using Gaussian09 to evaluate the Wiberg bond orders (with MM point charges included).

RESULTS

Changes in the Active Site upon Substrate Binding. Compared with other flavin-dependent amine oxidases, the active site of LSD1 is rather wide and spacious.[11] It has four major invaginations with distinct chemical properties for the specific binding of side chains on the substrate H3 tail. The first pocket contains the isoalloxazine ring of the FAD cofactor forming the main catalytic hydrophobic chamber of LSD1, while the other three are required for accommodating the histone tail adjacent to the substrate lysine (sLys).[25] In the first pocket, the isoalloxazine moiety is surrounded by the residues Arg310, Arg316, Val317, Gly330, Ala331, Met332, Val333, Phe538, Leu659, Asn660, Lys661, Trp695, Ser749, Ser760, Tyr761, and Glu801. These residues help with the exact positioning of the isoalloxazine moiety and the substrate that is required for catalysis. Among them, Arg310, Arg316, Lys661, Tyr761, and Glu801 may have different protonation states depending on the pH of the environment; the same applies to four residues on the H3 tail (H3R2, H3K4, H3K9, and H3K14). Since the active site of LSD1 is shielded against solvent access, substrate binding is expected to lead to changes in the acidity/basicity of the active-site residues. To check for possible changes in the protonation states of these residues upon substrate binding, we computed their acid dissociation constant (pK_a) values at the optimum-activity pH = 8.7 of LSD1[9] (Table 1) using H++ webserver.[53] Judging from the pK_a values in the absence of protein, the methylation site H3K4 (sLys, pK_a = 10.4) should be protonated (as well as the other lysines and arginine located on the H3 tail). Upon binding of the H3 tail to LSD1, H3K4 becomes more acidic (with the pK_a value changing from 10.4 to 6.5), suggesting that it will lose a proton to the microenvironment, while the other residues on the tail remain protonated. This is in line with the need of a

Table 1. Computed Acid Dissociation Constants (pK_a Values) of Selected Residues in Model Systems Containing Only the H3 Tail, Only LSD1, or Both of Them in a Complex, in the Absence and Presence of FAD and Crystal Water (at pH = 8.7)[a]

	without FAD			with FAD	
residue ID	only H3	only LSD1	LSD1 + H3	only LSD1	LSD1 + H3
R2	>12.0	–	>12.0	–	>12.0
H3K4 (sLys)	10.4	–	<0	–	6.5
H3K9	10.6	–	>12.0	–	11.1
H3K14	9.2	–	>12.0	–	>12.0
R310	–	>12.0	>12.0	>12.0	>12.0
R316	–	8.2	7.7	9.3	7.6
K661	–	11.2	7.5	7.4	3.4
Y761	–	>12.0	>12.0	>12.0	>12.0
E801	–	<0	<0	<0	<0

[a]Values in the left (right) panel were obtained without (with) FAD and crystal water molecules being included in the analysis (to show their effect on the pK_a values of the active-site residues).

dx.doi.org/10.1021/ja403582u | J. Am. Chem. Soc. 2013, 135, 13400–13413

deprotonated (neutral) amine site for proper oxidation, as has also been observed for other amine oxidases.[9,20]

The active-site residues surrounding the isoalloxazine moiety may act as a base and accept a proton from sLys. To identify a suitable active-site base, we calculated the pK_a values of these residues (Table 1). R316, K661, and E801 are predicted to have low pK_a values, but K661 is the only residue that is properly placed near the substrate and can be involved in proton transfer. The significant decrease in pK_a value upon binding of FAD and sLys to LSD1 suggests that the conserved K661 residue will lose a proton to bulk water via the microenvironment and can thereafter accept a proton from sLys. Although the exact mechanism for such a proton loss is not known, a breathing motion of the protein may assist this process, as has been proposed previously for another amine oxidase, DAAO.[54]

The residue Tyr761, which is part of the aromatic cage conserved in most amine oxidases and is in close proximity to sLys, was also considered as a potential active-site base. However, as evident from Table 1, Tyr 761 remains protonated after FAD and substrate binding, thus preventing the uptake of a proton. This supports the steric role of the conserved tyrosine residue for proper orientation of sLys (rather than a function as active-site base).

Protonation State of Substrate Lysine (sLys) and K661: MD Simulation Results. In order to gain deeper insight into the binding of the sLys in the active site, we performed a series of 20 ns MD simulations. We generated four different NVT ensembles by considering two different protonation states of K661 and sLys, which will be referred to as follows: (a) protonated sLys, sLys-NMe_2H^+; (b) deprotonated sLys, sLys-NMe_2; (c) protonated K661, K661-NH_3^+; and (d) deprotonated K661, K661-NH_2. In addition, the effect of mutating K661 into a methionine (K661M) was also investigated in two additional NVT ensembles (one for sLys-NMe_2H^+ and another one for sLys-NMe_2).

In the simulation of the protein with both lysines protonated (sLys-NMe_2H^+ and Lys661-NH_3^+), sLys remains far separated from the reactive center of flavin (i.e., the distance between FAD and sLys is more than 6 Å): sLys-NMe_2H^+ thus fails to bind, presumably because of the electrostatic repulsion in the active site. This underscores the necessity of Lys661 being deprotonated (neutral) for effective substrate binding. In the remaining five NVT ensembles with other protonation state combinations and with K661M, sLys is found to be located properly, close to and above the *re*-face of the isoalloxazine moiety of FAD, as required for amine oxidation. This positioning has also been observed for other amine oxidases.[17] There are two distinct binding modes of sLys at the reaction center (N5 of FAD) in these ensembles, depending on the chosen protonation state of sLys. These two orientations are termed "upward" and "downward" according to the alignment of the other (i.e. non-reacting) methyl group of sLys with respect to the C4a–N5 bond of FAD. They are depicted in Figure 3, along with our labeling convention for the atoms.

In the MD simulation with sLys-NMe_2H^+ and Lys661-NH_2, the protonated sLys prefers the downward orientation. It is linked to a water bridge consisting of three water molecules connected through H-bonds with the help of W695 (Figure 4). This three-water-bridge motif is an extension of the Lys–H_2O–N5 motif conserved in most amine oxidases, by incorporation of two additional water molecules. The water bridge is crucial since it connects sLys to FAD and K661, and thus provides a route for the proton shuttle from sLys to K661 in accord with the theoretical pK_a analysis. The stability of the water bridge is supported by the solvent-inaccessible design of the active site of LSD1, which is common to all amine oxidases.[17] W695 is located at the catalytic center (see Figure 4) and remains part of the water bridge during the entire MD simulation (see Supporting Information, Figure S1). It affords extra stabilization by hydrogen-bonding, and along with Y761 it controls the traffic of water molecules in the active site. The loss of LSD1 activity upon W695A mutation[25] can therefore be related to the resulting disruption of the water bridge. These findings support an important structural role for W695 in amine oxidation.

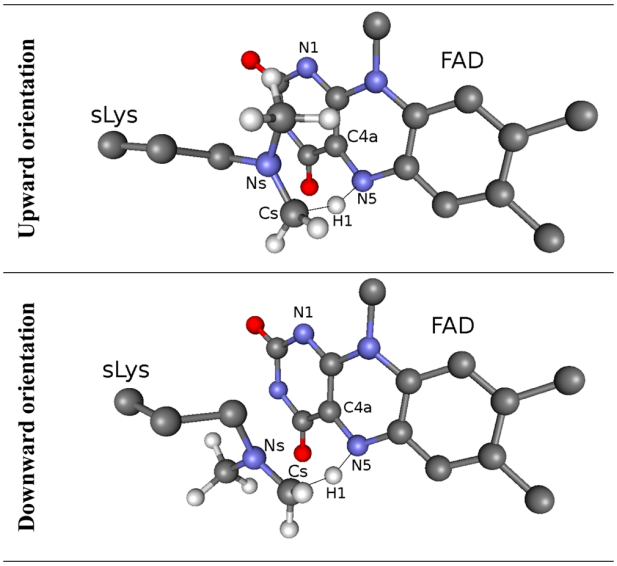


Figure 3. Ball-and-stick representations of the upward (top panel) and downward (bottom panel) orientations of sLys. The notation reflects the orientation of the non-reacting methyl group of sLys with respect to the N5–C4a bond of FAD. The structures show parts of the QM region from the QM/MM optimized **1-TS** geometries, with most of the hydrogen atoms, the deprotonated K661, and the three QM water molecules removed for clarity. Also included is the numbering of the relevant atoms (used in the text). The standard flavin numbering scheme has been adopted along with special atom labels for the substrate lysine (Cs and Ns). H1 is the hydride equivalent being transferred.

In the MD simulation with K661M and sLys-NMe_2H^+, we do not observe formation of the three-water-bridge motif in the K661M mutant. In this case, there is no residue that could serve as an H-bond donor/acceptor residue and connect to the isoalloxazine moiety of flavin via the conserved water molecule. Moreover, the methyl group of methionine in K661M will not lose a proton upon substrate binding and will thus not act as active-site base. The profound loss of enzymatic activity of LSD1 upon K661A mutation[18] can be ascribed to the lack of an effective proton shuttle from sLys to K661. This further supports the active-site base role of the conserved K661.

An already deprotonated sLys prefers to bind in upward orientation in the active site. In this orientation, the electrostatic interaction between two partially charged atoms (Ns of sLys and C4a of FAD) and the orbital interaction of the lone pair of sLys-Ns with the π^*-orbital (C4a–N5) of FAD help keeping the reactive partners for amine oxidation in the two subunits (N5 in FAD and Cs in sLys) in close contact. Selected average distances and angles derived from the NVT ensembles with Lys661-NH_3^+, Lys661-NH_2, and K661M are

dx.doi.org/10.1021/ja403582u | J. Am. Chem. Soc. 2013, 135, 13400–13413

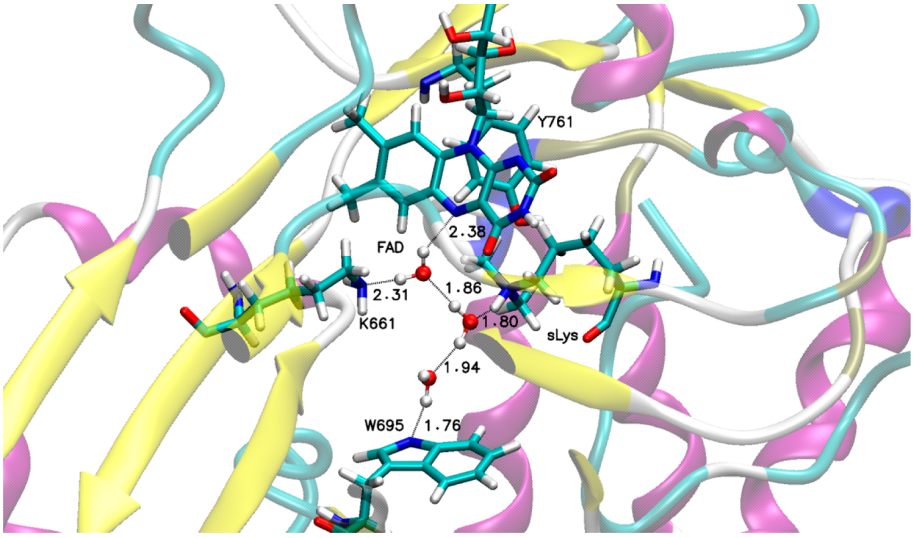


Figure 4. Typical snapshot showing the three-water-bridge motif in the downward orientation. Hydrogen bonds are marked with thin black lines; the corresponding distances are given in Å. The snapshot was taken from an MD simulation with sLys-NMe_2H^+ and Lys661-NH_2 (see text).

Table 2. Average Distances (Å) and Angles (degree) (with the Corresponding Standard Deviations) for Three Different NVT Ensembles with Deprotonated Substrate K4 (sLys)

structural property	deprotonated K661	protonated K661	K661M mutant
R(Ns−C4a)	3.23 ± 0.13	3.26 ± 0.14	3.14 ± 0.12
R(N5−Cs)	3.78 ± 0.26	3.84 ± 0.25	3.64 ± 0.23
θ(Cs−Ns−C4a−N5)	−34.0 ± 7.1	−16.8 ± 6.9	−32.4 ± 8.0
R(N5−H (H_2O))[a]	2.10 ± 0.34	1.98 ± 0.21	N/A

[a]H(H_2O) is the hydrogen of the conserved crystal water hydrogen-bonded to N5 of FAD.

compiled in Table 2. In all three cases, proper binding of sLys on top of the FAD plane is achieved, and the two subunits stay in close contact during the MD simulations. This suggests that the binding mode of deprotonated sLys with respect to FAD is affected significantly neither by the protonation state nor even by the mutation of K661. Therefore, if sLys has already lost its extra proton prior to binding in the active site (which is unlikely, given the usual pH of the environment), the protonation state of K661 is of no concern for proper binding of sLys.

These findings are in accordance with experimentally observed circular dichroism spectra and structural studies,[19] which indicate that K300M mutation in mPAO does not affect substrate binding and overall protein folding, even though the reaction rate is slowed down about 1400-fold. In contrast, the K661−H_2O−N5 motif cannot be formed in the K661M mutant (independent of the protonation state of sLys), whereas it is formed in wild-type LSD1 for both protonation states of K661, with a slight variation in the distance of this conserved lysine to the conserved water molecule (Table 2). Thus, our MD simulation results highlight the role of the K661 residue as an active-site base and the importance of the water bridge for mediating the proton shuttle between sLys and K661.

QM-Only and QM/MM Results. We carried out QM-only and QM/MM calculations to locate the stationary points of the two main proposed catalytic mechanisms (HT and SET, Figure 1). As starting geometries for QM/MM calculations, we used five different snapshots for each orientation, which were chosen randomly from two different MD simulation trajectories (i.e., Lys661-NH_2 with sLys-NMe_2H^+ for upward orientation and Lys661-NH_2 with sLys-NMe_2H^+ for downward orientation). The QM-only studies were performed on a model system consisting of sLys, the isoalloxazine ring of FAD, and two active-site water molecules. In a preliminary conformational analysis at the B3LYP level, we found that the components of the model system are completely free to move (and relax) due to the absence of constraints from the protein environment. As a consequence, the two distinct QM-only TSs for the HT pathway with downward and upward orientation differ by less than 0.5 kcal/mol in energy; for the sake of brevity, we limited our analysis at the QM-only level to the downward HT pathway. In the case of the SET mechanism, the QM-only optimizations yielded only one single TS (regardless of the chosen starting geometry). The gas-phase QM-only energetics obtained with four different DFT functionals are compiled in Table 3 and visualized in Figure S2. B3LYP yields activation barriers for the HT and SET mechanisms that are somewhat lower than those predicted by the other functionals (by 1−2, 1−8, and 2−5 kcal/mol compared to M06-2X, LC-ωPBE, and mPW1K, respectively). More importantly, however, the relative Gibbs free energies of the different stationary points exhibit the same trends for all four functionals. To limit the computational effort, we have therefore decided to apply only B3LYP in the further calculations.

Considering the crucial role of K661 and the three-water-bridge motif in the MD simulations (*vide supra*), the QM/MM calculations were performed with a standard QM region composed of sLys (side chain), the isoalloxazine moiety of FAD, K661, and the three bridging water molecules (72−73 QM atoms). Two different orientations of sLys (upward and

dx.doi.org/10.1021/ja403582u | J. Am. Chem. Soc. 2013, 135, 13400−13413

Table 3. QM-Only Relative Gibbs Free Energies (kcal/mol) for the Stationary Points (As Given in Figure 1) Evaluated Using Different DFT Functionals and the 6-31G* Basis Set[a]

	B3LYP-gas	B3LYP-water	M06-2X	LC-ωPBE	mPW1K
	Direct Hydride Transfer (HT; Singlet Manifold)				
^{1}RC	0.0	0.0	0.0	0.0	0.0
^{1}TS	31.0	26.7	33.0	32.2	36.4
^{1}PC	15.5	12.4	7.2	9.4	4.5
	Radical Mechanism (SET; Triplet Manifold)				
^{3}RC	35.1	30.2	35.6	42.5	37.5
^{3}TS	37.9	36.5	38.4	46.2	39.9
^{3}PC	29.8	26.9	27.7	36.4	25.5

[a]The present B3LYP/6-31G* results are in good agreement with those reported previously.[33]

downward, see Figure 3) were considered for modeling the stationary points of the two proposed reductive half-reaction mechanisms (HT and SET). Keeping in mind the importance of the protonation state of K661 for the binding of sLys in the active site (*vide supra*), the stationary points were modeled using different protonation states of K661, namely K661-NH_2 and K661-NH_3^+. In contrast, only deprotonated sLys was considered in the QM/MM calculations. The QM/MM relative energies and free energies for all possible combinations are compiled in Table 4 and visualized in Figure 5. All these values are computed as the average over five different sets of QM/MM calculations starting from five different snapshots taken randomly along the course of the corresponding NVT ensemble (see Supporting Information, Tables S1 and S2). Table 4 also contains the QM/MM energetics for the K661M mutant (as visualized in Figure S3), obtained as the average of two sets of calculations (see Table S3).

To check the effect of basis set extension, we also provide single-point QM(B3LYP-D/TZVPP)/MM relative energies computed at QM(B3LYP-D/6-31G*)/MM geometries (see Table S4). Similarly, we check the effect of expanding the standard QM region by including the residues Gly330, Met332, Val333, Thr335, Tyr761, and Val811, which form an extensive H-bond network to FAD and sLys (see Figure S4, upper panel). For this extended QM region (with 173−174 QM atoms depending on the protonation state of K661), we performed single-point QM(B3LYP-D/6-31G*)/MM energy calculations at the available optimized QM/MM geometries. The resulting energetics is given in Table S5. Evidently, the use of the larger basis set and of the extended QM region lowers the relative energies by ca. 2−3 kcal/mol, but the trends remain the same. Therefore, we shall focus in the following on the results of QM(B3LYP-D/6-31G*)/MM calculations with the standard QM region.

In the model systems with protonated K661, the TSs always have higher Gibbs free energies (relative to ^{1}**RC**) than their counterparts with deprotonated K661 (Table 4). For the sake of brevity, we will only discuss the results for the latter case when comparing the different pathways in upward and downward orientation. The effects of K661 protonation and K661M mutation will be addressed in a later section.

Hydride Transfer (HT) Mechanism. The HT mechanism involves the transfer of a proton and two electrons from sLys to FAD in a single concerted step (as H^- anion). It is generally considered as the most likely route for the rate-determining C−H bond oxidation step of LSD1-catalyzed lysine demethylation. A closed-shell description at the RKS level is well suited for this process that converts the reactant complex ^{1}**RC** via the transition state ^{1}**TS** to the product complex ^{1}**PC**. Hybrid QM/MM calculations predict free energy barriers of 20.9 ± 1.1 and 15.4 ± 1.0 kcal/mol for the HT mechanism with upward and downward orientation of sLys, respectively (Table 4). By contrast, QM-only calculations yield barriers of 31.0 kcal/mol in vacuum and 26.7 kcal/mol in water. The HT rate is thus increased dramatically by the protein environment.

Key structural properties for the stationary points are compiled in Tables 5 and 6 at the QM-only and QM/MM levels, respectively. Evidently, ^{1}**RC**, ^{1}**TS** and ^{1}**PC** are predicted to be more compact in the enzyme, both for the upward and downward orientation, i.e., the two reacting subunits (sLys and FAD) stay closer to each other in the protein environment. This effect of the environment is most pronounced for ^{1}**TS**. The corresponding optimized geometries are visualized in Figure 3 (QM/MM) and Figure S5 (QM). The QM/MM structures of ^{1}**TS** are quite similar for the upward and downward orientation, but rather different from the QM optimized structures. In the QM/MM TSs, the lysine tail is oriented parallel to the C4a−N5 π-bond in an optimum distance for the HT to the acceptor orbital. The optimized distances for the breaking Cs− H1 bond are 1.39 and 1.37 Å, whereas those for the forming N5−H1 bond are 1.24 and 1.25 Å, for the upward and downward orientation, respectively (Table 6), implying a late TS for HT pathway. The HT angle (N5−H1−Cs) is almost identical in both orientations (ca. 154°). The QM-only calculations also yield a late TS with the same HT angle (154°), but the bond distances for Cs−H1 (1.50 Å) and N5−H1 (1.16 Å) deviate appreciably. The Wiberg

Table 4. QM(B3LYP-D/6-31G*)/MM Relative Energies (Gibbs Free Energies in Parentheses) for the Stationary Points (see Figure 1) in Upward and Downward Orientation and for the K661M Mutant: Average Values and Standard Deviations (kcal/mol) over a Set of Five Snapshots

	deprotonated K661		protonated K661		
	upward	downward	upward	downward	K661M
	Direct Hydride Transfer (HT; Singlet Manifold)				
^{1}RC	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
^{1}TS	23.8 ± 1.93 (20.9 ± 1.14)	18.9 ± 0.95 (15.4 ± 0.98)	25.1 ± 2.05 (21.7 ± 2.12)	23.5 ± 2.97 (20.0 ± 2.75)	22.4 ± 1.0 (19.0 ± 1.1)
^{1}PC	2.05 ± 0.95 (2.42 ± 0.94)	−12.5 ± 3.57 (−12.1 ± 3.53)	13.5 ± 1.54 (13.7 ± 1.50)	15.8 ± 1.49 (15.6 ± 1.38)	1.80 ± 1.6 (1.60 ± 2.5)
	Radical Mechanism (SET; Triplet Manifold)				
^{3}RC	31.5 ± 2.28 (31.1 ± 2.28)	20.9 ± 1.89 (20.4 ± 1.70)	29.5 ± 2.63 (29.2 ± 2.66)	24.6 ± 1.85 (24.2 ± 1.82)	27.2 ± 0.5 (26.9 ± 0.5)
^{3}TS	48.0 ± 1.71 (44.2 ± 1.72)	37.1 ± 1.76 (33.2 ± 1.73)	49.3 ± 2.78 (45.6 ± 2.81)	44.7 ± 3.22 (41.1 ± 2.99)	48.2 ± 0.6 (44.4 ± 0.7)
^{3}PC	27.9 ± 1.38 (27.4 ± 1.37)	24.5 ± 2.27 (23.8 ± 2.36)	35.3 ± 2.88 (35.2 ± 2.85)	37.9 ± 3.14 (37.2 ± 3.37)	32.2 ± 0.6 (31.9 ± 0.7)

dx.doi.org/10.1021/ja403582u | J. Am. Chem. Soc. 2013, 135, 13400−13413

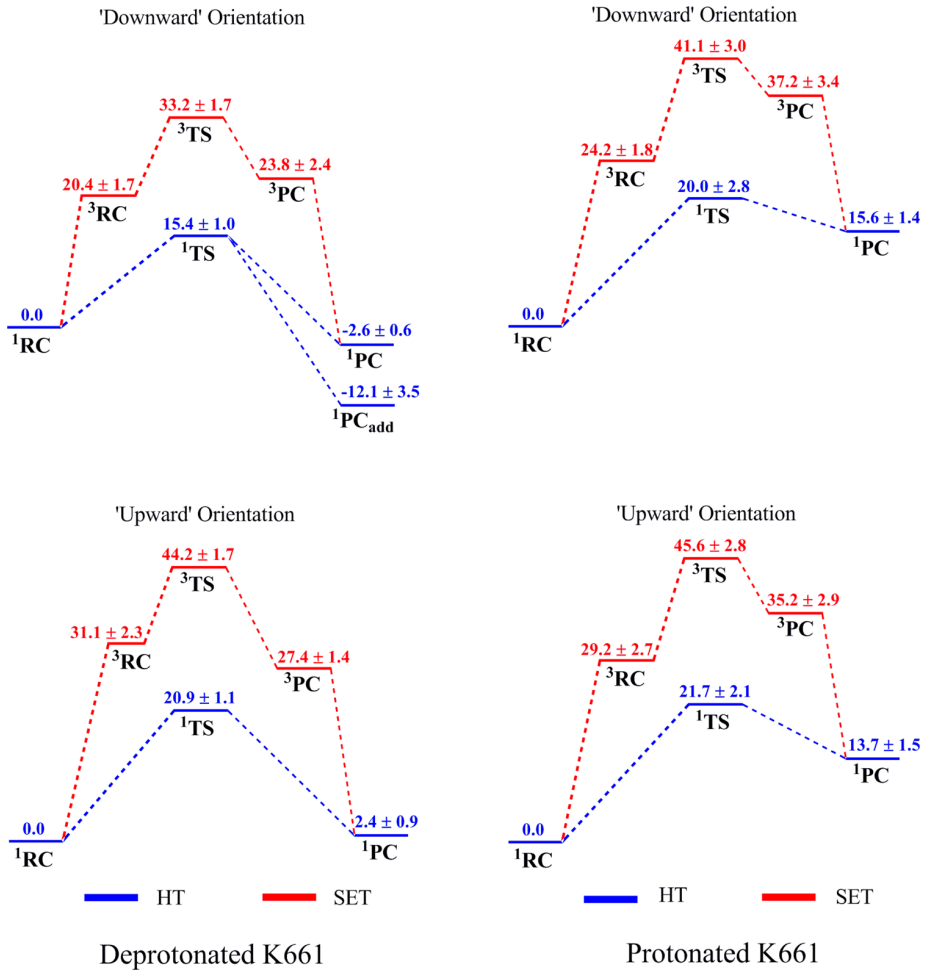


Figure 5. QM(B3LYP-D/6-31G*)/MM Gibbs free energy profiles (in kcal/mol) for HT and SET pathways obtained with deprotonated (left) and protonated (right) K661 and with downward and upward orientation. See Figure 1 for the definition of the stationary points. The profiles were computed with deprotonated sLys. The upper-left diagram includes the free energies both of the weakly interacting product complex ^{1}**PC** and the adduct ^{1}**PC**$_{add}$ (see text for details).

Table 5. Selected Distances (R, in Å), Angles, and Dihedral Angles (θ, in Degree) from QM-Only Optimizations of the Stationary Points at the B3LYP-D/6-31G* Level

	^{1}RC	^{1}TS	^{1}PC	^{3}RC	^{3}TS	^{3}PC
θ(Cs−H1−N5)	111.3	153.6	48.0	158.2	167.6	92.7
R(N5−H1)	3.00	1.16	1.05	2.96	1.37	1.04
R(Cs−H1)	1.10	1.50	2.21	1.10	1.35	3.38
R(Ns−C4a)	3.15	2.96	3.06	3.97	4.10	4.52
R(Cs−N5)	3.55	2.60	1.70	3.34	2.71	3.58
R(Cs−Ns)	1.46	1.35	1.37	1.45	1.39	1.46
R(C4a−N5)	1.30	1.38	1.46	1.37	1.37	1.37
θ(Ns−Cs−N5−C4a)	30.2	−0.6	46.7	−44.2	4.9	48.7

bond order ratio (Cs−H1 vs N5−H1) is around 0.85 in all ^{1}**TS** geometries (see Table S6).

As the HT reaction proceeds, the developing charges on the Ns$^{\delta-}$ (sLys) and C4a$^{\delta+}$ (FAD) atoms will provide some electrostatic stabilization, since the interacting centers C4a and Ns become sufficiently close at the TS (QM-only, 2.96 Å; QM/MM upward, 2.81 Å; QM/MM downward, 2.79 Å, indicating a stronger effect in the enzyme). Orbital interactions between the lone pair on sLys-Ns and the π^*-orbital (C4a−N5) of FAD may be important both in the reactant complex ^{1}**RC** and the transition state ^{1}**TS**. The HOMOs at the ^{1}**RC** and ^{1}**TS** geometries in both upward and downward orientation are shown in Figure 6. In the case of ^{1}**RC** (Figure 6, upper panel) with upward orientation, the lone pair at Ns of sLys is aligned such that it lies almost orthogonal to the plane of π-conjugation on the isoalloxazine ring of FAD. This stacked orientation of cofactor and substrate enhances the overlap between the orbitals of the two subunits that can stabilize this ^{1}**RC** species. By contrast, in downward orientation, ^{1}**RC** adopts a nonstacked alignment between the two reacting subunits that offers little

dx.doi.org/10.1021/ja403582u | J. Am. Chem. Soc. 2013, 135, 13400−13413

Table 6. Selected Distances (R, in Å), Angles, and Dihedral Angles (θ, in Degree) of Stationary Points with Upward and Downward Orientation (Deprotonated K661) Obtained from QM(B3LYP-D/6-31G*)/MM Optimizations

	upward orientation						downward orientation					
	^{1}RC	^{1}TS	^{1}PC^a	^{3}RC	^{3}TS	^{3}PC	^{1}RC	^{1}TS	^{1}PC^a	^{3}RC	^{3}TS	^{3}PC
θ(Cs−H1−N5)	112.4	154.6	54.6	98.9	152.8	81.2	99.9	153.6	48.1	120.7	155.4	102.6
R(N5−H1)	2.59	1.24	1.03	2.72	1.22	1.06	2.66	1.25	1.07	2.56	1.31	1.07
R(Cs−H1)	1.10	1.39	2.62	1.09	1.50	3.64	1.09	1.37	2.17	1.12	1.39	3.97
R(Ns−C4a)	2.69	2.81	2.94	2.84	3.11	3.08	2.74	2.79	2.92	3.16	3.20	3.64
R(Cs−N5)	3.18	2.57	2.19	3.08	2.64	3.63	3.04	2.55	1.66	3.28	2.64	4.33
R(Cs−Ns)	1.46	1.36	1.31	1.45	1.38	1.38	1.46	1.36	1.38	1.44	1.38	1.38
R(C4a−N5)	1.30	1.37	1.42	1.37	1.37	1.36	1.30	1.37	1.45	1.37	1.37	1.37
θ(Ns−Cs−N5−C4a)	−25.3	−21.7	−24.0	−26.4	−31.8	−34.5	10.2	3.9	46.2	−1.2	−14.3	−6.7

^{a}The column ^{1}PC contains the structural parameters of the weakly interacting product complex for upward orientation, and those of the adduct for downward orientation.

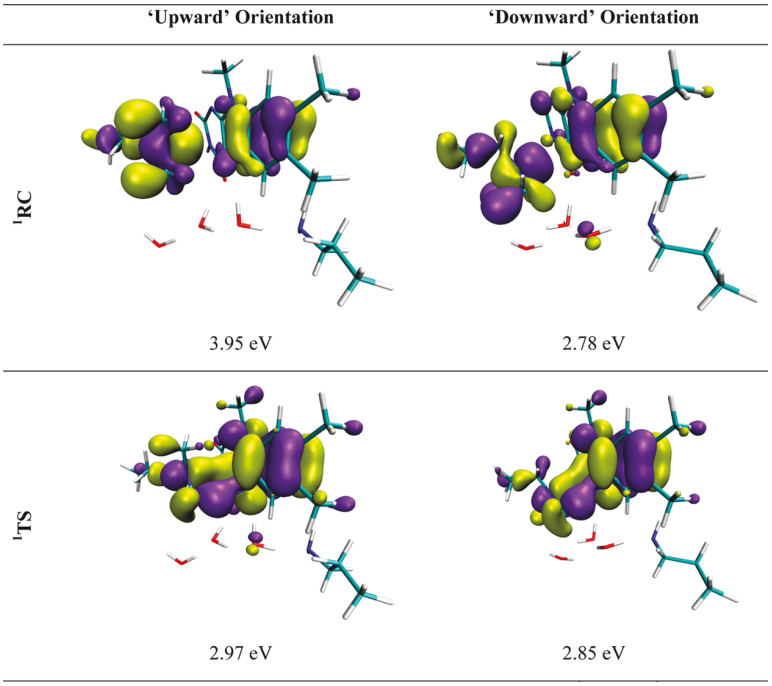


Figure 6. Highest occupied molecular orbital (HOMO) isosurfaces of ^{1}RC and ^{1}TS in upward and downward orientation. HOMO−LUMO gaps are given in eV. A contour value of 0.02 was chosen for creating the isosurfaces using VMD.

stabilization. On the other hand, in the ^{1}TS structures both for upward and downward orientation, the two reacting units assume a proper stacking orientation that allows enhanced orbital interactions. These findings may rationalize the lower activation barrier for the HT mechanism in downward orientation where ^{1}RC is less stabilized compared to upward orientation. To estimate the intrinsic magnitude of the associated energetic effect, TSs with different lysine orientations, which lack the C4a···Ns interaction, were modeled and optimized at the QM-only level: they were found to have higher energies (by 3−6 kcal/mol) than the reference QM-only TS.

The Cs···N5 interaction becomes crucial in the product complex ^{1}PC as the lone pair is now located on N5 rather than Ns. The Cs−N5 distance is computed to be 1.66 and 2.19 Å in downward and upward orientation, respectively. In the former case, a FAD-sLys adduct complex is formed, which is not possible in the latter case due to steric congestion. Consequently, ^{1}PC is significantly more stable in downward than in upward orientation, as can be seen from the reaction free energies (−12.1 ± 3.5 vs 2.4 ± 0.9 kcal/mol). Although the adduct formation in ^{1}PC in downward orientation is predicted to be a barrierless process that will occur readily, we tried to find the weakly interacting PC (with Cs−N5 distance of more than 2 Å), for direct comparison with the upward case. In two out of the five snapshots, by rearranging the water-bridge motif, we could locate a weakly interacting ^{1}PC minimum, with a reaction free energy of −2.63 kcal/mol (average value). Hence, the HT in downward orientation remains mildly exergonic even in this case (and does not become endergonic as in upward orientation). To study the interconversion of the two minima, we performed a one-dimensional relaxed potential energy scan for one of the snapshots using the Cs−N5 distance as the reaction coordinate. The resulting energy profile indicates facile formation of the adduct from the weakly interacting complex, with a barrier of less than 1 kcal/mol (see Figure S9).

Radical Single-Electron Transfer (SET) Mechanism. The radical SET mechanism involves intermediates (Figure 1). It starts with the activation of the αC−H bond of lysine by an

dx.doi.org/10.1021/ja403582u | J. Am. Chem. Soc. 2013, 135, 13400−13413

initial SET to the flavin to yield two monoradical intermediates (^{3}**RC**), namely the amine radical cation (**sLys**$^{\bullet+}$) and flavin semiquinone anion (**FAD**$^{\bullet-}$ or **Fl**$^{\bullet-}$). The subsequent rate-limiting step is the homolytic cleavage of the αC−H bond that generates the iminium cation (^{1}**PC**). It may proceed via two different paths. One is proton-coupled electron transfer (PCET), which is also called the direct hydrogen atom transfer path, whereas the other one involves consecutive transfers of a proton and an electron (second SET) from the amine to the flavin. In the latter case, the proton transfer leads to an intermediate pair of ammonium and **FADH**$^{\bullet-}$ radicals (^{3}**PC**), and the second SET step can occur directly or may also be mediated by adduct formation in the presence of an active-site radical.[28] To properly describe the open-shell species on the SET pathways, we employed the UKS treatment. As the two unpaired electrons on two different subunits of the model system can have the same or different spin, both the singlet and triplet manifolds were considered. In the case of the singlets, the UKS calculations as well as restricted open-shell Kohn−Sham (ROKS) calculations always converged to the closed-shell solutions as can be seen from the resulting energies (see Table S7) and spin densities (data not shown). Therefore, we only present the triplet UKS results for the SET mechanism; the stationary points are named accordingly, i.e., ^{3}**RC**, ^{3}**TS**, and ^{3}**PC**.

The first SET from **sLys** to **FAD**, yielding ^{3}**RC**, is predicted to be energetically demanding. QM-only calculations (without protein environment) give high barriers in the gas phase and in water (35.1 and 30.2 kcal/mol, respectively). We could not precisely determine the energy needed for intersystem crossing ($S_0 \rightarrow T_1$), which is however expected to be lower than that of the transition state ^{3}**TS** for the subsequent proton transfer from **sLys**$^{\bullet+}$ to **FAD**$^{\bullet-}$.[9] Indeed, the overall activation barrier for homolytic cleavage of the αC−H bond is 37.9 and 36.5 kcal/mol in the gas phase and in water, respectively; the energy lowering by the water environment is thus less than in the case of HT pathway (1.4 vs 4.3 kcal/mol). The generated radicals (^{3}**PC**) are also rather unstable, with energies of 29.8 and 26.9 kcal/mol, respectively, relative to ^{1}**RC**. The QM/MM calculations generally predict somewhat lower relative energies for ^{3}**RC**, ^{3}**TS**, and ^{3}**PC**, but the trends in the energetics are the same (Table 4). The protein environment is computed to lower the activation barrier of the SET pathway by ca. 5 kcal/mol when sLys is aligned in the downward orientation, but this barrier is still about twice as high as that for the HT pathway. In downward orientation, ^{3}**TS**, ^{3}**RC**, and ^{3}**PC** are all stabilized relative to ^{1}**RC**, which reflects the already discussed destabilization of ^{1}**RC** (see above). In upward orientation, the relative energies of these species are generally higher. Once the first SET is realized and the ^{3}**RC** species is formed, the subsequent proton transfer from **sLys**$^{\bullet+}$ to **FAD**$^{\bullet-}$ is predicted to have a relatively low barrier (with respect to ^{3}**RC**), both at the QM-only level and at the QM/MM level (see Tables 3 and 4). These low barriers are kinetically irrelevant, because the SET mechanism is inaccessible due to the high initial barrier for the formation of ^{3}**RC**.

Key structural features of the optimized stationary points from the QM-only and QM/MM calculations are presented in Tables 5 and 6, respectively. As for the HT pathway, the reactive core of the stationary points on the SET pathway is generally more compact in the protein environment. However, contrary to the HT case, the downward orientation leads to a significantly looser binding of the two subunits in all species (see R(Ns−C4a) in ^{3}**RC**, ^{3}**TS**, and ^{3}**PC**), which should alleviate any unfavorable interactions among the unpaired electron density distributed over the two subunits and may thus help to stabilize the downward orientation. In line with the latter, the products of the SET mechanism (^{3}**PC**) do not form an adduct complex, as opposed to the products of the HT mechanism (^{1}**PC**). In the transition state ^{3}**TS**, the proton-transfer angle θ(Cs−H1−N5) is 153−155° in both upward and downward orientation, and thus the same as in the HT case (^{1}**TS**), see Table 6. In the QM-only gas-phase calculations, this angle is significantly larger in ^{3}**TS** (168°, Table 5), presumably since sLys may move in the absence of protein constraints toward a more linear arrangement utilizing another acceptor orbital (i.e., the lone pair at N5). The protein environment enforces a transfer angle of 153−155° and the involvement of a singly occupied π acceptor orbital. This difference in the TS geometries is reflected in the proton transfer barriers of the QM-only (2.8 kcal/mol) and QM/MM calculations (13.1 and 12.8 kcal/mol in upward and downward orientation, respectively).

The spin densities computed for ^{3}**RC**, ^{3}**TS**, and ^{3}**PC** with upward and downward orientation of sLys are given in Figure S6. Evidently, the stationary points on the SET pathway generally contain sLys and FAD radicals as shown in Figure 1; in downward orientation, one of the conserved water molecules may also accommodate some unpaired electron density.

Effects of K661 and Y761 on the Reaction. In this section, we discuss the effects of the active-site residues on the LSD1-mediated amine oxidation reaction. Homologues of the active-site LSD1 residues K661 and Y761 are highly conserved among the members of the amine oxidase family and are expected to play a crucial role in the reactions catalyzed by amine oxidases. As noted in the MD simulations, K661 supports the formation of the three-water-bridge motif and the downward orientation with less stable reactants (sLys and FAD). Besides, theoretical pK_a analysis suggests that K661 may get deprotonated upon binding of sLys in the active site. To further augment our understanding of the effect of K661 protonation on the catalysis, we performed QM/MM calculations with K661 in its protonated form for both the downward and upward orientation of sLys. The resulting QM/MM energies are presented in Table 4. Compared with the case of deprotonated K661, protonated K661 yields the same trends in the energetics, but the activation barriers are higher (by 1−6 kcal/mol) for the HT and SET pathways in each orientation. Notably, the downward orientation of sLys yields lower barriers than the upward orientation for both pathways. The most dramatic change in the energetics is predicted for ^{1}**PC** and ^{3}**PC**, which are both destabilized with respect to the corresponding ^{1}**RC** species.

To identify possible causes of this destabilization of the PC species, some key structural properties of the stationary points from the QM/MM calculations with protonated K661 are compiled in Table S8. As can be seen from the Cs−N5 distances (3.14 and 2.98 Å), there is no adduct formation in ^{1}**PC** (not even in downward orientation) when K661 is protonated. Comparison of ^{1}**PC** geometries in downward orientation with deprotonated and protonated K661 (see Figure 7) reveals different H-bonding networks for the different protonation states. In ^{1}**PC** with protonated K661, the three involved water molecules are arranged such that they form an H-bond to N5 of FAD at one end and accept an H-bond from protonated K661 at the other end. As a consequence, the lone

dx.doi.org/10.1021/ja403582u | J. Am. Chem. Soc. 2013, 135, 13400−13413

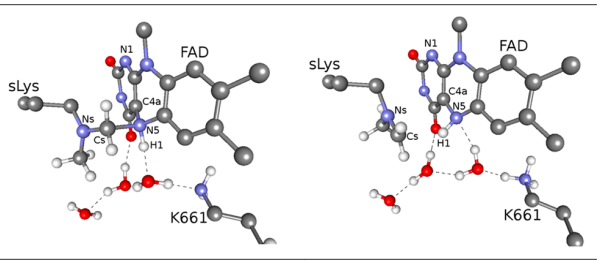


Figure 7. Comparison of product complexes ^{1}**PC** when K661 is deprotonated (left) and protonated (right). Adduct formation following hydride transfer depends on the protonation state of K661. H-bonds are indicated by black dashed lines. Some hydrogens are not shown for clarity.

electron pair on N5 is not freely available for adduct formation. These findings confirm that the deprotonated form of K661 is required for adduct formation between sLys and FAD, which is, in turn, needed to make the overall process exothermic.

Despite the fact that K661 plays a crucial role in the demethylation process by supporting H-bonding networks and acting as an active-site base, it is not expected to have a catalytic role in LSD1,[9] in contrast to the MPAO case.[19] To check the validity of this notion, we performed MD simulations for each of the two protonation states of sLys with K661 being mutated into methionine. As already discussed above, these runs for the K661M mutant do not show any binding of protonated sLys (Cs−N5 distance larger than 6 Å). On the other hand, in the unlikely event that sLys is already deprotonated initially, the MD run indicates that it may bind properly in front of *re*-face of FAD in a reactive distance with an upward-type orientation (Table 2). To analyze the effects of K661M mutation on the HT and SET energy profiles, we performed corresponding QM/MM calculations. Comparing the structural properties of the optimized stationary points for the wild-type enzyme in upward orientation and for the K661M mutant (see Tables 6 and S9) we generally find high similarity in the computed geometries, with particularly remarkable agreement in the case of ^{1}**TS** and ^{3}**TS**. Likewise, the calculated relative energies for the wild-type enzyme (upward orientation) and the mutant are in almost perfect agreement, within 1 kcal/mol, both for the HT and SET pathway (see Table 4). In addition, the QM/MM calculations (see Figure S6) do not give any unpaired electron density on the K661/M661 unit in any case (regardless of protonation state or sLys orientation), thus suggesting no role for this residue in electron transfers on the SET pathway. Taken together, these findings strongly support the notion that the catalytic mechanism of LSD1 is not dependent on the K661 residue. Therefore, the complete loss of LSD1 activity upon K661 mutation[18] can be ascribed to its role as active-site base that accepts a proton from protonated sLys prior to substrate binding.

In the MD runs with each of the six NVT ensembles, the other conserved residue, Y761, was found to stay in close contact with sLys (data not shown), in line with its anticipated role of orienting sLys in front of the flavin. In addition to this steric role, Y761 has also been considered as the active-site base and as the initial single electron acceptor in the radical mechanism.[1,23] We exclude both these possibilities as follows. To check the suitability of Y761 as active-site base, we performed a series of QM/MM calculations to compare the energy change upon proton transfer from sLys to Y761 and K661 (the only candidates for an active-site base). The proton transfer to deprotonated K661 is computed to be downhill by 9 kcal/mol. On the contrary, a minimum with protonated Y761 (Y-OH_2^+) could not be found as all optimization efforts ended up with back-transfer of the proton to sLys. If we assume Y761 to be deprotonated initially (Y-O^-) the proton transfer from sLys is calculated to be exothermic by 5−6 kcal/mol. However, this is a highly unrealistic assumption, for the following reasons: (a) theoretical pK_a analysis predicts Y761 to be neutral (Y-OH) after sLys binding; (b) no base stronger than tyrosine is present in the binding pocket to generate the tyrosyl anion; and (c) Y761F mutation does not lead to complete loss of LSD1 activity as in the case of K661A mutation.[25] Considering the possibility that Y761 might act as initial single electron acceptor in LSD1, we computed the QM/MM spin densities (see Figure S4) using an extended QM region that included Y761 and several neighboring residues. We find that neither Y761 nor any other of these residues can accommodate unpaired electron density, thus ruling out a role as single-electron acceptor that would stabilize the reactant species (^{3}**RC**) on the SET pathway.

DISCUSSION

C−H bond activation represents an elementary step of many enzymatic reactions and is often achieved by the use of metal-containing active sites capable of activating the inert C−H bond. By contrast, LSD1 catalyzes amine (αC−H bond) oxidation in a metal-free manner. Studies on various amine oxidases[17] have shown that the oxidative power of flavin can be significantly enhanced by the flavoprotein environment, making it potent enough to oxidize the C−H bond of a methyl group. Several mechanisms have been proposed in the literature for the rate-determining C−H bond cleavage step during amine oxidation, namely HT, SET, and adduct-forming mechanisms (involving carbanions and polar nucleophiles). The goal of this work is to clarify the mechanism of LSD1-catalyzed amine oxidation through MD simulations as well as QM-only and QM/MM calculations. Toward this goal, we also investigated the role of active-site LSD1 residues, namely K661, Y761, and W695.

The present computational results provide detailed insight into the mechanism of the reductive half-reaction (i.e., amine oxidation). Energetically (see Tables 3 and 4), the HT mechanism is clearly favored over the SET mechanism, for which the overall activation barrier is about twice as high. For the SET mechanism, we could not identify Tyr761 or any other active-site residue as candidate (instead of FAD) for accepting the initial electron from sLys in our QM/MM calculations. Therefore, the high energy required to form initial radicals (^{3}**RC**) and the resulting high overall barrier for the SET pathway are not diminished by the involvement of neighboring protein residues. This is not the case for MAO-A, which has been reported to form a tyrosyl radical that is in equilibrium with a flavosemiquinone radical: This tyrosyl-flavin radical pair is long-lived enough to be visible in EPR and ENDOR spectra.[23,24] The main reason for the absence of such a stable radical formation in the case of LSD1 is probably the non-covalent binding of FAD to LSD1, as opposed to the covalent binding in the case of MAO-A.

The present theoretical calculations introduce two distinct orientations (upward and downward) of the non-reacting methyl group of sLys with respect to the C4a−N5 bond of FAD. Our calculations predict an amenable barrier for C−H bond oxidation (via the HT pathway) for each of the two

dx.doi.org/10.1021/ja403582u | *J. Am. Chem. Soc.* 2013, 135, 13400−13413

orientations, as opposed to the rather high barrier of 32 kcal/mol reported previously.[39] The computations also show that the flavoprotein environment can keep sLys in the less stable downward orientation, which in turn reduces the activation barrier for the HT mechanism by about 5 kcal/mol. The computed activation barrier in upward orientation (20.9 kcal/mol) corresponds to a rate of 0.17 min^{-1}, which is in seemingly good agreement with the experimentally observed average turnover rate of LSD1 with the substrate dimethyl lysine (8.10 ± 0.20 min^{-1}).[8] The less stable downward orientation (with a barrier of 15.4 kcal/mol) leads to a much higher reaction rate of 1330 min^{-1}; this overestimate of the rate is consistent with the well-known tendency of B3LYP to underestimate barrier heights.[46,55] The water-bridge motif and the protein environment are found to be vital for supporting this less stable orientation.

The QM/MM-level ^{15}N KIEs with Wigner tunneling corrections do not provide any clear preference for the HT or SET mechanism in LSD1, like in the case of MTOX[33] (see Table S10 for results, computational details, and comparison to the MTOX results). The ^{15}N KIE values are reverse (less than unity) and in good agreement with the computed and observed values for MTOX.[33] They are insensitive to different sLys orientations, unlike the deuterium KIE values for sLys with trideuterated methyl groups, which are found to be more sensitive (see Table S10). Compared with the downward orientation, the computed deuterium KIE values for the upward orientation are smaller (for both mechanisms) and closer to experiment (4.0 ± 0.2). However, as tunneling effects for hydrogen are quite pronounced and cannot be fully accounted by the Wigner corrections, we refrain from conclusions based on such comparisons with experiment.

The MD simulations with different protonation states of K661 and sLys indicate that the protonation state of K661 is crucial for binding sLys in the active site. If K661 and sLys are both kept protonated in an NVT ensemble, sLys does not properly bind, presumably because of electrostatic repulsions; it binds only when either sLys or K661 or both are deprotonated. To accommodate sLys in the active site in its protonated form (sLys-NH_3^+), protonated K661 (K661-NH_3^+) needs to release its proton to the microenvironment in order to accept the extra proton from sLys. This initial proton transfer prior to the demethylation process is crucial, since it ensures the availability of the lone pair on the amine moiety of sLys for the oxidation step. It is computed to be quite facile, with a QM/MM activation barrier of 8.4 kcal/mol for the proton transfer from sLys-NH_3^+ to K661-NH_2. The proton shuttle between sLys and K661 is supported by an extended water-bridge motif that involves three active-site water molecules and deprotonated K661. In all MD simulations, the three water molecules remain in the space between sLys, FAD, and K661; the W695 tryptophan residue fortifies the water-bridge motif through hydrogen-bonding (Figure 4), which is essential as seen from the significant loss of LSD1 activity observed upon W695 mutation.[25] In the case of the K661M mutant, the absence of the water-bridge motif prevents an efficient proton transfer, which is probably the main reason for the experimentally observed complete loss of LSD1 activity.[18] The computed QM/MM energy profiles show that K661M mutation has almost no effect on the activation barriers of the HT and the SET mechanism. It thus appears that K661 has no catalytic role in LSD1, as opposed to its analogue (K300) in maize polyamine oxidase (MPAO).[19] Instead, K661 is likely to act as the active site-base that initially accepts a proton from sLys. Although the protonation state of K661 does not significantly affect the rates of the HT pathway, the QM/MM energy profiles indicate that it controls the stability of the product species (^{1}**PC**): protonated K661 prevents formation of the iminium-FADH adduct and leads to an endothermic process, whereas deprotonated K661 allows for such adduct formation in an exothermic process (in downward orientation). This again underscores the need for having deprotonated K661 in the active site before and after substrate binding.

The other conserved active-site residue, Y761, stays in close proximity to sLys during the MD simulations. It helps to align the substrate in front of the *re*-face of the isoalloxazine ring of FAD via steric (repulsive) interactions. Our current calculations suggest that Y761 does not act as an active-site base or an initial electron acceptor, which has sometimes been assumed to be its role as a member of the aromatic cage motif conserved in amine oxidases.[23] Therefore, given its mere steric role, Y761 would not seem indispensable for the amine oxidation process, which is line with the only partial reduction of LSD1 activity upon Y761F mutation.[25]

As an alternative to the HT and SET pathways, adduct-forming mechanisms have been proposed in the literature.[29] They involve concerted or nonconcerted heterolytic cleavage of the αC−H bond and nucleophilic attack of the lone pairs on the amino nitrogen atom of the substrate. Depending on the type of proton acceptor, there are two types of adduct-forming mechanisms. In the carbanion mechanism, an active-site residue abstracts a proton from the αC−H bond of sLys (Cs−H1), whereas in the polar nucleophilic mechanism the N5 atom of FAD makes a nucleophilic attack on the proton (H1); the C4a atom of FAD is the electrophilic center and prone to the nucleophilic attack from the Ns atom of sLys. Methyl C−H activation is a demanding process that requires a very strong base. In LSD1, K661 is the only available active-site base present in the crystal structure,[10] which is able to extract the extra proton from the amine site (Ns) of sLys, but it is not potent enough to activate the C−H bond. Therefore, the carbanion mechanism can be ruled out for LSD1. Likewise, although the polar nucleophilic mechanism is considered feasible for the amine oxidation mediated by MAO-A and MAO-B,[29,56] it is quite unlikely for LSD1 due to the bulkiness of the substrate (caused by the methyl groups bonded to Ns of sLys), which will destabilize any adduct species located on Ns because of unfavorable steric interactions with the flavin rings. Our QM/MM calculations indicate that such adduct species (by covalent bonding of Ns of sLys with C4a of FAD) are less stable than the species appearing along the HT pathway (by covalent bonding of Cs of sLys with N5 of FAD).

We also investigated the feasibility of some other conceivable pathways for amine oxidation. First, we tried to transfer a hydride equivalent from the αC−H bond of sLys to the N1 instead of the N5 position, both via HT and SET mechanisms. This reaction has to surmount significantly higher activation barriers in both cases (60.5 and 48.8 kcal/mol for HT and SET, respectively), presumably due to the steric shielding of the N1 position by nearby LSD1 residues (Thr810 and Val811). These residues are likely to stabilize the reduced flavin ($FADH^-$) formed via HT, as has been suggested for other flavoenzymes.[17] Moreover, the HT pathway was previously predicted to have a lower barrier in the gas phase if the N1 atom were protonated prior to the transfer.[33] However, our QM/MM calculations predict a very high barrier of ca. 90 kcal/mol for the

protonation of the N1 position, demonstrating that this process is not feasible in the protein environment.

CONCLUSION

Classical MD simulations as well as QM-only and QM/MM calculations were performed to elucidate the catalytic mechanism of the rate-determining amine oxidation step in LSD1-mediated demethylation of histone tail lysine. We find that the hydride transfer (HT) pathway is clearly favored over other proposed mechanisms, including the radical (or single-electron transfer, SET) route as well as the carbanion and polar-nucleophilic mechanisms. QM/MM calculations predict activation barriers in the range of 15−21 kcal/mol for the α-CH bond cleavage on the HT pathway, in reasonable agreement with the experimentally observed rates.

The various features of LSD1 that facilitate oxidative CH bond cleavage were analyzed in detail. It was shown that substrate lysine can assume two distinct orientations in the binding pocket relative to the isoalloxazine moiety of FAD, which are denoted as downward and upward based on the orientation of the non-reacting methyl group of sLys. The computed activation barriers for the HT and SET pathways with downward orientation of sLys are consistently lower than those with the alternative upward orientation. In both orientations, the lone pair on sLys lies almost orthogonal to the FAD-isoalloxazine ring, allowing for favorable n(sLys)−π*-(C4a−N5 of FAD) orbital interactions at the rate-limiting HT TS. This TS is further stabilized by electrostatic interactions with the surrounding protein environment, as can be seen from the barriers computed with (15−21 kcal/mol) and without (31 kcal/mol) the protein environment. The QM/MM calculations also provide an explanation why the SET pathway is disfavored in LSD1. In the SET mechanism, the first electron transfer from sLys to FAD generates a sLys-FAD radical pair (^{3}**RC**), which is much higher in energy than the closed-shell reactant complex (^{1}**RC**). This is mainly due to the inability of surrounding LSD1 residues (in particular Y761) to delocalize the unpaired electron density in the ^{3}**RC** species and to form a stable radical pair with FAD (replacing sLys as partner of FAD).

Several active-site LSD1 residues (K661, Y761, and W695) and the conserved water-bridge motif assist the catalysis in a number of ways: Y761 has only a steric effect in positioning the reaction partners properly; K661 acts as an active-site base; and the water bridge is crucial for promoting the proton transfer, with W695 shielding and stabilizing this bridge (in line with experimental point mutation studies). According to our calculations, K661 plays a crucial role in the LSD1-mediated demethylation of lysine substrates: K661 will get deprotonated as sLys (mostly present in the protonated form, sLys-NH_3^+) enters the binding pocket and will then act as the base to accept a proton from sLys, thus helping to "liberate" the Ns(sLys) lone pair that is required for the subsequent transfer of a hydride equivalent to FAD; moreover, deprotonated K661 is required to make the overall HT process exothermic by allowing adduct formation in the product species (^{1}**PC**). The studies on the K661M mutant provide further information: the computed barriers for the HT and SET pathways are essentially unaffected by the mutation, indicating that K661 does not have a direct catalytic influence in amine (sLys) oxidation. On the contrary, the active-site base role of K661 is supported by MD simulations showing that K661M mutation disrupts the water-bridge motif and thus prevents deprotonation of sLys-NH_3^+ after binding to LSD1; this will reduce LSD1 activity (as also confirmed experimentally).

To summarize, our results on LSD1 are consistent with the available experimental evidence and the basic results of a recent computational investigation.[39] The present study goes beyond previous work by providing comprehensive insight into the LSD1-mediated dimethylamine (lysine) oxidation at the molecular level, which may be helpful for designing novel inhibitor molecules suitable for controlling an abnormal demethylation/methylation balance.

ASSOCIATED CONTENT

Supporting Information

Detailed information on model system preparation, optimized geometries of various stationary points, various structural properties, QM/MM energies, energy profiles, Wiberg bond orders, spin densities, ^{2}H and ^{15}N KIEs, and visualization of the extended QM region and of frontier orbitals. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

thiel@kofo.mpg.de

Present Address

[†]M.P.: Centre for Excellence in Basic Sciences, Health Centre, University of Mumbai, Vidhyanagari Campus, Mumbai 400098, India

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

B.K. acknowledges Profs. Burak Erman and Ozlem Keskin for valuable discussions at an early stage of this study.

REFERENCES

(1) Forneris, F.; Battaglioli, E.; Mattevi, A.; Binda, C. *FEBS J.* **2009**, *276*, 4304−4312.
(2) Hu, P.; Zhang, Y. *J. Am. Chem. Soc.* **2006**, *128*, 1272−1278.
(3) Culhane, J. C.; Cole, P. A. *Curr. Opin. Chem. Biol.* **2007**, *11*, 561−568.
(4) Shi, Y.; Whetstine, J. R. *Mol. Cell* **2007**, *25*, 1−14.
(5) Yang, M.; Culhane, J. C.; Szewczuk, L. M.; Jalili, P.; Ball, H. L.; Machius, M.; Cole, P. A.; Yu, H. *Biochemistry* **2007**, *46*, 8058−8065.
(6) Paik, W.; Kim, S. *Biochem. Biophys. Res. Commun.* **1973**, *51*, 781−788.
(7) Shi, Y.-J.; Lan, F.; Matson, C.; Mulligan, P.; Whetstine, J.; Cole, P.; Casero, R. A.; Shi, Y. *Cell* **2004**, *119*, 941−953.
(8) Forneris, F.; Binda, C.; Dall'Aglio, A.; Fraaije, M. W.; Battaglioli, E.; Mattevi, A. *J. Biol. Chem.* **2006**, *281*, 35289−35295.
(9) Gaweska, H.; Henderson Pozzi, M.; Schmidt, D. M. Z.; McCafferty, D. G.; Fitzpatrick, P. F. *Biochemistry* **2009**, *48*, 5440−5445.
(10) Forneris, F.; Binda, C.; Adamo, A.; Battaglioli, E.; Mattevi, A. *J. Biol. Chem.* **2007**, *282*, 20070−20074.
(11) Chen, Y.; Yang, Y.; Wang, F.; Wan, K.; Yamane, K.; Zhang, Y.; Lei, M. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 13956−13961.
(12) Huang, J.; Sengupta, R.; Espejo, A. A. B.; Lee, M. M. G.; Dorsey, J. a J.; Richter, M.; Opravil, S.; Shiekhattar, R.; Bedford, M. T.; Jenuwein, T.; Berger, S. L. *Nature* **2007**, *449*, 105−109.
(13) Wang, J.; Hevi, S.; Kurash, J. K.; Lei, H.; Gay, F.; Bajko, J.; Su, H.; Sun, W.; Chang, H.; Xu, G.; Gaudet, F.; Li, E.; Chen, T. *Nat. Genet.* **2009**, *41*, 125−129.
(14) Cho, H.-S.; Suzuki, T.; Dohmae, N.; Hayami, S.; Unoki, M.; Yoshimatsu, M.; Toyokawa, G.; Takawa, M.; Chen, T.; Kurash, J. K.;

Field, H. I.; Ponder, B. a J.; Nakamura, Y.; Hamamoto, R. *Cancer Res.* **2011**, *71*, 655−660.
(15) Smith, B. C.; Denu, J. M. *Biochim. Biophys. Acta* **2009**, *1789*, 45−57.
(16) Lohse, B.; Kristensen, J. L.; Kristensen, L. H.; Agger, K.; Helin, K.; Gajhede, M.; Clausen, R. P. *Bioorg. Med. Chem.* **2011**, *19*, 3625−3636.
(17) Fraaije, M. W.; Mattevi, A. *Trends Biochem. Sci.* **2000**, *25*, 126−132.
(18) Lee, M. G.; Wynder, C.; Cooch, N.; Shiekhattar, R. *Nature* **2005**, *437*, 432−435.
(19) (a) Polticelli, F.; Basran, J.; Faso, C.; Cona, A.; Minervini, G.; Angelini, R.; Federico, R.; Scrutton, N. S.; Tavladoraki, P. *Biochemistry* **2005**, *44*, 16108−16120. (b) Fiorillo, A.; Federico, R.; Polticelli, F.; Boffi, A.; Mazzei, F.; Di Fusco, M.; Ilari, A.; Tavladoraki, P. *FEBS J.* **2011**, *278*, 809−821.
(20) Pozzi, M. H.; Gawandi, V.; Fitzpatrick, P. F. *Biochemistry* **2009**, *48*, 1508−1516.
(21) Pozzi, M. H. *Arch. Biochem. Biophys.* **2010**, *498*, 83−88.
(22) Bruckner, R. C.; Winans, J.; Jorns, M. S. *Biochemistry* **2011**, *50*, 4949−4962.
(23) Li, M.; Binda, C.; Mattevi, A.; Edmondson, D. E. *Biochemistry* **2006**, *45*, 4775−4784.
(24) Rigby, S. E. J.; Hynson, R. M. G.; Ramsay, R. R.; Munro, A. W.; Scrutton, N. S. *J. Biol. Chem.* **2005**, *280*, 4627−4631.
(25) Stavropoulos, P.; Blobel, G.; Hoelz, A. *Nat. Struct. Mol. Biol.* **2006**, *13*, 626−632.
(26) Silverman, R. B.; Hoffman, S. J.; Catus, W. B. *J. Am. Chem. Soc.* **1980**, *102*, 7126−7128.
(27) Silverman, R. B. *Acc. Chem. Res.* **1995**, *28*, 335−342.
(28) Silverman, R. B. *Prog. Brain Res.* **1995**, *106*, 23−31.
(29) Miller, J. R.; Edmondson, D. E. *Biochemistry* **1999**, *38*, 13670−13683.
(30) Gaweska, H.; Fitzpatrick, P. F. *BioMol. Concepts* **2011**, *2*, 365−377.
(31) Scrutton, N. S. *Nat. Prod. Rep.* **2004**, *21*, 722−730.
(32) Fitzpatrick, P. F. *Arch. Biochem. Biophys.* **2010**, *493*, 13−25.
(33) Ralph, E. C.; Hirschi, J. S.; Anderson, M. A; Cleland, W. W.; Singleton, D. A; Fitzpatrick, P. F. *Biochemistry* **2007**, *46*, 7655−7664.
(34) Schmidt, D. M. Z.; McCafferty, D. G. *Biochemistry* **2007**, *46*, 4408−4416.
(35) McCann, A. E.; Sampson, N. S. *J. Am. Chem. Soc.* **2000**, *122*, 35−39.
(36) Chen, Z.; Zhao, G.; Martinovic, S. *Biochemistry* **2005**, 15444−15450.
(37) Forneris, F.; Binda, C.; Vanoni, M. A.; Mattevi, A.; Battaglioli, E. *FEBS Lett.* **2005**, *579*, 2203−2207.
(38) Karasulu, B.; Keskin, O.; Erman, B. In *5th International Symposium on Health Informatics and Bioinformatics*; IEEE: Antalya, 2010; pp 197−205.
(39) Kong, X.; Ouyang, S.; Liang, Z.; Lu, J.; Chen, L.; Shen, B.; Li, D.; Zheng, M.; Li, K. K.; Luo, C.; Jiang, H. *PLoS One* **2011**, *6*, e25444.
(40) Brooks, B. R.; Brooks, C. L., III; MacKerell, A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M. *J. Comput. Chem.* **2009**, *30*, 1545−1615.
(41) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. *J. Comput. Phys.* **1977**, *23*, 327−341.
(42) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Jr., J. A. M.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*, Revision A.02; Gaussian Inc.: Pittsburgh, PA, 2009.
(43) Becke. *J. Chem. Phys.* **1993**, *98*, 5648−5652.
(44) Zhao, Y.; Truhlar, D. G. *Theor. Chem. Acc.* **2007**, *120*, 215−241.
(45) Vydrov, O. A; Scuseria, G. E. *J. Chem. Phys.* **2006**, *125*, 234109.
(46) Lynch, B. J.; Fast, P. L.; Harris, M.; Truhlar, D. G. *J. Phys. Chem. A* **2000**, *104*, 4811−4815.
(47) Sherwood, P.; de Vries, A. H.; Guest, M. F.; Schreckenbach, G.; Catlow, C. R. A.; French, S. A.; Sokol, A. A.; Bromley, S. T.; Thiel, W.; Turner, A. J.; Billeter, S.; Terstegen, F.; Thiel, S.; Kendrick, J.; Rogers, S. C.; Casci, J.; Watson, M.; King, F.; Karlsen, E.; Sjovoll, M.; Fahmi, A.; Schäfer, A.; Lennartz, C. *J. Mol. Struct. THEOCHEM* **2003**, *632*, 1−28.
(48) Ahlrichs, R.; Bär, M.; Baron, H. -P.; Bauernschmitt, S.; Böcker, S.; Ehrig, M.; Eichkorn, K.; Elliott, S.; Furche, F.; Haase, F.; Häser, M.; Horn, H.; Huber, C.; Huniar, U.; Kattannek, M.; Kölmel, C.; Kollwitz, M.; May, K.; Ochsenfeld, C.; Öhm, H.; Schäfer, A.; Schneider, U.; Treutler, O.; von Arnim, M.; Weigend, F.; Weis, P.; Weiss, H. *TURBOMOLE*, v. 6.3; 2011.
(49) Grimme, S. *J. Comput. Chem.* **2006**, *27*, 1787−1799.
(50) Smith, W.; Forester, T. R. *J. Mol. Graphics* **1996**, *14*, 136−141.
(51) de Vries, A. H.; Sherwood, P.; Collins, S. J.; Rigby, A. M.; Rigutto, M.; Kramer, G. J. *J. Phys. Chem. B* **1999**, *103*, 6133−6141.
(52) Billeter, S. R.; Turner, A. J.; Thiel, W. *Phys. Chem. Chem. Phys.* **2000**, *2*, 2177−2186.
(53) Gordon, J. C.; Myers, J. B.; Folta, T.; Shoja, V.; Heath, L. S.; Onufriev, A. *Nucleic Acids Res.* **2005**, *33*, W368−371.
(54) Harris, C.; Pollegioni, L.; Ghisla, S. *Eur. J. Biochem.* **2001**, *268*, 5504−5520.
(55) Lingwood, M.; Hammond, J. R.; Hrovat, D. A.; Mayer, J. M.; Borden, W. T. *J. Chem. Theory Comput.* **2006**, *2*, 740−745.
(56) Edmondson, D. E.; Binda, C.; Mattevi, A. *Arch. Biochem. Biophys.* **2007**, *464*, 269−276.

APPENDIX B

Amine Oxidation Mediated by N-Methyltrytophan Oxidase: Computational Insights into the Mechanism, the Role of Active-Site Residues and Covalent Flavin Binding

Bora Karasulu, and Walter Thiel

ACS Catal., **2015,** 5, pp 1227-1239.

ACS Catalysis

Research Article

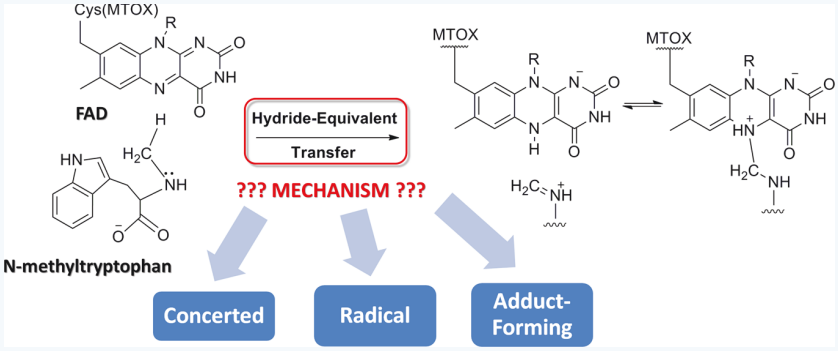
pubs.acs.org/acscatalysis

Amine Oxidation Mediated by *N*-Methyltryptophan Oxidase: Computational Insights into the Mechanism, Role of Active-Site Residues, and Covalent Flavin Binding

Bora Karasulu and Walter Thiel*

Max-Planck-Institut für Kohlenforschung, Kaiser-Wilhelm-Platz 1, 45470 Mülheim, Germany

Supporting Information



ABSTRACT: Amine oxidation, a process widely utilized by flavoprotein oxidases, is the rate-determining step in the three-step demethylation of *N*-methyltryptophan (NMT) catalyzed by *N*-methyltryptophan oxidase (MTOX), which employs a covalently bound flavin adenine dinucleotide (FAD) as cofactor. For the required transfer of a hydride ion equivalent, three pathways (direct/concerted, radical, and adduct-forming/polar nucleophilic) have been proposed, without a consensus on which one is commonly used by amine oxidases. We combine theoretical pK_a analysis, classical molecular dynamics (MD) simulations, and pure quantum mechanics (QM) and hybrid QM/molecular mechanics (QM/MM) calculations to provide molecular-level insights into the catalytic mechanism of NMT oxidation and to analyze the role of MTOX active-site residues and covalent FAD incorporation for NMT binding and oxidation. The QM(B3LYP-D2/6-31G(d))/CHARMM results clearly favor a direct concerted hydride transfer (HT) mechanism involving anionic NMT as the reactive species. On the basis of classical canonical MD simulations and QM/MM calculations of wild-type MTOX and two mutants (K341Q and H263N), we propose that the K341 residue acts as an active-site base and electrostatically, whereas H263 and Tyr249 only support substrate alignment. Covalent FAD binding leads to a more bent isoalloxazine moiety, which facilitates the binding of anionic NMT but increases the catalytic activity of FAD only slightly.

KEYWORDS: *flavin adenine dinucleotide (FAD), N-methyltryptophan (NMT), biocatalysis, quantum mechanics/molecular mechanics (QM/MM), molecular dynamics, flavoprotein oxidase*

INTRODUCTION

N-Methyltryptophan oxidase (MTOX) catalyzes the oxidative demethylation of a modified amino acid, *N*-methyl-L-tryptophan (*N*-methyltryptophan, NMT), which yields tryptophan, formaldehyde, and hydrogen peroxide as products. Although MTOX can also act on other similar amines (sarcosine, carbinol amines, and thioglycolate), it displays the highest activity with its genuine substrate, NMT (k_{cat} = 4600 min^{-1} and $k_{cat,app}$ = 990 min^{-1}).[1,2] MTOX is a member of the amine oxidase family that also includes monomeric sarcosine oxidase (MSOX), heterotetrameric sarcosine oxidase (TSOX), pipecolare oxidase (PIPOX), and fructosyl amino acid oxidase. Among the members of this family, MTOX shares the highest structural homology (43% sequence identity) with MSOX[3] that is, allegedly, the best characterized member of this family.[4] MTOX and MSOX possess highly analogous reactive pockets. Like all members in this family, they covalently bind a flavin through a conserved cysteine residue (MTOX/MSOX: Cys308/Cys315).[2] Both MTOX and MSOX contain a flavin adenine dinucleotide (FAD) in the form of a covalently bound [8α-(*S*-cysteinyl)FAD] complex. The covalent linkage prevents premature dissociation of an oxidized FAD from MSOX, prior to amine oxidation.[4] Along with the highly electropositive arrangement of the immediate protein environment (containing highly basic residues), the covalent binding of FAD has been proposed to afford a sizable enhancement of the reduction

Received: October 31, 2014
Revised: January 10, 2015

DOI: 10.1021/cs501694q
ACS Catal. 2015, 5, 1227−1239

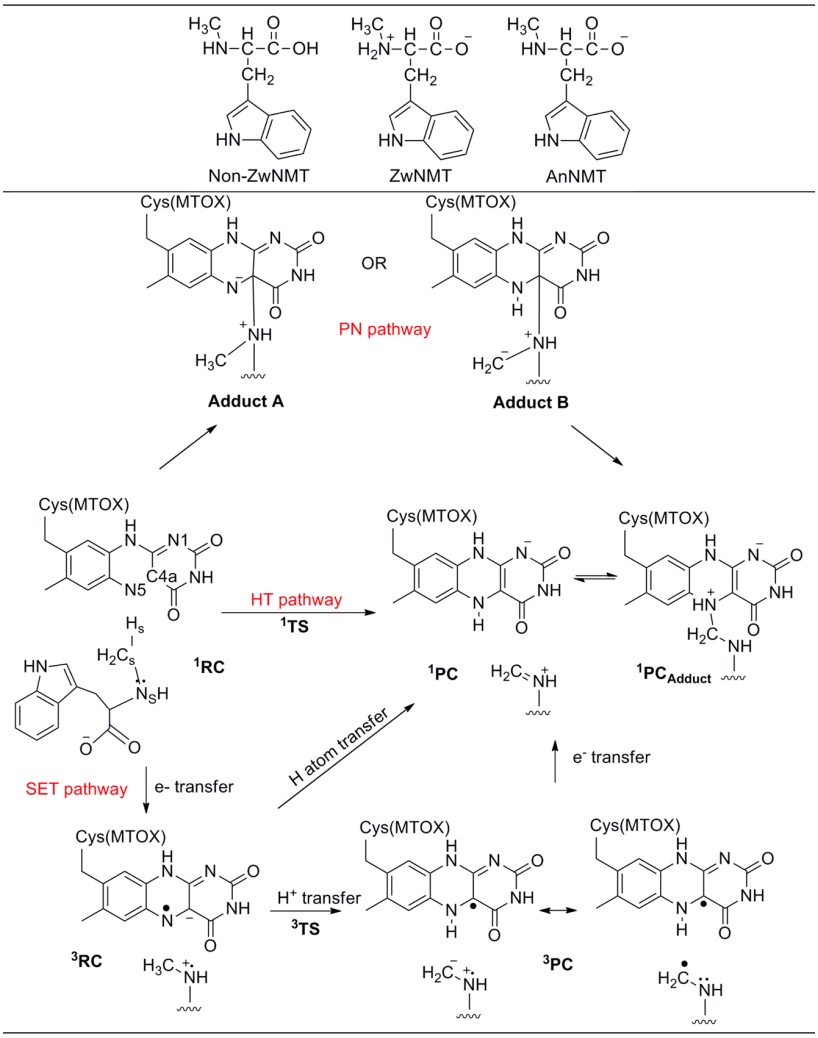


Figure 1. (top) Three forms of *N*-methyltryptophan (NMT): nonzwitterionic (Non-ZwNMT), zwitterionic (ZwNMT), and anionic (AnNMT). (bottom) Polar nucleophilic (PN), direct hydride transfer (HT), and radical (SET) pathways proposed for the amine oxidation step of NMT demethylation catalyzed by MTOX, shown for AnNMT as a representative substrate. In the bottom panel, middle left, labels are included for mechanistically relevant atoms. Abbreviations: RC/PC, reactant/product complex; TS, transition state.

potential of FAD (estimated to ca. 120 mV) in comparison to the noncovalently bound counterparts.[4,5]

Studies on the pH-dependent activity and absorption spectroscopy of MTOX[6,7] and MSOX[8] indicate binding of the substrate in a less reactive neutral form (as zwitterionic or nonzwitterionic species, ZwNMT and Non-ZwNMT; see Figure 1), which is dominant in solution at pH 8.0. Given the optimum working conditions of MTOX (pH 8.0),[1] the substrate NMT is likely ionized within the protein environment by an active-site base to yield the more reactive anionic form (AnNMT; Figure 1). Anionic NMT forms a charge transfer (CT) complex with an oxidized flavin that serves as charge acceptor.[6−8] This complex is thought to be stabilized by pairs of basic residues (MSOX/MTOX: R49/R48-K265/K259 and R52/R51-K348/K341 on the *Si* and *Re* faces of flavin, respectively).[4,8] The deprotonation is believed to occur after binding in the reactive pocket and before the rate-determining C−H cleavage.[7] Two active-site residues in MSOX (Tyr317 and His269) were considered to act as base, but this could be ruled out through mutagenesis studies.[9,10] Deprotonation in MSOX could occur via a putative proton relay system involving N5 of FAD, conserved K265 (homologous to K259 in MTOX), and several nearby water molecules.[9,11] However, K259 mutation in MTOX yielded only a 60-fold rate decrease in the reductive half-reaction, compared with a 2300-fold decrease in the reaction of fully reduced flavin with oxygen, and hence K259 was assigned as the main oxygen activation site.[6] Considering the strong basic properties and the close proximity to FAD-N5 and the backbone of NMT, the K341/K348 residues (MTOX/MSOX) are promising candidates for this role, in line with the almost complete loss of amine oxidation activity in corresponding mutants of MTOX (K341Q, K341R, and K341L).[2]

Catalytic Mechanisms for Oxidative Demethylation of Amines. In general, amine demethylation involves the removal of a methyl group from the N terminus of a given amine substrate via a redox process. Shi et al.[12] proposed a three-step catalytic mechanism for histone (H3K4met) demethylation that can be generalized to amine demethylation mediated by other flavoenzyme oxidases. The first step is the activation or

DOI: 10.1021/cs501694q
ACS Catal. 2015, 5, 1227−1239

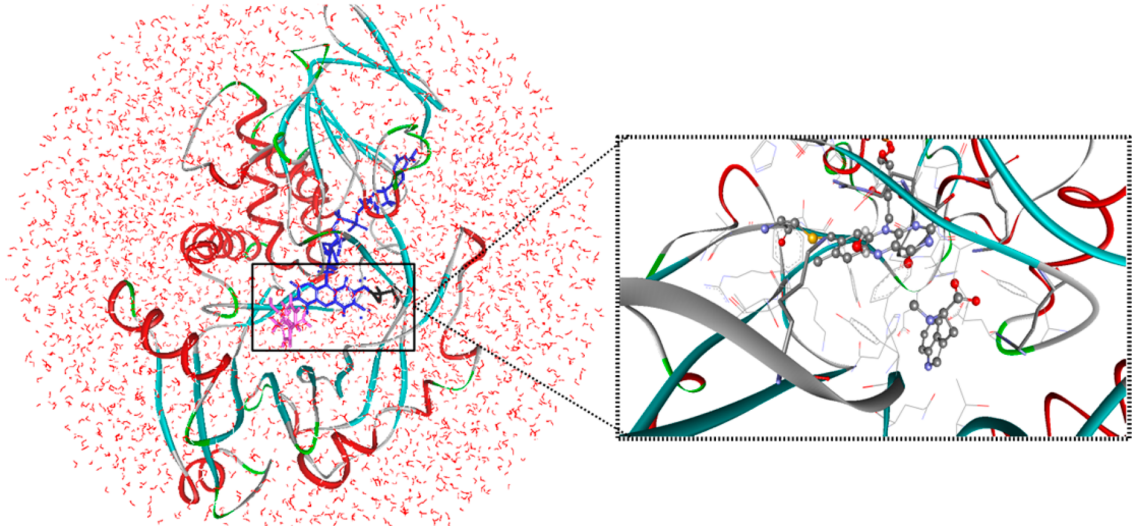


Figure 2. Representative snapshot of the simulation system before the MD simulations: FAD, Cys308, and substrate NMT are shown in stick representations in blue, black, and purple, respectively. In the enlarged active-site view, the atoms in the ball-and-stick representation are colored according to atom type (gray, carbon; white, hydrogen; blue, nitrogen; red, oxygen; yellow, sulfur).

cleavage of the α-CH bond in the methyl group of the amine substrate, followed by an oxidation via transfer of a hydride equivalent. Primary deuterium isotope studies indicate that the cleavage of the chemically inert α-CH bond is the rate-limiting step in different amine oxidases.[7,13] Amine oxidation leads to a two-electron reduction of the flavin (FAD in MTOX), which is then (in an oxidative half-reaction) reoxidized by molecular oxygen, with formation of hydrogen peroxide. This reoxidation is a complementary process that proceeds with very high rates,[6,14] presumably via a modified ping-pong mechanism in MTOX.[15] In the subsequent two steps of demethylation, the formed iminium intermediate is hydrolyzed, and the resulting carbinol amine spontaneously rearranges to yield formaldehyde and the demethylated amine.

The rate-limiting first step of amine demethylation requires the transfer of a hydride equivalent from the amine substrate to FAD. The two electrons and the proton can be transferred in different order, which gives rise to several possible mechanisms: namely, a direct hydride transfer (HT), a radical mechanism via single-electron transfer (SET),[16−18] and an adduct-forming polar nucleophilic mechanism.[19] These mechanisms are depicted in Figure 1 for the case of anionic NMT oxidation. The simplest one is the direct concerted transfer of the two electrons and the proton as a hydride anion from the α-carbon atom of the substrate to FAD. In the radical and adduct-forming mechanisms, there are intermediate species.

Numerous published studies favor different mechanisms for different amine oxidases (see refs 20−25 for comprehensive reviews). ^{15}N kinetic isotope effect measurements and computations on MTOX did not provide clear support for either the HT or SET mechanism.[7,26] The SET mechanism was supported on the basis of cyclopropyl inhibitor studies on LSD1,[27] but this was later challenged[13] by virtue of the fact that some other flavoprotein oxidases, for which hydride transfer is considered the most likely mechanism, are also inactivated by cyclopropyl inhibitors.[28,29] The absence of any conceivable intermediate (i.e., a flavin or amine radical species) in kinetic, EPR, ENDOR, and other spectroscopic studies on MTOX[1,30] as well as other amine oxidases[11,13,31] was proposed to support the HT mechanism. However, the failure to detect any intermediates was considered to be inconclusive by others,[16,26] in view of the following scenario. The initial SET process could be reversible but not rate-limiting, with the back-transfer being much faster, which would lead to low concentrations of the short-lived radical intermediates that could be undetectable by experimental techniques.

These conflicting experimental findings on the catalytic mechanism of an amine oxidation call for a theoretical study to unravel molecular-level details that are not directly observable in experiments. Among the published computational studies on amine oxidases, only a few cover some aspects within the scope of the current work. A pure quantum mechanical (QM) investigation on MTOX-mediated amine oxidation addressed a simple model system consisting of a truncated isoalloxazine moiety of the flavin and dimethylamine ($NH(CH_3)_2$) as the substrate.[26] The computed DFT(B3LYP) energies were found to favor the direct hydride mechanism over the radical-SET mechanism for MTOX.[26] Another QM-only DFT(M06-2X) investigation was carried out[32] on monoamine oxidase (MAO), which binds FAD covalently through a cysteine residue. Vianello et al.[32] proposed a two-step hydride transfer scheme as the working mechanism for MAO-B, involving an adduct between reduced flavin and the imine species resulting from the transfer of a hydride equivalent. QM/molecular mechanical (QM/MM) studies on enzymatic demethylation are still missing for MTOX, but a few are available for other similar flavoprotein amine oxidases. In recent QM/MM work,[33] we addressed lysine-specific demethylase 1 (LSD1), an oxidase that binds a FAD noncovalently and works with aliphatic methyllysine substrates. Direct hydride transfer was shown to be the most feasible mechanism in LSD1, the formation of the imine-FAD adduct was confirmed (through a N5−C4a bond, similar to the situation in MAO[32]), and the lysine residue in the conserved Lys-H_2O-N5 motif (K661, analogous to K259 in MTOX) was identified as the active-site base responsible for ionizing the protonated substrate. Another recent QM/MM study considered human MAO-B with benzylamine as substrate

DOI: 10.1021/cs501694q
ACS Catal. 2015, 5, 1227−1239

and supported a concerted asynchronous polar nucleophilic mechanism.[34]

In this article, we report the results from classical MD simulations as well as QM-only and QM/MM calculations on the amine oxidation step of *N*-methyltrytophan demethylation mediated by MTOX. The objective is to probe the catalytic mechanism in the amine oxidation step and to gain detailed molecular-level insight into the role of the active-site MTOX residues. We elaborate on the role of H263 and K341 as potential active-site bases that may ionize the enzyme−substrate complex. We complete our discussion by addressing the role of covalent binding of the cofactor FAD during substrate binding and amine oxidation.

COMPUTATIONAL METHODOLOGY

The simulations reported here are based on the currently only available crystal structure of MTOX (PDB code 2UZZ[2]) from *Escherichia coli* in complex with FAD, which is covalently bound via Cys308 and acts as an oxidative agent for NMT. The crystal structure lacks crystal water molecules due to its low resolution, 3.2 Å, and does not contain a substrate (NMT) in the active site. Therefore, the substrate was modeled into the binding pocket through flexible docking procedures using Vina[35] (see the Supporting Information for further details). The docking of NMT resulted in a structure with a good binding affinity (−7.6 kcal/mol), which agrees well with the proposed structural model for NMT in MTOX (prepared using dimethylglycine in MSOX as template)[2] in terms of H-bonding and π−cation interactions. In addition, one crucial water molecule was added manually, which forms the conserved Lys-H_2O-N5 motif proposed for MTOX[6] and other similar oxidases (such as MSOX[36]). The simulation system consisted of ca. 18000 atoms (see Figure 2). It was neutralized and solvated in a water droplet (with a radius of 35 Å around the center of mass of its main components: substrate, flavin, and enzyme) in analogy to our previous study.[33] More detailed information on the setup procedure is given in the Supporting Information.

The classical MD simulations were performed with CHARMM[37] at a constant temperature (300 K) using a time step of 1 fs. The MD production runs were long enough (20 ns) to provide good starting points for the subsequent QM/MM calculations. The MM force field parameters for the three forms of the nonstandard NMT substrate were prepared by modifying the standard Trp parameters. Force field parameters for FAD were adopted from the literature[38] and modified to account for the covalent binding of Cys308 at the C8-methyl position. CHARMM22 parameters[39] and the TIP3P model[40] were used for all standard residues and for water molecules, respectively. The protonation states of the protein residues at the optimal pH of MTOX (pH 8.0)[1] were determined with the PROPKA software.[41]

QM-only calculations were carried out using density functional theory (DFT) and the Gaussian09 program suite.[42] For benchmarking purposes, four different functionals (B3LYP[43] and B3LYP-D2 with Grimme-type dispersion corrections,[44] M06-2X,[45] LC-ωPBE,[46] and ωB97xD[47]) were utilized in combination with the 6-31G(d) or TZVP basis set. Geometry optimizations and the following vibrational analysis were done either in the gas phase or in water described by an implicit solvent model (CPCM[48]), without any constraints in the optimizations. Intrinsic reaction coordinate (IRC) calculations were used for locating the reactant and product complexes (RC and PC), starting from the transition state (TS) connecting them.

Hybrid QM/MM studies of the full simulation system were performed with the ChemShell program suite.[49] The QM part of the system was treated at the DFT level (B3LYP-D2 with Grimme-type dispersion corrections[44]) using Gaussian09. The MM calculations were handled by the DL_POLY code[50] implemented in ChemShell using the aforementioned force-field parameters. The QM/MM treatment employed electrostatic embedding in combination with the charge-shift scheme,[51] and the atoms at the QM/MM boundary were treated by the link-atom approach.[49] QM/MM geometry optimizations were carried out with the hybrid delocalized internal coordinate (HDLC) optimizer[52] implemented in ChemShell. Starting geometries for optimizations were taken from several snapshots of the canonical MD ensembles, and only residues within 15 Å of FAD were optimized in order to reduce the computational burden and also to retain the overall protein structure. The optimized structures were subject to numerical force constant calculations in ChemShell to determine the vibrational modes and to characterize the optimized stationary points (i.e., one single negative eigenvalue of the corresponding Hessian matrix for TS, none for minima). Gibbs free energies (ΔG) and other thermodynamic properties were evaluated using the standard rigid-rotor harmonic-oscillator approximation. Theoretical reaction rates were approximated from the relevant Gibbs free energy barrier $\Delta G^{\ddagger}$ (the free energy of activation) in the usual manner using the Arrhenius equation.

In both QM-only and QM/MM calculations, the ground-state singlet and the lowest triplet states were described using restricted and unrestricted Kohn−Sham (RKS and UKS) treatments, respectively. In addition, UKS calculations were performed for putative open-shell singlet species to check whether they may yield an open-shell (radical-type) configuration with energy lower than that of the closed-shell configuration. The residues included in the QM parts were truncated at appropriate sp^3-hybridized carbon atoms whenever possible. To be specific, FAD was represented by a lumiflavin comprised of the isoalloxazine ring, with a cut at the C1′−N10 bond (thus excluding the side chain), whereas Cys308 was truncated at the backbone. Lysine (K341, when included) was truncated at its C_β−C_γ bond (without the backbone part).

RESULTS AND DISCUSSION

Changes in the Active Site upon Substrate Binding. Theoretical acid dissociation constants (pK_a) were computed with PROPKA.[41] The resulting values are compiled for important active-site residues and NMT in Table 1. The pK_a analysis suggests that, at the optimum working conditions of MTOX (pH 8.0),[1] the stepwise addition of subunits (FAD and NMT) does not cause any change in the protonation states of key active-site residues. In contrast, isolated NMT is predicted to be in its zwitterionic form like other amino acids, whereas it is most likely ionized and assumes the anionic form in the binding pocket, as also indicated by the experimental pH profiles of MTOX with sarcosine as substrate.[7]

Judging from the computed pK_a values, the potential candidates as an active-site base that may receive a proton from the substrate amine prior to oxidation are the neutral lysines (K218, K259, K341) and the conserved H^+ exchanger (His263, protonated at the N_E position). In view of their close proximity to the backbone of the substrate NMT,[2] His263 and

DOI: 10.1021/cs501694q
ACS Catal. 2015, 5, 1227−1239

Table 1. Computed pK_a Values at Selected NMT Atoms and MTOX Residues in Different Model Systems: Isolated NMT and MTOX as well as MTOX-FAD and MTOX-FAD-NMT Complexes

atom/residue ID	only NMT	only MTOX[a]	MTOX-FAD[a]	MTOX-FAD-NMT[a]
NMT-N_{ter}	10.0			4.5
NMT-C_{ter}	5.0			−1.9
NMT-N_{ar}	4.5			2.6
Arg48		11.3	10.3	10.3
Arg51		11.4	11.3	13.1
Lys218		8.0	7.8	7.6
Tyr249		13.1	14.0	14.1
Lys259		5.9	5.8	5.7
His263		3.1	2.5	0.3
Lys341		6.2	5.6	7.7

[a]With a single crystal water molecule (see Computational Methodology).

K341 seem most suitable for this role. K259 is part of the Lys-H_2O-N5 motif conserved across amine oxidases and has already been shown to be the site of oxygen activation in MTOX.[6] The homologue of His263 in MSOX (H269) apparently does not act as an active-site base, since the turnover rate is only twice as slow as that in the H269N mutant.[10] Given the high structural homology between MSOX and MTOX,[3] one would not expect His263 to be the active-site base in MTOX either; however, this still needs to be proven. Unlike His263, which is essentially deprotonated before and after substrate binding (pK_a values of 2.5 and 0.3; Table 1), an equilibrium of the protonated and unprotonated forms is predicted for K341 (pK_a values of 5.6 and 7.7) that is slightly shifted toward the unprotonated form after substrate binding.

In contrast, the other previously nominated candidate, Tyr249, remains protonated/neutral in all cases (see Table 1), disabling it from taking up protons and acting as a base. This suggests that Tyr249 mainly has a steric role in substrate binding, which is consistent with the fact that mutation of its homologue in MSOX (Tyr317) only leads to a 20-fold decrease in the maximum rate of the reductive half-reaction.[9]

Dynamics of Substrate Binding. To cover all possible binding scenarios, we simulated 12 different NVT ensembles, which arise from combining the two conceivable protonation states of K259 and K341 with the three forms of the substrate NMT, while keeping K218 and His263 always unprotonated/neutral. We also investigated the K341Q mutant, with two protonation states of K259, to further analyze the role of K341. Some characteristic features extracted from each ensemble are compiled in Table S1 in the Supporting Information for direct comparison. The pK_a analysis does not give a clear preference between the nonzwitterionic or zwitterionic forms of neutral NMT (before ionization) in the active site. In the following, we thus consider both possible binding scenarios for neutral NMT on the basis of the NVT ensembles of wild-type MTOX.

In the case of zwitterionic NMT, the binding mode of the substrate is mainly determined by the protonation state of K341 (see Table S1 in the Supporting Information). If K341 is deprotonated, NMT remains rather far away in an unreactive position (R(Cs−N5) = 5.5−7.5 Å). NMT-NH_2 is in H-bonding distance with Gln247 (via O_E) and also in close contact (a) with the COO^- terminus and K341 (via bulk water) for deprotonated K259 or (b) with His263 for protonated K259. On the other hand, if K341 is protonated, the two subunits, FAD and NMT, are packed more closely (R(Cs−N5) = 3.4−4.2 Å), which is mainly due to the enhanced H-bond donor character of K341. In this case, NMT-NH_2 is engaged in a H-

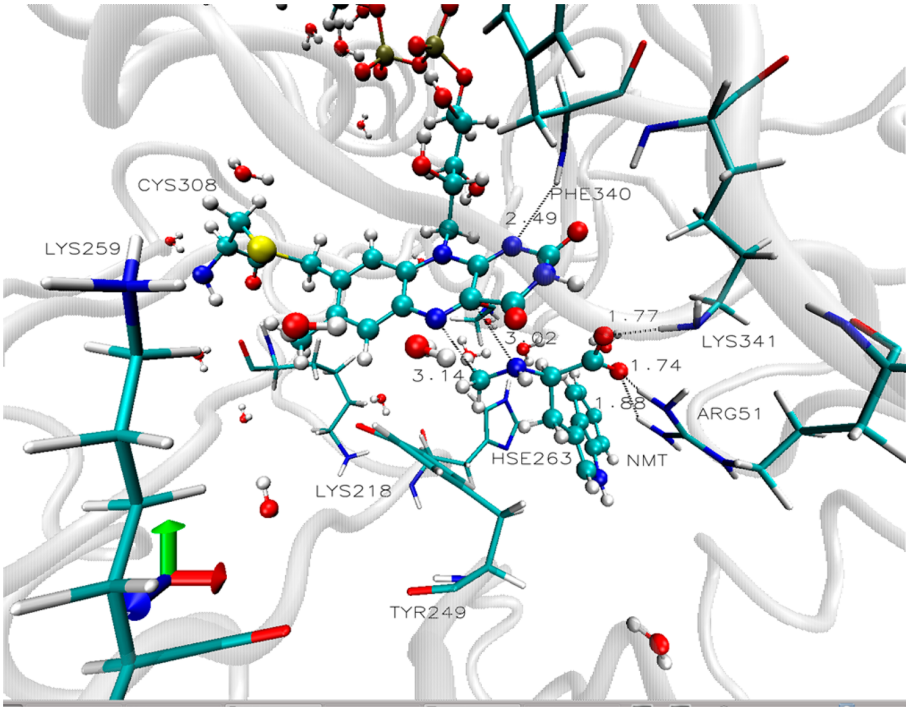


Figure 3. Typical snapshot from the NVT ensemble for AnNMT (K259, protonated; K341, deprotonated). Depicted are FAD and NMT (ball and stick), important MTOX residues (stick), and the MTOX structure (cartoon).

DOI: 10.1021/cs501694q
ACS Catal. 2015, 5, 1227−1239

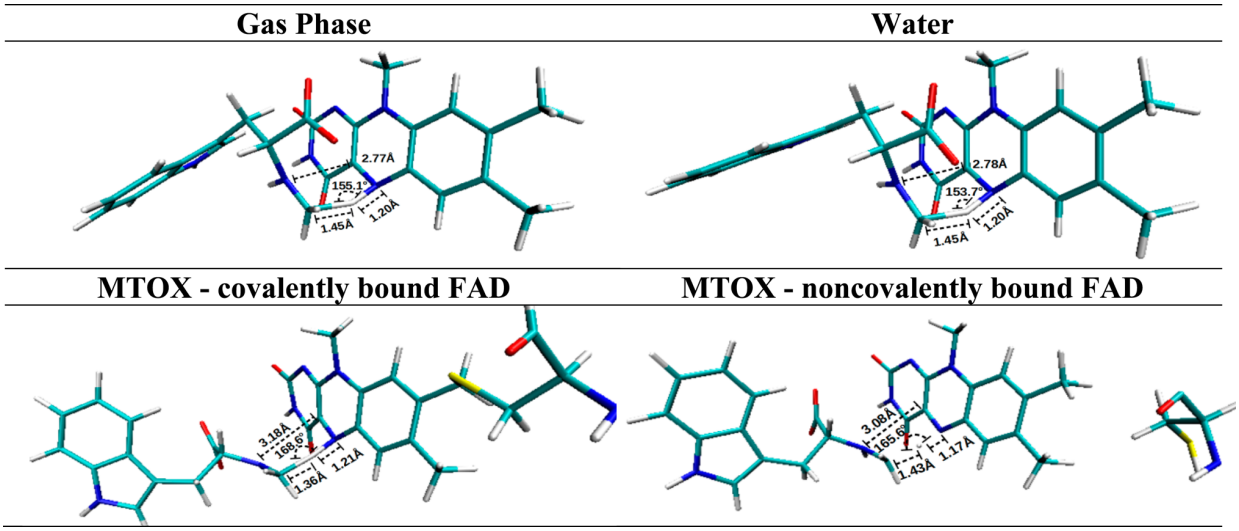


Figure 4. Geometries of the transition state for the HT pathway (^{1}**TS**) in different environments: (top) QM(B3LYP-D2/6-31G(d)) structures; (bottom) QM/MM structures (QM region only). Key structural parameters are included; more data are given in Tables S6 and S7 in the Supporting Information. See Figure 1 for labels and the text for the model system definitions.

bonding network with Gly337 (via backbone-O), His263, and bulk water.

Nonzwitterionic NMT is found to exhibit proper and stable binding only when K259 is protonated (see Table S1 in the Supporting Information). In this case, the N-terminus of NMT has no H-bonding partner, and K341 is the main H-bonding partner of the C terminus in line with its designated role as active-site base. The protonation state of K341 determines the other partner in the H-bonding network (Arg51-Tyr54 or Leu49) as well as the compactness of the reactive centers on FAD and NMT, with unprotonated K341 providing a more compact binding (Table S1).

The H-bonding networks encountered in different NVT trajectories suggest that the extra proton on the N-terminus of ZwNMT can be delivered to (a) the nearby bases (His263 and K341), (b) the N1 atom on FAD via Gly337, or (c) the C terminus of NMT (with conversion to the nonzwitterionic form). In contrast, the extra proton on the C terminus of Non-ZwNMT can be delivered easily to K341, which is always located in H-bonding distance of Non-ZwNMT. The NVT ensembles show the formation of a proton relay system involving Tyr54, His338, and K341 (see Figure S1 in the Supporting Information), while His263 has direct access to bulk water in each of the NVT ensembles. Hence, both K341 and His263 are suitable to act as a base.

On the other hand, the putative proton relay system proposed for MSOX[9,11] via a proton shuttle to N5 or K259 (direct or via water molecules) seems unlikely in MTOX. This is because such a proton shuttle is not seen in any of the NVT ensembles with a neutral NMT form, except for the special case of Non-ZwNMT with both K259 and K341 being unprotonated. In this latter case, a direct H bond between N5 and COOH of NMT is formed for very short periods, making this interaction fleeting and unreliable. However, a direct proton shuttle (via water molecules and without including N5) between the N terminus of NMT and K259 is not found in any of these NVT trajectories, as opposed to the case of its homologue (K661) in LSD1.[33] Taken together, these findings are congruent to the proposed role of K259 in oxygen activation rather than in amine oxidation, as deduced from experimental mutation studies.[6]

A stable binding of AnNMT, the anticipated reactive form, is achieved only in one case: when K259 is protonated and K341 is unprotonated (see Figure 3). In the corresponding MD trajectory, two conditions are satisfied that were previously shown[33,34] to be required to facilitate the transfer of a hydride equivalent, namely the close proximity of the reactive centers (N5 and Cs, average distance 3.37 Å) and a nonbonded Ns−C4a interaction (average distance 3.09 Å) involving the Ns lone pair and the π^* orbital (C4a−N5). The H-bond donors (K341, R51, and Y54) appear to be important for the close packing of the two subunits. Moreover, this MD trajectory sustains the N5-H_2O-Lys motif that is conserved across amine oxidase families,[6] with an expansion of the standard motif by an extra water molecule.

The microenvironment in MTOX has a delicate electrostatic balance, and hence changes in the active site may result in partial or complete failure of AnNMT binding. To quote a specific example, when K218 is protonated, we could not obtain a proper reactive binding of the substrate (i.e., R(N5−Cs) > 6.0 Å) for any combination of the K259 and K341 protonation states. On the other hand, if His263 were fully protonated (on both N_D and N_E) in contrast to the pK_a analysis (*vide supra*), there would be a loose yet reactive binding of the substrate (R(N5−Cs) = ca. 3.59 Å). The trajectory for the H263N mutant, in contrast, features three different arrangements, with R(N5−Cs) distances of ca. 4.1, 5.6, and 7.0 Å, respectively (see Figure S2 in the Supporting Information), which suggests that any reactive binding in the H263N mutant (R(N5−Cs) = ca. 4.1 Å) is clearly not as persistent as in the wild type (see Figure S2). This supports the role of His263 (His269 in MSOX) of aligning the substrate in a reactive position via π-stacking interactions.

The K341Q mutant appears to bind the substrate in a reactive orientation (albeit somewhat looser, R(N5−Cs) = ca. 3.31 Å; Figure S2 in the Supporting Information). The H-bonding partners of the NMT/C terminus differ slightly in the mutant and the wild type: R51 and K341/Q341 are common to

DOI: 10.1021/cs501694q
ACS Catal. 2015, 5, 1227−1239

Table 2. QM-Only Gibbs Free Relative Energies (in kcal/mol) of the Stationary Points (Figure 1) for Three Different Forms of NMT, Computed at the B3LYP-D2/6-31G(d) Level in the Gas Phase and Water[a]

	HT pathway (singlet manifold)			SET pathway (singlet manifold)			SET pathway (triplet manifold)		
NMT type	^{1}RC	^{1}TS	^{1}PC[b]	^{3}RC	^{3}TS	^{3}PC	^{3}RC	^{3}TS	^{3}PC
				Gas Phase					
nonzwitterionic	0.0	41.3 (435*i*)	22.1	*c*	41.0 (784*i*)	*c*	54.0	60.7 (335*i*)	37.5
zwitterionic[d]	0.0								
anionic	0.0	31.0 (1102*i*)	19.1	*c*	*c*	*c*	25.1	36.3 (1090*i*)	36.2
				Water					
nonzwitterionic	0.0	40.7 (693*i*)	18.7	*c*	*c*	*c*	46.9	52.2 (141*i*)	39.0
zwitterionic	0.00	58.3 (771*i*)	23.9	*c*	54.0 (918*i*)	48.8	34.6	60.0 (1995*i*)	46.9
anionic	0.0	26.5 (1160*i*)	14.7	*c*	*c*	*c*	29.2	31.7 (924*i*)	29.2

[a]The complete list of results for the different density functionals and basis sets can be found in Table S4 in the Supporting Information. [b]An adduct with a Cs−N5 covalent bond is formed for the nonzwitterionic and zwitterionic forms of NMT, but not for the anionic form. [c]Yields the closed-shell configuration. [d]Yields the nonzwitterionic form.

both, whereas Tyr54 in wild-type MTOX is replaced with FAD-N3 in K341Q. The latter replacement is likely the main reason for the slightly increased separation between the two reactive centers. These findings suggest that K341 is not vital for the proper orientation of the COO^- end of AnNMT, as another neutral amino acid could provide the required steric effects as well.

Reductive Activity of Different NMT Forms. We performed QM-only (gas phase and water) and QM/MM (protein environment) calculations to locate the stationary points of the three proposed amine oxidation pathways (PN, HT, and SET; Figure 1). During these calculations, we separately considered the three forms (nonzwitterionic, zwitterionic, and anionic) of the NMT substrate. For sampling purposes, we used several snapshots from different NVT ensembles of each NMT form as the starting geometries for QM/MM calculations. As discussed earlier, a reactive binding of NMT (i.e., with low Cs−N5 distance) can be achieved only for specific combination(s) of the protonation states of K259 and K341 for each NMT form (see Table S1 in the Supporting Information). The combinations used in the corresponding NVT ensembles are documented in Tables S2 and S3 in the Supporting Information, along with some key features of these NVT ensembles. Despite serious attempts, we could not locate the adducts A and B (PN pathway, Figure 1) as stable minima, neither in QM-only nor in QM/MM calculations. This is likely due to van der Waals repulsions caused by the bulky methyl group flanking the Ns atom in NMT, which will then prevent the formation of a covalent Ns−C4a bond. Therefore, from now on, we will only consider the two feasible pathways (HT and SET).

The QM-only calculations were done on a model system, which contains the isoalloxazine moiety of FAD (taken as lumiflavin, including the methyl group bound to N10) and one of the three forms of the substrate NMT (with a total of 60−61 atoms depending on the NMT form considered; see Figure 4, top panel). Benchmark calculations using four density functionals (performed in the gas phase and water with Non-ZwNMT; Table S4 in the Supporting Information) indicate that the Gibbs relative free energies from different functionals may differ noticeably, but the trends in these energies do not depend on the choice of the functional. Likewise, the use of a larger basis, TZVP, has only a small effect on the predicted Gibbs relative free energies (0.1−5 kcal/mol, see Table S4: B3LYP and B3LYP-D2, Non-ZwNMT in the gas phase) and, more importantly, does not alter the qualitative picture. Comparison of the energetics from B3LYP and B3LYP-D2 (Table S4) demonstrates that accounting for dispersion corrections has a more significant effect on the predicted activation barriers (raised or lowered by 2−10 kcal/mol) for the HT and SET mechanisms of all NMT forms. This is not surprising, since the model system consists of two aromatic subunits with a high degree of π stacking that leads to strong dispersion interactions. Previous QM/MM studies have shown that the inclusion of dispersion effects can significantly improve the accuracy of calculated barriers for enzymatic reactions[53,54] and that B3LYP/MM calculations without dispersion corrections may yield qualitatively wrong reaction energetics.[55] Given this situation, we will focus on the B3LYP-D2/6-31G(d)-based results in the remainder of this article. The corresponding free energies are documented in Table 2 for all three NMT forms in the gas phase and in water.

It is evident from the activation barriers for both the HT and SET pathways in the gas phase and in water that the anionic form of NMT is more reactive than the two neutral forms (Table 2). In water, the preference for the anionic form is more clear-cut than in the gas phase, probably because of the enhanced stabilization of the CT-complex species by the polar environment. For analogous reasons, the activation barriers on the HT and SET pathways are further reduced for anionic NMT in water by about 5 kcal/mol. Interestingly, in both the singlet and triplet manifold, the SET pathway is more favored for the neutral NMT forms than for anionic NMT, which prefers the HT pathway. The associated barriers, however, are close to each other in energy, thus preventing a conclusive verdict. In addition, other density functionals (except M06-2X) also predict a TS species on the broken-symmetry singlet SET pathway corresponding to a singlet-coupled radical pair of nonzwitterionic NMT and FAD, in the gas phase and in water (Table S4 in the Supporting Information).

In the gas phase, the zwitterionic form of NMT is not stable and spontaneously converts back to the nonzwitterionic form, whereas the two charged centers are properly stabilized in water. However, the transfer of a hydride equivalent is significantly less facile in the zwitterionic than in the nonzwitterionic form. The reactant complex (^{1}**RC**) with the zwitterionic form is only about 0.5−2.3 kcal/mol less stable than that with the nonzwitterionic form; thus, the strong decrease of the amine oxidation rate in water can be ascribed to the high destabilization of the TS species. The low energy difference of the two forms in water also points to a high rate of the interconversion between them.

DOI: 10.1021/cs501694q
ACS Catal. 2015, 5, 1227−1239

Table 3. QM(B3LYP-D2/6-31G(d))/MM Relative Energies (Gibbs Relative Free Energies in Parentheses) for the Stationary Points (Figure 1) of the HT and SET Mechanisms for Different NMT Forms with Noncovalently or Covalently Bound FAD: Average Values and Standard Deviations (kcal/mol) over a Set of Five Snapshots[a]

	HT pathway (singlet manifold)			SET pathway (triplet manifold)		
NMT type	^{1}RC	^{1}TS	^{1}PC	^{3}RC	^{3}TS	^{3}PC
			Covalently Bound FAD			
nonzwitterionic	0.00	41.3 ± 2.54	18.5 ± 4.14	44.2 ± 0.77	55.8 ± 2.24	36.0 ± 1.15
	(0.00)	(38.8 ± 2.57)	(20.8 ± 4.24)	(41.7 ± 0.82)	(51.0 ± 2.15)	(34.0 ± 1.36)
zwitterionic	0.00	65.6 ± 2.11	37.7 ± 3.65	39.2 ± 4.66	75.5 ± 1.03	46.7 ± 3.12
	(0.00)	(63.3 ± 2.32)	(39.2 ± 3.84)	(37.5 ± 4.41)	(72.4 ± 2.15)	(43.4 ± 3.10)
anionic (wild type)	0.00	24.1 ± 2.46	2.78 ± 2.11	27.7 ± 4.56	37.9 ± 2.80	30.8 ± 4.57
	(0.00)	(21.3 ± 2.27)	(5.04 ± 1.89)	(26.0 ± 4.00)	(34.1 ± 2.82)	(29.3 ± 4.66)
anionic (K341Q)	0.00	30.7 ± 0.30	15.1 ± 2.98	26.0 ± 3.12	43.5 ± 1.99	39.2 ± 2.78
	(0.00)	(28.4 ± 0.33)	(16.0 ± 2.70)	(24.2 ± 3.10)	(40.0 ± 2.09)	(37.8 ± 2.84)
anionic (H263N)	0.00	36.4 ± 5.73	5.56 ± 8.84	21.4 ± 6.12	49.1 ± 6.43	34.1 ± 4.31
	(0.00)	(33.5 ± 5.51)	(5.42 ± 8.10)	(19.3 ± 6.41)	(45.1 ± 6.09)	(32.2 ± 3.67)
			Noncovalently Bound FAD			
nonzwitterionic	0.00	34.8 ± 2.41	6.06 ± 1.85	38.4 ± 2.47	46.3 ± 4.9	28.3 ± 6.98
	(0.00)	(32.9 ± 2.50)	(9.40 ± 1.81)	(35.9 ± 2.73)	(45.4 ± 3.6)	(26.5 ± 6.9)
anionic	0.00	25.9 ± 3.09	2.71 ± 1.28	25.7 ± 2.74	36.2 ± 2.95	26.7 ± 2.59
	(0.00)	(23.3 ± 3.32)	(5.56 ± 1.46)	(23.5 ± 2.58)	(32.0 ± 2.92)	(24.5 ± 2.72)

[a]Results for the K341Q and H263N mutants are also given.

The amine oxidation process is likely affected by the immediate protein surrounding. Acting as a bridge between MTOX and FAD, Cys308 is expected to have a strong influence. Therefore, in order to capture the potential effects of the covalent binding of FAD on the reaction, the QM region was expanded in QM/MM calculations by inclusion of Cys308 (71−72 QM atoms; see Figure 4, bottom-left inset). The resulting B3LYP-D2/MM relative energies and free energies for the HT and SET pathways are compiled in Table 3 for different NMT forms as well as for different binding modes (noncovalent or covalent) of the cofactor FAD. Table 3 also contains the QM/MM results for the K341Q and H263N mutants. Evidently, AnNMT is the most reactive form on both the HT and SET pathways in line with expectations, and we will thus focus on AnMNT in our further comparisons. All available energy values associated with AnNMT are visualized in Figure 5, and the corresponding QM-only and QM/MM geometries with AnNMT are shown in Figure S3 in the Supporting Information.

To check the effect of basis set extension on QM/MM energetics, we evaluated for a single snapshot (with AnNMT) the single-point QM(B3LYP-D2/TZVP)/MM relative energies at the QM(B3LYP-D2/6-31G(d))/MM geometries. Likewise, we also checked the effect of expanding the QM region by including all of the active-site residues that are in close contact (within 5 Å) with the substrate NMT: viz., L49, I50, R51, Y54, G55, E56, T239, Q247, Y248, Y249, H263, E316, H338, and K341 (for a total of 299 QM atoms, visualized in Figure S4 in the Supporting Information). The single-point QM(B3LYP-D2/6-31G(d))/MM relative energies were computed using the expanded QM region at the geometries available from the standard QM region. For a direct comparison, all the computed energies are compiled in Table S5 in the Supporting Information. Inspection of Table S5 shows that extension of the basis set slightly lowers the computed relative energies (minor changes up to 3 kcal/mol). The effect of expanding the QM region is somewhat more pronounced (changes of 1−5 kcal/mol, with a trend to alleviate the activation barriers for the HT and SET mechanisms), which is likely due to the improved treatment of nonbonded interactions. However, in both cases, the overall qualitative picture does not change. Therefore, and also considering the computational efforts involved, we will primarily discuss the QM/MM results obtained with the standard QM region and 6-31G(d) basis in the following (unless explicitly stated otherwise).

For the amine oxidation in the MTOX environment, the anionic form of NMT is preferred over the neutral forms (Table 3), as in the gas phase and in water (Table 2). Likewise, the zwitterionic form again yields the lowest oxidation rates, regardless of which pathway is considered (HT or SET). Different levels of reactivity of the different NMT forms may be attributed to the degree of compactness of the resulting FAD-NMT complex. As is evident from Table S2, AnNMT provides the most compact reactive core in the reactant complex, with a smaller separation between the reactive centers (R(N5−Cs) = 3.09 ± 0.13 Å). AnNMT also provides an enhanced nonbonded Ns−C4a interaction (R(Ns−C4a) = 3.37 ± 0.17 Å). On the other hand, ZwNMT leads to a looser binding of the two subunits, mainly due to H bonding of the N terminus of NMT to Gly337 (unlike the Non-ZwNMT and AnNMT forms). This keeps Ns rather far away from the N5−C4a locus of FAD and greatly weakens the Ns−C4a interaction (4.25 ± 0.31 Å). In contrast, the N5−Cs distance takes an intermediate value (3.51 ± 0.20 Å) between those for Non-ZwNMT and AnNMT (4.07 ± 0.26 and 3.37 ± 0.17 Å), whereas the alignment of C4a−N5 and Ns−Cs bonds differs significantly in ZwNMT and in the two other forms (θ(Ns−Cs−N5−C4a) = −3° vs 48°). We note that the alignment of the C4a−N5 and Ns−Cs bonds is highly important, as it affects the strength of the nonbonded C4a−Ns interaction.

Mechanistic Insights into the HT and SET Pathways. The HT mechanism involves the transfer of one proton and two electrons from NMT to FAD in a single concerted step (as hydride anion). We make use of a closed-shell description at the restricted Kohn−Sham (RKS) level for this process that converts the reactant complex ^{1}**RC** via the transition state ^{1}**TS** to the product complex ^{1}**PC**. The radical-SET mechanism, in contrast, proceeds through open-shell intermediates (Figure 1).

DOI: 10.1021/cs501694q
ACS Catal. 2015, 5, 1227−1239

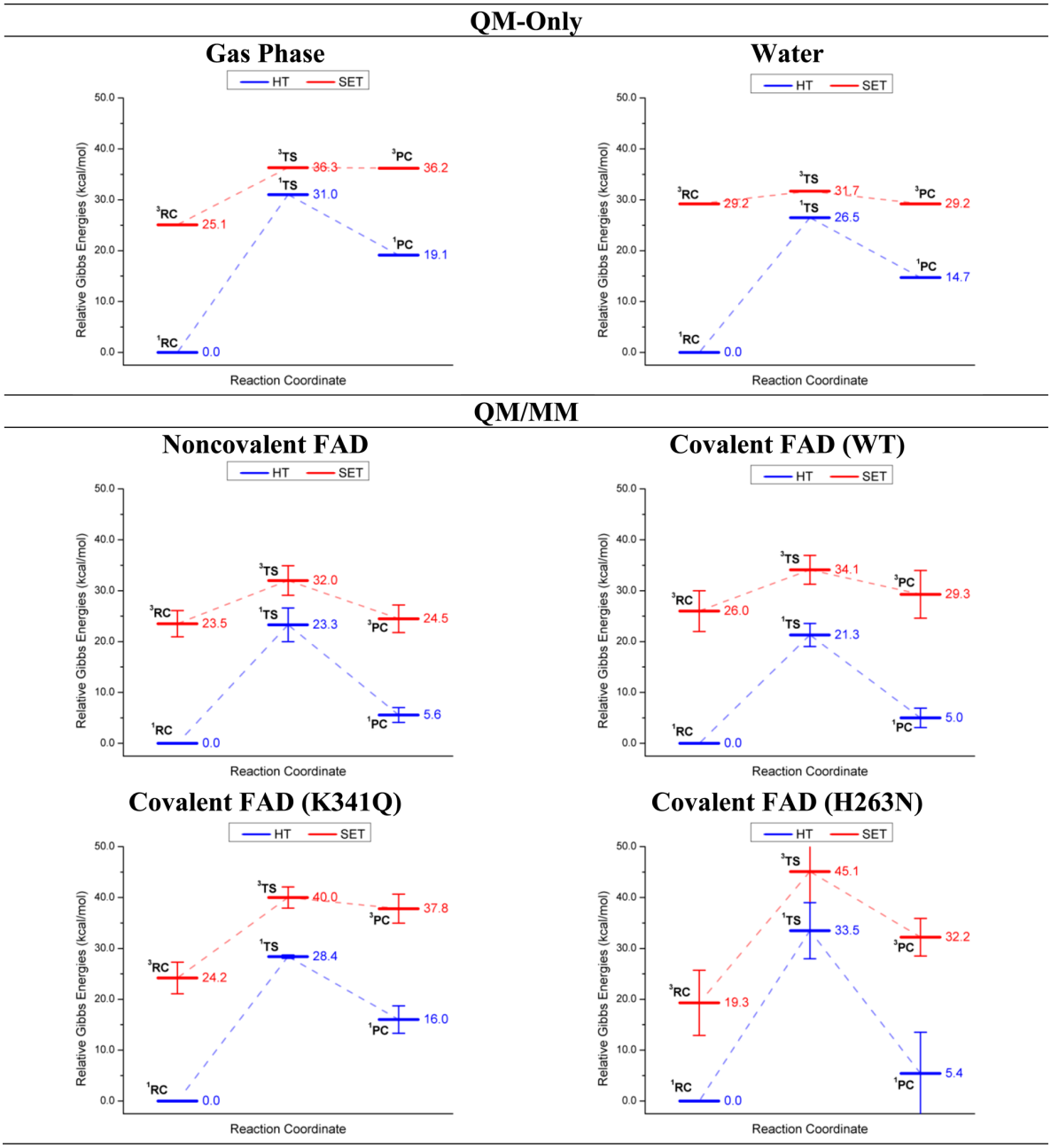


Figure 5. QM(B3LYP-D2/6-31G(d))-only and QM/MM Gibbs free energy profiles (in kcal/mol) of AnNMT for HT and SET pathways in different media and in MTOX with different model systems (see text for definitions). See Figure 1 for the definition of the stationary points.

The single-electron transfer from the α-C−H bond of NMT to FAD generates two radical intermediates (^{3}**RC**): namely the flavin semiquinone anion (**FAD**$^{\bullet-}$) and amine radical cation (**NMT**$^{\bullet+}$). The subsequent rate-limiting step is the homolytic cleavage of the α-C−H bond, which leads to the iminium cation (^{1}**PC**). The latter may undergo different reactions: (a) proton-coupled electron transfer (PCET) or direct hydrogen atom transfer and (b) stepwise transfer of a proton and an electron (second SET) from the amine to FAD. In the latter case, the initial proton transfer yields an intermediate pair of ammonium and FADH$^{\bullet-}$ radicals (^{3}**PC**), and the second SET step can occur directly or may also be mediated by adduct formation in the presence of an active-site radical.[16,18] The open-shell species on the SET pathways were described by the unrestricted Kohn−Sham (UKS) method. As the two unpaired electrons on the two separate subunits of the model system can have the same or different spin, we considered both the singlet and triplet manifolds. In contrast to the QM-only case (gas phase and water; Table 2), the broken-symmetry UKS calculations always converged to the closed-shell solution at the QM/MM level for any of the NMT substrate forms (as is evident from the computed energies, $\langle S^2 \rangle$ values, and spin densities; data not shown). The same was found when using restricted open-shell Kohn−Sham (ROKS) calculations for the QM region. Therefore, only triplet SET values are included in Table 3 and Figure 5.

DOI: 10.1021/cs501694q
ACS Catal. 2015, 5, 1227−1239

Hybrid QM/MM calculations on MTOX with a covalently bound FAD and the most reactive NMT form (AnNMT; see Figure 5) predict a free energy barrier of 21.3 ± 2.3 kcal/mol for the HT mechanism (approximate rate 0.087 min^{-1}), whereas the QM-only calculations yield barriers of 31.0 kcal/mol in vacuum and 26.5 kcal/mol in water. The protein environment thus increases the HT rate dramatically in comparison to water (by 4 orders of magnitude). In contrast, the MTOX environment increases the corresponding activation barrier for the SET mechanism in comparison to the polar solvent (34.1 ± 2.8 vs 31.7 kcal/mol), thus lowering the SET rate by 2 orders of magnitude. The computed SET rate in MTOX is approximately 3.5×10^{-11} min^{-1}, rendering the HT pathway clearly favorable at physiological temperatures.

To probe the possible reasons why the most reactive AnNMT form prefers the HT pathway in all media, we compiled key structural features in the QM-only and QM/MM geometries (averages over five snapshots) of the stationary points for the two pathways in Tables S6 and S7 in the Supporting Information. For the HT pathway, we note a gradual increase in the compactness of the reacting subunits in the reactant complexes (^{1}**RC**), as is evident from the Cs−N5 distances (gas phase, 3.82 Å; water, 3.33 Å; MTOX, 3.01 Å). Thus, the increase in the HT rates on going from the gas phase to the protein environment can be linked to the enhanced interaction of the reactive sites (N5 and Cs). This is reminiscent of the increase in the computed HT rates when going from neutral to anionic NMT that can also be attributed to the tighter binding in the latter case (see Reductive Activity of Different NMT Forms).

Likewise, the environment mainly affects the separation of two reacting centers in the product complex (^{1}**PC**). The ^{1}**PC** species is a weakly interacting complex in the gas phase and in water (R(Cs−N5) = 2.08 and 2.06 Å, respectively) but forms an adduct in the protein environment (1.69 Å). This has evident consequences for the energetics (Figure 5): the ^{1}**PC** adduct is more stable than the weakly interacting ^{1}**PC** species (relative free energies: 5.0 ± 1.9 vs 14.7 kcal/mol), even though the reaction remains endergonic. The ^{1}**PC** adduct contains a covalent bond between the α-carbon atom of the amine substrate and the N5 atom of FAD, as also found for other amine oxidases.[32,33]

In the mechanism of amine demethylation,[12] the initial amine oxidation is followed by the hydrolysis of the formed iminium cation. To prepare for this second step, the covalent bond in the FAD−imine adduct (^{1}**PC**) needs to be broken to yield a weakly interacting complex with a separate iminium cation. In the case of MAO-B, a two-step mechanism has been proposed for the dissociation of the initial adduct,[32] which involves the transfer of a proton from the flanking nitrogen in the (protonated) substrate (here called Ns) via two active-site water molecules located on the *Re* face of FAD. This proton transfer was suggested to facilitate the breaking of the covalent N5−Cs bond, which generates the final product complex between the reduced flavin FADH anion and the iminium cation. In the case of MTOX, active-site water molecules are not present in any NVT ensemble with any NMT form (see Table S2 in the Supporting Information), which makes the proposed two-step mechanism questionable. However, instead of water, the Gly337, Phe340, and K341 residues (all in H-bonding distance to N1 and O2) might serve as potential mediators for the proton transfer to the N1 atom of FAD.

To investigate the feasibility of a hydride equivalent transfer in the case of FAD protonated at N1, we generated the corresponding starting structures from either one of two neutral NMT forms (ZwNMT or Non-ZwNMT) by manually shifting the proton from NMT-Ns or NMT-COOH, respectively, and then computing reaction profiles for the HT pathway. In the case of ZwNMT, the resulting energetics at optimized geometries (Table S8) show that the overall barrier is indeed reduced by ca. 8 kcal/mol after N1 protonation, but it is still very high (55.7 kcal/mol). In addition, the intermediate ^{1}**RC** species with a FAD protonated at N1 is much less stable (by 45.7 kcal/mol) in comparison to that with unprotonated N1. In the case of Non-ZwNMT, this ^{1}**RC** species (with protonated N1) is destabilized less (by 18.0 kcal/mol, relative to the ^{1}**RC** species with unprotonated N1); we have not computed the barrier to its formation by proton transfer to N1. Given the overall low energy required for the favored HT pathway with AnNMT and unprotonated FAD-N1 (22.1 kcal/mol), a mechanism involving N1 protonation prior to HT seems quite improbable. Furthermore, N1 protonation after HT would require a neutral form of the substrate to react, which can be excluded because the anionic form is much more reactive (see Reductive Activity of Different NMT Forms). On the basis of these findings, we propose that the formed ^{1}**PC** adduct will react spontaneously: (a) by reprotonation of the backbone of the imine product to convert it back to neutral NMT, (b) by protonation of the N1 atom of FAD after the hydride transfer, or (c) as part of a larger collective protein movement, such as the release of the substrate out of the active site (on a longer time scale).

In contrast to the situation in the ^{1}**RC** and ^{1}**PC** species, the MTOX environment does not have a significant effect on the distance between the reactive centers in ^{1}**TS** (R(Cs−N5) = 2.55−2.58 Å). However, it notably affects the hydride transfer angle (θ(Cs−H1−N5)) and the alignment of the C4a−N5 and Cs−Ns bonds (θ(Ns−Cs−N5−C4a)) (see Tables S6 and S7 in the Supporting Information and Figure 4). The active-site residues in MTOX enforce a more linear path for the hydride transfer in comparison to the gas phase and water (θ(Cs−H1−N5) = 169° vs 154−155°), We note in this connection that MTOX also employs an HT arrangement more linear than that of LSD1[33] and MAO-B[34] (θ(Cs−H1−N5) = 150−155°).

On the radical-SET pathway, the first SET from NMT to FAD creates a radical pair (^{3}**RC**), which has already been shown to be energetically demanding in LSD1.[33] Although we did not compute the energy required for intersystem crossing ($S_0 \rightarrow T_1$, conversion of ^{1}**RC** to ^{3}**RC**), it is expected to be lower than the barrier of the rate-limiting homolytic cleavage of the α-C−H bond (i.e., ^{3}**TS** of the SET pathway).[13,26] The resulting radical pair is computed to be of high energy in the gas phase and water (25.1 and 29.2 kcal/mol, respectively). There is no significant stabilization by the MTOX environment (26.0 ± 4.0 kcal/mol), so that the radical pair remains an unfavorable intermediate prior to the homolytic cleavage of the α-C−H bond on the SET pathway. Given the rather low barrier for the subsequent proton transfer in MTOX (^{3}**RC** → ^{3}**TS**: 8.1 kcal/mol), the radical pair can be converted rather easily to radical products (^{3}**PC**) that are also unstable (29.3 ± 4.7 kcal/mol). However, this low proton transfer barrier is kinetically irrelevant, as the SET process is inaccessible due to the high overall barrier (^{3}**TS**: 34.1 ± 2.8 kcal/mol).

To check whether the active-site residues in MTOX may contribute to the stabilization of the radical pair (^{3}**RC**), we

DOI: 10.1021/cs501694q
ACS Catal. 2015, 5, 1227−1239

computed the QM/MM energies and spin densities with an extended QM region (vide supra). It is evident from the spin densities (Figure S5 in the Supporting Information) that a radical pair of FAD and NMT is formed as expected for the SET pathway, since none of the neighboring MTOX residues take up part of the unpaired electron density (neither in ^{3}**RC** nor in the other stationary points on the SET pathway). The QM/MM results with the extended QM region lower the relative energy of ^{3}**RC** only slightly (by ca. 4.0 kcal/mol; Table S5 in the Supporting Information), presumably due to the QM treatment of active-site H-bonding interactions.

Judging from the key structural parameters (Tables S6 and S7 in the Supporting Information), the protein environment makes the reactive core in the stationary points of the SET pathway more compact (as in the case of the HT pathway). On the other hand, a product adduct is not formed on the SET pathway (^{3}**PC**) in any medium, as opposed to the HT pathway (^{1}**PC**; see the Cs−N5 distances in Tables S6 and S7). Attempts to enforce a ^{3}**PC** adduct ended up with extremely unstable species (ca. 80 kcal/mol), which is as expected, since two subunits with unpaired electrons of the same spin will tend to stay apart rather than form an adduct.

Catalytic Role of His263 and Lys341. To further augment our understanding of the role of the two most likely active-site base candidates, we extended the QM/MM calculations for wild-type MTOX by including either H263 or K341 in the QM region (data not shown), and we also investigated the H263N and K341Q mutants (Table 3 and Figure 5). Including H263 or K341 in the QM part leads to only minor reductions of the HT and SET barriers (by 1 and 2−3 kcal/mol, respectively), and Mulliken population analysis shows almost no changes in the charge distribution in the reactive core (Cs, N5, Ns, and C4a) of ^{1}**RC**, ^{1}**TS**, and ^{1}**PC** (data not shown). However, inclusion of K341 in the QM part does affect the charges of the oxygen atoms at the C terminus, reflecting the electrostatic influence of K341. Furthermore, on the SET pathway, no spin density is found on these residues in ^{3}**RC**, ^{3}**TS**, and ^{3}**PC**. Taken together, all of these findings imply that His263 and Lys341 do not have a catalytic role in MTOX.

The computed QM/MM energies (Figure 5) indicate a substantial decrease in the HT and SET rates in both the H263N and K341Q mutants, in comparison to wild-type MTOX. As discussed earlier, the H263N mutation does not support a stable AnNMT binding, since R(N5−Cs) remains generally larger than 5 Å in MD simulations (Figure S2 in the Supporting Information). In contrast, the K341N mutation weakens the binding only slightly, with R(N5−Cs) increasing by 0.10−0.15 Å (Table S2 in the Supporting Information). Even though we deliberately chose snapshots for QM/MM from the H263N MD trajectory with the closest possible contact between two reactive sites, the resulting HT and SET barriers are much higher in H263N than in K341Q. In this regard, QM/MM calculations suggest that His263 is more important for substrate alignment in MTOX than in MSOX, in which the mutant of the H263 homologue (H269N) has only a 2-fold slower turnover rate.[10] This may be related to the different nature of the substrates for MTOX (*N*-methyltryptophan) and MSOX (sarcosine). In particular, sarcosine lacks the aromatic indole side chain that could provide π-stacking and π−cation interactions with His263 and Arg51, respectively. In contrast, the current QM/MM results indicate only a partial reduction of the HT rate in K341Q in comparison to wild-type MTOX. This is likely due to electrostatic (stabilization of COO^-) effects of K341 rather than to steric effects (substrate binding; see MD results in Dynamics of Substrate Binding). The current QM/MM results do not rule out an active-site base role for K341, but they also do not directly support this possibility.

Effects of FAD Covalent Binding on Substrate Binding and NMT Oxidation. It has been proposed that the covalent binding of the flavin cofactor in MSOX and similar oxidases enhances protein stability (by avoiding the loss of a weakly bound oxidized cofactor) and improves enzymatic activity (by increasing the flavin redox potential and aiding substrate binding).[4,5] To probe the molecular basis for the catalytic role of FAD covalent binding, we created an artificial model system, with FAD noncovalently bound to MTOX, and compared it to wild-type MTOX. Specifically, we capped the Cys308-S and FAD-8α-methyl ends with hydrogen atoms, while the rest of the system was kept intact (73−74 atoms in the QM part; see Figure 4 for visualization). We repeated our standard computational procedure using this model system and the three forms of the NMT substrate. We present key features of the resulting NVT trajectories in Table S3 in the Supporting Information, the computed QM/MM energies in Table 3 and Figure 5, and key structural properties in Table S7 in the Supporting Information.

We first note that there is no stable binding of AnNMT in any of the NVT trajectories with noncovalently bound FAD, whereas ZwNMT and Non-ZwNMT bind properly and even more compactly in comparison to their counterparts in the covalently bound FAD case (see R(Ns−C4a) and R(N5−Cs) values in Tables S2 and S3 in the Supporting Information). For further analysis, we inspected how the planarity of FAD is affected by its binding mode. To this end, we monitored six dihedral angles within the isoalloxazine moiety of FAD (θ(C9−C8−C7−C6), θ(C9−C9a−C5a−C6), θ(N10−C9a−C5a−N5), θ(N10−C10a−C4a−N5), θ(N1−C10a−C4a−C4), θ(N1−C2−N3−C4)) throughout the NVT trajectories of different NMT forms (Table S9 in the Supporting Information). Judging from the sum of the average absolute deviations from the reference values at planarity, the isoalloxazine moiety is closer to planarity in the case of noncovalent FAD binding (42−43° vs 61−68° for covalent FAD binding). It is (almost) perfectly planar in FAD without the protein environment (data not shown). These findings are in accord with previous arguments that the nonplanarity of FAD is the main factor in modulating the reduction potential of MSOX and other oxidases.[24] However, the current MD simulations do not provide evidence that MTOX enforces FAD to assume a "butterfly bent" structure, as proposed for TMADH.[56] This may be due to the differences in the binding sites of FAD in the two oxidases, which impose different constraints on the isoalloxazine moiety.

Since AnNMT does not bind properly in our artificial model system with noncovalently bound FAD, we assumed in our QM/MM calculations that the substrate binds in the Non-ZwNMT form, which is computed to afford the tightest binding (Table S3 in the Supporting Information) and that, being in close contact with unprotonated K341, NMT then donates its extra proton to the C terminus right before the transfer of the hydride equivalent. The QM/MM energies (Table 3 and Figure 5) indicate that the barrier for the HT pathway is slightly higher (by 2.0 kcal/mol) with noncovalently (rather than covalently) bound FAD. This increase is of the same magnitude as the previous estimate for the increase in reduction potential of

DOI: 10.1021/cs501694q
ACS Catal. 2015, 5, 1227−1239

MSOX due to covalent FAD binding (120 mV ≈ 2.8 kcal/mol, assuming a one-electron transfer, on the basis of binding affinities of Cys315 mutants).[4] In line with this minor change in the energetics, the key structural features in the optimized stationary points of the HT pathway show only slight variations (see, for instance, R(Cs−N5), R(Ns−C4a), θ(Cs−H1−N5), and θ(Ns−Cs−N5−C4a) in Table S7 in the Supporting Information).

CONCLUSIONS

Using results from theoretical pK_a analysis, classical MD simulations, and QM-only and QM/MM calculations, we present a comprehensive molecular-level analysis of the catalytic mechanism of the NMT oxidation catalyzed by MTOX, with special attention to the role of several active-site residues and covalent FAD binding. Our QM-only (gas phase and water) and QM/MM (protein environment) calculations predict that the anionic form (AnNMT) is far more reactive than the neutral forms of NMT, as expected in view of the optimum working conditions of MTOX (pH 8.0).[1] The QM-only and QM/MM results with AnNMT clearly favor the HT pathway over the radical/SET and adduct forming mechanisms, in accord with the previous computational studies on other flavoprotein oxidases.[26,32−34] In comparison to the protein-free environment (gas phase and water), MTOX expedites the HT rates markedly by lowering the activation barrier (21.3 ± 2.3 kcal/mol, corresponding to a rate of 0.087 min^{-1}), in reasonable agreement with the experimentally observed rates.[1,2] According to our computational results, the SET pathway requires the initial formation of a rather unstable radical pair (^{3}**RC**) and encounters an even higher activation barrier thereafter. The MTOX environment stabilizes neither the radical pair nor the transition state so that the SET pathway remains inaccessible in the enzyme. The adduct-forming (polar-nucleophilic) mechanism is ruled out on the grounds that the adducts formulated for this route (with an Ns−C4a bond; see Figure 1) are not stable intermediates and readily dissociate during QM-only and QM/MM optimizations.

On the HT pathway, the product species (^{1}**PC**) is computed to be an adduct of NMT-imine and FAD with a covalent Cs−N5 bond, congruent with the previous findings for other amine oxidases (MAO-B[32] and LSD1[33]). To break this adduct complex, a two-step HT mechanism was proposed in the case of MAO-B,[32] the second step of which involves protonation of the FAD at N1, with the proton coming from the N terminus of the amine substrate. Our MD simulations and QM/MM calculations do not support this mechanism in MTOX, for two reasons: (1) the mechanism requires that a neutral NMT species (ZwNMT or Non-ZwNMT) rather than AnNMT is oxidized, which is not feasible (computed barriers of 39 and 63 kcal/mol), and (2) the two water molecules needed to connect NMT-N and FAD-N1 to mediate the proton transfer are missing in MTOX. To check for an alternative variant of this mechanism, we presumed that Gly337 mediates the proton transfer, given its close proximity to the backbone of NMT and to FAD-N1 in the NVT ensembles. Hence, we checked the influence of an initial N1 protonation on NMT oxidation and found that, under these circumstances, oxidation may occur only for Non-ZwNMT, although it is much less facile than for the standard HT pathway with AnNMT. In view of these findings, we propose that the breakup of the adduct complex requires either reprotonation of the NMT backbone by the active-site base or protonation of FAD at N1.

Compliant with experimental studies on pH-dependent MTOX activity,[6,7] our computational findings indicate binding of neutral NMT in the active site, which is then ionized by an active-site base to give the reactive anionic form. Among the potential active-site bases proposed for MTOX (or MSOX that has high structural similarity[3]), we eliminate Tyr249 because of its location (MD simulations) and its protonation state (theoretical pK_a analysis), which is consistent with the only partial loss of activity upon mutation of its homologue in MSOX (Y317F).[9] MD simulations and QM/MM results suggest a steric role for His263 in orienting the substrate into the reactive position via π-stacking interactions; His263 appears to play a more eminent role in MTOX than in MSOX, considering the only 2-fold lower rates upon mutation of its homologue in MSOX,[10] which may be linked to the fact that MSOX genuinely works on a smaller aliphatic substrate. Our computational results indicate that K341 acts electrostatically and as an active-site base, which is consistent with the almost complete loss of MTOX activity (95% or 250-fold) upon K341Q mutation that is apparently not due to the 10% loss in FAD incorporation.[2]

MTOX employs FAD as the oxidizing agent, which is covalently bound to Cys308 via its C8 methyl group. Using an artificial theoretical model, with FAD noncovalently bound to MTOX, we scrutinized the role of the FAD binding mode. Our MD simulations show that a stable binding of anionic NMT is not likely in the case of noncovalent FAD binding. This is probably related to the more planar structures of the isoalloxazine ring that are seen in the MD simulations with noncovalent FAD. Flavin nonplanarity was proposed to be the reason for enhanced redox potentials of amine oxidases with a covalently bound flavin.[24,56] Our QM/MM results indicate a minor increase in the activation barriers for the HT pathway (by 2.0 kcal/mol) in the case of noncovalent binding.

ASSOCIATED CONTENT

Supporting Information

The following file is available free of charge on the ACS Publications website at DOI: 10.1021/cs501694q.

Computational details, characteristic features from NVT ensembles, complete set of QM-only results (gas phase and water) of three NMT forms with four density functionals, effect of extending basis sets and QM part on QM/MM energies, key structural properties from QM-only and QM/MM geometries, effect of protonating the FAD-N1 position on the HT and SET pathways, average deviation from planarity of the isoalloxazine ring in noncovalently and covalently bound FAD, proton relay system via K341, Cs−N5 distances in selected MD trajectories, and visualization of QM-only and QM/MM geometries, of the extended QM region, and of spin densities (PDF)

AUTHOR INFORMATION

Corresponding Author

*E-mail for W.T.: thiel@kofo.mpg.de.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

B.K. thanks Dr. Jan Götze, Dr. Mahendra Patil, and Dr. Mario Barbatti for valuable discussions in the course of this study.

DOI: 10.1021/cs501694q
ACS Catal. 2015, 5, 1227−1239

REFERENCES

(1) Khanna, P.; Schuman Jorns, M. *Biochemistry* **2001**, *40*, 1451−1459.
(2) Ilari, A.; Bonamore, A.; Franceschini, S.; Fiorillo, A.; Boffi, A.; Colotti, G. *Proteins* **2008**, *71*, 2065−2075.
(3) Koyama, Y.; Ohmori, H. *Gene* **1996**, *181*, 179−183.
(4) Hassan-Abdallah, A.; Zhao, G.; Jorns, M. S. *Biochemistry* **2006**, *45*, 9454−9462.
(5) Heuts, D. P. H. M.; Scrutton, N. S.; McIntire, W. S.; Fraaije, M. W. *FEBS J.* **2009**, *276*, 3405−3427.
(6) Bruckner, R. C.; Winans, J.; Jorns, M. S. *Biochemistry* **2011**, *50*, 4949−4962.
(7) Ralph, E. C.; Fitzpatrick, P. F. *Biochemistry* **2005**, *44*, 3074−3081.
(8) Wagner, M. A.; Trickey, P.; Chen, Z. W.; Mathews, F. S.; Jorns, M. S. *Biochemistry* **2000**, *39*, 8813−8824.
(9) Zhao, G.; Jorns, M. S. *Biochemistry* **2005**, *44*, 16866−16874.
(10) Zhao, G.; Song, H.; Chen, Z.-W.; Mathews, F. S.; Jorns, M. S. *Biochemistry* **2002**, *41*, 9751−9764.
(11) Zhao, G.; Jorns, M. *Biochemistry* **2006**, *45*, 5985−5992.
(12) Shi, Y.-J.; Lan, F.; Matson, C.; Mulligan, P.; Whetstine, J.; Cole, P.; Casero, R. A.; Shi, Y. *Cell* **2004**, *119*, 941−953.
(13) Gaweska, H.; Henderson Pozzi, M.; Schmidt, D. M. Z.; McCafferty, D. G.; Fitzpatrick, P. F. *Biochemistry* **2009**, *48*, 5440−5445.
(14) Forneris, F.; Binda, C.; Aglio, A. D.; Fraaije, M. W.; Battaglioli, E.; Mattevi, A.; Dall'Aglio, A. *J. Biol. Chem.* **2006**, *281*, 35289−35295.
(15) Khanna, P.; Jorns, M. *Biochemistry* **2001**, *40*, 1441−1450.
(16) Silverman, R. B. *Acc. Chem. Res.* **1995**, *28*, 335−342.
(17) Silverman, R. B. R.; Hoffman, S. S. J.; Catus, W. B. W. *J. Am. Chem. Soc.* **1980**, *102*, 7126−7128.
(18) Silverman, R. B. *Prog. Brain Res.* **1995**, *106*, 23−31.
(19) Miller, J. R.; Edmondson, D. E. *Biochemistry* **1999**, *38*, 13670−13683.
(20) Forneris, F.; Battaglioli, E.; Mattevi, A.; Binda, C. *FEBS J.* **2009**, *276*, 4304−4312.
(21) Culhane, J. C.; Cole, P. A. *Curr. Opin. Chem. Biol.* **2007**, *11*, 561−568.
(22) Fraaije, M. W.; Mattevi, A. *Trends Biochem. Sci.* **2000**, *25*, 126−132.
(23) Gaweska, H.; Fitzpatrick, P. F. *BioMol. Concepts* **2011**, *2*, 365−377.
(24) Scrutton, N. S. *Nat. Prod. Rep.* **2004**, *21*, 722−730.
(25) Fitzpatrick, P. F. *Arch. Biochem. Biophys.* **2010**, *493*, 13−25.
(26) Ralph, E. C.; Hirschi, J. S.; Anderson, M. A.; Cleland, W. W.; Singleton, D. A.; Fitzpatrick, P. F. *Biochemistry* **2007**, *46*, 7655−7664.
(27) Schmidt, D. M. Z.; McCafferty, D. G. *Biochemistry* **2007**, *46*, 4408−4416.
(28) Mccann, A. E.; Sampson, N. S. *J. Am. Chem. Soc.* **2000**, *122*, 35−39.
(29) Chen, Z. Z.; Zhao, G.; Martinovic, S.; Jorns, M. S.; Mathews, F. S. *Biochemistry* **2005**, *44*, 15444−15450.
(30) Wagner, M. A.; Jorns, M. S. *Biochemistry* **2000**, *39*, 8825−8829.
(31) Forneris, F.; Binda, C.; Vanoni, M. A.; Mattevi, A.; Battaglioli, E. *FEBS Lett.* **2005**, *579*, 2203−2207.
(32) Vianello, R.; Repič, M.; Mavri, J. *Eur. J. Org. Chem.* **2012**, *2012*, 7057−7065.
(33) Karasulu, B.; Patil, M.; Thiel, W. *J. Am. Chem. Soc.* **2013**, *135*, 13400−13413.
(34) Abad, E.; Zenn, R.; Kästner, J. *J. Phys. Chem. B* **2013**, *117*, 14238−14246.
(35) Trott, O.; Olson, A. J. *J. Comput. Chem.* **2010**, *31*, 455−461.
(36) Zhao, G.; Bruckner, R. C.; Jorns, M. S. *Biochemistry* **2008**, *47*, 9124−9135.
(37) Brooks, B.; Brooks, C. *J. ...* **2009**, *30*, 1545−1615.
(38) Luo, G.; Andricioaei, I.; Xie, X. S.; Karplus, M. *J. Phys. Chem. B* **2006**, *110*, 9363−9367.
(39) MacKerell, A. D.; Bashford, D.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiórkiewicz-Kuczera, J.; Yin, D.; Karplus, M. *J. Phys. Chem. B* **1998**, *102*, 3586−3616.
(40) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926−935.
(41) Olsson, M.; Sondergaard, C.; Rostkowski, M.; Jensen, J. *J. Chem. Theory Comput.* **2011**, *7*, 525−537.
(42) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09, Revision D.01*, Gaussian Inc., Wallingford, CT, 2013.
(43) Becke, A. *J. Chem. Phys.* **1993**, *98*, 5648−5652.
(44) Grimme, S. *J. Comput. Chem.* **2006**, *27*, 1787−1799.
(45) Zhao, Y.; Truhlar, D. G. *Theor. Chem. Acc.* **2007**, *120*, 215−241.
(46) Vydrov, O. A.; Scuseria, G. E. *J. Chem. Phys.* **2006**, *125*, 234109.
(47) Chai, J.-D.; Head-Gordon, M. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6615−6620.
(48) Cossi, M.; Rega, N.; Scalmani, G.; Barone, V. *J. Comput. Chem.* **2003**, *24*, 669−681.
(49) Sherwood, P.; de Vries, A. H.; Guest, M. F.; Schreckenbach, G.; Catlow, C. R. A.; French, S. A.; Sokol, A. A.; Bromley, S. T.; Thiel, W.; Turner, A. J.; Billeter, S.; Terstegen, F.; Thiel, S.; Kendrick, J.; Rogers, S. C.; Casci, J.; Watson, M.; King, F.; Karlsen, E.; Sjøvoll, M.; Fahmi, A.; Schaefer, A.; Lennartz, C. *J. Mol. Struct. (THEOCHEM)* **2003**, *632*, 1−28.
(50) Smith, W.; Forester, T. R. R. *J. Mol. Graph.* **1996**, *14*, 136−141.
(51) De Vries, A. H.; Sherwood, P.; Collins, S. J.; Rigby, A. M.; Rigutto, M.; Kramer, G. J. *J. Phys. Chem. B* **1999**, *103*, 6133−6141.
(52) Billeter, S. R. S.; Turner, A. A. J.; Thiel, W. *Phys. Chem. Chem. Phys.* **2000**, *2*, 2177−2186.
(53) Van der Kamp, M. W.; Mulholland, A. J. *Biochemistry* **2013**, *52*, 2708−2728.
(54) Lonsdale, R.; Harvey, J. N.; Mulholland, A. J. *J. Phys. Chem. Lett.* **2010**, *1*, 3232−3237.
(55) Lawan, N.; Ranaghan, K. E.; Manby, F. R.; Mulholland, A. J. *Chem. Phys. Lett.* **2014**, *608*, 380−385.
(56) Trickey, P.; Basran, J.; Lian, L.; Chen, Z.; Barton, J. D.; Sutcliffe, M. J.; Scrutton, N. S.; Mathews, F. S. *Biochemistry* **2000**, *39*, 7678−7688.

DOI: 10.1021/cs501694q
ACS Catal. 2015, 5, 1227−1239

APPENDIX C

Photoinduced Intra-molecular Charge Transfer in an Electronically Modified Flavin Derivative: Roseoflavin

Bora Karasulu, and Walter Thiel

J. Phys. Chem. B, **2015,** 119, pp 928-943.

THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

pubs.acs.org/JPCB

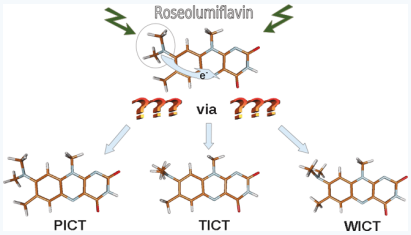
Photoinduced Intramolecular Charge Transfer in an Electronically Modified Flavin Derivative: Roseoflavin

Bora Karasulu and Walter Thiel*

Max-Planck-Institut für Kohlenforschung, Kaiser-Wilhelm-Platz 1, 45470, Mülheim, Germany

Supporting Information

ABSTRACT: The photophysical properties of a push−pull flavin derivative, roseoflavin (RoF), are investigated in different surroundings at the molecular level, with focus on intramolecular charge transfer (ICT). Time-dependent density functional theory (TD-DFT, CAM-B3LYP functional) and DFT-based multi-reference configuration interaction (DFT/MRCI) are used to compute excited-state energies and one-electron properties of a truncated RoF model, roseolumiflavin (RoLF). Solvent effects are taken into account implicitly by the conductor-like polarizable continuum model and explicitly through a microsolvation scheme. In the gas phase, the calculations predict no crossing between the lowest locally excited (LE) and charge-transfer (CT) states upon twisting the dimethylamine donor group relative to the plane of the isoalloxazine acceptor moiety, whereas this crossing is found to be facile in solution (i.e., in water or benzene). Crossing of the LE and CT states facilitates ICT, which is the main cause of the fluorescence quenching and dual fluorescence character experimentally observed for roseoflavin in solution. The barrier for the ICT process is computed to be lower in water than in benzene, consistent with the enhanced ICT rates observed in more polar solvents. We present a detailed study of the molecular mechanism of the photoinduced ICT process in RoLF. For a typical donor−acceptor chromophore, three such mechanisms are discussed in the literature, which differ in the alignment of the donor and acceptor planes, namely, planar ICT (PICT), perpendicular-twisted ICT (TICT), and wagging ICT (WICT). Our theoretical results suggest that the TICT mechanism is favored in RoLF.



1. INTRODUCTION

Flavin derivatives have been used as active-site probes for various types of enzymes, since they display different spectral, chemical, and mechanistic properties compared with natural flavins.[1,2] In particular, flavin derivatives with red-shifted absorption were studied to check whether the cellular processes induced by flavin photodynamics can be carried out at lower energies.[3] Hence, the photophysical behavior of a myriad of flavin derivatives was thoroughly investigated in previous theoretical and experimental work (see, for instance, refs 1 and 3−7 and references therein). The current calculations focus on the photophysical properties of an electronically modified riboflavin anologue, roseoflavin (8-dimethylamino-8-demethyl-D-riboflavin, RoF, see Figure 1). Roseoflavin is a blue-light receptor. It differs from riboflavin (RF) in that the methyl group at the C8 position of the isoalloxazine ring is replaced by a dimethylamino (DMA) group, with the side chain (R in Figure 1) remaining intact. This structural modification affects various photophysical and biochemical properties. RoF is an antagonist of RF (also known as vitamin B2) and acts as an antibiotic against Gram-positive bacteria.[8,9] Contrary to RF, RoF shows interesting photodynamics, as it undergoes structural changes upon photoexcitation that mainly involve the torsion of its DMA group.[10] The optical properties of riboflavin and roseoflavin in various solvents and protein environments were the subject of a number of theoretical[4,11−21] and experimental studies.[2,10,12,20,22−31]

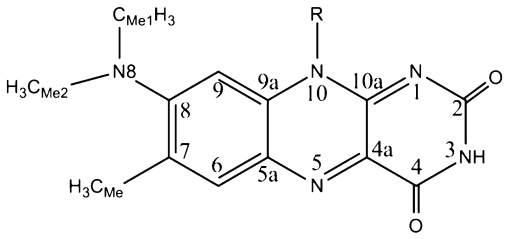


Figure 1. Chemical structure of roseoflavin (RoF). RoF differs from wild-type riboflavin by the replacement of the methyl group bound to C8 with a DMA group. R is the ribityl chain in RoF. It may be different in roseoflavin derivatives; e.g., for roseolumiflavin (RoLF), R = methyl, for RoFMN, R = ribityl-5′-phosphate, and for RoFAD, R = ribityl-(9-adenosyl)-pyrophosphate.

Previous experimental studies have established that both the structural and photophysical properties of RoF differ significantly from those of natural RF.[2,10,12,23,28] The first band in the adsorption spectrum of RoF in water is red-shifted ($\lambda_{max,RoF}$ = 500 nm, $\lambda_{max,RF}$ = 445 nm)[12,26] and more intense compared to RF, as evident from the reported oscillator strengths (f = 0.475 vs f = 0.197).[10] The second absorption

Special Issue: William L. Jorgensen Festschrift

Received: June 19, 2014
Revised: September 1, 2014
Published: September 12, 2014

band of RoF is weak and blue-shifted (320 nm), whereas RF displays a strong peak at 380 nm.[12,26] The second-lowest absorption of RoF is suggested to be hidden under the main first absorption band, in contrast to RF and some other flavin derivatives.[12] Finally, the third adsorption band of RoF is very strong and seen in the UV region at 257 nm.[12] This third peak is red-shifted compared to RF.[26] The absorption and emission spectral properties for isolated RF and RoF in a variety of solvents are summarized in ref 10.

At pH 7, RoF has a very low fluorescence quantum yield (Φ_F = 3.2 × 10^{-4}) around 540 nm,[12] while neutral RF shows very strong fluorescence (Φ_F = 0.26).[26] The quenching of fluorescence in RoF has been associated with its dual fluorescence character (i.e., fast and slow components in the fluorescence spectrum) stemming from its ability for internal charge transfer that is missing in RF.[10] Accordingly, the polarity of the solvent is one of the main determinants of the extent of charge transfer in RoF.[10,32] Solutions of RoF are red in polar solvents, indicating enhanced internal charge transfer, and yellow to orange with a strong green fluorescence in nonpolar solvents that tend to suppress internal charge transfer.[10,29,32,33]

Donor−acceptor (D−A, or push−pull) molecules form a large family of chromophores. They contain an electron donor and an acceptor moiety that are interconnected by a single bond, e.g., as in aliphatic amines or aminobenzonitriles. An extensive review of D−A molecules can be found in ref 34. As a common characteristic, D−A molecules can undergo an ICT process after photoexcitation. ICT involves the transfer of charge density from the donor to the acceptor and the formation of a zwitterionic photoadduct accompanied by structural rearrangement. Three different theoretical models have been proposed to describe the ICT mechanism and the resulting zwitterionic forms, which differ in the relative orientation of the planes of the donor and acceptor groups (see Figure 2): (a) a planar intramolecular charge transfer (PICT) yielding the quinoid form,[35] (b) a twisted intramolecular charge transfer (TICT) leading to a perpendicular-twisted species,[36] and (c) a wagging intramolecular charge transfer (WICT) generating a zwitterion with a pyramidal donor group.[37] The planes of the donor and the acceptor moieties are (close to) parallel in PICT and (almost) perpendicular to each other in TICT, while in WICT the plane of the donor adopts a pyramidal orientation by bending out of the plane of the acceptor group. Until now, there is no consensus on whether one of the three different mechanisms is generally able to explain the ICT process in typical D−A molecules (and if so, which one). In this context, RoF has the characteristics of a D−A molecule, with the DMA group acting as the donor and the isoalloxazine ring as the acceptor. Hence, an ICT process in RoF may be expected to lead to one (or more) of the three zwitterionic RoF photoproducts (see Figure 2). In previous experimental work on RoF,[10] only PICT and TICT were anticipated, without any clear preference.[10] In a later low-temperature fluorescence decay study on RoF, Zirak et al.[27] favored the TICT over the PICT mechanism.

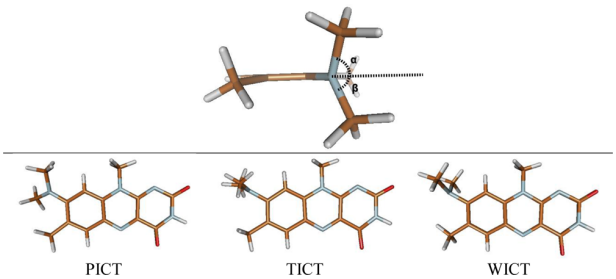


Figure 2. (top) Definition of the dihedral angles, α = C9−C8−N8−CMe1 and β = C9−C8−N8−CMe2, that describe the torsion of the DMA group with respect to the isoalloxazine plane (see text for details). (bottom) Planar (PICT), perpendicular-twisted (TICT), and wagging (WICT) forms of the zwitterionic photoproducts of roseolumiflavin.

Only rather few quantum chemical studies on RoF (or closely related analogues) are available in the literature. In early work by Pill-Soon Song et al.,[12] the transition dipole moment vectors were calculated at the ground-state minimum using an SCF-MO CI method. The first theoretical electronic absorption, CD, and MCD spectra of RoF were determined by Shiga et al.[11] at the semiempirical Pariser−Parr−Pople level. Choe et al.[13] performed DFT/B3LYP calculations on the absorption spectra of a roseoflavin-like model system (8-NH_2-lumiflavin, with an NH_2 group instead of the DMA substituent) in the gas phase and in water, but they did not cover the structural and electronic changes following ICT. Merz et al.[15] used an *ab initio* method (CC2) in combination with QM/MM techniques to investigate roseoflavin photophysics in the gas phase, in water, and in a flavoprotein environment (BLUF domain) to answer the question of why BLUF photoreceptors with roseoflavin cofactors lose their biological functionality. The scarcity of previous theoretical work in this area calls for further computational studies to achieve a more detailed understanding of roseoflavin photophysics at the molecular level.

In this contribution, we present quantum-chemical calculations on excited-state properties of an RoF model system in vacuum, benzene, and water, with special focus on the molecular mechanism of the ICT process. We employ time-dependent density functional theory (TD-DFT)[38,39] and DFT-based multireference configuration interaction (DFT/MRCI)[40] to analyze the minimum-energy geometries, the one-electron properties, and the conical intersections at the LE/CT crossing seams. Our calculations provide molecular-level explanations on why the WICT form is not a feasible ICT photoproduct of RoF, why the TICT mechanism is preferred over PICT, and why polar solvents enhance the ICT rates.

2. COMPUTATIONAL METHODS

The photophysical properties of flavins are mainly affected by the isoalloxazine part, as shown by the absorption and emission spectra of flavin mononucleotide (FMN), riboflavin, and lumiflavin.[6,7] While the ribityl side chain may undergo intramolecular interactions with the N1 atom in solution,[41] it is mainly responsible for the stabilization of the flavin inside the protein environment. Therefore, and in consideration of the required computational effort, we chose to model roseoflavin by roseolumiflavin (RoLF), in which the ribityl chain is replaced by a methyl group. This model system contains 36 atoms and is depicted in Figure 2.

The minimum-energy geometries of the ground state and low-lying electronically excited states were obtained using the Gaussian 09 software.[42] Geometry optimizations were carried out at the DFT level using B3LYP/6-31G(d)[43] and CAM-B3LYP/6-31G(d)[44] for the ground state and linear-response TD-DFT[45−47] for the excited states; here, CAM-B3LYP was chosen due to its documented success in describing the excited

dx.doi.org/10.1021/jp506101x | J. Phys. Chem. B 2015, 119, 928−943

states of local, Rydberg, and CT character in a benchmark with 18 different molecules.[48] Starting from input geometries resembling the designated PICT, TICT, or WICT zwitterions, the corresponding stationary points of the relevant excited states were located using the nonadiabatic optimization algorithm in Gaussian 09 that allows for monitoring the state being followed. The optimized structures were checked for imaginary frequencies to ensure that they represent true minima. No symmetry constraints were applied in geometry optimizations (C_1 point group, with unconstrained torsion of the DMA group).

Solvent effects were taken into account using an implicit continuum solvent model, the conductor-like polarizable continuum model (CPCM),[49,50] as implemented in Gaussian 09. Dielectric constants of $\varepsilon = 78.00$ and $\varepsilon = 2.27$ were chosen for water and benzene, respectively.[19] For aqueous solution, an alternative approach is to treat nearby water molecules explicitly, in order to allow for a more realistic description of hydrogen bonding around the hydrophilic part of roseoluminflavin. In the spirit of previous studies on flavin derivatives,[4,17,19,51] we considered microsolvation in combination with a continuum treatment of bulk solvent, with the isoalloxazine moiety of roseolumiflavin being surrounded by four explicit water molecules (see the next section for details). We investigated microsolvation only for water but not for benzene. In the case of benzene, the specific interactions with RoLF are expected to be weaker, and their explicit modeling did not seem worthwhile in view of the very high computational costs.

Multireference methods are in principle well suited to describe different types of excited states, including CT states and states with double-excitation character. With this in mind, we performed single-point calculations using the DFT-based multireference configuration interaction (DFT/MRCI) method[40] to determine vertical excitation energies, oscillator strengths, and dipole moments of the electronically excited states at the DFT-optimized geometries. DFT/MRCI was chosen because of its good performance in the computation of vertical excitation energies of flavins (typical errors of 0.2 eV)[19] and in benchmarks for a large set of valence excited states of organic molecules with $\pi\pi^*$ and $n\pi^*$ character.[52] The molecular orbitals (MOs) for the DFT/MRCI calculations were generated at the DFT(BHLYP/TZVP) level using TURBOMOLE (version 6.3).[53] DFT/MRCI calculations were performed without symmetry (C_1 point group) using the standard parameters[40] to compute the 20 lowest roots. When required, solvent effects were included implicitly during the generation of the Kohn−Sham MOs by using the conductor-like screening model (COSMO) implemented in TURBOMOLE.[54]

3. RESULTS AND DISCUSSION

Throughout this paper, we make use of a pair of dihedral angles, θ(C9−C8−N8−C_{Me1}) and θ(C9−C8−N8−C_{Me2}), (labeled α and β, respectively, see the top panel of Figure 2), to measure the twist of the DMA group with respect to the isoalloxazine plane. Generally, the isoalloxazine ring remains essentially planar in each relaxed geometry, whereas the orientation of DMA may vary significantly, and the methyl groups at the C7 and N10 positions adapt accordingly. Therefore, we will use α and β as the main descriptors for characterizing different geometries. In specific cases, we will also address some of the bond lengths for more detailed comparisons (*vide infra*).

We classify the stationary points on the excited-state surfaces into three groups with different orientation (i.e., PICT, TICT, and WICT), in accordance with the three possible zwitterionic forms (Figure 2). The members of the same group have similar geometries with low root-mean-square deviations (RMSD values) and almost equal dihedral angles (α/β). The typical values of the dihedral angles for each group depend to some extent on the solvent being present and on the density functional being used in the optimization (see Tables S1 and S4 in the Supporting Information for detailed data).

3.1. Assessing the Performance of B3LYP and CAM-B3LYP for Excited-State Properties. Minimizations using B3LYP and CAM-B3LYP lead to rather similar geometries for the various states of interest in RoLF (see the Supporting Information) and are thus inconclusive with regard to performance assessment. Larger differences are found for excited-state energies. Figure 3 shows the vertical transition energies from TD-B3LYP/6-31G(d), TD-CAM-B3LYP/6-31G(d), and DFT/MRCI calculations for ground-state (GS), PICT, and TICT species. The GS, PICT, and TICT minima of RoLF were located at the B3LYP and CAM-B3LYP levels, and the resulting optimized geometries were then used in single-point TD-DFT and DFT/MRCI calculations to determine the excited-state energies. The corresponding frontier Kohn−Sham orbitals are presented in Figure S1 (Supporting Information). For a given species, the DFT/MRCI energies are essentially identical at B3LYP and CAM-B3LYP geometries (Figure 3), in line with the minute geometry differences (see the Supporting Information, part A).

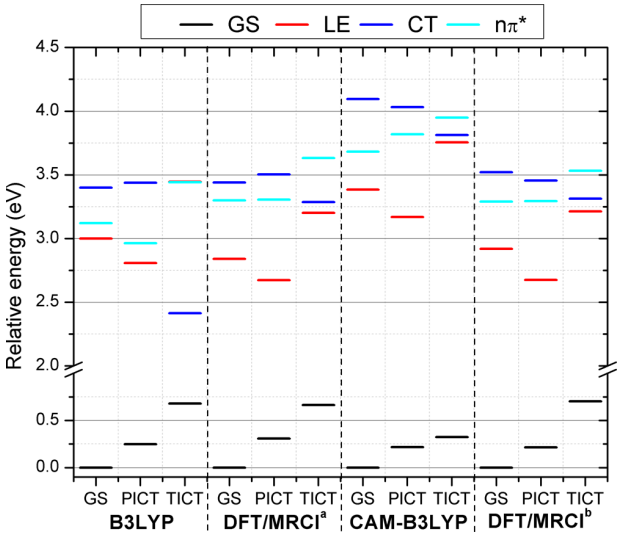


Figure 3. Comparison of the GS, LE, CT, and $n\pi^*$ state energies of RoLF from TD-B3LYP/6-31G(d), TD-CAM-B3LYP/6-31G(d), and DFT/MRCI calculations at the GS, PICT, and TICT minima. B3LYP and DFT/MRCI[a] energies at B3LYP optimized geometries (left two columns) and CAM-B3LYP and DFT/MRCI[b] energies at CAM-B3LYP optimized geometries (right two columns) are given relative to the GS minimum energy at the same level of theory. A break was applied on the *y*-axes (1−1.75 eV). The $n\pi^*$ and LE states in the B3LYP/TICT column are degenerate.

A closer inspection of Figure 3 reveals that the three applied methods give the most similar energy level schemes at the GS

dx.doi.org/10.1021/jp506101x | J. Phys. Chem. B 2015, 119, 928−943

Table 1. DFT/MRCI Singlet and Triplet Adiabatic Excitation Energies (ΔE), Oscillator Strengths ($f(r)$), and Dipole Moments ($|\mu(r)|$) of Several Low-Lying States Calculated at Several Minima of Roseolumiflavin in the Gas Phase[a]

order and types of states		electronic structure of the given state at GS minimum geometry		type of the minimum geometry: GS (planar)	LE state (PICT)	CT state (TICT)	T1 min (planar)
				Singlet			
GS	$1^1A'$	ground state (96%)	ΔE	0.00 eV	0.22 eV	0.70 eV	0.37 eV
			$\|\mu(r)\|$	10.64 D	11.37 D	11.07 D	11.94 D
^{1}LE	$2^1A'$	(90%) $\pi_H \rightarrow \pi_L^*$	ΔE	2.92 eV	2.68 eV	3.21 eV	2.82 eV
			$f(r)$	0.50	0.38	0.27	0.44
			$\|\mu(r)\|$	13.03 D	11.83 D	11.20 D	12.46 D
$^1n\pi^*$	$3^1A'$	(45%) $n_{H-4} \rightarrow \pi_L^*$	ΔE	3.29 eV	3.30 eV	3.53 eV	3.34 eV
		(31%) $n_{H-3} \rightarrow \pi_L^*$	$f(r)$	0.00	0.00	0.00	0.00
			$\|\mu(r)\|$	7.19 D	3.81 D	2.16 D	4.29 D
^{1}CT[b]	$5^1A'$	(77%) $\pi_{H-1} \rightarrow \pi_L^*$	ΔE	3.52 eV	3.46 eV	3.31 eV	3.54 eV
			$f(r)$	0.15	0.23	0.04	0.17
			$\|\mu(r)\|$	19.84 D	19.44 D	26.77 D	15.64 D
				Triplet			
^{3}LE	$1^3A'$	(87%) $\pi_H \rightarrow \pi_L^*$	ΔE	2.20 eV	2.02 eV	2.55 eV	2.05 eV
			$\|\mu(r)\|$	11.97 D	10.81 D	8.94 D	11.12 D
^{3}CT	$3^3A'$	(66%) $\pi_{H-1} \rightarrow \pi_L^*$	ΔE	2.97 eV	2.89 eV	3.23 eV	3.01 eV
			$\|\mu(r)\|$	15.04 D	16.44 D	24.55 D	15.52 D

[a]Energies are given relative to the ground-state minimum. [b]Please note that the CT state is the S_2 state at the TICT minimum and the S_4 state at the other geometries.

Table 2. DFT/MRCI Singlet and Triplet Adiabatic Excitation Energies (ΔE), Oscillator Strengths ($f(r)$), and Dipole Moments ($|\mu(r)|$) of Several Low-Lying States Calculated at Several Minima of Roseolumiflavin in Water[a]

order and types of states		electronic structure of the given state at GS minimum geometry		type of the minimum geometry: GS (planar)	LE state (PICT)	CT state (TICT)	T1 min (planar)
				Singlet			
GS	$1^1A'$	ground state (95%)	ΔE	0.00 eV	0.09 eV	0.67 eV	0.16 eV
			$\|\mu(r)\|$	17.28 D	16.37 D	20.75 D	22.04 D
^{1}LE	$2^1A'$	(84%) $\pi_H \rightarrow \pi_L^*$	ΔE	2.76 eV	2.58 eV	3.30 eV	2.64 eV
			$f(r)$	0.70	0.76	0.24	0.80
			$\|\mu(r)\|$	22.00 D	24.09 D	15.58 D	24.14 D
^{1}CT	$3^1A'$	(83%) $\pi_{H-1} \rightarrow \pi_L^*$	ΔE	3.16 eV	3.02 eV	2.73 eV	3.11 eV
			$f(r)$	0.01	0.00	0.01	0.00
			$\|\mu(r)\|$	22.90 D	22.95 D	33.04 D	23.27 D
$^1n\pi^*$	$4^1A'$	(76%) $n_{H-3} \rightarrow \pi_L^*$	ΔE	3.57 eV	3.61 eV	3.93 eV	3.67 eV
			$f(r)$	0.00	0.00	0.00	0.00
			$\|\mu(r)\|$	15.74 D	18.77 D	14.66 D	20.08 D
				Triplet			
^{3}LE	$1^3A'$	(89%) $\pi_H \rightarrow \pi_L^*$	ΔE	2.05 eV	1.85 eV	2.77 eV	1.85 eV
			$\|\mu(r)\|$	20.93 D	24.31 D	22.52 D	24.77 D
^{3}CT	$2^3A'$	(84%) $\pi_{H-1} \rightarrow \pi_L^*$	ΔE	2.69 eV	2.60 eV	2.66 eV	2.72 eV
			$\|\mu(r)\|$	21.93 D	23.27 D	25.06 D	23.53 D

[a]Energies are given relative to the ground-state minimum.

minimum (e.g., in terms of the energy gaps between low-lying excited states). They yield the same state order: LE is the S_1 state, whereas the CT state is considerably higher in energy (S_5, S_4, and S_5 for B3LYP, CAM-B3LYP, and DFT/MRCI, respectively). All other states below the CT state have $n\pi^*$ character, with very low oscillator strength (note that only the lowest $n\pi^*$ state is shown in Figure 3). All three methods give the same state order at the PICT geometry as well. When going from the PICT to the TICT geometry, the lowest $n\pi^*$ state is destabilized and moves above the CT state. At the TICT geometry, the CT state is stabilized, while the LE state is destabilized by the torsion of the DMA group which decreases the π-overlap between the DMA substituent and the isoalloxazine ring, in line with the principle of minimum overlap in the TICT mechanism.[36] More striking is the change in the state order at the B3LYP level at the TICT geometry: the CT state is overstabilized by B3LYP and moves below the LE state. In the case of CAM-B3LYP and DFT/MRCI, the LE and CT states are close to each other at the TICT geometry, with the LE state still being slightly lower in energy. Considering the experimentally observed decrease of the ICT rate in RoF with decreasing solvent polarity,[10,32] an internal charge transfer upon photoexcitation seems unlikely in vacuum, which is consistent only with the CAM-B3LYP and DFT/MRCI results.

dx.doi.org/10.1021/jp506101x | J. Phys. Chem. B 2015, 119, 928−943

Table 3. DFT/MRCI Singlet and Triplet Adiabatic Excitation Energies (ΔE), Oscillator Strengths ($f(r)$), and Dipole Moments ($|\mu(r)|$) of Several Low-Lying States Calculated at Several Minima of Roseolumiflavin in Benzene[a]

order and types of states		electronic structure of the given state at GS minimum geometry		type of the minimum geometry: GS (planar)	LE state (PICT)	CT state (TICT)	T1 min (planar)
				Singlet			
GS	$1^1A'$	ground state (94%)	ΔE	0.00 eV	0.16 eV	0.64 eV	0.26 eV
			$\|\mu(r)\|$	12.91 D	14.25 D	12.44 D	15.54 D
^{1}LE	$2^1A'$	(79%) $\pi_H \rightarrow \pi_L^*$	ΔE	2.88 eV	2.67 eV	3.27 eV	2.77 eV
			$f(r)$	0.57	0.50	0.28	0.62
			$\|\mu(r)\|$	15.59 D	14.86 D	12.20 D	16.56 D
^{1}nπ^*	$3^1A'$	(41%) $n_{H-3} \rightarrow \pi_L^*$	ΔE	3.34 eV	3.42 eV	3.81 eV	3.48 eV
		(36%) $n_{H-4} \rightarrow \pi_L^*$	$f(r)$	0.00	0.00	0.00	0.00
			$\|\mu(r)\|$	11.27 D	9.33 D	7.26 D	9.63 D
^{1}CT	$4^1A'$	(76%) $\pi_{H-1} \rightarrow \pi_L^*$	ΔE	3.39 eV	3.25 eV	3.05 eV	3.31 eV
			$f(r)$	0.10	0.18	0.00	0.13
			$\|\mu(r)\|$	21.38 D	22.08 D	29.11 D	20.71 D
				Triplet			
^{3}LE	$1^3A'$	(91%) $\pi_H \rightarrow \pi_L^*$	ΔE	2.16 eV	2.00 eV	2.62 eV	2.00 eV
			$\|\mu(r)\|$	14.80 D	14.67 D	10.96 D	16.12 D
^{3}CT	$2^3A'$	(85%) $\pi_{H-1} \rightarrow \pi_L^*$	ΔE	2.87 eV	2.76 eV	3.04 eV	2.89 eV
			$\|\mu(r)\|$	18.30 D	19.24 D	29.15 D	18.49 D

[a]Energies are given relative to the ground-state minimum.

The LE, CT, and nπ^* state energies are generally predicted to be higher by CAM-B3LYP than by DFT/MRCI. The first vertical absorption of RoLF is calculated at 391 and 449 nm by CAM-B3LYP and DFT/MRCI, respectively. The DFT/MRCI result is closer to the experimental value (500 nm) for RoF as expected. In the following, we will make use of CAM-B3LYP geometries and energies in combination with single-point DFT/MRCI calculations.

3.2. Roseolumiflavin Photophysics in Different Environments. In this section, we analyze the photophysical properties of roseolumiflavin in the gas phase, in water, and in benzene. Water and benzene were chosen as solvents to investigate the influence of solvent polarity. In order to keep track of the changes occurring on the GS, LE, and CT state surfaces upon twisting of the DMA group after photon absorption, we optimized the corresponding GS, PICT, TICT, and T_1 minima at the (TD-)CAM-B3LYP level.

Tables 1−3 list the adiabatic excitation energies, oscillator strengths, dipole moments, and contributions from the leading configurations for the lowest singlet and triplet states, as obtained from single-point DFT/MRCI calculations at the GS, PICT, TICT, and T_1 minima in the gas phase, water, and benzene, respectively. For a qualitative inspection of the results, the adiabatic energies relative to the GS minimum energy are depicted in Figure 4 for all relevant states in the gas phase, water, and benzene. The corresponding frontier Kohn−Sham orbitals (computed at the BHLYP/TZVP level at the aforementioned minima) are shown in Figures S2−S4 of the Supporting Information. Generally, the CT states in Tables 1−3 are characterized by a dipole moment that is significantly higher compared to the ground state. The LE states have high oscillator strength, whereas the nπ^* states have almost zero oscillator strength. The assignment of LE, CT, or nπ^* character was always verified by an inspection of the frontier Kohn−Sham orbitals that contribute to the dominant configuration(s).

3.2.1. Gas Phase. At the ground-state minimum geometry, the lowest singlet excited state (S_1) is an LE state with high oscillator strength (f = 0.50) that is dominated by the $\pi_H \rightarrow \pi_L^*$ transition (see Figure S2, Supporting Information, for the MOs). The CT state (S_4) mainly involves the $\pi_{H-1} \rightarrow \pi_L$ transition and lies 0.60 eV above the LE state. It can also be regarded to be of mixed LE/CT character, since it has a moderately high oscillator strength (f = 0.15) and a high dipole moment. These features arise from the mixed character of HOMO−1, with n-components from the nitrogen lone pair of the DMA group and π-components from the isoalloxazine ring (Figure S2, Supporting Information). The S_3 and S_4 states have nπ^* character (f = 0.00) and are complementary to each other in the sense that they are well described as single excitations from the oxygen lone pairs on O2 and O4 to the lowest π^* orbital ($n_{H-3} \rightarrow \pi^*$ and $n_{H-4} \rightarrow \pi^*$, see Figure S2, Supporting Information).

Upon photoexcitation, the LE state is populated in the Franck−Condon (FC) region. The ultrafast rearrangement of the electronic configuration is followed by the slower relaxation of the nuclei leading to the PICT minimum on the LE surface. This relaxation lowers the energy by a rather small amount (0.24 eV, see Table 1), in accord with the slight torsion of the DMA group. The CT state is also stabilized but less so (by 0.06 eV), probably due to a slight enhancement in the n-character of HOMO−1 (involved in the leading CT configuration). The nπ^* state is apparently not affected by this slight torsion of the DMA group, since it preserves its place in the state order and its relative energy remains essentially the same.

The TICT orientation is reached by a large twist (ca. 90°) of the DMA group. It corresponds to the minimum of the CT state. In going from the PICT to the TICT minimum, both the GS and LE states are destabilized, whereas the CT state is further stabilized and becomes the S_2 state, remaining slightly above the LE S_1 state (by 0.1 eV). Hence, the two states do not cross, in agreement with the expectation that ICT is not facile in the gas phase (see section 3.1).[10,29,32,33] Due to the perpendicular orientation of the DMA group with respect to the isoalloxazine plane, the π-systems of the isoalloxazine moiety and the DMA group do not overlap. Accordingly, the HOMO−1 is mainly localized on the DMA group (see Figure

dx.doi.org/10.1021/jp506101x | J. Phys. Chem. B 2015, 119, 928−943

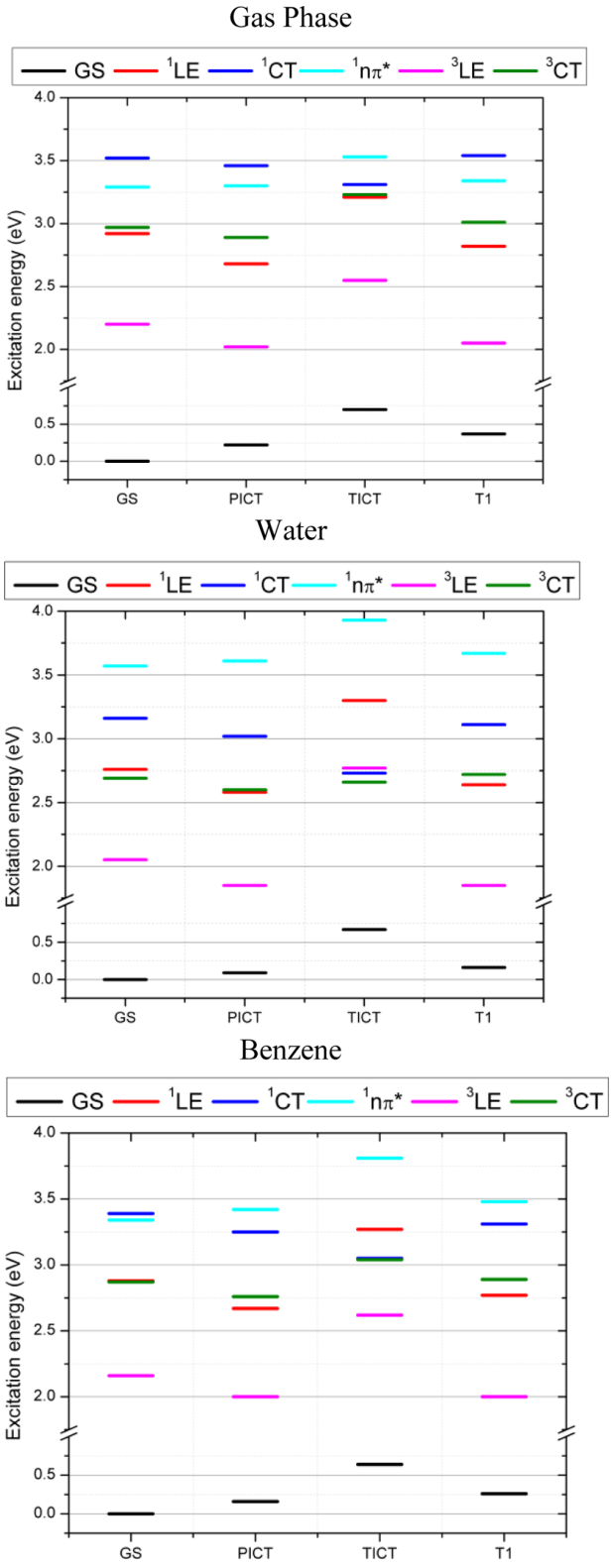


Figure 4. Adiabatic energies of the lowest states of roseolumiflavin in the gas phase, water, and benzene from single-point DFT/MRCI calculations at CAM-B3LYP minimum geometries, with CPCM corrections for water or benzene. Energies are given relative to the GS minimum (see Tables 1−3 for corresponding values). A break was applied on the *y*-axes (1−1.75 eV).

S2, Supporting Information, TICT row), which boosts the nπ^* character of the CT state. Therefore, the CT state has a high dipole moment (26.8 D) and almost zero oscillator strength because of the absence of the $\pi\pi^*$ character. Upon DMA torsion, the lowest nπ^* state is destabilized (by 0.23 eV) and crosses with the highly stabilized CT state, as in previous CC2 calculations.[15] However, contrary to the CC2 results, it does not cross with the GS or LE state at the DFT/MRCI level.

Triplet states can likewise be assigned as LE or CT states. At the ground-state minimum, the T_1 state has LE character, while the T_3 state has CT character (as indicated by its high dipole moment). Contrary to the singlet CT state, the triplet CT state is destabilized at the TICT geometry, where these CT states (now S_2 and T_2) become almost degenerate (ΔE = 0.08 eV). The degeneracy of low-lying singlet and triplet states at the perpendicular-twisted geometry is an integral part of the TICT mechanism.[55] This near-degeneracy is a prerequisite for triplet formation in the gas phase by intersystem crossing provided that there is sufficient spin−orbit coupling. However, triplet formation in the gas phase has not yet been observed experimentally for roseoflavin.

The structural properties (e.g., the DMA dihedral angles) and the excited-state properties (i.e., vertical excitation energies, dipole moments, and oscillator strengths) are very similar at the lowest triplet minimum (LE, T_1) and the singlet PICT geometry (LE, S_1). This further supports the assignment of the T_1 state as an LE state.

3.2.2. Water. Compared with the gas phase, the major difference of the photophysics of roseolumiflavin in water is the crossing of the LE and CT surfaces upon torsion of the DMA group with respect to the isoalloxazine plane. This enables the population of the CT state with concomitant depopulation of the LE state and results in an intramolecular charge transfer, as observed experimentally.[10,25,27,32] As evident from Table 2 and Figure 4, the energetic order of the singlet (as well as triplet) LE and CT states is reversed when the DMA group is twisted into the TICT orientation, and one may thus expect that there is a conical intersection seam connecting the two surfaces. Unlike the gas phase, the nπ^* state lies well above the CT state at any given conformation, which prevents any possible crossing with the CT state; it is somewhat destabilized upon DMA torsion. Compared with the gas phase (Table 1), we also note a general trend (except for the nπ^* state) of a lowering of the computed adiabatic excitation energies at each of the listed minimum geometries (Table 2), indicating a stabilization of the highly polar excited states by interactions with the polar solvent.

Further analysis of the data in Table 2 reveals the following points. At the ground-state minimum, the LE state is still the lowest singlet excited state (S_1) in water, but it has a higher oscillator strength (f = 0.70) compared to the gas phase. At the DFT/MRCI level, the LE state mainly involves a $\pi_H \rightarrow \pi_L^*$ transition which requires 2.76 eV (450 nm). The absorption peak is blue-shifted compared to the experimental band maximum at 2.46 eV (503 nm).[10] The deviation is still within the usual error margin (ca. 0.2−0.3 eV) of the DFT/MRCI method.[5] We also note that a more thorough comparison would need to account for the vibrational progressions in the experimental absorption spectra.[56]

With the stabilization brought about by solvent polarity, the CT state is now the S_2 state at the ground-state geometry of RoLF in water, but it still mainly involves the $\pi_{H-1} \rightarrow \pi_L^*$ transition. The energy separation of the LE and CT states is

dx.doi.org/10.1021/jp506101x | J. Phys. Chem. B 2015, 119, 928−943

smaller in water (0.40 eV) than in the gas phase (0.60 eV), presumably due to the stronger stabilization of the polar CT state by the polar solvent. We note, however, that the dipole moments of the LE and CT states differ much less in water (22.0 D vs 22.9 D) than in the gas phase (13.0 D vs 19.9 D). Hence, the polarity of both states is increased in the polar solvent environment (less so in the intrinsically more polar CT state). The assignment of the CT state of roseolumiflavin in water is still clear-cut considering the dominant excitation, the dipole moment, and the transition dipole moment vector (pointing from the donor to the acceptor group as in the gas phase, data not shown). The nπ^* state, on the other hand, has a dipole moment in water that is somewhat smaller than that of the ground state (15.74 D vs 17.28 D), reminiscent of the situation in the gas phase (7.19 D vs 10.64 D).

The frontier Kohn−Sham orbitals computed at the GS minimum in water (Figure S3, Supporting Information, top panel) slightly differ from those in the gas phase. The HOMO is delocalized over the three rings of the isoalloxazine moiety, whereas the HOMO−1 is localized more on the DMA group (still with some overlap of the π-systems of the donor and acceptor moieties). The CT state is essentially devoid of LE character in water (as opposed to the gas phase, *vide supra*), which is reflected in the almost zero oscillator strength. The nπ^* state arises from excitations involving the same MOs as in the gas phase (Figures S2 and S3, Supporting Information), with the $n_{H-3} \rightarrow \pi^*$ excitation being more dominant in water (see Tables 1 and 2).

As in the gas phase, the LE state relaxes after the vertical photoexcitation to the PICT geometry, which is the minimum on the LE surface. This corresponding energy reduction (0.16 eV) is somewhat smaller than that in the gas phase (0.24 eV), presumably, because the LE state is already stabilized by the polar solvent so that the small torsion of the DMA group provides less additional stabilization. As a result of the minor geometry changes when going from the GS to the PICT minimum, the variation of the Kohn−Sham orbitals (Figure S2, Supporting Information, middle panel) is minute and hard to detect.

From the PICT minimum, roseoflavin can follow two paths: forward or backward. In the latter case, RoF relaxes back to the GS either by nonradiative or radiative transitions. The nonradiative route requires conformational rearrangement of the DMA group, whereas the radiative route involves fluorescence from the LE state, which constitutes the fast component ($\tau_{F,f}$ = 1.42 ps) of the experimentally observed two-component radiative decay.[10] This vertical fluorescence peak is calculated to be at ca. 499 nm (2.49 eV), while the experimental one is at 536 nm (2.32 eV).[10] On the forward path, roseoflavin has to overcome a low barrier to populate the CT state through a conical intersection (CIn) seam, where the LE and CT state surfaces are crossing. This involves an intramolecular charge transfer and nuclear relaxation toward the TICT orientation (the minimum on the CT surface). A detailed analysis of the CIn seam will be given later in section 3.4 (*vide infra*).

In the TICT orientation, the CT state is energetically lower than the LE state at the DFT/MRCI level (Table 2). Besides, the oscillator strength of the LE state is reduced substantially (f = 0.24), and the corresponding dipole moment is also lower than in the PICT orientation. On the contrary, the dipole moment of the CT state rises to 33.0 D, clearly reflecting the charge separation created by ICT. Inspection of the Kohn−Sham MOs (Figure S3, Supporting Information) rationalizes these changes in oscillator strength and dipole moment. In contrast to the PICT geometry, the LE and CT states at the TICT geometry involve $\pi_{H-1} \rightarrow \pi_L^*$ and $\pi_H \rightarrow \pi_L^*$ transitions, respectively. There are changes both in the character of the frontier MOs and in the composition of the excited states. At the TICT geometry, the HOMO is localized on the DMA group, with no contribution from the isoalloxazine moiety; there is minimum overlap between the π-systems of the donor and acceptor, in accordance with the proposed principle requirements of the TICT mechanism.[36] The LUMO is delocalized over the isoalloxazine ring so that the $\pi_H \rightarrow \pi_L^*$ transition will involve charge migration from the donor DMA group to the acceptor isoalloxazine group, in line with the elevated dipole moment of the CT state.

At the DFT/MRCI level, the PICT minimum is lower in energy than the TICT minimum by 3.57 kcal/mol (0.15 eV), and thus, it is the more stable zwitterionic form. However, the energy difference is small so that the interchange between the two conformations seems quite probable. This opens a path for nonradiative return from the CT to the LE state at the TICT minimum through intramolecular charge recombination. Roseolumiflavin may also return to the ground state by a radiative path: The fluorescence at the TICT orientation constitutes the slow component ($\tau_{F,s}$ = 1.32 ns) of the measured biexponential decay.[10] DFT/MRCI predicts this fluorescence to occur at ca. 600 nm in water. No experimental value is available for comparison, since no separate peak was recorded in water for the slow component of the emission spectrum. Time and spectrally resolved fluorescence measurements would be needed to check whether this peak is hidden under the broad fluorescence of the LE state.

Photoinduced ICT is likely the reason for the strong quenching of fluorescence (low fluorescence quantum yield) of roseoflavin in water, as also proposed for a flavin-based dye dyad.[17] A low rate of radiative emission from the CT state was observed experimentally in water,[10] which is consistent with an internal diabatic electron transfer (i.e., ICT) and a low oscillator strength of the CT state. DFT/MRCI indeed predicts a very small oscillator strength for the CT state of roseolumiflavin in water (f = 0.01) and a somewhat higher value in the gas phase (f = 0.04), where ICT is not feasible.

As in the gas phase, the two lowest triplet states in water can be assigned as LE (T_1) and CT (T_2) states, based on the composition of these states and their dipole moments. As in the singlet case, the triplet CT state is stabilized by the presence of the polar solvent. The triplet LE and CT states are further stabilized by the slight torsion of the DMA group in the PICT orientation, while the perpendicular-twisted TICT orientation of DMA causes a significant destabilization of both triplet states such that they become almost degenerate with the singlet CT state (gaps of 0.04 and 0.07 eV). This should facilitate triplet formation by intersystem crossing. Experimentally, in aqueous solution at pH 7, the phosphorescence quantum yield has indeed been reported to be higher for RoF than for RF.[27]

3.2.3. Benzene. Benzene is a nonpolar but polarizable solvent. As expected, the predicted photophysical characteristics of roseolumiflavin in benzene are intermediate between those in the gas phase and in water. The computed excited-state energies (see Table 3 and Figure 4) generally lie between the values in the gas phase and in water, with trends resembling more the gas phase. Judging from the reversed order of the singlet LE and CT states at the TICT geometry, RoLF will be able to populate the CT state through a conical intersection

dx.doi.org/10.1021/jp506101x | J. Phys. Chem. B 2015, 119, 928−943

also in benzene, like in water, and thus also support photoinduced intramolecular charge transfer. In contrast, the nπ* state is considerably destabilized upon DMA torsion and thus crosses with the CT state, like in the gas phase.

The singlet and triplet LE and CT states have the same energetic order in benzene and water, and the composition of these states is also similar in both solvents. The Kohn−Sham frontier MOs (Figure S3, Supporting Information) differ only very slightly in benzene and water. Some minor quantitative variations in the HOMO may be related to the decreased oscillator strength of the LE state at the GS and PICT minima in benzene.

The computed dipole moments in benzene are generally somewhat larger than in the gas phase but show the same trends (see Tables 1 and 3). The highest deviations from the computed dipole moments in water are found for the singlet and triplet LE states at the PICT geometry and for the triplet LE and CT states at the TICT geometry. In the latter case, the triplet CT has a dipole moment in water that is almost 3 times larger than the one in benzene or in the gas phase. Correspondingly, the triplet states at the TICT geometry have the same energetic order in benzene and in the gas phase but not in water.

3.2.4. Explicit Solvation by Water Molecules. We studied microsolvation effects using a model with four water molecules that are H-bonded to O2, N3, O4, and N5 of RoLF (Figure 5).

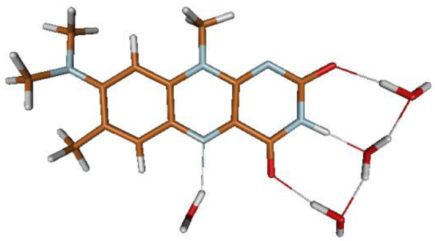


Figure 5. Microsolvation model for RoLF with four water molecules. Hydrogen bonds are depicted with thin lines. Shown is the CAM-B3LYP/6-31G(d) ground-state minimum.

The minima of the ground state and the lowest excited states were optimized at the CAM-B3LYP/6-31G(d) level either in the gas phase or embedded in a continuum solvent (CPCM, water) for the sake of comparison. These optimized geometries were employed in subsequent single-point DFT/MRCI calculations with the SVP basis (instead of TZVP) to limit the computational effort. Test calculations on the isolated roseolumiflavin system (without explicit water molecules) gave deviations in the excitation energies of less than 0.03 eV when switching from the TZVP to the SVP basis.

DFT/MRCI vertical transition energies, oscillator strengths, and dipole moments at the GS minima for this model are compiled in Table 4 along with the corresponding data for isolated RoLF in the gas phase, water, and benzene. Results for dihedral angles are given in Table S4 (Supporting Information). Judging from these data, the addition of four explicit water molecules has a significant effect on the computed geometries and absorption spectrum. Inspection of the Kohn−Sham frontier orbitals shows that there is no qualitative change for LUMO and HOMO−1 (data not shown), while the HOMO (see Figure S5, Supporting Information) is notably affected by the addition of explicit water molecules, with a shift of electron density from the donor DMA group to the acceptor isoalloxazine ring. Both the LE and CT states are stabilized by ca. 0.12 eV at the GS minimum compared to the isolated RoLF, which translates into a bathochromic shift of the first absorption peak from 442 to 463 nm. At the PICT geometry, the addition of the explicit water molecules also stabilizes both the LE and CT states but to a lesser extent (by ca. 0.06 eV). At the TICT geometry, the CT state is stabilized more strongly than the LE state (0.21 eV vs 0.06 eV). This suggests that the use of models with explicit microsolvation is advisable to capture differential stabilization effects on states of different polarity, which may be important for describing ICT processes.

When the microsolvation model is used by itself, without a continuum description of bulk solvent, the computed absorption spectrum is blue-shifted, and the results resemble those from pure gas phase calculations (Table 4). The experimental excitation energies are better reproduced when including a continuum solvent treatment. Incorporating explicit

Table 4. Gas Phase and COSMO-Corrected DFT/MRCI Vertical Excitation Energies, Oscillator Strengths, and Dipole Moments of the Singlet LE and CT States of RoLF[a]

geometry optimization method	DFT/MRCI details	state	ΔE (eV)	λ (nm)	$f(r)$	$\lvert\mu(r)\rvert$ (D)
only RoLF	vacuum (ε = 1)	LE	2.92	425	0.50	13.03
gas phase CAM-B3LYP		CT	3.52	352	0.15	19.84
only RoLF	COSMO (water, ε = 78)	LE	2.76	450	0.70	21.98
CPCM (water) + CAM-B3LYP		CT	3.16	393	0.01	22.90
RoLF + 4 water	vacuum (ε = 1)[b]	LE	2.85	435	0.56	13.96
gas phase CAM-B3LYP		CT	3.39	366	0.14	20.48
RoLF + 4 water	COSMO (water, ε =78)[b]	LE	2.68	463	0.78	24.85
CPCM (water) + CAM-B3LYP		CT	3.12	397	0.01	23.55
	experimental (water)[c]	LE	2.47	503	0.48	
		CT				
only RoLF	COSMO (benzene, ε = 2.27)	LE	2.88	431	0.57	15.59
CPCM (benzene) + CAM-B3LYP		CT	3.39	366	0.10	21.38
	experimental (benzene)[c]	LE	2.55	486	0.35	
		CT				

[a]The ground-state geometries were optimized at the CAM-B3LYP level in the gas phase, in water (CPCM), in benzene (CPCM), and for the microsolvation model (CPCM). [b]DFT/MRCI calculations with the SVP basis; in all other cases, the TZVP basis was used. [c]Experimental values for RoLF in water and benzene taken from ref 10.

dx.doi.org/10.1021/jp506101x | J. Phys. Chem. B 2015, 119, 928−943

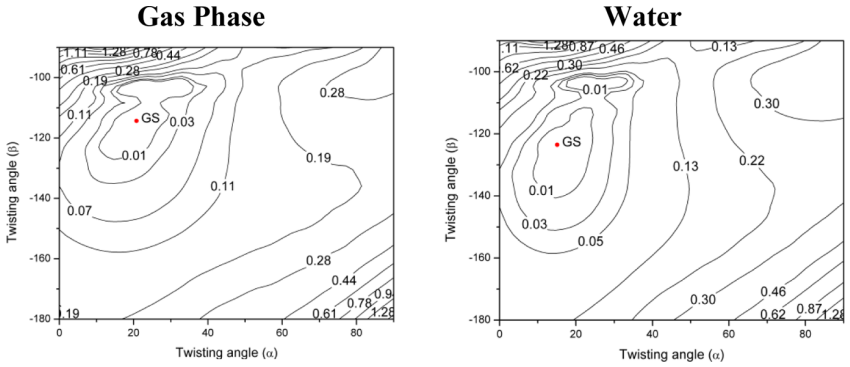


Figure 6. Energy contour map (in eV) of 2D relaxed PES scans of the ground state with respect to the dihedral angles (α and β, Figure 2) in (left) the gas phase and (right) water. The red points with the label "GS" indicate the α/β pairs for the fully optimized stationary points.

and implicit solvent effects through the combination of a microsolvation model and a continuum solvent model thus appears to be the preferred approach to compute realistic absorption spectra using DFT/MRCI.

3.3. Structural Changes upon Photoexcitation in Vacuum and Water. In general, the structural changes following the absorption of a green-light photon can be divided into two groups: (a) bond length alternation in the isoalloxazine ring and (b) twist of the DMA group and the flanking methyl groups bound to N7 and N10. Here we focus on the latter point, noting that the optimized values of the relevant bond lengths of the various mimima of RoLF in the gas phase and in water are documented in Tables S5 and S6 (Supporting Information) (with the C4a−N5 bond showing the largest variations). To analyze the structural changes in terms of the twist of the DMA group, we present two-dimensional (2D) relaxed potential energy surface (PES) scans for the GS, LE, and CT states. The PES of a given state was scanned at the CAM-B3LYP/6-31G(d) level with respect to the dihedral angles for DMA torsion (α and β, Figure 2) in 10° steps. The energies were obtained from constrained geometry optimizations, in which only the two scanning variables were kept frozen, following the gradients of the state of interest. Some points could not be computed, since the corresponding dihedral constraints gave rise to nonphysical (severely hindered) geometries. These points were interpolated/extrapolated using the thin plate spline gridding method implemented in the Origin 8.5 software.

We start the discussion with the ground state of RoLF before photon absorption. The contour maps of 2D PES scans of the ground state in (left) vacuum and (right) water are shown in Figure 6. Energies are given relative to the ground-state minimum. The bottom left origin of the plots corresponds to the perfect in-plane orientation of the two DMA methyl groups with respect to the isoalloxazine plane, while the perpendicular-twisted (TICT) orientation of the two planes is reached at the top right of the contour map.

As seen in Figure 6, the planar conformation of the DMA group is favored energetically in the ground state, while the twisted conformation is disfavored. In terms of (α/β), the ground-state minima in the gas phase (20°/−117°) and water (16°/−126°) deviate somewhat from the perfect planar orientation (0°/−180°), because of the steric hindrance created by the methyl group at the C7 position. The energy suddenly increases when the two DMA methyl groups adopt a perpendicular arrangement with respect to each other (0°/−90° or 90°/−180°) because of the increased van der Waals repulsion between the methyl groups. At the ground-state minimum, the DMA and isoalloxazine planes are more coplanar in water than in the gas phase, which can be attributed to the enhanced π-overlap of the donor and acceptor groups. This is consistent with the enhanced oscillator strength of the optically bright LE state in water that is found both experimentally and computationally (compared with benzene).

For further analysis, the LE and CT state potential energy surfaces were scanned in an analogous manner. The 2D contour maps of the relative energies and the energy differences between the LE and CT states are depicted in Figure 7 as a function of the scanned dihedral angles. The oscillator strength and dipole moment of each state were monitored to keep track of the character of the low-lying excited states. By assigning LE and CT character to the states with highest oscillator strength and dipole moment, respectively, we managed to distinguish between the LE and CT surfaces even when the energy order of these states changed due to crossings. In an alternative visualization, 3D representations of the 2D PES scans are presented in Figure S6 (Supporting Information). Missing points on the surfaces in Figures 7 and S6 (Supporting Information) were again interpolated/extrapolated using the thin plate spline gridding method implemented in Origin 8.5.

In the gas phase (Figure 7, left panel), the LE surface has the global minimum and a local maximum centered at dihedral angles of 20°/−120° and 90°/−90°, respectively. The global minimum from the scan is close to the fully relaxed PICT minimum (at 21°/−117°), whereas the local maximum represents the TICT orientation. Although not covered by the contour map, the mirror image of the PICT minimum (−20°/120°) will also be a minimum with the same energy. The contour map shows that the twisting of the DMA group from planar to perpendicular-twisted orientation destabilizes the LE state. Quantitatively, the PICT minimum has a relative energy of 3.2 eV, while the TICT maximum on the LE surface is higher in energy by ca. 0.3 eV. Contrary to the LE case, the gas phase CT surface has its minimum in the TICT orientation (87/−87°) with a relative energy of 3.8 eV, with another local minimum in the PICT orientation that is only slightly higher in energy by ca. 0.06 eV. Hence, the TICT orientation is favored for the CT state in the gas phase, but the interchange between two orientations should occur very fast due to the low energy barrier. In spite of the destabilization of the LE state and the stabilization of the CT state upon twisting the DMA group, the two surfaces do not cross in the gas phase because of the large

dx.doi.org/10.1021/jp506101x | J. Phys. Chem. B 2015, 119, 928−943

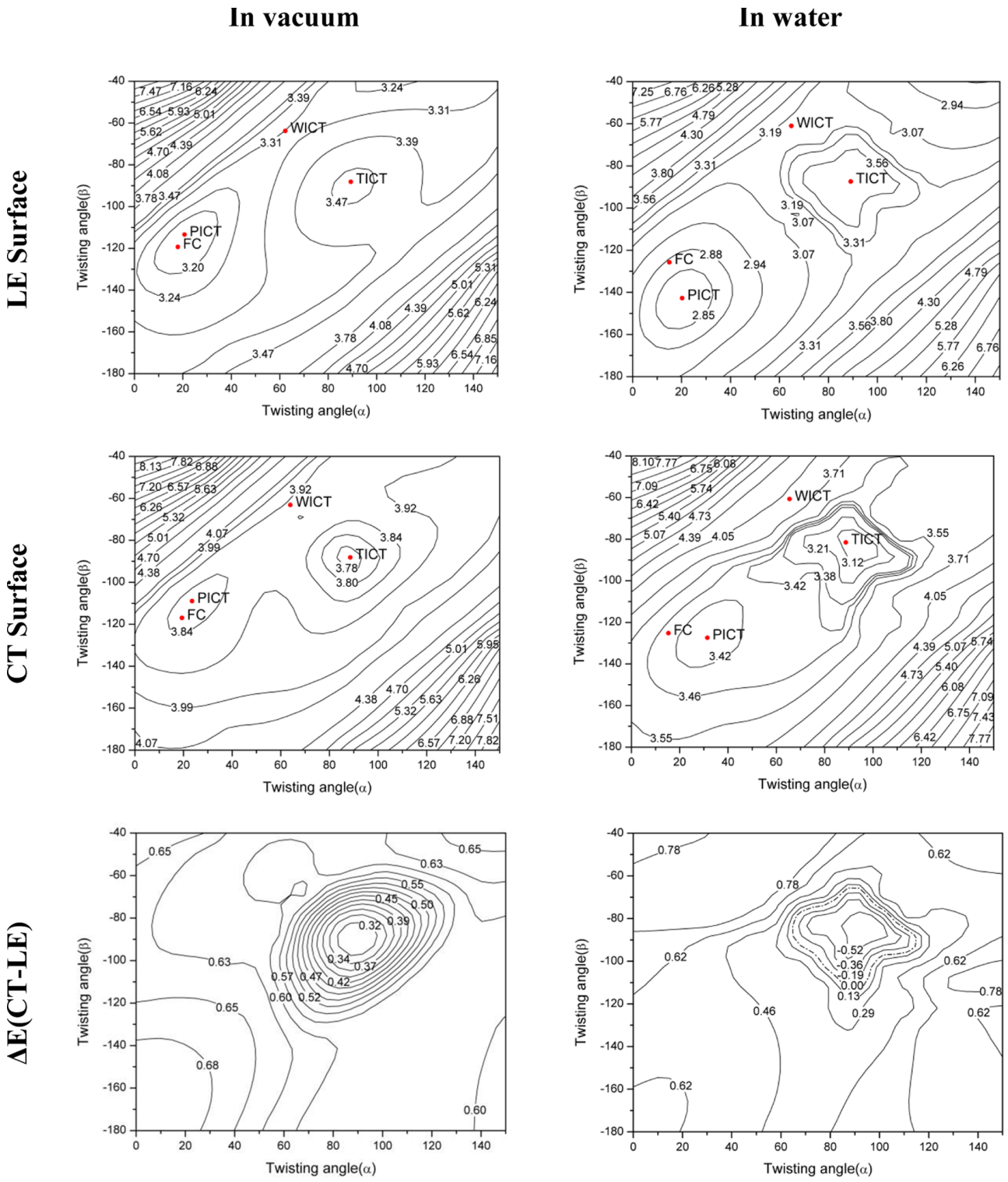


Figure 7. Energy contour map (in eV) of 2D relaxed PES scans of the (top panel) LE and (middle panel) CT states with respect to the dihedral angles (α and β, see Figure 2). Bottom panel: Energy difference of the LE and CT states. The α/β pairs for the fully optimized stationary points for the PICT, TICT, and WICT species on the LE and CT surfaces and the projections of the Franck−Condon point on each surface are indicated by red points with corresponding labels. Geometries were obtained at the CAM-B3LYP/6-31G(d) level in the gas phase (left column) and with CPCM corrections in water (right column) by following LE or CT state gradients. In the water case, the interpolated conical intersection seam is shown as a dash-dot line; there is no such seam in the gas phase.

energy gap between them. This can easily be seen in the 3D PES representation in Figure S6 (Supporting Information) (top panel). Likewise, comparison of the 2D contour maps (Figures 6 and 7) clearly suggests that thermal relaxation is not possible,

dx.doi.org/10.1021/jp506101x | J. Phys. Chem. B 2015, 119, 928−943

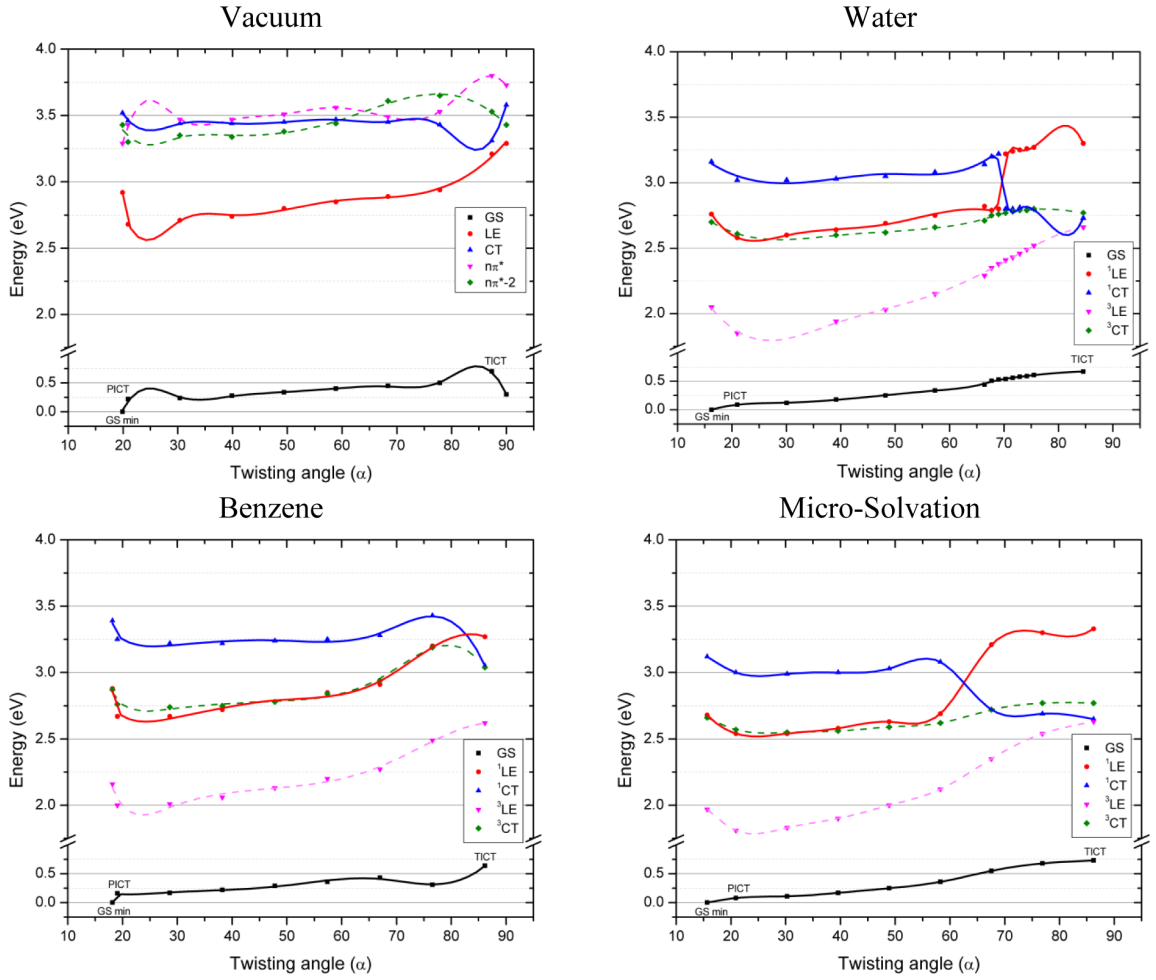


Figure 8. DFT/MRCI adiabatic potential energy curves of the ground state and the lowest excited states as a function of twisting angle (α) obtained using the relaxed linear interpolation geometries between the PICT and TICT minima. Energies are given relative to the corresponding ground-state minimum. Each point was optimized with constraints (fixed values of the dihedral angles α and β) on the S_1 surface at the CAM-B3LYP level in the gas phase, water, and benzene as well as by using the microsolvation model (see section 3.2.4). A break was applied on the y-axes (1−1.75 eV).

since the GS surface does not cross any of the LE, CT, and nπ* surfaces. This finding contrasts with previous CC2 results, which indicate that there is such a relaxation pathway via LE/nπ* and nπ*/GS conical intersections.[15]

In water (Figure 7, right panel), the LE surface has the global minimum at the PICT orientation (21°/−143°) and a maximum at the TICT orientation (86°/−86°), like in the gas phase. However, the location of the PICT minimum and the energies at the PICT and TICT geometries differ to a great extent in water and in the gas phase. Water stabilizes the PICT minimum by about 0.4 eV and destabilizes the TICT maximum by 0.1 eV. Likewise, there are large differences between the CT surfaces in water and in the gas phase. In the latter case, the local PICT minimum on the CT surface is located at ca. 30°/−130°, while the "global" TICT minimum is still close to 90°/−90°. The polar solvent stabilizes the PICT-like and TICT geometries on the CT surface significantly (by 0.4 and 0.7 eV, respectively). Furthermore, the energy difference between the TICT minimum on the CT surface and the PICT minimum on the LE is diminished from ca. 0.6 eV to ca. 0.3 eV in the presence of the polar water solvent. Upon twisting the DMA group, the polar CT state is stabilized more strongly by the polar water solvent and drops below the LE state in energy. As clearly seen from the 3D PES representation (Figure S6, Supporting Information, bottom panel), this leads to a crossing of the LE and CT surfaces close to the TICT orientation. Finally, as in the gas phase, comparison of the contour maps for the ground state and the LE or CT states does not reveal any GS/LE or GS/CT conical intersections in water, again contrary to previous CC2 results that predict a GS/CT conical intersection.[15]

Interestingly, a stationary point can be located on the LE surface in the dihedral angle range around 65°/−65°, both in the gas phase and in water. The corresponding geometries have relative energies of ca. 3.3 and 3.2 eV, respectively. Geometry optimizations starting from different points, picked randomly close to this region, end up at a stationary point at (65.2°/−65.2°), which corresponds to the WICT orientation (i.e., the third zwitterionic form). However, normal-mode analysis reveals that this stationary point is a saddle point with two imaginary-frequency modes. Hence, this WICT species is not a stable zwitterionic form. The WICT mechanism has been considered as a possible ICT mechanism for donor−acceptor molecules in the literature.[34] However, the present TD-DFT calculations suggest that the WICT species is only a stationary

dx.doi.org/10.1021/jp506101x | J. Phys. Chem. B 2015, 119, 928−943

point that connects the two stable zwitterionic forms on the LE surface: PICT and TICT.

3.4. DFT/MRCI Investigation of the LE/CT Conical Intersections. For a complete picture of the photophysics of roseolumiflavin, it is important to locate and characterize the LE/CT conical intersections. The change in the energetic order of the LE and CT states at the CAM-B3LYP and DFT/MRCI levels suggests that there should be a conical intersection seam between two surfaces in water and benzene but not in the gas phase.

Judging from the CAM-B3LYP 2D scan of the LE and CT surfaces (Figure S6, Supporting Information), there is clearly no crossing of the LE and CT surfaces in the gas phase. On the other hand, the PES scans in water reveal a region that seems like an isolated island around the TICT orientation, where the LE state is higher in energy than the CT state. The border of this region (i.e., the CIn seam) is marked in the 2D contour map of the LE−CT energy difference as a dash-dot line (Figure 7, bottom right panel). The minimum-energy crossing point (MECP) on this CIn seam is found in the region (60°/−90° to 70°/−100°) at an energy of ca. 3.26 eV relative to the ground-state minimum. This MECP resembles the TICT rather than the PICT orientation. This may be taken as another indication that roseolumiflavin favors the TICT over the PICT mechanism for the ICT process, as proposed on the basis of a low-temperature decay analysis.[27] In this context, we note that the MECP on the CIn seam lies 0.41 eV above the PICT minimum and 0.14 eV above the TICT minimum. Although these values are not accurate as the MECP has only been roughly located, one may still conclude that roseolumiflavin has to overcome a higher barrier for reaching the CT surface (ICT) than for returning back to the LE surface.

It is well established that TD-DFT methods perform poorly for systems with static correlation, e.g., for almost degenerate states.[57] They may thus not describe CIn seams properly, which can be characterized more accurately by multireference methods. Therefore, we checked the CIn results from CAM-B3LYP by DFT/MRCI calculations. Because of the extreme computational costs, we could only perform a one-dimensional scan instead of a two-dimensional one. We generated six equidistant points by linear interpolation of the dihedral angles α and β (LIDA) between the PICT and TICT minimum. These six starting geometries were optimized at the CAM-B3LYP level, by relaxing all degrees of freedom except the two dihedral angles and following the gradients on the lowest excited singlet state (S_1). Single-point DFT/MRCI calculations were carried out at the relaxed geometries obtained in the gas phase, in water, and in benzene, to determine the vertical excitation energies, dipole moments, and oscillator strengths. The corresponding energies and the best-fit curves (using the cubic B-spline method in Origin 8.5) for the lowest excited states are plotted in Figure 8. The excitation energies at the corresponding ground-state minimum are also given along with the PICT and TICT minima. The LE and CT states could be tracked even in the case of a change in the energetic order by considering their oscillator strengths and dipole moments.

The DFT/MRCI minimum-energy potential curves obtained in the gas phase (Figure 8, top left panel) do not reveal any conical intersection between the LE and CT states, in line with the CAM-B3LYP results. The CT state is the S_4 state for planar geometries, whereas the S_2 and S_3 states have nπ^* character (denoted as nπ^* and nπ^*-2), as evident from the major contributing excitations and the almost zero oscillator strengths. The CT state is stabilized upon twisting the DMA group and becomes the S_2 state at the TICT orientation. However, the lack of a polar environment limits the degree of CT stabilization, and the energetic order does not change even though the LE and CT states become almost degenerate. As the energy gap between the states becomes quite small, we wanted to be sure that further twisting of the DMA group does not induce a state crossing. Therefore, we added an extrapolation point with dihedral angles of 90°/−90° (only in the gas phase, last point on the corresponding curves), at which the energy gap between the LE and CT states was found to increase again (confirming that there is no crossing in the gas phase).

The DFT/MRCI level minimum-energy potential curves in water (Figure 8, top right panel) predict a conical intersection at the geometry where the LE and CT states have almost equal oscillator strengths (f = 0.19). To refine the estimate of the CIn geometry, we applied the LIDA-based procedure again (see above) defining six extra linear-interpolation points between the two geometries adjacent to the crossing. The best fit of the DFT/MRCI energies at these extra points provides a CIn geometry with dihedral angles of 70°/−100°, which resembles the TICT rather than the PICT orientation (see Figure S7, Supporting Information, for a comparison of the geometries). Even after the attempted refinement, the predicted CIn geometry is only a rough estimate, since the DFT/MRCI calculations give nondegenerate LE and CT state energies that differ by 0.4 eV. This is due to the fact that CAM-B3LYP geometries were used in the single-point DFT/MRCI calculations and that the linear-interpolation approach itself is just an approximation. Predicting more reliable CIn geometries would require proper DFT/MRCI-based geometry optimizations, which are impractical without analytic gradients (not yet available). We note, however, that the DFT/MRCI energy barriers for reaching the CT state surface (i.e., ICT) and for the transfer back to the LE surface are 0.44 and 0.29 eV, respectively, which agree well with those from CAM-B3LYP (0.41 and 0.14 eV). This confirms that one should expect a higher rate for back transfer (from CT to LE) than for ICT (from LE to CT).

In a less polar solvent like benzene (see Figure 8, bottom left panel), the CIn seam is located much closer to the TICT geometry and thus occurs at higher energy than in water. The computed MECP has dihedral angles of 83°/−89° and lies 3.27 eV above the GS minimum. The close structural resemblance to the TICT minimum (86°/−87°) can be ascribed to the smaller stabilization of the charge-separated CIn species by the less polar benzene solvent, as suggested previously.[10] Since the CT state energy is higher than in water, the LE/CT crossing can only be achieved by further twisting of the DMA group, which further minimizes the overlap between the π-systems of the donor and acceptor. The computed energy barriers for ICT and back-transfer are 0.60 and 0.22 eV, respectively, implying a diminished ICT rate and an enhanced back-transfer rate compared with water, which is in line with the experimental findings.[10]

To account for the effect of explicit intermolecular hydrogen bonding on the CIn seam, we again applied the linear-interpolation approach with the previously introduced micro-solvation model for water (*vide supra*). The minimum-energy LE and CT potential curves reveal a MECP with dihedral angles of 63°/−108° and an energy of 2.92 eV relative to the GS minimum; the LE and CT states have almost equal oscillator strength (f = 0.21) at the MECP. Comparing the

dx.doi.org/10.1021/jp506101x | J. Phys. Chem. B 2015, 119, 928−943

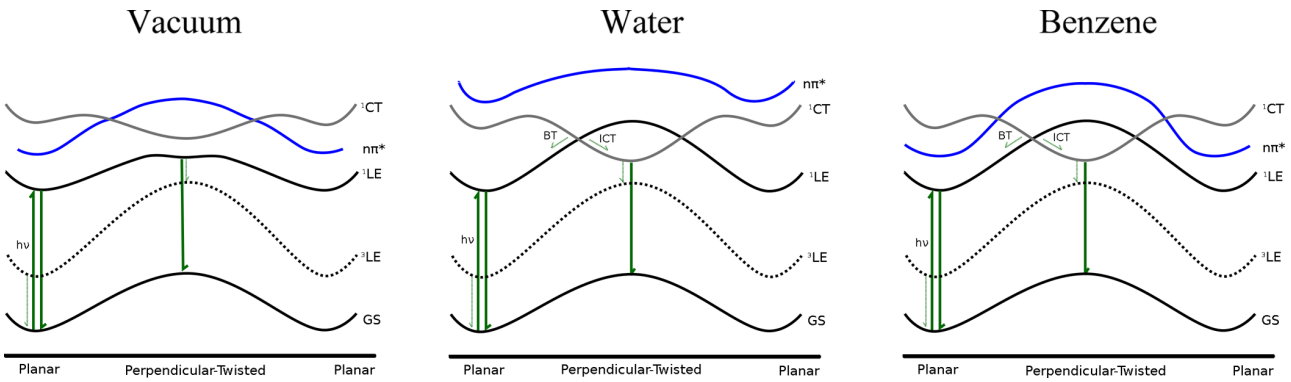


Figure 9. Overall photophysical behavior of roseolumiflavin in vacuum, water, and benzene based on CAM-B3LYP and DFT/MRCI calculations. BT and ICT denote back-transfer and intramolecular charge transfer.

energies of the MECP and the PICT and TICT minima yields estimates for the barriers to ICT and back-transfer of 0.38 and 0.27 eV, respectively. The inclusion of explicit water molecules in our microsolvation model thus notably affects the predicted MECP geometry and lowers both the computed MECP energy and the related barriers for the ICT and back-transfer processes. The higher stabilization of the charge-separated species at the CIn seam can be attributed to specific hydrogen bonding with the explicitly treated water molecules.

The minimum-energy surfaces of the two lowest triplet states were also tracked using the linear-interpolation approach for all model systems. The corresponding curves are depicted in Figure 8 as well (denoted as ^{3}LE and ^{3}CT, respectively, indicating the resemblance to their singlet counterparts), except for the gas phase where we considered it more important to include the nπ* and nπ*-2 states, in order to show their crossings with the CT state. The ^{3}LE (T_1) state always lies between the ground state and the S_1 state during the entire DMA twist. In water, it becomes almost degenerate with the singlet CT state close to the TICT minimum, whereas it remains far below the S_1 state in the gas phase and in benzene. The ^{3}CT (T_2) state stays close to the S_1 state in the gas phase, water, and benzene, with several crossings between these states during the course of DMA torsion (see Figure 8). On the other hand, the linear-interpolation approach does not reveal any crossing between the ground state and any of the singlet or triplet excited states in any of the systems considered.

Finally, we address the changes in the DFT/MRCI dipole moments of the ground state and the lowest singlet excited states during the DMA twist, as obtained from the linear-interpolation procedure (see Figure S8, Supporting Information). The ground-state dipole moments generally show a decrease upon DMA twisting, which is more pronounced in the more polar solvent (water). The S_1 dipole moments slightly decrease in the gas phase when going from the PICT to the TICT minimum, but they become much larger in water because of the gradual increase in the CT character of the S_1 state upon DMA twisting. Conversely, in benzene, the S_1 dipole moment first decreases gradually with the DMA twist until reaching the LE/CT crossing, after which it rises suddenly because of the sudden increase in the CT character of the S_1 state. Compared with water, the S_1 dipole moment starts to increase at a later stage of DMA twisting (i.e., only close to the TICT orientation), because the CT state is stabilized less efficiently in benzene. The S_2 dipole moments increase significantly upon DMA twisting in the gas phase but decrease considerably both in water and in benzene. These trends are again caused by changes in the character of the S_2 state. The torsion of the DMA group strengthens its CT character in the gas phase, whereas the opposite is true in water and benzene. In analogy to the findings for the S_1 state, the changes in the S_2 dipole moments are gradual in water and sudden in benzene.

4. CONCLUSIONS

Roseoflavin (RoF) is a derivative of native riboflavin (RF), in which the methyl group at the C8 position is replaced by a DMA group. This leads to different biological and photophysical properties of these two flavins. As opposed to RF, RoF is a push−pull (D−A) chromophore that undergoes a fast intramolecular charge transfer (ICT) in polar solvents. Consequently, a biexponential decay spectrum and fluorescence quenching are observed for RoF[10] but not for RF. Besides, RoF undergoes structural changes upon photoexcitation, which mainly involve the twisting of the donor DMA group with respect to the acceptor isoalloxazine moiety. Like for other D−A chromophores, the photodynamics of RoF was proposed to proceed through PICT,[35] TICT,[36] or WICT[37] mechanisms, leading to one of the corresponding zwitterionic forms (Figure 2) after ICT.

In this study, we investigated the photophysical properties of roseoflavin by quantum-chemical calculations on a simplified model system (roseolumiflavin, RoLF) in the gas phase, in water (using implicit and explicit solvent models), and in benzene, to cover environments of different polarity. We scanned the relevant potential surfaces of the LE and CT states with respect to the twist of the DMA group at the CAM-B3LYP and DFT/MRCI levels, in order to reveal possible crossings that enable the population of the CT state and consequently the ICT process. We also scrutinized the conical intersection seams and MECP geometries at these computational levels.

On the basis of the present results, we summarize the overall photophysical behavior of RoLF in the gas phase and in solution (water and benzene) in Figure 9. Photoexcitation via green-light absorption preferentially populates the LE state (S_1) and not the CT state, which has a low oscillator strength. The LE state quickly relaxes from the Franck−Condon region to the planar PICT minimum. The torsion of the DMA group is energetically feasible in the LE state. It leads to a destabilization of the LE state because of the diminishing π-overlap between the DMA and isoalloxazine moieties and to a stabilization of the CT state accompanied by an increase of its dipole moment (relative to the ground state). In solution (especially in water),

dx.doi.org/10.1021/jp506101x | J. Phys. Chem. B 2015, 119, 928−943

the polar CT state is stabilized by the environment more efficiently than the LE state, and it thus becomes lower in energy than the LE state at the perpendicular-twisted (TICT) orientation. In the gas phase, there is no such stabilization by the environment, and hence, the LE state always remains below the CT state. Accordingly, we could locate a conical intersection seam between the LE and CT states only in water and benzene. Our calculations thus provide a feasible ICT pathway via conical intersections only in water and benzene but not in the gas phase. These results are in agreement with the experimental observation of alleviated intramolecular charge transfer rates of roseoflavin in nonpolar solvents.[2,10,22,28,32]

The MECP geometries from CAM-B3LYP and DFT/MRCI are similar in benzene and water and more closely resemble the TICT rather than PICT geometries. This clearly suggests that RoLF utilizes the TICT mechanism to achieve the intramolecular charge transfer, in line with a proposal based on low-temperature fluorescence decay measurements for RoF.[27] The present calculations do not support a WICT mechanism, since such species are found to correspond to metastable zwitterions that interconnect PICT and TICT species. In line with experimental findings,[10] the DFT/MRCI energy barriers for the ICT and BT processes affirm diminished ICT and enhanced BT rates in benzene as compared to water.

The DFT/MRCI calculations on the lowest triplet states show that they also have LE and CT character and thus form counterparts to the singlet states. They are predicted to become almost degenerate with the S_1 state at the TICT geometry in the gas phase, water, and benzene, as previously proposed within the framework of the TICT mechanism.[55] This could enable population of the triplet states provided that the spin−orbit coupling (SOC) is sufficiently strong. Experimentally, the quantum yield of triplet formation in aqueous solution has indeed been found to be higher for RoF than for RF. SOC calculations as well as further experimental studies would be useful for further investigating this issue.

On the methodological side, we validated the performance of the CAM-B3LYP and DFT/MRCI methods for describing the charge-transfer characteristics of roseolumiflavin (and presumably of similar compounds). A microsolvation scheme with several explicit water molecules combined with a continuum model for bulk solvent was found to be most appropriate for computing the absorption spectrum and excited-state properties of roseolumiflavin in water.

■ ASSOCIATED CONTENT

Supporting Information

Assessment of the performance of B3LYP and CAM-B3LYP for excited-state structural properties; CAM-B3LYP excited-state energies and geometries (dihedral angles α/β and bond lengths); overlay of optimized geometries; frontier Kohn−Sham orbitals; 3D representation of the LE and CT surfaces with respect to the dihedral angles (α/β); dipole moments of the ground state and low-lying excited states; choice of microsolvation model; color maps of energy contour plots. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: thiel@kofo.mpg.de.

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Dr. M. R. Silva-Junior for his help in the initial stages of the project. B.K. would also like to acknowledge Dr. J. Götze and Dr. M. Patil for valuable discussions.

■ REFERENCES

(1) Ghisla, S.; Massey, V. New Flavins for Old: Artificial Flavins as Active Site Probes of Flavoproteins. *Biochem. J.* **1986**, *239*, 1−12.
(2) Mathes, T.; Vogl, C.; Stolz, J.; Hegemann, P. In Vivo Generation of Flavoproteins with Modified Cofactors. *J. Mol. Biol.* **2009**, *385*, 1511−1518.
(3) Marian, C. M.; Setsuko, N.; Rai-Constapel, V.; Karasulu, B.; Thiel, W. Photophysics of Flavin Derivatives Absorbing in the Blue-Green Region: Thioflavins as Potential Cofactors of Photoswitches. *J. Phys. Chem. B* **2014**, *118*, 1743−1753.
(4) Salzmann, S.; Martinez-Junza, V.; Zorn, B.; Braslavsky, S. E.; Mansurova, M.; Marian, C. M.; Gärtner, W. Photophysical Properties of Structurally and Electronically Modified Flavin Derivatives Determined by Spectroscopy and Theoretical Calculations. *J. Phys. Chem. A* **2009**, *113*, 9365−−9375.
(5) Salzmann, S.; Silva-Junior, M. R.; Thiel, W.; Marian, C. M.; M, C. Influence of the LOV Domain on Low-Lying Excited States of Flavin: A Combined Quantum-Mechanics/Molecular-Mechanics Investigation. *J. Phys. Chem. B* **2009**, *113*, 15610−15618.
(6) Sun, M.; Moore, T. A.; Song, P. S. Molecular Luminescence Studies of Flavines. I. Excited States of Flavines. *J. Am. Chem. Soc.* **1972**, *94*, 1730−1740.
(7) Bowd, A.; Byrom, P.; Hudson, J.; Turnbull, J. Excited States of Flavine Coenzymes-III. Fluorescence and Phosphorescence Emissions. *Photochem. Photobiol.* **1968**, *8*, 1−10.
(8) Kasai, S.; Kubo, Y.; Yamanaka, S.; Hirota, T.; Sato, H.; Tsuzukida, Y.; Matusi, K. Anti-Riboflavin Activity of 8N-Alkyl Analogues of Roseoflavin in Some Gram-Positive Bacteria. *J. Nutr. Sci. Vitaminol.* **1978**, *24*, 339−350.
(9) Pedrolli, D. B.; Jankowitsch, F.; Schwarz, J.; Langer, S.; Nakanishi, S.; Frei, E.; Mack, M. Riboflavin Analogs as Antiinfectives: Occurrence, Mode of Action, Metabolism and Resistance. *Curr. Pharm. Des.* **2013**, *19*, 2552−2560.
(10) Zirak, P.; Penzkofer, A.; Mathes, T.; Hegemann, P. Photo-Dynamics of Roseoflavin and Riboflavin in Aqueous and Organic Solvents. *Chem. Phys.* **2009**, *358*, 111−122.
(11) Shiga, K.; Nishina, Y.; Ohmine, I.; Horiike, K.; Kasai, S.; Matsui, K.; Watari, H.; Yamano, T. A Study of the Absorption, Circular Dichroism and Magnetic Circular Dichroism Spectra of a Flavin Derivative: The Π-Electronic Structure of 8-Amino-8-Demethyl-D-Riboflavin. *J. Biochem.* **1980**, *87*, 281−287.
(12) Song, P.-S.; Walker, E. B.; Vierstra, R. D.; Poff, K. L. Roseoflavin as a Blue Light Receptor Analog: Spectroscopic Characterization. *Photochem. Photobiol.* **1980**, *32*, 393−398.
(13) Choe, Y. Y.-K.; Nagase, S.; Nishimoto, K. Theoretical Study of the Electronic Spectra of Oxidized and Reduced States of Lumiflavin and Its Derivative. *J. Comput. Chem.* **2007**, *28*, 727−739.
(14) Klaumünzer, B.; Kröner, D.; Saalfrank, P. (TD-)DFT Calculation of Vibrational and Vibronic Spectra of Riboflavin in Solution. *J. Phys. Chem. B* **2010**, *114*, 10826−10834.
(15) Merz, T.; Sadeghian, K.; Schütz, M. Why BLUF Photoreceptors with Roseoflavin Cofactors Lose Their Biological Functionality. *Phys. Chem. Chem. Phys.* **2011**, *13*, 14775−14783.
(16) Sadeghian, K.; Bocola, M.; Schütz, M. A QM/MM Study on the Fast Photocycle of Blue Light Using Flavin Photoreceptors in Their

Light-Adapted/active Form. *Phys. Chem. Chem. Phys.* **2010**, *12*, 8840−8846.
(17) Sadeghian, K.; Schütz, M. On the Photophysics of Artificial Blue-Light Photoreceptors: An Ab Initio Study on a Flavin-Based Dye Dyad at the Level of Coupled-Cluster Response Theory. *J. Am. Chem. Soc.* **2007**, *129*, 4068−4074.
(18) Neiss, C.; Saalfrank, P.; Parac, M.; Grimme, S. Quantum Chemical Calculation of Excited States of Flavin-Related Molecules. *J. Phys. Chem. A* **2003**, *107*, 140−147.
(19) Salzmann, S.; Tatchen, J.; Marian, C. M. The Photophysics of Flavins: What Makes the Difference between Gas Phase and Aqueous Solution? *J. Photochem. Photobiol., A* **2008**, *198*, 221−231.
(20) Vdovin, A.; Slenczka, A.; Dick, B. Electronic Spectroscopy of Lumiflavin in Superfluid Helium Nanodroplets. *Chem. Phys.* **2013**, *422*, 195−203.
(21) Domratcheva, T.; Udvarhelyi, A.; Shahi, A. Computational Spectroscopy, Dynamics, and Photochemistry of Photosensory Flavoproteins. In *Flavins and Flavoproteins:Methods and Procedures*; Weber, S., Schleicher, E., Eds.; Springer: New York, 2014; pp 191−228.
(22) Tyagi, A.; Penzkofer, A.; Mathes, T.; Hegemann, P. Photophysical Characterisation and Photo-Cycle Dynamics of LOV1-His Domain of Phototropin from Chlamydomonas Reinhardtii with Roseoflavin Monophosphate Cofactor. *J. Photochem. Photobiol., B* **2010**, *101*, 76−88.
(23) Matsui, K.; Kasai, S. Roseoflavin, Nekoflavin, Schizoflavin. In *Chemistry and Biochemistry of Flavoenzymes*; Müller, F., Ed.; CRC Press: Boca Raton, FL, 1991; pp 105−120.
(24) Sikorska, E.; Khmelinskii, I. V.; Prukała, W.; Williams, L.; Patel, M.; Worrall, D. R.; Bourdelande, J. L.; Koput, J.; Sikorski, M. Spectroscopy and Photophysics of Lumiflavins and Lumichromes. *J. Phys. Chem. A* **2004**, 1501−1508.
(25) Tyagi, A.; Zirak, P.; Penzkofer, A.; Mathes, T.; Hegemann, P.; Mack, M.; Ghisla, S. Absorption and Emission Spectroscopic Characterisation of 8-Amino-Riboflavin. *Chem. Phys.* **2009**, *364*, 19−30.
(26) Drössler, P.; Holzer, W.; Penzkofer, A.; Hegemann, P. pH Dependence of the Absorption and Emission Behaviour of Riboflavin in Aqueous Solution. *Chem. Phys.* **2002**, *282*, 429−439.
(27) Zirak, P.; Penzkofer, A.; Mathes, T.; Hegemann, P. Absorption and Emission Spectroscopic Characterization of BLUF Protein Slr1694 from Synechocystis Sp. PCC6803 with Roseoflavin Cofactor. *J. Photochem. Photobiol., B* **2009**, *97*, 61−70.
(28) Tyagi, A.; Penzkofer, A.; Mathes, T.; Hegemann, P. Photo-Degradation Behaviour of Roseoflavin in Some Aqueous Solutions. *Chem. Phys.* **2010**, *369*, 27−36.
(29) Shinkai, S.; Kameoka, K.; Honda, N.; Ueda, K.; Manabe, O.; Lindsey, J. Spectral and Reactivity Studies of Roseoflavin Analogs: Correlation between Reactivity and Spectral Parameters. *Bioorg. Chem.* **1986**, *14*, 119−133.
(30) Kasai, S.; Miura, R.; Matsui, K. Chemical Structure and Some Properties of Roseoflavin. *Bull. Chem. Soc. Jpn.* **1975**, *48*, 2877−2880.
(31) Penzkofer, A. Photoluminescence Behavior of Riboflavin and Lumiflavin in Liquid Solutions and Solid Films. *Chem. Phys.* **2012**, *400*, 142−153.
(32) Shinkai, S.; Kameoka, K.; Honda, N.; Ueda, K.; Manabe, O. Reactivity Studies of Roseoflavin Analogues: A Correlation between Reactivity and Absorption Maxima. *J. Chem. Soc., Chem. Commun.* **1985**, 673−674.
(33) Shinkai, S. Remote Control of Flavin Reactivities by an Intramolecular Crown Ring Serving as a Metal Binding Site: Relationship between Spectral Properties and Dissociation of the 8-Sulfonamide Group. *Bioorg. Chem.* **1987**, *15*, 269−282.
(34) Grabowski, Z. R.; Rotkiewicz, K.; Rettig, W. Structural Changes Accompanying Intramolecular Electron Transfer: Focus on Twisted Intramolecular Charge-Transfer States and Structures. *Chem. Rev.* **2003**, *103*, 3899−4032.
(35) Zachariasse, K. Photo-Induced Intramolecular Charge Transfer and Internal Conversion in Molecules with a Small Energy Gap between S1 and S2. Dynamics and Structure. *J. Photochem. Photobiol., A* **1997**, *105*, 373−383.
(36) Rotkiewicz, K.; Grellmann, K.; Grabowski, Z. Reinterpretation of the Anomalous Fluorescense of Pn, N-Dimethylamino-Benzonitrile. *Chem. Phys. Lett.* **1973**, *19*, 315−318.
(37) Zachariasse, K. A.; von der Haar, T.; Hebecker, A.; Leinhos, U.; Kuhnle, W. Intramolecular Charge Transfer in Aminobenzonitriles: Requirements for Dual Fluorescence. *Pure Appl. Chem.* **1993**, *65*, 1745−1750.
(38) Furche, F. On the Density Matrix Based Approach to Time-Dependent Density Functional Response Theory. *J. Chem. Phys.* **2001**, *114*, 5982.
(39) Furche, F.; Ahlrichs, R. Adiabatic Time-Dependent Density Functional Methods for Excited State Properties. *J. Chem. Phys.* **2002**, *117*, 7433−7447.
(40) Grimme, S.; Waletzke, M. A Combination of Kohn−Sham Density Functional Theory and Multi-Reference Configuration Interaction Methods. *J. Chem. Phys.* **1999**, *111*, 5645−5655.
(41) Wolf, M. M. N.; Schumann, C.; Gross, R.; Domratcheva, T.; Diller, R. Ultrafast Infrared Spectroscopy of Riboflavin: Dynamics, Electronic Structure, and Vibrational Mode Analysis. *J. Phys. Chem. B* **2008**, *112*, 13424−13432.
(42) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; et al. *Gaussian 09*, revision A.02; Gaussian, Inc.: Wallingford, CT, 2009.
(43) Becke, A. Density-Functional Thermochemistry. III. The Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98*, 5648−5652.
(44) 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.
(45) Parr, R.; Yang, W. *Density-Functional Theory of Atoms and Molecules*; Oxford University Press: Oxford, U.K., 1994.
(46) Runge, E.; Gross, E. K. U. Density-functional Theory for Time-dependent Systems. *Phys. Rev. Lett.* **1984**, *52*, 997−1000.
(47) Dreuw, A.; Head-Gordon, M. Single-reference ab Initio Methods for the Calculation of Excited States of Large Molecules. *Chem. Rev.* **2005**, *105*, 4009−4037.
(48) Peach, M. J. G.; Benfield, P.; Helgaker, T.; Tozer, D. J. Excitation Energies in Density Functional Theory: An Evaluation and a Diagnostic Test. *J. Chem. Phys.* **2008**, *128*, 044118.
(49) Cossi, M.; Rega, N.; Scalmani, G.; Barone, V. Energies, Structures, and Electronic Properties of Molecules in Solution with the C-PCM Solvation Model. *J. Comput. Chem.* **2003**, *24*, 669−681.
(50) 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.
(51) Klaumünzer, B.; Kröner, D.; Lischka, H.; Saalfrank, P. Non-Adiabatic Excited State Dynamics of Riboflavin after Photoexcitation. *Phys. Chem. Chem. Phys.* **2012**, 8693−8702.
(52) Silva-Junior, M. R.; Schreiber, M.; Sauer, S. P. A.; Thiel, W. Benchmarks for Electronically Excited States: Time-Dependent Density Functional Theory and Density Functional Theory Based Multireference Configuration Interaction. *J. Chem. Phys.* **2008**, *129*, 104103.
(53) *TURBOMOLE V6.3 2011*, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989−2007, TURBOMOLE GmbH, since 2007; available from http://www.turbomole.com.
(54) Schäfer, A.; Klamt, A.; Sattel, D.; Lohrenz, J. C. W.; Eckert, F. COSMO Implementation in TURBOMOLE: Extension of an Efficient Quantum Chemical Code towards Liquid Systems. *Phys. Chem. Chem. Phys.* **2000**, *2*, 2187−2193.
(55) Grabowski, Z. R.; Rotkiewicz, K.; Siemiarczuk, A. Dual Fluorescence of Donor-Acceptor Molecules and the Twisted Intramolecular Charge Transfer (TICT) States. *J. Lumin.* **1979**, *18*, 420−424.
(56) Götze, J.; Karasulu, B.; Thiel, W. Computing UV/vis Spectra from the Adiabatic and Vertical Franck-Condon Schemes with the Use

of Cartesian and Internal Coordinates. *J. Chem. Phys.* **2013**, *139*, 234108.
(57) Kaduk, B.; Van Voorhis, T. Conical Intersections Using Constrained Density Functional Theory-Configuration Interaction. *J. Chem. Phys.* **2010**, *133*, 061102.

APPENDIX D

Photophysics of Flavin Derivatives Absorbing in the Blue-Green Region: Thioflavins as Potential Cofactors of Photoswitches

Christel M. Marian, Setsuko Nakagawa, Vidisha Rai-Constapel, Bora Karasulu, and Walter Thiel

J. Phys. Chem. B, **2014**, 118 (7), pp 1743-1753

THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

pubs.acs.org/JPCB

Photophysics of Flavin Derivatives Absorbing in the Blue-Green Region: Thioflavins As Potential Cofactors of Photoswitches

Christel M. Marian,*[,†] Setsuko Nakagawa,[‡] Vidisha Rai-Constapel,[§] Bora Karasulu,[∥] and Walter Thiel[∥]

[†]Institute of Theoretical and Computational Chemistry, Heinrich Heine University Düsseldorf, Universitätsstr. 1, D-40225 Düsseldorf, Germany

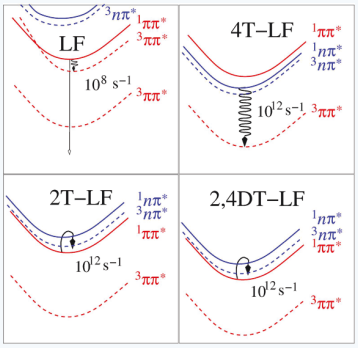
[‡]Department of Human Life and Environment, Kinjo Gakuin University, Nagoya 463 8521, Japan

[§]Institute of Theoretical and Computational Chemistry, Heinrich Heine University Düsseldorf, Universitätsstr. 1, D-40225 Düsseldorf, Germany

[∥]Max-Planck-Institut für Kohlenforschung, Kaiser-Wilhelm-Platz 1, 45470 Mülheim an der Ruhr, Germany

Supporting Information

ABSTRACT: The purpose of this study was to find flavin derivatives with absorption maxima in the blue-green region of the visible spectrum that might be used as alternative cofactors in blue-light photoreceptors. To this end, the vertical absorption spectra of eight lumiflavin-related compounds were calculated by means of quantum chemical methods. The compounds differ from lumiflavin by the subsitution of an S atom for an O atom at the 2- and/or 4-positions of the isoalloxazine core, the substitution of an N atom for a CH group in the 6- and/or 9-positions, or an extension of the π system at the 7- and 8-positions. For the three most promising compounds, 2-thio-lumiflavin, 4-thio-lumiflavin, and 2,4-dithio-lumiflavin, the quantum chemical investigations were extended to include geometry relaxations in the excited states, rates for spin-forbidden transitions and an estimate of spectral shifts brought about by polar protic environments. We find these thiocarbonyl compounds to have very promising excited-state properties. They absorb in the blue-green wavelength regime around 500 nm, i.e., substantially red-shifted with respect to lumiflavin that is the cofactor of natural blue-light photoreceptors. Their triplet quantum yields are predicted to be close to unity while their triplet lifetimes are long enough to enable bimolecular photochemical reactions. The combination of these properties makes the thioflavins potentially suitable candidates as cofactors in biomimetic photoswitches.



INTRODUCTION

Recent years have seen a boost of studies on flavin photochemistry and physics especially due to their decisive role in blue light-mediated signal transduction in bacteria, archaea, plants and fungi.[1−5] Light, oxygen, and voltage-sensitive (LOV) domains are protein photosensors that bind a flavin mononucleotide (FMN) cofactor (Figure 1) non-covalently. Upon absorption of blue light, the LOV domain undergoes a photocycle. The primary step after light absorption hereby involves a rapid decay of the excited singlet state to the lowest excited triplet state via an intersystem crossing (ISC) mechanism. In a second step a metastable covalent adduct of a nearby cysteine residue and the isoalloxazine framework of the chromophore is generated resulting in a structural signal that is transduced to an effector domain.

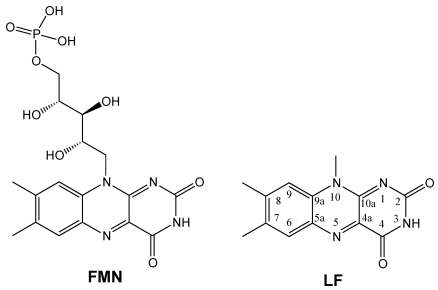


Figure 1. Chemical structures of flavin mononucleotide (FMN) and lumiflavin (LF) with IUPAC atomic labels.

LOV domains can be fused to cellular effector proteins to create artificial light-sensitive variants enabling photochemical regulation of protein function.[5−8] A genetically encoded photoactivatable Rac fused to a LOV domain from phototropin has even been used to control the motility of living cells.[9] In some applications, however, there is a need for red-shifted photochemical tools with enhanced light-sensitivity.[10] It could thus be interesting to replace the cofactor of natural blue-light receptors by a flavin derivative with reduced photoexcitation energy (thereby minimizing the radiation damage of the surrounding biological material) while keeping the photocycle intact.

Received: October 2, 2013
Revised: January 29, 2014
Published: January 30, 2014

Riboflavin-5′-phosphate, also known as FMN (Figure 1), can be divided into two parts: the isoalloxazine (benzol[*g*]pteridine-2,4(3H,10H)-dione) core ring that is the blue-light sensitive part and the ribophosphyl chain that is required for the signal transduction. In lumiflavin (LF) the ribophosphyl chain of FMN in 10-position has been replaced for a methyl group. Since the photophysical and photochemical behavior of flavins is dominated by the isoalloxazine core, experimental absorption and emission spectra of FMN and LF are very similar. In contrast, electronic modification of the isoalloxazin core, such as the substitution of a CH group for the nitrogen atom in 1- or 5-positions, does not only influence the absorption characteristics, but has also strong impact on the triplet formation and the photocycle in LOV domains.[11−14]

In this work, the photophysics of eight LF-related compounds (Figure 2) has been investigated. The idea was

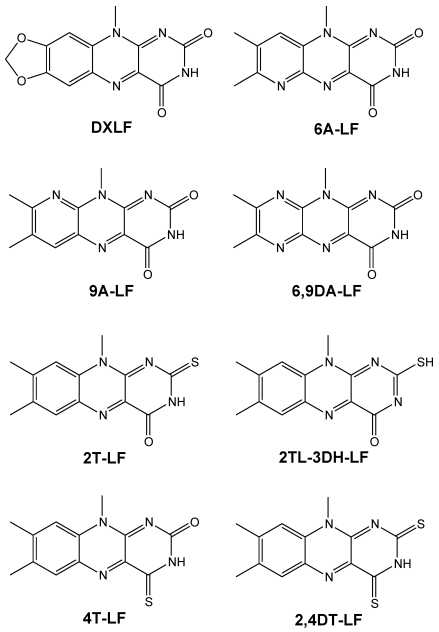


Figure 2. Chemical structures of flavin derivatives studied in this work.

to modify the π system by either replacing species in the isoalloxazine core by isovalent atoms or by exchanging substituents for more extended ones. In doing so, electrostatic properties and spatial limitations in the LOV pocket have been taken into consideration. The methyl groups attached to the 7- and 8-positions of the native LF can, for instance, be replaced by a trifluoromethyl, mercapto, methoxy, or hydroxy group but not by very bulky substituents.[15,16] At the hydrophilic side of the LOV pocket, the pteridine moiety is surrounded by several hydrogen bond donors and acceptors. It may thus be assumed that the hydrogen bond network plays an essential role in the photocycle. Here, we report calculations on a derivate in which a −O−CH_2-O- bridge has been attached to the carbon atoms in 7- and 8-positions, respectively, thus forming a dioxolane ring (DXLF). Furthermore, we have substituted CH groups for N atoms in the benzene moiety of the isoalloxazine core (6A-LF, 9A-LF, and 6,9DA-LF). With regard to the pteridine moiety, we investigate the photophysical effects occurring when the carbonyl groups in 2- and/or 4-positions are replaced by thiocarbonyl groups (2T-LF, 4T-LF, and 2,4DT-LF). In addition, an iminothiol tautomer of 2T-LF, 2TL-3DH-LF, has been studied.

The most promising compounds, i.e., the thiolumiflavins, have been investigated in more detail. The first absorption bands of 2-thio-lumiflavin (2T-LF) and 4-thio-lumiflavin (4T-LF) are known to be red-shifted with respect to LF.[17] In contrast, the location of the dark states ($n\pi^*$ states, triplet $\pi\pi^*$ states) that are crucial for the photophysics and photochemistry of the compounds are widely unknown. Starting with the isolated chromophores, the impact of solvent effects on the excited-state properties is evaluated. ISC and phosphorescence rate constants are determined to judge both the accessibility and the lifetime of the T_1 state that is supposed to be the photochemically active state in LOV domains of biological blue-light photoreceptors. Our results will be compared to corresponding data for LF that has been the subject of recent studies in our group.[11,18,19]

METHODS AND COMPUTATIONAL DETAILS

Geometry optimizations of the ground state and the lowest triplet states were carried out at the level of (unrestricted) density functional theory ((U)DFT) using the Turbomole 6.3[20] program package. For all calculations, we employed the standard TZVP basis set from the Turbomole library.[21] The B3LYP functional[22,23] was used for optimizing the molecular geometries. Unless otherwise noted C_s symmetry constraints were imposed on the ground and excited state geometries of the isolated chromophores. Electronically excited singlet and triplet state geometries were determined at the level of time-dependent DFT[24] (TDDFT).[25] To ensure that the resulting geometries correspond to true minima of the potential energy hypersurface (PEH), harmonic vibrational frequencies were calculated numerically with the program SNF.[26]

Vertical electronic excitation energies and dipole moments were obtained from subsequent single-point calculations using the combined density functional theory/multireference configuration interaction (DFT/MRCI) method of Grimme and Waletzke.[27] In this method, the MRCI expansion is built up in a one-particle basis of Kohn−Sham orbitals employing the BHLYP functional[28,29] and matrix elements are scaled using empirically determined parameters. Typically, this method yields excitation energies with errors below 0.2 eV.[27,30] Technical details of the DFT/MRCI calculations were chosen to be identical to those in the original work of Salzmann et al. on LF.[18]

ISC rate constants have been determined perturbationally, making use of the Fermi golden rule ansatz. In this ansatz, the decay rate of an initially populated Ψ_{S_a},v_{aj} vibronic state via ISC to a quasicontinuum of final vibronic states $\Psi^{\alpha}_{T_b},\{v_{bk}\}$ is given by

$$k_{\mathrm{ISC}} = \frac{2\pi}{\hbar}\sum_{\alpha}\sum_{k}|\langle\Psi_{S_a}, v_{aj}|\hat{\mathcal{H}}_{\mathrm{SO}}|\Psi^{\alpha}_{T_b}, v_{bk}\rangle|^2\delta(E_{aj} - E_{bk}) \quad (1)$$

Typically, eq 1 is expanded into a Taylor series around an appropriately chosen reference point q_0, e.g., the equilibrium geometry of the initial state. Furthermore, the Condon approximation

$$k^{\mathrm{FC}}_{\mathrm{ISC}} \approx \frac{2\pi}{\hbar}\sum_{\alpha}|\langle\Psi_{S_a}|\hat{\mathcal{H}}_{\mathrm{SO}}|\Psi^{\alpha}_{T_b}\rangle|^2_{q_0}\sum_{k}|\langle v_{aj}|v_{bk}\rangle|^2\delta(E_{aj} - E_{bk}) \quad (2)$$

dx.doi.org/10.1021/jp4098233 | J. Phys. Chem. B 2014, 118, 1743−1753

has been employed, i.e., it is assumed that the coupling between the initial and the final states can be separated into an electronic and a vibrational factor. This assumption is usually fulfilled in good approximation for El-Sayed allowed ISC, e.g., $^1n\pi^* \rightsquigarrow {}^3\pi\pi^*$ and $^1\pi\pi^* \rightsquigarrow {}^3n\pi^*$. Electronic spin−orbit matrix elements (SOMEs) between the DFT/MRCI wave functions in eq 2 have been computed using the spin−orbit coupling kit (Spock).[31,32] For reasons of efficiency, the one-center mean-field approximation to the Breit−Pauli Hamiltonian has been used for the description of the spin−orbit coupling.[33,34] It has been shown that the accuracy of this approximation lies within better than 5% of the full treatment.[35,36] The vibrational contributions to the rate in eq 2 have been determined using a time-dependent approach.[37] Briefly, a Fourier transform representation of delta function appearing in eq 2 is employed.

$$\delta(E_{aj} - E_{bk}) = \int_{-\infty}^{+\infty} e^{it(E_{aj}-E_{bk})}\, dt \tag{3}$$

In the harmonic oscillator model and making use of the Condon approximation, the ISC rate can then be determined by numerical integration of

$$k_{ISC}^{corr} = |\langle\Psi_S|\hat{\mathcal{H}}_{SO}|\Psi_T\rangle|^2 \int_{-\infty}^{\infty} dt G\,(t) e^{it(\Delta E_{ST}^0 + \frac{1}{2}Tr\Omega_S)} \tag{4}$$

where $G(t)$ is a generating function

$$G(t) = 2^{N/2}\sqrt{\frac{\det(S^{-1}\Omega_S\Omega_T)}{\det(J^\dagger\Omega_T BJ + \Omega_S)\det(J^\dagger\Omega_T B^{-1}J + \Omega_S)}}$$
$$\exp(D^\dagger(\Omega_T BJ(J^\dagger\Omega_T BJ + \Omega_S)^{-1}J^\dagger\Omega_T B - \Omega_T B)D)$$

that contains diagonal matrices Ω_S, Ω_T, S, and B with elements $(\Omega_S)_{ii} = \omega_{Si}$, $(\Omega_T)_{ii} = \omega_{Ti}$, $S_{ii} = \sinh(i\omega_{Ti}t)$, $B_{ii} = \tanh(i\omega_{Ti}t/2)$. Herein, ω_{Si} and ω_{Ti} represent the harmonic vibrational frequencies of the initial and final states, respectively. The mass-weighted normal modes of the final state are related to their counterparts of the initial state by the Duschinsky transformation[38] $Q_T = JQ_S + D$ where J is the Duschinsky rotation matrix and D the displacement vector. The Fourier transform of a Gaussian function of width 10 cm^{-1} was used to damp the oscillations of the time-correlation function. For its numerical integration, we have chosen an interval of 12 ps and 2^{18} grid points. Despite the inherent approximations and the incertitude of the terms entering eq 2, we consider the order of magnitude of the calculated rate constants to be reliable. For further details on the theory and the methods for calculating spin−orbit coupling and ISC see, for instance, a recent review article.[39]

For second-order spin-forbidden properties, such as phosphorescence rates, the convergence of sum-over-state perturbation expressions is known to be very slow.[40] For this reason, we use the DFT/MRSOCI, a multireference spin−orbit configuration interaction program,[41] to compute spin−orbit coupled DFT/MRCI wave functions for this purpose. Since we are interested only in the $T_1 \rightarrow S_0$ radiative transition, it is sufficient to determine the four lowest roots of the MRSOCI matrix. To avoid numerical inaccuracies, a tight convergence threshold of 10^{-8} E_h was imposed in the Davidson procedure of the DFT/MRSOCI step.

Dipole transition matrix elements and oscillator strengths were evaluated in their length forms at the DFT/MRCI (fluorescence) or DFT/MRSOCI (phosphorescence) levels.

To estimate spectral shifts due to electrostatic interaction in polar solvents we employed the conductor-like screening model (COSMO) which is implemented in the TURBOMOLE package.[42,43] Relative permittivities of ε_r = 78 (water), ε_r = 33 (acetonitrile), and ε_r = 4 were used. A value of ε_r = 4 corresponds to a typical protein environment[44,45] and is also close to the permittivity of chloroform $\varepsilon_r \approx 4.8$ at room temperature. When COSMO was applied, the MRCI expansion was built up from the one-particle basis of COSMO optimized Kohn−Sham orbitals. Because of technical reasons, C_1 symmetry had to be used for all DFT/MRCI calculations involving COSMO. For both, singlet and triplet multiplicity, 20 roots were computed. Since COSMO cannot properly model hydrogen bonding, the effects of hydrogen bonding in aqueous solution were mimicked by microhydration. For this purpose, we placed four water molecules next to the hetero atoms of the isoalloxazine ring and optimized the ground state without symmetry constraints. In a further approach, the two previous solvation models were combined.

RESULTS AND DISCUSSION

Vertical excitation spectra in vacuum of all compounds are collected in Tables S1 and S2 of the Supporting Information (SI). It is seen that the aza substitution in the benzene moiety of the isoalloxazine ring does not alter the essential absorption characteristics with respect to LF. In all three cases, the first bright transition is found at 2.94−2.97 eV, i.e., these derivates absorb at much too short wavelengths (417−422 nm) to be interesting in the present context. Even a red shift of 0.1−0.15 eV due to solvent−solute interaction in water, as may be expected from the native compound, will not place the absorption in these molecules to the desired wavelength regime. However, 6A-LF, 9A-LF, and 6,9DA-LF might be interesting from a different point of view: Their triplet $n\pi^*$ states are located at lower energy than in LF which in turn could help increasing the ISC rates in solution. For a detailed discussion of solvent effects on the intermediate triplet formation in flavins the reader is referred to former work.[18]

The absorption wavelength of the first band of DXLF is shifted into the desired direction. In the gas phase, we obtain a vertical excitation energy of 2.84 eV (436 nm) for the first $^1(\pi \rightarrow \pi^*)$ transition. Adiabatically the excitation energy is lowered to 2.60 eV, i.e., we expect the band origin to lie at 477 nm. Solvent effects stabilize the upper state slightly so that the band origin (488 nm) could fit into the required wavelength window. In this compound, we encounter a different problem, however. Its first triplet $n\pi^*$ state that is essential for a fast ISC mechanism employing either direct or vibronic spin−orbit coupling is energetically not easily accessible, neither in the gas phase nor in solution. The iminothiol tautomer of 2T-LF, 2TL-3DH-LF, exhibits a seemingly ideal first absorption band with origin at 511 nm (2.43 eV) and maximum at 462 nm (2.68 eV) in vacuum. These values change only marginally upon solvation in water, in very good agreement with experimental values measured for the corresponding methylthiolether.[17] However, Dudley et al. report this 2-methylthiolether (and the corresponding 4-methylthiolether) to be chemically instable. According to these authors, they are easily protonated in 1-position with subsequent fast hydrolysis to yield LF.[17]

On these grounds, only the spectra and the photophysics of the thionyl compounds are presented and discussed in more detail in the following. For comparison, results of native LF have been taken from the work of Salzmann et al.[11,18]

dx.doi.org/10.1021/jp4098233 | J. Phys. Chem. B 2014, 118, 1743−1753

2-Thio-lumiflavin. *Vacuum.* Geometry optimization of the electronic ground state yields a planar nuclear arrangement of the ring system. Bond lengths of the ground and excited state structures are shown in Table S3 of the SI. Besides the obvious fact that the C_2═S bond is considerably longer than the corresponding C_2═O bond (167 vs 121 pm), only slight changes of the geometry parameters are observed when comparing 2T-LF with LF. As a result of the weaker double bond character of the C_2═S bond, the neighboring C_2–N_1 and C_2–N_3 bonds slightly contract with respect to LF. All other geometry parameters are nearly unchanged.

With regard to electronic structure, the introduction of the thionyl group in 2-position is found to have a pronounced effect. At the molecular orbital (MO) level, the highest occupied molecular orbital (HOMO) is a π orbital (π_H) that exhibits a large amplitude on the sulfur center (Figure 3)

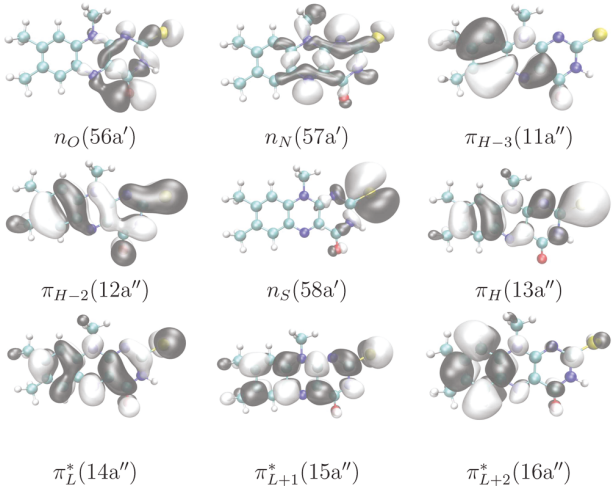


Figure 3. Frontier orbitals of 2T-LF at the optimized ground-state geometry. (Isovalue = 0.02.)

whereas the π system in the HOMO of LF is delocalized over the whole ring system (Figure 4 in ref 11). The corresponding π-MO of 2T-LF is only the third-highest occupied MO (π_{H-2}). The second-highest occupied MO of 2T-LF is mainly comprised of the nonbonding in-plane 3p orbital of sulfur (n_S). Other occupied MOs that are important for the characterization of the low-lying electronic states are the nonbonding in-plane orbital on oxygen (n_O) and a σ*-type orbital with large contributions of in-plane p orbitals on the N_1 and N_5 atoms which we denote by n_N. The three lowest unoccupied molecuar orbitals (LUMOs) are similar in shape and energetic order in 2T-LF and LF. While the LUMO (π^*_L) as well as the next higher MO (π^*_{L+1}) are delocalized over the ring system, π^*_{L+2} concentrates in the benzene ring. It is thus the antibonding counterpart of the π_{H-3} orbital.

The different shapes and energetic orders of the MOs in 2T-LF and LF are reflected in the electronic spectra. In Table 1, vertical absorption energies of 2T-LF are shown. In the singlet manifold, the lowest excited state corresponds to a $^1(n_S \rightarrow \pi^*_L)$ excitation with negligible oscillator strength. The strong first absorption (oscillator strength $f(r) \approx 0.51$) is associated with the second excited singlet state $^1(\pi_H \rightarrow \pi^*_L)$. In both cases, the transition is accompanied by an electron transfer from the sulfur region to the delocalized π system. Thus, the electric dipole moments of the two lowest-lying electronically excited

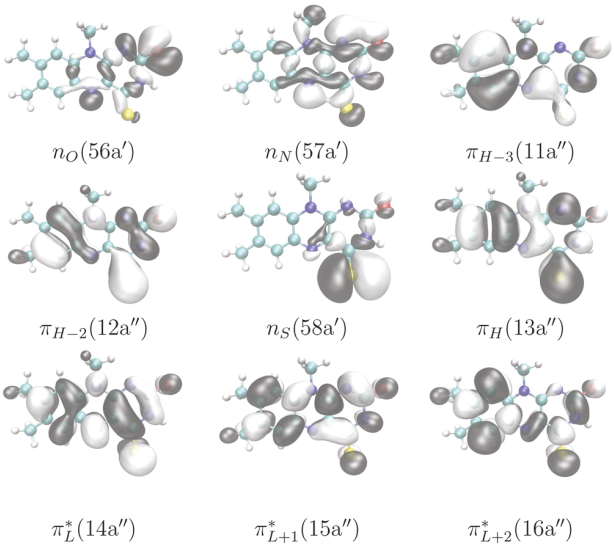


Figure 4. Frontier orbitals of 4T-LF at the optimized ground-state geometry. (Isovalue = 0.02.)

Table 1. Vertical Excitation Spectra of 2T-LF in the Gas Phase and in Water Solution

		gas phase		water solution		
state	electronic structure	ΔE [eV]	$f(r)$	ΔE [eV]	λ_{max} [a]	$f(r)$
$1^1A'$	ground state	0.00		0.00		
$1^1A''$	$n_S \rightarrow \pi^*_L$	1.80	0.0000	2.55		0.0018
$2^1A'$	$\pi_H \rightarrow \pi^*_L$	2.35	0.5056	2.50	2.52 eV (492 nm)	0.5265
$2^1A''$	$n_N \rightarrow \pi^*_L$	3.22	0.0017	3.49		0.0093
$3^1A'$	$\pi_{H-2} \rightarrow \pi^*_L$	3.26	0.0890	3.23	3.18 eV (390 nm)	0.1724
$1^3A'$	$\pi_H \rightarrow \pi^*_L$	1.68		1.91		
$1^3A''$	$n_S \rightarrow \pi^*_L$	1.74		2.49		
$2^3A'$	$\pi_{H-2} \rightarrow \pi^*_L$	2.71		2.64		
$3^3A'$	$\pi_{H-3} \rightarrow \pi^*_L$	2.92		2.95		

[a] Absorption maxima in 0.1 N phosphate buffer (pH 7), ref 17.

states are smaller than that of the ground state (see Table S4 of the SI for details), with consequences for the differential stabilization by polar solvents (see below). Excitations from the n_N MO are located in a similar energy region as in LF. A medium strong $^1(\pi_{H-2} \rightarrow \pi^*_L)$ transition and a stong transition of mixed $^1(\pi \rightarrow \pi^*)$ origin are found in the UVA region of the absorption spectrum. In the triplet manifold the order of states is reversed with respect to the singlets. Because of the larger exchange interaction between π and π^* MOs as compared to $n\pi^*$ exchange interactions, the $^3(\pi_H \rightarrow \pi^*_L)$ state constitutes the T_1 ($1^3A'$) state but its energy separation to the $^3(n_S \rightarrow \pi^*_L)$ is small. Both triplet states are located energetically below the S_1 ($1^1A''$) state in the vertical excitation spectrum and are thus expected to play an important role in its decay dynamics.

Upon geometry relaxation, the S_1, S_2 ($2^1A'$), T_1 and T_2 ($1^3A''$) states are stabilized by about 0.3–0.4 eV while the order of states is preserved (Table 2). Their PEHs have true minima at C_s symmetric structures and their geometry parameters are found to be very similar (Table S3 of the SI). This is a consequence of the fact that both the n_S and π_H orbitals from which the excitations originate are dominated by nonbonding

dx.doi.org/10.1021/jp4098233 | J. Phys. Chem. B 2014, 118, 1743–1753

Table 2. Adiabatic DFT/MRCI Excitation Energies [eV] of 2T-LF in the Gas Phase and in Water Solution

state	electronic structure	gas phase	water solution[a]
$1^1A''$	$n_S \rightarrow \pi_L^*$	1.44	2.19
$2^1A'$	$\pi_H \rightarrow \pi_L^*$	1.95	2.10
$1^3A'$	$\pi_H \rightarrow \pi_L^*$	1.39	1.62
$1^3A''$	$n_S \rightarrow \pi_L^*$	1.40	2.15

[a]Values were obtained by adding the solvent shift at the S_0 geometry to the adiabatic excitation energy of the isolated molecule.

sulfur in-plane and out-of-plane 3p orbitals, respectively. The only substantial differences between both minimum nuclear arrangements are found for the N_5-C_{5a}, $C_{5a}-C_{9a}$, and C_7-C_8 bonds where the π_H MO exhibits some electron density whereas the n_S orbital does not. Adiabatic excitation energies are presented in Table 2, information on the energies of further states at the excited-state minima can be found in Table S4 of the SI. It appears that the geometry optimization at the TDDFT level employing the B3-LYP functional slightly overshoots for the singlet excited states since the lowest DFT/MRCI energy of the S_2 state is found at the TDDFT-optimized S_1 minimum and not at the S_2 minimum. In the corresponding triplet states, T_1 and T_2, the differences of the geometry parameters are less pronounced. Major changes with respect to the ground-state minimum geometry involve the C_2-S bond which is elongated by about 8 pm and the neighboring C_2-N_1 and C_2-N_3 bonds which are shortened by 3−5 pm. As may be expected from the bonding and antibonding characteristics of the π_L^* MO, the C_4-C_{4a} and $C_{4a}-C_{10a}$ bonds contract in the excited states whereas the neighboring $C_{4a}-N_5$ bond is elongated.

To summarize, in contrast to LF, the $^1(\pi_H \rightarrow \pi_L^*)$ state is not the lowest-lying excited singlet state of 2T-LF in the gas phase. We find the dark $^1(n_S \rightarrow \pi_L^*)$ state approximately 0.5 eV below. Fluorescence will thus be very weak or absent in this compound in the gas phase or apolar solvents. Due to the larger exchange splitting in $\pi\pi^*$ states as compared to $n\pi^*$ states, the $^3(\pi_H \rightarrow \pi_L^*)$ and $^3(n_S \rightarrow \pi_L^*)$ minima are nearly degenerate.

Solvent Effects. The influence of the solvent on the minimum geometries is found to be rather small. As may be seen from Table S3 of the SI most of the bond lengths experience changes below 1 pm when the states are reoptimized for a cluster containing four explict water molecules. Due to hydrogen bond formation, the $C_4{=}O$ bond is elongated by 1.6 pm in the singlet states and by 1.8 pm in the T_1 state. The minimum of the T_2 state of the water cluster could not be optimized at the TDDFT/B3-LYP level. The increase of the $C_4{=}O$ bond lengths goes along with a shortening of the neighboring N_3-C_4 and C_4-C_{4a} bonds.

The lowest-lying absorption band of natural flavins is known to be nearly unaffected by the polarity or the proticity of the solvent.[46] In contrast, the corresponding band of 2T-LF exhibits negative solvatochromism. The reason for the different solvatochromism of the lowest $^1(\pi \rightarrow \pi^*)$ excitation in 2T-LF and LF is the partial charge transfer from $3p_\pi$ at the sulfur center to the ring system in 2T-LF which is missing in the corresponding natural flavin. For example, the absorption maximum of 2-thio-tetra-O-acetylriboflavin shifts from 510 nm in $CHCl_3$ to 496 nm in aqueous solution at pH 7.[17] This trend is corroborated by the results of our calculations. In the microsolvated cluster, the first bright $^1(\pi_H \rightarrow \pi_L^*)$ transition experiences a slight bathochromic shift (see Table S5 of the SI for details). However, as soon as the electrostatic interaction with the polarizable continuum is switched on, a hypsochromic shift sets in. In our best water model with four explicit water molecules surrounded by a polarizable continuum of relative permittivity $\varepsilon_r = 78$, the blue shift of the absorption band with respect to the gas phase amounts to about 0.15 eV (Table 1). Despite the negative solvatochromism, the $^1(\pi_H \rightarrow \pi_L^*)$ state is the lowest-lying singlet state in the Franck−Condon (FC) spectrum of 2T-LF in aqueous solution. Its computed vertical excitation energy (2.50 eV, 495 nm) agrees excellently with the maximum of the first absorption band in 0.1 molar phosphate buffer at pH 7 (492 nm).[17]

The huge hypsochromic shift of the $^1(n_S \rightarrow \pi_L^*)$ state (ca. 0.75 eV) is caused primarily by the interaction of the highly polarizable sulfur lone pair with the polar surrounding (ε_r = 78). The proticity of the solvent and the concomitant stabilization of the n_O orbital due to hydrogen bonding, which was found to be substantial in LF, is less important for the photophysics of 2T-LF. With respect to its electronic structure, the S_1 state of LF has its correspondence in the S_3 state (S_4 in vacuum). Again, we observe excellent agreement of its computed vertical excitation energy (3.23 eV, 383 nm) with the maximum of the second absorption band at 390 nm.[17] This is true even for the strong absorption band with a maximum at 316 nm. According to our calculations, this band originates from the $S_0 \rightarrow S_7$ transition in aqueous solution (3.98 eV, 311 nm) which is dominated by the $^1(\pi_H \rightarrow \pi_{L+1}^*)$ excitation.

The binding pocket of the LOV domain provides a similar hydrogen bonding network as the four water molecules in our model but a significantly less polar environment than bulk water.[19] The 2T-LF·$4H_2O$ cluster surrounded by a polarizable continuum of relative permittivity $\varepsilon_r = 4$ may therefore serve as a realistic model for 2T-LF in the LOV domain. In this environment, the lowest excited singlet state of 2T-LF is predicted to originate from the $^1(n_S \rightarrow \pi_L^*)$ excitation (see Table S5 of the SI for details).

Triplet Formation. Inspection of Table 2 shows that, from an energetic point of view, the T_1 and T_2 states are possible candidates for efficient ISC from S_1 in the gas phase and in solution. Since the S_1 state is of $^1(n_S \rightarrow \pi_L^*)$ type in the gas phase and of $^1(\pi_H \rightarrow \pi_L^*)$ type in solution, SOMEs are given for all relevant ISC combinations in Table 3. From the qualitative El-Sayed rules,[47] electronic spin−orbit coupling is expected to be strong only when a change of orbital angular momentum

Table 3. 2T-LF: Calculated Spin−Orbit Matrix Elements $|\langle i|\hat{H}_{SO}|f\rangle|$ [cm^{-1}] Evaluated at the Minimum of the Initial Singlet State, Adiabatic Electronic Energy Differences ΔE^{ad} [cm^{-1}], and ISC Rate Constants k_{ISC} [s^{-1}] in the Gas Phase and Water Solution (See Text)

channel		SOME	rate
$i \rightsquigarrow f$	ΔE^{ad}	$\lvert\langle i\lvert\hat{H}_{SO}\rvert f\rangle\rvert$	k_{ISC}
$1(n_S \rightarrow \pi_L^*) \rightsquigarrow {}^3(\pi_H \rightarrow \pi_L^*)_x$		160.5	
$1(n_S \rightarrow \pi_L^*) \rightsquigarrow {}^3(\pi_H \rightarrow \pi_L^*)_y$		4.4	
$1(n_S \rightarrow \pi_L^*) \rightsquigarrow {}^3(\pi_H \rightarrow \pi_L^*)_z$		0.0	
$1(\pi_H \rightarrow \pi_L^*) \rightsquigarrow {}^3(n_S \rightarrow \pi_L^*)_x$		135.8	
$1(\pi_H \rightarrow \pi_L^*) \rightsquigarrow {}^3(n_S \rightarrow \pi_L^*)_y$		2.3	
$1(\pi_H \rightarrow \pi_L^*) \rightsquigarrow {}^3(n_S \rightarrow \pi_L^*)_z$		0.0	
$\{S_1 \rightsquigarrow T_1\}_{vac}$	409		$\sim 5 \times 10^{12}$
$\{S_1 \rightsquigarrow T_2\}_{wat}$	−433		$\sim 7 \times 10^{11}$[a]

[a]At 298 K.

dx.doi.org/10.1021/jp4098233 | J. Phys. Chem. B 2014, 118, 1743−1753

takes place. For this reason, direct ISC is expected to be fast for $^1(n_S \rightarrow \pi_L^*) \rightsquigarrow {}^3(\pi_H \rightarrow \pi_L^*)$ and for $^1(\pi_H \rightarrow \pi_L^*) \rightsquigarrow {}^3(n_S \rightarrow \pi_L^*)$. In these states, a substantial part of the unpaired spin-density is located at the sulfur center. Hence, SOMEs are considerably larger than for the parent compound where the spin density in the S_1 $^1(\pi_H \rightarrow \pi_L^*)$ state is delocalized.

As noted above, the geometry parameters of the low-lying excited-state minima are quite similar. This means that we expect the rate constant to obey the energy gap law for the weak coupling case,[48] i.e., to decrease exponentially with increasing energy separation. When computing these rates, we have assumed that the electronic SOMEs and the shapes of the PEHs do not change substantially by solvent effects. The energy gap between the S_1 and T_1 minima is small in the vacuum (below 500 cm^{-1}) and increases from ~3500 cm^{-1} in a protic, but only slightly polar medium to ~4600 cm^{-1} in water. Accordingly, the fast $^1(n_S \rightarrow \pi_L^*) \rightsquigarrow {}^3(\pi_H \rightarrow \pi_L^*)$ ISC in vacuum proceeding at about 5×10^{12} s^{-1} slows down by more than 2 orders of magnitude to about 1×10^{10} s^{-1}. In aqueous solution, this process will presumably play a minor role because the $^1(n_S \rightarrow \pi_L^*)$ is only the second excited singlet in that medium.

In contrast, the energy gap between the S_2 and T_2 states shrinks in polar solvents. We obtain a rate constant of 5×10^{11} s^{-1} for the $^1(\pi_H \rightarrow \pi_L^*) \rightsquigarrow {}^3(n_S \rightarrow \pi_L^*)$ ISC in vacuum, but internal conversion (IC) to the $^1(n_S \rightarrow \pi_L^*)$ state followed by ISC to the $^3(\pi_H \rightarrow \pi_L^*)$ will be the prevailing decay channel after photoexcitation in the gas phase. In water solution, we obtain a small negative (~−400 cm^{-1}) energy gap for minima of the $^1(\pi_H \rightarrow \pi_L^*)$ and $^3(n_S \rightarrow \pi_L^*)$ pair of states in water. Since the zero-point vibrational energies are similar in both potential wells, the ISC would thus be forbidden at 0 K. At room temperature and assuming a Boltzmann population of the vibrational levels in the initial electronic state, an ISC rate of ~7×10^{11} s^{-1} is obtained for $^1(\pi_H \rightarrow \pi_L^*) \rightsquigarrow {}^3(n_S \rightarrow \pi_L^*)$. The radiationless decay to the triplet manifold is much faster than a radiative transition back to the ground state which occurs at rate of ~6×10^7 s^{-1}. Moreover, since conical intersections with the ground state potential energy surface have not been found in the vicinity of the S_1 minimum, we may suppose that internal conversion to the electronic ground state cannot compete either with fast intersystem crossing. Hence, we expect a triplet quantum yield close to unity in 2T-LF and fluorescence to be quenched, both in vacuum and solution.

Triplet Lifetime. Due to the heavy atom effect, thiocarbonyl compounds are expected to exhibit more pronounced phosphorescence and thus an intrinsically shorter triplet lifetime than their carbonyl counterparts. This has indeed been observed in many cases.[49] In the present context, the question arises whether the T_1 state of 2T-LF is sufficiently long-lived to undergo bimolecular photochemical reactions, in particular the bond formation with the neighboring cystein residue in the LOV domain.

In the gas phase, the T_1 $^3(\pi_H \rightarrow \pi_L^*)$ state splits into two lower-lying near degenerate sublevels under the influence of spin−orbit coupling and a third sublevel approximately 11 cm^{-1} above. The two lower fine-structure substates exhibit substantial contributions from the T_2 $^3(n_S \rightarrow \pi_L^*)$ state owing to the large SOME and the small energy separation between these states. In contrast, the coefficients of singlet configuration state functions in the spin−orbit mixed wave functions of these substates are tiny. Likewise, the triplet admixture to the electronic ground state wave function is marginal. Our calculations predict the phosphorescence lifetime of these sublevels at cryogenic temperatures to be approximately 20 s. The higher-lying T_1 fine-structure component mixes strongly with the S_1 $^1(n_S \rightarrow \pi_L^*)$ state. Intensity borrowing from the (relatively weak) $S_1 \rightarrow S_0$ emission yields a phosphorescence rate constant of about 10 s^{-1} for this level, corresponding to a lifetime of 100 ms. Assuming equal populations of the three sublevels yields a high-temperature limit of 3.6 s^{-1} for the radiative decay rate, corresponding to an average T_1 phosphorescence lifetime of 280 ms.

In aqueous solution, the probability for a phosphorescence decay of the T_1 state is expected to be even smaller than in the gas phase because of the hypsochromic shifts of the perturbing $(n_S \rightarrow \pi_L^*)$ states. We therefore conclude that the T_1 state of 2T-LF is sufficiently long-lived to undergo photochemical reactions.

4-Thio-lumiflavin. *Vacuum.* Also the 4T-LF isomer exhibits a planar nuclear arrangement of the ring system in the electronic ground state. Selected bond lengths are displayed in Table S6 of the SI. Again, only slight changes of the geometry parameters are observed when comparing 4T-LF with LF. In the neighborhood of C_4═S bond, the N_3−C_4 and C_4−C_{4a} bonds slightly contract with respect to LF. All other geometry parameters are nearly unchanged.

The frontier MOs of 4T-LF resemble those of 2T-LF both with respect to energetic order and shape when the oxygen and sulfur centers are interchanged (Figure 4). The π_H and π_{L+2}^* MOs in 4T-LF just appear somewhat more delocalized. In the singlet manifold, the lowest excited state corresponds to a $^1(n_S \rightarrow \pi_L^*)$ excitation with negligible oscillator strength (Table 4).

Table 4. Vertical Excitation Spectra of 4T-LF in the Gas Phase and in Water Solution

		gas phase		water solution		
state	electronic structure	ΔE [eV]	$f(r)$	ΔE [eV]	λ_{max} [a]	$f(r)$
$1^1A'$	ground state	0.00		0.00		
$1^1A''$	$n_S \rightarrow \pi_L^*$	1.72	0.0000	2.40	2.34 (530 nm) [b]	0.0000
$2^1A'$	$\pi_H \rightarrow \pi_L^*$	2.62	0.3767	2.57	2.51 (494 nm)	0.3952
$3^1A'$	$\pi_{H-2} \rightarrow \pi_L^*$	3.07	0.0610	3.11	3.14 (368 nm)	0.0528
$2^1A''$	$n_N \rightarrow \pi_L^*$	3.16	0.0019	3.40		0.0016
$1^3A''$	$n_S \rightarrow \pi_L^*$	1.61		2.29		
$1^3A'$	$\pi_H \rightarrow \pi_L^*$	1.84		1.99		
$2^3A'$	$\pi_{H-2} \rightarrow \pi_L^*$	2.44		2.49		
$2^3A''$	$n_N \rightarrow \pi_L^*$	2.83		3.07		

[a] Absorption maxima in 0.1 N phosphate buffer (pH 7), ref 17. [b] Shoulder in absorption spectrum in 0.1 N phosphate buffer (pH 7), ref 17.

Its excitation energy is comparable to the corresponding state in 2T-LF while its dipole moment is even more reduced with respect to the electronic ground state (see Table S7 of the SI). Like in 2T-LF, the optically bright $^1(\pi_H \rightarrow \pi_L^*)$ excitation is associated with the $S_2 \rightarrow S_0$ transition. However, despite its higher excitation energy, its oscillator strength ($f(r) \approx 0.38$) is smaller than the corresponding transition in 2T-LF. We find the electric dipole moment of S_2 to be nearly unchanged with respect to the electronic ground state. Hence, we expect no or small solvatochromism of this band in polar solvents (see Section). The S_3 and S_4 states have reversed order in the FC spectrum of both isomers. While the $^1(n_N \rightarrow \pi_L^*)$ excitation is

dx.doi.org/10.1021/jp4098233 | J. Phys. Chem. B 2014, 118, 1743−1753

located in a similar energy range as in 2T-LF and LF, the medium strong $^1(\pi_{H-2} \rightarrow \pi_L^*)$ transition is found at substantially lower excitation energies (3.07 eV) in 4T-LF than in 2T-LF (3.26 eV) and in LF (3.84 eV).[18]

The order of states in the triplet manifold is the same as in the singlet manifold but the energy gap between the $^3(n_S \rightarrow \pi_L^*)$ and $^3(\pi_H \rightarrow \pi_L^*)$ is considerably smaller than between the corresponding singlet states. In contrast to 2T-LF, the $^3(\pi_S \rightarrow \pi_L^*)$ excitation yields the lowest triplet state in 4T-LF in the gas phase. In the FC region, even a third triplet state, $^3(\pi_{H-2} \rightarrow \pi_L^*)$, is located energetically below the bright S_2 state and might thus be important for the photophysics of 4T-LF.

The largest geometrical changes between the ground- and excited-state geometries (Table S6 of the SI) are observed in the pteridine-ring. The similarities between the excited-state geometries are much less pronounced than in the 2T-LF isomer. This may be a consequence of the fact that the π_H MO is not so strongly localized at the sulfur center in 4T-LF as the n_S orbital. The N_1-C_2 bond, for instance, is shortened in the S_1 state with regard to the electronic ground state whereas it is elongated in the T_2 state. Particularly the extremely large change of the C_2-N_3 bond length in the S_2 state by almost 15 pm catches the eye. Again, it appears that the geometry distortions in the S_2 state are too pronounced at the TDDFT(B3-LYP) level, since the DFT/MRCI energy at the S_2 geometry (2.44 eV) is higher by 0.11 eV than at the S_1 and T_1 minimum nuclear arrangements (Table S7 of the SI). TDDFT(B3-LYP) yields a slightly out-of-plane distorted minimum nuclear arrangement in the latter T_1 state, the deviations of dihedral angles from planarity being below 1°. It turns out, however, that the DFT/MRCI method favors the planar structure by 38 cm^{-1}. The geometry parameters of the T_3 minimum are characterized by a simultaneous elongation of the $C_2=O$ and $C_4=S$ as may be expected from the density distribution in the π_{H-2} MO. In addition, the $C_{4a}-N_5$ is stretched by about 4 pm, characteristic of the occupation of the π_L^* MO.

The adiabatic energy difference between the S_1 and S_2 states of 4T-LF amounts to nearly 1 eV (Table 5) and is thus substantially larger than in the 2T-LF isomer. The T_1 and T_2 states are stabilized by about 0.3 eV upon geometry relaxation. Both are located energetically below the S_1 and S_2 states in the gas phase. The S_2 and T_3 minima are found to be near degenerate. This fact may gain importance in polar solvents where the $^{1,3}(n_S \rightarrow \pi_L^*)$ states are expected to experience strong blue shifts.

Table 5. Adiabatic DFT/MRCI Excitation Energies [eV] of 4T-LF in the Gas Phase and in Water Solution

state	electronic structure	gas phase	water solution[a]
$1^1A''$	$n_S \rightarrow \pi_L^*$	1.47	2.13
$2^1A'$	$\pi_H \rightarrow \pi_L^*$	2.44	2.39
$1^3A''$	$n_S \rightarrow \pi_L^*$	1.35	2.03
$1^3A'$	$\pi_H \rightarrow \pi_L^*$	1.55	1.70
$2^3A'$	$\pi_{H-2} \rightarrow \pi_L^*$	2.33	2.38

[a]Values were obtained by adding the solvent shift at the S_0 geometry to the adiabatic excitation energy of the isolated molecule.

Solvent Effects. It might be expected that the largest change of the geometry parameters with respect to the vacuum structure would occur for the $C_2=O$ bond due to the formation of strong hydrogen bonds. This is not what is observed, however, in the electronic ground state of the 4T-LF·$4H_2O$ cluster. Interestingly, the $C_4=S$ bond is elongated by 2 pm while the neighboring N_3-C_4 bond shrinks by 1.6 pm. All other bond length changes due to water solvation are below 1 pm. On the basis of our results for 2T-LF·$4H_2O$ (see above) we have refrained from optimizing the excited state structures of 4T-LF·$4H_2O$. Instead, adiabatic energies obtained in vacuum will be combined with solvent shifts computed at the ground state geometry to estimate adiabatic energies in solution.

The solvent effects on the vertical excitation spectrum of 4T-LF (Table S8 of the SI) are less pronounced than those of the 2T-LF isomer. From the analysis of the electric dipole moments this was to be expected for the $\pi_H \rightarrow \pi_L^*$ excitations, but the smaller blue shift of the $n_S \rightarrow \pi_L^*$ excited states of 4T-LF came as a surprise. The $^{1,3}(n_S \rightarrow \pi_L^*)$ states experience hypsochromic shifts of ~0.4 eV in aqueous solution. The differential solvent effects are thus not sufficient to push the $^1(\pi_S \rightarrow \pi_L^*)$ minimum energetically above the $^1(\pi_H \rightarrow \pi_L^*)$ minimum. The $^1(n_S \rightarrow \pi_L^*)$ state therefore remains to be the first excited singlet state even in aqueous solution (Table 4). Tentatively, we attribute the shoulder at 530 nm (2.34 eV) observed in the experimental absorption spectrum[17] to the $S_1 \rightarrow S_0$ electronic transition for which we obtain a value of 2.40 eV in aqueous solution. The maximum of this absorption band at 494 nm (2.51 eV)[17] fits well to the computed vertical excitation energy of the $S_2 \rightarrow S_0$ transition (2.57 eV). Likewise, we assign the second absorption band with maximum at 368 nm (3.14 eV)[17] to the $S_3 \rightarrow S_0$ transition (theoretical value 3.11 eV) which is dominated by the $^1(\pi_{H-2} \rightarrow \pi_L^*)$ excitation.

In the triplet manifold, the energy gap between the corresponding $^3(n_S \rightarrow \pi_L^*)$ and $^3(\pi_H \rightarrow \pi_L^*)$ minima in vacuum is much smaller than between the corresponding singlet states. In this case, the interaction with the polar environment is sufficient to reverse the energetic order of the triplet states. Thus, in water the T_1 state is characterized by a $^3(\pi_H \rightarrow \pi_L^*)$ electronic structure with the T_2 $^3(n_S \rightarrow \pi_L^*)$ state close by. The T_3 state experiences a small blue shift with respect to vacuum, but remains in the same energy range as the S_2 state.

Intersystem Crossing. Since the S_1 state of 4T-LF clearly exhibits $^1(n_S \rightarrow \pi_L^*)$ character in vacuum and polar protic environments (Table 5), we have refrained from computing rates for the $^1(\pi_H \rightarrow \pi_L^*)$ decay via ISC. Instead, it is assumed that a fast internal conversion to the S_1 state takes place. In vacuum and at low temperatures, the $^1(n_S \rightarrow \pi_L^*) \nrightarrow {}^3(\pi_H \rightarrow \pi_L^*)$ channel is energetically not available (Table 6). Our calculations place the $^3(\pi_H \rightarrow \pi_L^*)$ state adiabatically 615 cm^{-1} above the minimum of the S_1 state. In principle, the $^1(n_S \rightarrow \pi_L^*)$

Table 6. 4T-LF: Calculated Spin−Orbit Matrix Elements $|\langle i|\hat{H}_{SO}|f\rangle|$ [cm^{-1}] Evaluated at the Minimum of the Initial Singlet State, Adiabatic Electronic Energy Differences ΔE^{ad} [cm^{-1}], and ISC Rate Constants k_{ISC} [s^{-1}] in the Gas Phase and Water Solution (See Text)

channel		SOME	rate
$i \nrightarrow f$	ΔE^{ad}	$\lvert\langle i\lvert\hat{H}_{SO}\rvert f\rangle\rvert$	k_{ISC}
$^1(n_S \rightarrow \pi_L^*) \nrightarrow {}^3(\pi_H \rightarrow \pi_L^*)_x$		−73.1	
$^1(n_S \rightarrow \pi_L^*) \nrightarrow {}^3(\pi_H \rightarrow \pi_L^*)_y$		−141.7	
$^1(n_S \rightarrow \pi_L^*) \nrightarrow {}^3(\pi_H \rightarrow \pi_L^*)_z$		0.0	
$\{S_1 \nrightarrow T_2\}_{vac}$	−615		$\sim 1.5 \times 10^{12}$[a]
$\{S_1 \nrightarrow T_1\}_{wat}$	3663		$\sim 1.3 \times 10^{12}$

[a]At 298 K.

dx.doi.org/10.1021/jp4098233 | J. Phys. Chem. B 2014, 118, 1743−1753

state of 4T-LF could decay radiationlessly to the $^3(n_S \rightarrow \pi_L^*)$ state via ISC. According to the El-Sayed rules, this transition is forbidden because it is not accompanied by a change of orbital angular momentum, but it could borrow intensity by vibronic coupling with the nearby $(\pi_H \rightarrow \pi_L^*)$ states. Vibronic spin−orbit coupling has not been taken into account in this work. Once the much faster El-Sayed allowed ISC is energetically available, vibronic spin−orbit coupling is assumed to play a minor part. Population of higher vibrational levels in the S_1 potential well at room temperature leads to very fast $^1(n_S \rightarrow \pi_L^*) \rightsquigarrow ^3(\pi_H \rightarrow \pi_L^*)$ decay at the picosecond time scale. In aqueous solution, where the energy separation between these states is larger, the $^1(n_S \rightarrow \pi_L^*) \rightsquigarrow ^3(\pi_H \rightarrow \pi_L^*)$ ISC proceeds at about 1.3×10^{12} s^{-1}. Hence, we expect a triplet quantum yield close to unity in 4T-LF and fluorescence to be quenched in solution. In the gas phase, ISC from the lowest singlet state is predicted to be an activated process.

We have refrained from computing phosphorescence lifetimes because similar orders of magnitude are expected as for 2T-LF.

2,4-Dithio-lumiflavin. *Vacuum.* Also the 2,4DT-LF molecule exhibits a planar nuclear arrangement of the ring system in the electronic ground state. (For geometry parameters see Table S9 of the SI.) Similar trends as in the monothiosubstituted compounds are observed: Introduction of the C═S groups is accompanied by a slight contraction of the neighboring C−N and/or C−C bonds while all other geometry parameters are nearly unchanged with respect to the parent compound.

The HOMO and HOMO-1 of 2,4DT-LF (Figure 5) are very similar to those of 2T-LF while HOMO-2 and HOMO-3

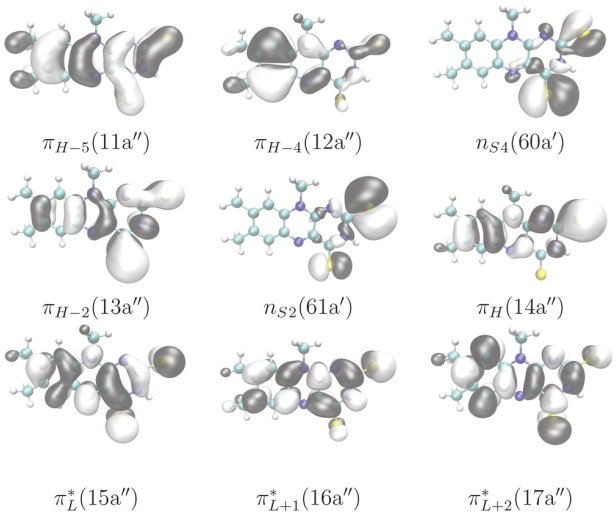


Figure 5. Frontier orbitals of 2,4DT-LF at the optimized ground-state geometry. (Isovalue = 0.02.)

resemble the HOMO and HOMO-1 of 4T-LF. Occupation of the $3p_\pi$ and n orbitals at the S_4 center is thus energetically preferred over the corresponding orbitals at the S_2 center. The π_{H-4} MO with high electron density in the benzene ring plays only a minor role for the lowest-lying electronic states of 2,4DT-LF. Interestingly, σ-type orbitals with lone-pair character at the nitrogen atoms are not among the six highest occupied MOs of 2,4DT-LF. The LUMO, LUMO+1, and LUMO+2 exhibit qualitatively the same shape as in all lumiflavin derivatives investigated in this work.

In accord with the energetic order of the MOs, $(n_N \rightarrow \pi_L^*)$ states are not among the low-lying excited states of 2,4DT-LF (Table 7). In vacuum, the first two excited singlet states are

Table 7. Vertical Excitation Spectra of 2,4DT-LF in the Gas Phase and in Water Solution

		gas phase		water solution	
state	electronic structure	ΔE [eV]	$f(r)$	ΔE [eV]	$f(r)$
$1^1A'$	ground state	0.00		0.00	
$1^1A''$	$n_{S2} \rightarrow \pi_L^*$	1.57	0.0000	2.22	0.0001
$2^1A''$	$n_{S4} \rightarrow \pi_L^*$	1.83	0.0000	2.45	0.0045
$2^1A'$	$\pi_H \rightarrow \pi_L^*$	2.20	0.3839	2.32	0.3952
$3^1A'$	$\pi_{H-2} \rightarrow \pi_L^*$	2.52	0.2519	2.73	0.1742
$1^3A''$	$n_{S2} \rightarrow \pi_L^*$	1.49		2.13	
$1^3A'$	$\pi_H \rightarrow \pi_L^*$	1.62		1.77	
$2^3A''$	$n_{S4} \rightarrow \pi_L^*$	1.77		2.40	
$2^3A'$	$\pi_{H-2} \rightarrow \pi_L^*$	1.81		2.20	

linear combinations of $(n_{S2} \rightarrow \pi_L^*)$ and $(n_{S4} \rightarrow \pi_L^*)$ excitations. The transitions carry only tiny electric dipole oscillator strengths and hence are expected to be dark. In contrast, strong absorption is predicted for the $S_3 \leftarrow S_0$ and $S_4 \leftarrow S_0$ transitions dominated by the $^1(\pi_H \rightarrow \pi_L^*)$ and $^1(\pi_{H-2} \rightarrow \pi_L^*)$ excitations, respectively. We found a nonplanar minimum structure in the T_1 $^3(n_S \rightarrow \pi_L^*)$ state of 2,4DT-LF. The deviations from planarity are minimal (i.e., < 1°), however. The $^3(\pi_H \rightarrow \pi_L^*)$ configuration dominates the T_2 state in vacuum, while the second $^3(n_S \rightarrow \pi_L^*)$ state has moved to the T_3 PEH. Otherwise, the electronic structures of the low-lying singlet and triplet states resemble each other closely (Table 7 and Table S10 of the SI).

Solvent Effects. Vertical excitation energies for various solvent environments are displayed in Table S11 of the SI. To our knowledge, experimental absorption spectra are not available for this compound or related ones. In aqueous solution of 2,4DT-LF (Table 7), the first intense absorption maximum is predicted to occur at 2.32 eV (534 nm). It originates from the $^1(\pi_H \rightarrow \pi_L^*)$ excitation. In the vertical spectrum, the lowest singlet excited state is still due to the $^1(n_S \rightarrow \pi_L^*)$ transition, despite a huge blue shift of the transition wavelength with respect to the gas-phase value. Adiabatically, the order of states appears to be reversed because the relaxation is found to be larger in the $^1(\pi_H \rightarrow \pi_L^*)$ state. Assuming relaxation and solvent effects to be additive, an estimated value of 1.95 eV is obtained for the adiabatic $^1(\pi_H \rightarrow \pi_L^*)$ excitation energy (Table 8). The minimum of the $^1(n_S \rightarrow \pi_L^*)$ PEH is estimated to lie approximately 0.08 eV higher at about 2.03 eV.

Table 8. Adiabatic DFT/MRCI Excitation Energies [eV] of 2,4DT-LF in the Gas Phase and in Water Solution

state	electronic structure	gas phase	water solution[a]
$1^1A''$	$n_{S2} \rightarrow \pi_L^*$	1.38	2.03
$2^1A''$	$n_{S4} \rightarrow \pi_L^*$	1.63	2.25
$2^1A'$	$\pi_H \rightarrow \pi_L^*$	1.83	1.95
$1^3A''$	$n_{S2} \rightarrow \pi_L^*$	1.33	1.97
$1^3A'$	$\pi_H \rightarrow \pi_L^*$	1.38	1.53

[a]Values were obtained by adding the solvent shift at the S_0 geometry to the adiabatic excitation energy of the isolated molecule.

dx.doi.org/10.1021/jp4098233 | J. Phys. Chem. B 2014, 118, 1743−1753

Due to the close proximity of these two singlet states, their vibronic interaction is supposed to be substantial. For the second $^1(n_S \rightarrow \pi_L^*)$ state, an adiabatic excitation energy of 2.25 eV is predicted.

A polar protic environment changes the order of triplet states with respect to the gas phase. The lowest triplet state clearly arises from the $^3(\pi_H \rightarrow \pi_L^*)$ in water solution. The lower of the two $^3(n_S \rightarrow \pi_L^*)$ states is found to be almost degenerate with the $^1(\pi_H \rightarrow \pi_L^*)$ state under these conditions (Table 8).

Intersystem Crossing. An overview over the calculated ISC rates of photoexcited 2,4DT-LF can be obtained from Table 9.

Table 9. 2,4DT-LF: Calculated Spin−Orbit Matrix Elements $|\langle i|\hat{H}_{SO}|f\rangle|$ [cm^{-1}] Evaluated at the Minimum of the Initial Singlet State, Adiabatic Electronic Energy Differences ΔE^{ad} [cm^{-1}], and ISC Rate Constants k_{ISC} [s^{-1}] in the Gas Phase and Water Solution (See Text)

channel		SOME	rate
$i \rightsquigarrow f$	ΔE^{ad}	$\lvert\langle i\lvert\hat{H}_{SO}\rvert f\rangle\rvert$	k_{ISC}
$^1(n_S \rightarrow \pi_L^*) \rightsquigarrow {}^3(\pi_H \rightarrow \pi_L^*)_x$		−156.2	
$^1(n_S \rightarrow \pi_L^*) \rightsquigarrow {}^3(\pi_H \rightarrow \pi_L^*)_y$		−1.4	
$^1(n_S \rightarrow \pi_L^*) \rightsquigarrow {}^3(\pi_H \rightarrow \pi_L^*)_z$		0.0	
$^1(\pi_H \rightarrow \pi_L^*) \rightsquigarrow {}^3(n_S \rightarrow \pi_L^*)_x$		−126.1	
$^1(\pi_H \rightarrow \pi_L^*) \rightsquigarrow {}^3(n_S \rightarrow \pi_L^*)_y$		12.2	
$1(\pi_H \rightarrow \pi_L^*) \rightsquigarrow {}^3(n_S \rightarrow \pi_L^*)_z$		0.0	
$\{S_1 \rightsquigarrow T_2\}_{vac}$	−35		~8.1 × 10^{12a}
$\{S_1 \rightsquigarrow T_2\}_{wat}$	−94		1 × 10^{12a}

[a]At 298 K.

In the gas phase, $S_1 \rightsquigarrow T_1$ ISC is El-Sayed forbidden. Due to the near-degeneracy of the S_1 and T_2 states, we find the El-Sayed allowed nonradiative decay of the S_1 state to proceed at the subpicosecond time scale from vibrationally excited levels of the initial electronic state. Accordingly, a triplet yield close to 1.0 is expected at room temperature.

In water solution, the situation is reversed. Here, the $^1(\pi_H \rightarrow \pi_L^*)$ excitation constitutes the S_1 state that is nearly degenerate with the $^3(n_S \rightarrow \pi_L^*)$ T_2 state. At room temperature, we obtain a rate constant of 1×10^{12} s^{-1} for the $S_1 \rightsquigarrow T_1$ ISC. Since the rate for a radiative decay of the S_1 state is more than 3 orders of magnitude smaller than k_{ISC}, fluorescence cannot compete with triplet formation.

SUMMARY AND CONCLUSIONS

In this work, we have investigated the absorption characteristics and excited-state decay mechanisms of various lumiflavin derivatives using advanced quantum chemical methods. We find three thiocarbonyl compounds, the monosubstituted 2T-LF and 4T-LF as well as the disubstituted 2,4DT-LF, to have promising photophysical properties. They absorb in the blue-green wavelength regime around 500 nm, i.e., substantially red-shifted with respect to lumiflavin that is the cofactor of natural blue-light photoreceptors. While in lumiflavin fluorescence and triplet formation compete in polar protic solvents, we expect the triplet quantum yields to be close to unity in the thioflavins. Their first singlet excited state is predicted to decay efficiently via ISC at the picosecond time scale. Their lowest triplet states, on the other hand, are sufficiently long-lived to enable bimolecular photochemical reactions. The combination of these properties makes the thioflavins potentially suitable candidates as cofactors in biomimetic photoswitches.

Figure 6 provides a schematic overview over the primary decay mechanisms following photoexcitation of lumiflavin and

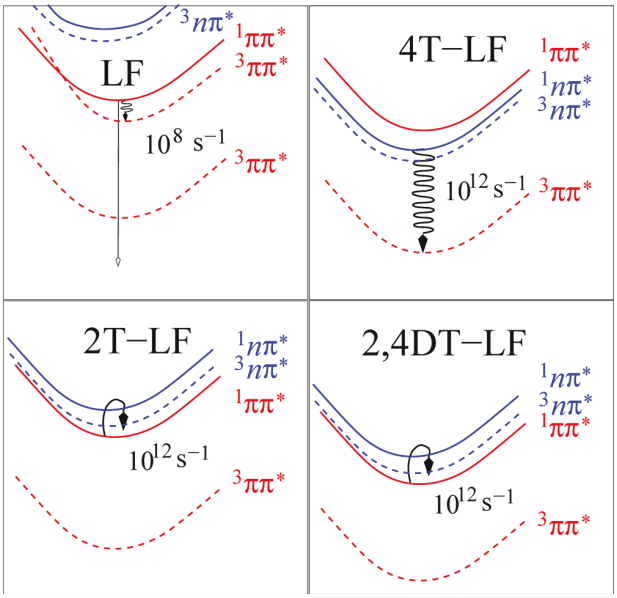


Figure 6. Schematic resume of the triplet formation mechanisms and intersystem crossing rates in water solution in lumiflavin and the thio-lumiflavins studied in this work.

the thioflavins in water solution. In previous work on LF[18] it was shown that vibronic spin−orbit coupling has to be invoked to reach the triplet manifold from the initially populated $^1(\pi_H \rightarrow \pi_L^*)$ state. Fluorescence that takes place at a rate of $k_F \approx 5 \times 10^7$ s^{-1} can thus compete with the El-Sayed forbidden ISC proceeding at a rate of $k_{ISC} \approx 10^8$ s^{-1}. We find two major causes that substantially accelerate ISC in thioflavins. The heavy-atom effect increases the size of the spin−orbit matrix elements substantially. At least equally important is the energetic stabilization of the first singlet and triplet $n\pi^*$ states. In 4T-LF, the $^1(n_S \rightarrow \pi_L^*)$ state constitutes the S_1 state even in solution. Photoexcitation to the optically bright S_2 state is followed by fast internal conversion to S_1 and an El-Sayed allowed ISC to the $^3(\pi_H \rightarrow \pi_L^*)$ state at a rate of $k_{ISC} \approx 10^{12}$ s^{-1}. In contrast, the optically bright $^1(\pi_H \rightarrow \pi_L^*)$ is the lowest singlet state in aqueous solutions of 2T-LF. Due to the close proximity of the $^3(n_S \rightarrow \pi_L^*)$ state, fluorescence will be efficiently quenched. Our calculations predict the triplet formation to be an activated process proceeding with a rate of $k_{ISC} \approx 10^{12}$ s^{-1} at room temperature. Interestingly, the introduction of a second sulfur atom in the 4-position in 2,4DT-LF lowers the excitation energies of the first and second singlet and triplet states, but it does not alter the decay paths as compared to 2T-LF. Finally, the $^3(\pi_H \rightarrow \pi_L^*)$ character of the T_1 state that is common to all four compounds is responsible for the long triplet lifetime.

The combination of a high triplet quantum yield and a long triplet lifetime is a favorable prerequisite for initiating a photocycle in LOV domains. In contrast to other flavin binding pockets such as BLUF, no aromatic amino acids are found close to the cofactor in LOV domains. Thus, quenching of the excitation by excitation energy or electron transfer is less probable in the latter. Rather, in natural blue-light receptors featuring LOV domains the cofactor is stabilized by a hydrogen-bonding network of the surrounding amino acids. In the course

dx.doi.org/10.1021/jp4098233 | J. Phys. Chem. B 2014, 118, 1743−1753

of the photocycle, the thiol hydrogen of a neighboring cystein migrates to the N_5 center and a covalent bond is formed between the cystein sulfur and the carbon atom in C_{4a} position of the isoalloxazine core. As mentioned already in the introduction, LOV domains can be fused to cellular effector proteins to create artificial blue light-sensitive variants enabling photochemical regulation of protein function. For optogenetic control, sensitization by blue light might cause too much harm. Hence, scientists are looking worldwide for triggers that can be stimulated at longer wavelengths. The three thioflavins considered in this work achieve this objective. In addition, the chemical behavior of LF should best be mimicked by 2T-LF since the substitution of a sulfur for a carbonyl oxygen in 2-position of FMN is expected to have only minor influence on the electron density distributions at the rather distant reactive N_5 and C_{4a} centers of the isoalloxazine core. Hence, among the present LF variants, we believe 2T-LF to be the most promising artificial cofactor. Studies investigating the photophysics and photochemistry of thioflavins in the protein environment of LOV domains are currently being undertaken in our laboratories.

■ ASSOCIATED CONTENT

Supporting Information

Vertical excitation spectra in vacuum of all compounds at their respective ground-state equilibrium geometries; selected bond lengths of the ground and excited state structures, vertical excitation energies at ground- and excited-state geometries, and environment effects on the excitation energies of all thiolumiflavins. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: Christel.Marian@uni-duesseldorf.de. Tel.: +49 211 8113210.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

Financial support by the Deutsche Forschungsgemeinschaft through MA 1051/12-1 is gratefully acknowledged.

■ REFERENCES

(1) Losi, A. Flavin-Based Blue-Light Photosensors: A Photobiophysics Update. *Photochem. Photobiol.* **2007**, *83*, 1283−1300.
(2) Briggs, W. R.; et al. The Phototropin Family of Photoreceptors. *Plant Cell* **2001**, *13*, 993−997.
(3) Gomelsky, M.; Klug, G. BLUF: A Novel FAD-Binding Domain Involved in Sensory Transduction in Microorganisms. *Trends Biochem. Sci.* **2002**, *10*, 497−500.
(4) Herrou, J.; Crosson, S. Function, Structure and Mechanism of Bacterial Photosensory LOV Proteins. *Nat. Rev. Microbiol.* **2011**, *9*, 713−723.
(5) Losi, A.; Gärtner, W. Old Chromophores, New Photoactivation Paradigms, Trendy Applications: Flavins in Blue Light-Sensing Photoreceptors. *Photochem. Photobiol.* **2011**, *87*, 491−510.
(6) Riggsbee, C. W.; Deiters, A. Recent Advances in Photochemical Control of Protein Function. *Trends Biotechnol.* **2010**, *28*, 468−475.
(7) Pastrana, E. Optogenetics: Controlling Cell Function with Light. *Nat. Methods* **2011**, *8*, 24−25.
(8) Drepper, T.; Krauss, U.; zu Berstenhorst, S. M.; Pietruszka, J.; Jaeger, K. E. Lights on and Action! Controlling Microbial Gene Expression by Light. *Appl. Microbiol. Biotechnol.* **2011**, *90*, 23−40.
(9) Wu, Y. I.; Frey, D.; Lungu, O. I.; Jaehrig, A.; Schlichting, I.; Kuhlman, B.; Hahn, K. M. A Genetically Encoded Photoactivatable Rac Controls the Motility of Living Cells. *Nature* **2009**, *461*, 104−108.
(10) Kramer, R. H.; Fortin, D. L.; Trauner, D. New Photochemical Tools for Controling Neuronal Activity. *Curr. Opin. Neurobiol.* **2009**, *19*, 1−9.
(11) Salzmann, S.; Martinez-Junza, V.; Zorn, B.; Braslavsky, S. E.; Mansurova, M.; Marian, C. M.; Gärtner, W. Photophysical Properties of Structurally and Electronically Modified Flavin Derivatives Determined by Spectroscopy and Theoretical Calculations. *J. Phys. Chem. A* **2009**, *113*, 9365−9375.
(12) Hecht, S.; Richter, G.; Bacher, A.; Joshi, M.; Römisch, W.; Greiner, G.; Frank, R.; Weber, S.; Eisenreich, W.; Fischer, M. Photocycle of a blue light receptor LOV2 domain reconstituted with 5-deazaFMN. In *Flavins and Flavoproteins 2005; Proceedings of the 15th International Symposium on Flavins and Flavoproteins*; Nishino, T., Miura, R., Tanokura, M., Fukui, K., Eds.; ARchiTect Inc.: Tokyo, 2005; pp 569−574.
(13) Silva-Junior, M. R. Quantum Mechanical/Molecular Mechanics Study of Electronically Excited States and Assessment of Methods for Calculating Vertical Excitation Energies. Ph.D. Thesis; Heinrich Heine Universität Düsseldorf, 2011.
(14) Silva-Junior, M. R.; Mansurova, M.; Gärtner, W.; Thiel, W. Photophysics of Structurally Modified Flavin Derivates in the Blue-Light Photoreceptor Ytva: A Combined Experimental and Theoretical Study. *ChemBioChem* **2013**, *14*, 1648−1661.
(15) Macheroux, P.; Bornemann, S.; Ghisla, S.; Thorneley, R. N. F. Studies with Flavin Analogs Provide Evidence that a Protonated Reduced FMN Is the Substrate-Induced Transient Intermediate in the Reaction of Escherichia Coli Chorismate Synthase. *J. Biol. Chem.* **1996**, *271*, 25850−25858.
(16) Mansurova, M.; Simon, J.; Salzmann, S.; Marian, C. M.; Gärtner, W. Spectroscopic and Theoretical Study on Electronically Modified Chromophores in LOV Domains: 8-Bromo- and 8-Trifluoromethyl-Substituted Flavins. *ChemBioChem* **2013**, *14*, 645−654.
(17) Dudley, K. H.; Ehrenberg, A.; Hemmerich, P.; Müller, F. Spektren und Strukturen der am Flavin-Redoxsystem beteiligten Partikeln. Studien in der Flavinreihe IX [1]. *Helv. Chim. Acta* **1964**, *47*, 1354−1382.
(18) Salzmann, S.; Tatchen, J.; Marian, C. M. The Photophysics of Flavins: What Makes the Difference Between Gas Phase and Aqueous Solution? *J. Photochem. Photobiol. A* **2008**, *198*, 221−231.
(19) Salzmann, S.; Silva-Junior, M. R.; Thiel, W.; Marian, C. M. Influence of the LOV Domain on Low-Lying Excited States of Flavin: A Combined Quantum-Mechanics/Molecular-Mechanics Investigation. *J. Phys. Chem. B* **2009**, *113*, 15610−15618.
(20) TURBOMOLE V6.3 2011, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989−2007, TURBOMOLE GmbH, since 2007; available from http://www.turbomole.com.
(21) Schäfer, A.; Huber, C.; Ahlrichs, R. Fully Optimized Contracted Gaussian Basis Sets of Triple Zeta Valence Quality for Atoms Li to Kr. *J. Chem. Phys.* **1994**, *100*, 5829−5835.
(22) Becke, A. D. Density-Functional Thermochemistry. 3. The Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98*, 5648−5652.
(23) 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.
(24) Bauernschmitt, R.; Ahlrichs, R. Treatment of Electronic Excitations within the Adiabatic Approximation of Time-Dependent Density Functional Theory. *Chem. Phys. Lett.* **1996**, *256*, 454−464.
(25) Furche, F.; Ahlrichs, R. Adiabatic Time-Dependent Density Functional Methods for Excited State Properties. *J. Chem. Phys.* **2002**, *117*, 7433−7447.
(26) Kind, C.; Reiher, M.; Neugebauer, J. SNF *Version 2.2.1: A Program Package for Numerical Frequency Analyses*; Universität Erlangen, 1999−2002.

(27) Grimme, S.; Waletzke, M. A Combination of Kohn-Sham Density Functional Theory and Multi-Reference Configuration Interaction Methods. *J. Chem. Phys.* **1999**, *111*, 5645−5655.
(28) Becke, A. D. A New Mixing of Hartree-Fock and Local Density Functional Theories. *J. Chem. Phys.* **1993**, *98*, 1372−1377.
(29) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron Density. *Phys. Rev.* **1988**, *B 37*, 785−789.
(30) Silva-Junior, M. R.; Schreiber, M.; Sauer, S. P. A.; Thiel, W. Benchmarks for Electronically Excited States: Time-Dependent Density Functional Theory and Density Functional Theory Based Multireference Configuration Interaction. *J. Chem. Phys.* **2008**, *128*, 104103/1−14.
(31) Kleinschmidt, M.; Tatchen, J.; Marian, C. M. Spin−Orbit Coupling of DFT/MRCI Wavefunctions: Method, Test Calculations, and Application to Thiophene. *J. Comput. Chem.* **2002**, *23*, 824−833.
(32) Kleinschmidt, M.; Marian, C. M. Efficient Generation of Matrix Elements of One-Electron Spin-Orbit Operators. *Chem. Phys.* **2005**, *311*, 71−79.
(33) Hess, B. A.; Marian, C. M.; Wahlgren, U.; Gropen, O. A Mean-Field Spin−Orbit Method Applicable to Correlated Wavefunctions. *Chem. Phys. Lett.* **1996**, *251*, 365−371.
(34) AMFI is an atomic spin-orbit integral program written by Schimmelpfennig, B., University of Stockholm, 1996.
(35) Tatchen, J.; Marian, C. M. On the Performance of Approximate Spin−Orbit Hamiltonians in Light Conjugated Molecules: The Fine-Structure Splitting of HC_6H^+, NC_5H^+, and NC_4N^+. *Chem. Phys. Lett.* **1999**, *313*, 351−357.
(36) Danovich, D.; Marian, C. M.; Neuheuser, T.; Peyerimhoff, S. D.; Shaik, S. Spin−Orbit Coupling Patterns Induced by Twist and Pyramidalization Modes in C_2H_4: A Quantitative Study and a Qualitative Analysis. *J. Phys. Chem. A* **1998**, *102*, 5923−5936.
(37) Etinski, M.; Tatchen, J.; Marian, C. M. Time-Dependent Approaches for the Calculation of Intersystem Crossing Rates. *J. Chem. Phys.* **2011**, *134*, 154105−1−154105−9.
(38) Duschinsky, F. The Importance of the Electron Spectrum in Multi Atomic Molecules Concerning the Franck-Condon Principle. *Acta Physicochim.* **1937**, *7*, 551−566.
(39) Marian, C. M. Spin−Orbit Coupling and Intersystem Crossing in Molecules. *WIREs Comput. Mol. Sci.* **2012**, *2*, 187−203.
(40) Langhoff, S. R.; Kern, C. W. Molecular Fine Structure. In *Modern Theoretical Chemistry, ed. by H. F. Schaefer III*; Plenum: New York, 1977; Vol. *4*, pp 381−437.
(41) Kleinschmidt, M.; Tatchen, J.; Marian, C. M. Spock.CI: A Multireference Spin−Orbit Configuration Interaction Method for Large Molecules. *J. Chem. Phys.* **2006**, *124*, 124101−1−124101−17.
(42) Klamt, A.; Schüürmann, G. COSMO - A New Approach to Dielectric Screening in Solvents with Explict Expressions for the Screening Energy and its Gradient. *J. Chem. Soc., Perkin Trans.* **1993**, *2*, 799−805.
(43) Schäfer, A.; Klamt, A.; Sattel, D.; Lohrenz, J.; Eckert, F. COSMO Implementation in TURBOMOLE: Extension of an Efficient Quantum Chemical Code towards Liquid Systems. *Phys. Chem. Chem. Phys.* **2000**, *2*, 2187−2193.
(44) Mertz, E. L.; Krishtalik, L. I. Low Dielectric Response in Enzyme Active Site. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 2081−2086.
(45) Siegbahn, P. E.; Himo, F. The Quantum Chemical Cluster Approach for Modeling Enzyme Reactions. *WIREs Comput. Mol. Sci.* **2011**, *1*, 323−336.
(46) Sung, M.; Moore, T. A.; Song, P.-S. Molecular Luminescence Studies of Flavines. I. Excited States of Flavines. *J. Am. Chem. Soc.* **1972**, *94*, 1730−1740.
(47) El-Sayed, M. A. The Triplet State: Its Radiative and Nonradiative Properties. *Acc. Chem. Res.* **1968**, *1*, 8−16.
(48) Englman, R.; Jortner, J. Energy Gap Law for Radiationless Transitions in Large Molecules. *Mol. Phys.* **1970**, *18*, 145−164.
(49) Maciejewski, A.; Steer, R. P. The Photophysics, Physical Photochemistry, and Related Spectroscopy of Thiocarbonyls. *Chem. Rev.* **1993**, *93*, 67−98.

APPENDIX E

Computing UV/Vis Spectra from the Adiabatic and Vertical Franck-Condon Schemes with the use of Cartesian and Internal Coordinates

Jan P. Götze*, Bora Karasulu* and Walter Thiel
*Equal authors

J. Chem. Phys., **2013**, 139, 234108

THE JOURNAL OF CHEMICAL PHYSICS **139**, 234108 (2013)

Computing UV/vis spectra from the adiabatic and vertical Franck-Condon schemes with the use of Cartesian and internal coordinates

Jan P. Götze,[a] Bora Karasulu,[a] and Walter Thiel
Max-Planck-Institut für Kohlenforschung, Kaiser-Wilhelm-Platz 1, 45470 Mülheim an der Ruhr, Germany

(Received 26 July 2013; accepted 26 November 2013; published online 17 December 2013)

We address the effects of using Cartesian or internal coordinates in the adiabatic Franck-Condon (AFC) and vertical Franck-Condon (VFC) approaches to electronic spectra. The adopted VFC approach is a simplified variant of the original approach [A. Hazra, H. H. Chang, and M. Nooijen, J. Chem. Phys. **151**, 2125 (2004)], as we omit any contribution from normal modes with imaginary frequency. For our test molecules ranging from ethylene to flavin compounds, VFC offers several advantages over AFC, especially by preserving the properties of the FC region and by avoiding complications arising from the crossing of excited-state potential surfaces or from the failure of the harmonic approximation. The spectral quality for our target molecules is insensitive to the chosen approach. We also explore the effects of Duschinsky rotation and relate the need for internal coordinates to the absence of symmetry elements. When using Duschinsky rotation and treating larger systems without planar symmetry, internal coordinates are found to outperform Cartesian coordinates in the AFC spectral calculations. © *2013 AIP Publishing LLC*. [http://dx.doi.org/10.1063/1.4844055]

INTRODUCTION

In recent years, the UV/vis properties of large molecules have met increasing interest, for example, because of the activities in the fields of photovoltaics[1,2] and optogenetics.[3] Linear response (LR) theory is a practical method to obtain optical spectra in a computationally efficient way, from the energy differences between electronic states usually evaluated at the ground-state minimum geometry. Time-dependent density functional theory (TD-DFT) is currently the most popular method for computing LR spectra.[4] However, there are several well-known drawbacks to LR-TD-DFT, for example, the lack of multi-excitation character[5–9] or the sometimes strong dependence of the results on the chosen density functional.

Theoretical Franck-Condon (FC) spectra can be directly compared to experiment.[10–14] They take into account both the ground- and excited-state vibrational levels, but there are also approaches that only consider the ground-state vibrational distribution (see Ref. 15 for an example). The computation of FC spectra requires the overlap between the ground- and excited-state vibrational wave functions, which affects the absorption or fluorescence intensity of the vibrational bands. In this work, we will use the time-independent FC method (TI-FC),[16–21] but time-dependent approaches (TD-FC) are available as well.[22] For both TI-FC and TD-FC, one can employ Duschinsky rotation,[23,24] which allows for mode mixing by constructing the excited-state modes from a linear combination of the ground-state modes.

Regardless of the chosen approach for computing the FC spectra, there is the question of how to represent the underlying potential energy surfaces (PES). For the ground state, the harmonic approximation is straightforward and generally quite good, as excitation events usually start from a geometry close to the ground-state minimum. The excited-state PES, however, may change character along its gradient by crossing with other surfaces. Even if there is an excited-state minimum, it needs to be geometrically close to the ground-state minimum (the FC point) to ensure that the harmonic approximation describes the FC region well. Otherwise one may resort to a proper anharmonic treatment, which is however quite costly and therefore only feasible for small molecules.[25–31]

For large systems, it thus seems attractive to use a vertical Franck-Condon (VFC) approach,[29] in which an excited-state structure is generated by extrapolating from the FC point with a single Newton-Raphson step. By definition, this geometry would correspond to a minimum on a perfectly harmonic excited-state PES and will therefore be called extrapolated minimum (EM). In reality, the EM structure will generally not be a stationary point, and the extrapolated energies and derivatives are not those of the excited-state PES. On the other hand, it might indeed be better to calculate the FC spectrum using the EM structure rather than the true minimum of the PES[29] (i.e., using the VFC approach rather than an adiabatic FC approach abbreviated as AFC). Results from VFC and AFC calculations have recently been compared in some detail.[32]

In this article, we present a slightly simplified VFC variant in both Cartesian and internal coordinate representation. The problem of using Duschinsky rotation[23] with imaginary-frequency normal modes will be treated in the simplest conceivable way, by ignoring the corresponding eigenvectors. This allows the use of the very same coordinates for computing the ground- and excited-state Hessians, which in turn enables us to use internal coordinates with an orthogonal Duschinsky rotation matrix, as opposed to the case where coordinates have to be constructed anew at an adiabatic excited-state PES minimum.[33] We discuss how much the choice of coordinate system influences the results, and we present

[a]B. Karasulu and J. P. Götze contributed equally to this work.

comparisons between the results from VFC and AFC calculations for target molecules ranging from ethylene to flavin compounds.

EXCITED-STATE PROPERTIES

In this section, we will outline what needs to be modified in the VFC method when employing internal instead of Cartesian coordinates in the normal-mode representation. First and foremost, we will show that the shift vector in orthonormal coordinates along the target excited-state surface ($\mathbf{\Delta Q}_t$) is identical for both coordinate systems.

The determination of $\mathbf{\Delta Q}_t$ in Cartesian coordinates has been described before, e.g., by Petrenko and Neese[34] in a time-dependent context. In the following, we will therefore only survey the core points of the Cartesian approach and mainly address the differences that arise when using internal coordinates. All required geometry optimizations and frequency calculations were performed with Gaussian09,[35] using the B3LYP[36] or CAM-B3LYP[37] functionals and either the 6-311++G(d,p) or 6-31G(d) basis sets.[38]

At the optimized ground-state geometry represented by set of Cartesian coordinates for N atoms ($\boldsymbol{\xi}_0$), the analytic Cartesian gradient $\boldsymbol{V}^C$ and Hessian $\boldsymbol{H}^C$ of the target state t are given by

$$\frac{\partial E_t}{\partial \boldsymbol{\xi}} = \boldsymbol{V}^C, \tag{1}$$

$$\frac{\partial^2 E_t}{\partial \boldsymbol{\xi}^2} = \boldsymbol{H}^C. \tag{2}$$

The ground-state equilibrium geometry will generally not be a stationary point for the target state t. Diagonalization of the Cartesian Hessian $\boldsymbol{H}^C$ will thus mix some rotational motion into the vibrational modes and thus affect the eigenvalues of the latter (i.e., the vibrational frequencies). To alleviate the associated problems, we perform the vibrational analysis in internal coordinates and then transform back to Cartesian coordinates (as explained below), but alternative techniques could also be applied, such as Eckart orientation[39] or quaternion treatment.[40]

In the VFC method, the reference geometry of the target excited state is generated by performing a single Newton-Raphson step along the normal coordinates. In our VFC variant, we freeze modes that do not have a positive eigenvalue of the Hessian, as the resulting step would be uphill in energy. This is a simplification of the original VFC approach,[29] in which such modes are treated as well by using a scan with perpendicular coordinates being frozen. We will discuss the merits and shortcomings of our simplification, whenever appropriate.

According to Hooke's law, the gradient at the FC point $\boldsymbol{x}_{FC}$ is related to the displacement $\mathbf{\Delta x}$,

$$\boldsymbol{V}_{FC} = -\boldsymbol{k}\mathbf{\Delta x}, \tag{3}$$

where $\boldsymbol{x}$ can be any kind of coordinate set and $\boldsymbol{k}$ is the corresponding vector of force constants. In harmonic approximation, the displacement coordinates $\mathbf{\Delta Q}_t$ of the target state are generated from mass-weighted normal-mode Cartesian displacement coordinates $\mathbf{\Delta q}$,

$$\mathbf{\Delta}\,\boldsymbol{Q}_t^C = \boldsymbol{L}_t^{C,T}\,\mathbf{\Delta q} = \boldsymbol{L}_t^{C,T}\sqrt{\boldsymbol{m}}(\boldsymbol{\xi}_t - \boldsymbol{\xi}_0), \tag{4}$$

where T denotes the matrix transpose. The columns of the matrix $\boldsymbol{L}_t{}^C$ are the orthonormal Cartesian eigenvectors of the Hessian (the normal modes), and $\boldsymbol{m}$ is a vector containing the atomic masses associated with each Cartesian coordinate in $\boldsymbol{\xi}$.

A representation of the normal modes in internal coordinates ($\boldsymbol{L}_t{}^I$) is known to be advantageous when one needs to describe complex curvilinear motions.[41] A non-redundant set of internal coordinates can be defined in terms of a standard (arbitrary) Z-matrix.[41] In the following, unless noted otherwise we adopt the procedural conventions and the matrix notation of Ref. 41. The normal modes can be converted from Cartesian to non-redundant internal coordinates as follows:

$$\boldsymbol{L}^I = \boldsymbol{B}''\boldsymbol{m}^{-1/2}\boldsymbol{L}^C, \tag{5}$$

$$\boldsymbol{B}'' = \boldsymbol{G}'^{-1/2}\boldsymbol{B}', \tag{6}$$

$$\boldsymbol{G}' = \boldsymbol{B}'\boldsymbol{m}^{-1}\boldsymbol{B}'^{T}. \tag{7}$$

Using primes to label matrices referring to non-redundant internal coordinates,[41] $\boldsymbol{G}'$ denotes the Wilson $\boldsymbol{G}$ matrix.[42] The Wilson $\boldsymbol{B}'$ matrix transforms the 3N Cartesian coordinate displacements $\mathbf{\Delta}\boldsymbol{\xi}$ to the set $\mathbf{\Delta}\boldsymbol{S}'$ of 3N-6 (5) non-redundant internal coordinate displacements,

$$\mathbf{\Delta}\boldsymbol{S}' = \boldsymbol{B}'\mathbf{\Delta}\boldsymbol{\xi}. \tag{8}$$

The gradients in internal coordinates $\boldsymbol{S}'$ are obtained from the mass-weighted Cartesian gradients

$$\boldsymbol{V}^I = \frac{\partial E_t}{\partial \boldsymbol{S}'} = \boldsymbol{G}'^{-1/2}\boldsymbol{B}''\boldsymbol{m}^{-1}\boldsymbol{V}^C. \tag{9}$$

The Hessian $\boldsymbol{H}^{MWI}$ in mass-weighted internal coordinates is related to the Hessian $\boldsymbol{H}^{MWC}$ in mass-weighted Cartesian coordinates by the following transformation:

$$H_{ij}^{MWI} = B_{ij}'' m_i^{-1/2} H_{ij}^{MWC} m_j^{-1/2} B_{ji}'', \tag{10}$$

$$H_{ij}^{MWC} = \frac{H_{ij}^C}{\sqrt{m_i m_j}}. \tag{11}$$

This transformation is approximate because it neglects contributions involving gradient terms; it is exact at stationary points with zero gradient.[43] We have confirmed numerically for one of the currently studied flavin molecules that this approximation is not critical, since the eigenvalues of the Cartesian and internal Hessians generally agree very well [mostly within 1%, see Table S1 in the supplementary material[44]]. We also find that spectra from the Cartesian VFC Hessian are essentially identical to those from the internal coordinate VFC Hessian (see Results and Discussion section) showing that this approximation is justified for our purposes.

The normal modes $\boldsymbol{L}_t{}^I$ are obtained as the eigenvectors of the Hessian $\boldsymbol{H}^{MWI}$ in mass-weighted internal coordinates. The Newton-Raphson displacements in orthonormal internal

coordinates are then given by

$$\Delta \boldsymbol{Q}_t^I = \boldsymbol{L}^{I,T} \boldsymbol{G}'^{-1/2} \Delta \boldsymbol{S}'. \tag{12}$$

Both $\Delta \boldsymbol{Q}_t{}^C$ and $\Delta \boldsymbol{Q}_t{}^I$ are orthonormal displacements from the FC point along the target PES. Within the VFC scheme, we compute $\Delta \boldsymbol{Q}_t$ first by performing a Newton-Raphson step and then construct the associated displacements in Cartesian or internal coordinates. Hence, even though the resulting EM geometries will be different in real space, their displacements from the FC point are identical in the corresponding orthonormal spaces. The resulting real space geometry is therefore of no concern. We thus have

$$\Delta \boldsymbol{Q}_t^I = \Delta \boldsymbol{Q}_t^C = \Delta \boldsymbol{Q}_t. \tag{13}$$

Therefore, Eq. (3) can be used for any coordinate set to construct $\Delta \boldsymbol{Q}_t$ by projecting the mass-weighted forces in any representation onto the corresponding normal modes

$$\boldsymbol{L}_t^{C,T} \boldsymbol{m}^{-1/2} \boldsymbol{V}^C = \boldsymbol{L}_t^{I,T} \boldsymbol{G}'^{1/2} \boldsymbol{V}^I = \boldsymbol{V}^Q = -\boldsymbol{k} \Delta \boldsymbol{Q}_t. \tag{14}$$

Each element of $\Delta \boldsymbol{Q}_t$ belongs to a normal mode n, depending on the force constants $\boldsymbol{k}$ and the gradients $\boldsymbol{V}^Q$ projected onto the normal coordinates. The force constants are given by the eigenvalues $\boldsymbol{\lambda}$ of the mass-weighted Hessian in 3N-6 (5) mass-weighted internal coordinates ($\boldsymbol{k} = \boldsymbol{\lambda}$).

At the FC point, one often has to face the problem of imaginary frequencies in the target excited state. As already mentioned, we discard modes with negative λ before using them in any computation. We note in this context that a common strategy in the calculation of FC spectra is to focus on the vibrational modes that affect the spectral shape most strongly.[45,46] Our approach may be regarded as a different kind of mode selection, in which we include all modes except those that are technically problematic and are not expected to influence the spectrum too much. An alternative would be the independent mode displaced harmonic oscillator model (IMDHO[12,47]) that uses the ground-state vibrational modes in a VFC-type treatment and thus disregards the differences in the shapes of the ground-state and excited-state PES, which might also cause artifacts in the computed spectra (see the section on FC Spectra). In our opinion, the omission of imaginary-frequency modes may be preferable, as the computed spectra should remain realistic up to the point where such modes become crucial (e.g., in a photodissociation spectrum[48]). A superior approach is to perform explicit PES scans along each imaginary-frequency normal mode to account for anharmonicity, as has been done in the VFC treatment of ethylene by Hazra, Chang, and Nooijen.[29] This is clearly beyond the scope of this article, considering the large molecules addressed below. For extended systems, we also note that the number of imaginary-frequency modes will remain small and represent a decreasing fraction of the full set of normal modes with increasing molecular size. Moreover, the FC integrals connected to such modes are expected to be small, at least at lower temperatures, since the corresponding excited-state potential minimum is expected to be quite far away from the FC point and the associated nuclear-motion wave function is thus likely to be broad.

The excited-state energy at the displaced EM geometry is given in the VFC scheme as follows:

$$E_{EM} = E_0 + E_{vert,t} - \frac{1}{2} \boldsymbol{\lambda} \Delta \boldsymbol{Q}_t^2, \tag{15}$$

where E_0 and $E_{vert,t}$ are the ground-state energy and the vertical excitation energy at the FC point, respectively, and modes with negative λ are again omitted in the last term.

FC SPECTRA

A slightly modified local version of the FCClasses program was used for the computation of the TI-FC spectra.[17,49] Except for the excited state shift vector $\Delta \boldsymbol{Q}_t$ and the normal modes, it requires some further input, namely, the ground-state eigenvalues $\boldsymbol{\lambda}_0$ and normal modes $\boldsymbol{L}_0$ (obtained by standard normal mode analysis), the displacement in orthonormal space for the ground state normal modes ($\Delta \boldsymbol{Q}_0$), and the Duschinsky matrix $\boldsymbol{J}$.[23] The latter is constructed for Cartesian coordinates by

$$\boldsymbol{J}^C = \boldsymbol{L}_0^{C,T} \boldsymbol{L}_t^C \tag{16}$$

and for internal coordinates by

$$\boldsymbol{J}^I = \boldsymbol{L}_0^{I,T} \boldsymbol{L}_t^I. \tag{17}$$

As $\boldsymbol{J}$ should be a square matrix, we may need to delete some columns from $\boldsymbol{L}_0$ to allow proper mapping (if imaginary modes in the excited state have been discarded). Thus, we remove those normal modes from the ground-state set that have the highest similarity to the excited-state imaginary modes (i.e., the largest Duschinsky matrix element). We can then construct $\Delta \boldsymbol{Q}_0$ by

$$\Delta \boldsymbol{Q}_0 = -\boldsymbol{J}^T \Delta \boldsymbol{Q}_t. \tag{18}$$

Since we have computed $\boldsymbol{J}$ for identical coordinate sets in the ground and excited state, it is orthogonal before the eventual deletion of columns and rows. After this deletion, $\boldsymbol{J}$ is no longer orthogonal but the operation of Eq. (18) remains valid. Since we have two different $\boldsymbol{J}$ matrices, one for each coordinate set, we may expect the corresponding displacements $\Delta \boldsymbol{Q}_0$ to differ as well (namely, $\Delta \boldsymbol{Q}_0{}^C$ or $\Delta \boldsymbol{Q}_0{}^I$). This is one issue that we seek to evaluate in this article, namely, if and how the choice of coordinate system affects the FC spectra with Duschinsky rotation and the corresponding computed ground-state displacements. For comparison, we also present VFC spectra for which we have set $\boldsymbol{J}$ to unity (after construction of $\Delta \boldsymbol{Q}_0$) and $\boldsymbol{\lambda}_t$ to $\boldsymbol{\lambda}_0$, which corresponds to a slightly modified version of the IMDHO model.[47] For the true IMDHO model (i.e., for the AFC case), we set $\Delta \boldsymbol{Q}_0$ equal to $\Delta \boldsymbol{Q}_\mathbf{t}$ as well. We also apply a zero-point energy correction to the IMDHO spectra based on the differences in $\boldsymbol{\lambda}_t$ and $\boldsymbol{\lambda}_0$, to avoid spectral shifts due to overestimated zero-point energies in the excited state.

Further details on the procedures for calculating spectra in the TI-FC scheme using FCClasses can be found elsewhere.[17–20,49] In terms of parameters, we computed up to 10^9 integrals and up to 60 quanta for C1 and C2 classes. We included all states with a Boltzmann population of at least 0.0001 at a temperature of 0.01 K. The calculated stick spectra

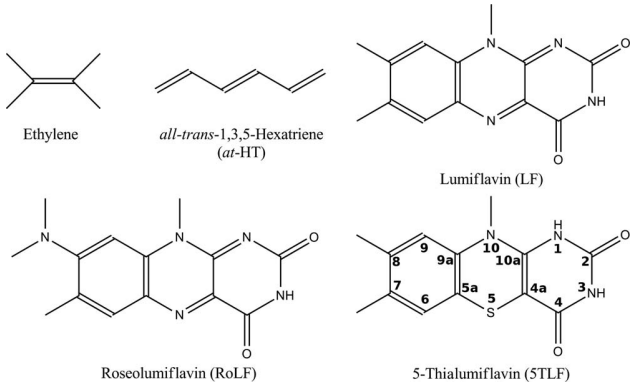


FIG. 1. Test molecules used in this work. Note that RoLF and 5TLF are of C_1 symmetry in all studied geometries, while the others are of D_{2h}, C_{2h} or C_s ground-state symmetry (ethylene, *at*-HT, or LF, respectively).

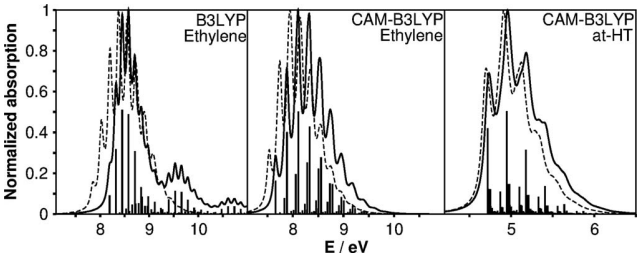


FIG. 2. S_1 TI-VFC spectra (solid lines and sticks) of ethylene and *all-trans*-1,3,5-hexatriene (*at*-HT). Dashed lines: modified IMDHO approximation. Cartesian and internal coordinate system results are identical (internal coordinate results shown). For VFC B3LYP spectra of ethylene, the number of integrals had to be reduced to 10^6 due to program limitations. Stick spectra in arbitrary units.

were convoluted to regular spectra by Lorentzian broadening using a HWHM of 0.05 eV. Our test systems are shown in Figure 1. They include two linear alkenes, namely, ethylene and *all-trans*-1,3,5-hexatriene (*at*-HT), as well as several flavin compounds presented thereafter. Additionally, the supplementary material[44] contains two aromatic model systems that allow for a more detailed analysis of the findings for roseoflavin.

We have adapted the FCClasses code to feed in our computed energies, as well as $\mathbf{\Delta Q}_0$, $\mathbf{\Delta Q}_t$, and $\boldsymbol{J}$ at the required instances. The $\boldsymbol{G}$ and $\boldsymbol{B}$ matrices are generated by our own code. Many of currently adopted procedures were inspired by the DUSHIN program, which is also used to produce the corresponding shift vectors and Duschinsky matrices during the computation of AFC spectra in internal coordinates; this program was kindly provided by Professor Reimers.[41]

RESULTS AND DISCUSSION

Ethylene and 1,3,5-hexatriene

Ethylene is a small molecule, for which the use of curvilinear coordinates shows significant effects[50] that may also affect the VFC spectra. We compare the results obtained with the CAM-B3LYP[37] and B3LYP[36,51] density functionals using the 6-311++G(d,p) basis.[38] The supplementary material[44] contains VFC results analogous to those from the work of Hazra, Chang, and Nooijen[29] for comparison. We restrict our analysis here to the spectra obtained from DFT calculations.

Another widely used model system[24,45] is *at*-HT. Compared to ethylene, we reduced the basis set to 6-31G(d) and only used a single functional (CAM-B3LYP) for *at*-HT. This will be the standard setup for all other calculations presented in this article. Note that, for both ethylene and *at*-HT, no excited-state minimum structure could be obtained, which prevents the calculation an AFC spectrum. Using the VFC scheme, however, we were able to obtain spectra, which are shown in Figure 2 (with numerical details given in Table I).

Regarding the quality of the computed ethylene VFC spectra, CAM-B3LYP provides better agreement than B3LYP with experiment, which displays a broadband between 7 and 9 eV.[52] Similarly good agreement is found for *at*-HT, which experimentally has a band starting at 4.5 eV and extending up to 5.5 eV and higher energies.[53] The VFC spectrum of ethylene agrees well with the ones reported previously using internal coordinates in the AFC scheme.[50]

TABLE I. Details of the VFC spectra and the extrapolated minima (EM). All spectra computed using CAM-B3LYP and 6-31G(d) basis, unless noted otherwise. n_{imag} is the number of imaginary modes. RMSD values between EM and FC point. Cartesian RMSD (Å) after structure alignment; internal RMSD in terms of bonds (Å)/angles (deg)/dihedrals (deg) from a redundant set of internal coordinates. Det($\boldsymbol{J}$) before deletion of imaginary modes is 1. C is spectral yield (ratio obtained/expected FC integral).

Molecule	State	n_{imag}	ΔE(EM-FC)	Coord	$RMSD^{cart}$	$RMSD^{int}$	Det($\boldsymbol{J}$)	C (IMDHO[a])
Ethylene[b] (CAM-B3LYP)	S_1	2	−0.68 eV	Cart	0.03	0.06/1.81/0.00	0.758	1.00 (1.00)
				Int	0.04	0.06/0.84/1.56	0.758	1.00 (1.00)
Ethylene[b] (B3LYP)	S_1	5	−0.43 eV	Cart	0.10	0.04/13.86/0.00	0.920	1.00 (1.00)
				Int	0.07	0.05/6.66/12.39	0.920	1.00 (1.00)
at-HT	S_1	3	−0.35 eV	Cart	0.03	0.03/1.32/0.02	0.736	1.00 (1.00)
				Int	0.04	0.03/1.97/3.30	0.736	1.00 (1.00)
Riboflavin	S_1	2	−0.26 eV	Cart	0.02	0.02/1.31/0.15	0.459	0.937 (0.999)
				Int	0.02	0.02/1.32/0.15	0.459	0.937 (0.999)
Roseoflavin	S_{11}	2	−0.19 eV	Cart	0.20	0.04/1.87/7.05	0.424	0.684 (0.001)
				Int	0.19	0.01/1.12/7.08	0.421	0.770 (0.002)
5-thiaflavin	S_8	5	−0.28 eV	Cart	0.08	0.02/0.94/3.23	0.178	0.693 (0.724)
				Int	0.06	0.02/1.01/4.21	0.178	0.697 (0.729)

[a]Modified IMDHO with $\mathbf{\Delta Q}_t \neq -\mathbf{\Delta Q}_0$ to avoid full equivalence of Cartesian and internal approaches.
[b]6-311++G(d,p) basis.

Table I allows for a quantitative comparison for both coordinate sets. The spectral yield is equal to 1 for ethylene and *at*-HT (in all cases), indicating that the best spectrum possible has been obtained within the TI-FC framework and the chosen parameters. Both coordinate representations are equivalent in the case of ethylene and *at*-HT: they give the same VFC spectra and the same determinant of the Duschinsky matrix (which deviates from unity because of the deletion of imaginary modes and their ground-state counterparts in our VFC scheme).

For ethylene, the root-mean-square deviations (RMSD) between the FC point and the EM structure are smaller for CAM-B3LYP than for B3LYP (Table I). However, despite showing smaller RMSD values, the CAM-B3LYP structures are energetically more distant from the FC point. This can be explained when looking at the number of imaginary modes: CAM-B3LYP shows only 2 (out of 12) imaginary modes at the FC point, compared with 5 for B3LYP. The CAM-B3LYP EM structure is therefore constructed from displacement along 10 modes, and only 7 for B3LYP. Thus, the energetic shift should be less pronounced for B3LYP, as can indeed be seen from Table I (−0.43 vs. −0.68 eV). The effect of excluding more imaginary modes can also be noted when comparing to the IMDHO spectra: The IMDHO spectra from B3LYP and CAM-B3LYP are similar (although energetically shifted), while the spectra excluding the imaginary modes have distinct differences. Therefore, the choice of density functional can also affect the spectral shape in our approach through the number of imaginary-frequency modes that are present.

The trends observed for ethylene are also valid for *at*-HT. The spectra show no dependence on the choice of coordinates, and the changes between the dihedral angles at the FC point and the EM structure are much larger when using internal rather than Cartesian coordinates in *at*-HT. This is similar to the ethylene case.

Finally, we note that ethylene and *at*-HT share an important feature: Even though their excited-state behavior is largely governed by internal rotation around the central bond, the corresponding forces cancel each other when starting from the planar FC point due to symmetry. This applies to many small molecules with rotational degrees of freedom that have at least C_s symmetry. In the following, we will therefore focus on larger flavin compounds, preferably without symmetry. In the supplementary material,[44] we address, in the context of a more detailed analysis of the flavin results, two other asymmetric aromatic systems, which were found to be unsuited for the calculation of FC spectra regardless of the chosen approach due to large $\mathbf{\Delta}\boldsymbol{Q}_0$ and $\mathbf{\Delta}\boldsymbol{Q}_t$ elements.

Flavin compounds

We now turn our attention to three larger molecules of biological interest, all belonging to the family of flavins. The molecular models treated here are actually the smaller lumiflavin variants, which are only missing a ribose chain that is not involved in the photophysics. We address riboflavin (modeled as lumiflavin, LF), roseolumiflavin (RoLF), and 5-thialumiflavin (5TLF). For each of them, we selected an exemplary target state to examine three different issues: (i) for LF/S_1, we do not expect any effects as the molecule has C_s symmetry; (ii) for RoLF/S_{11}, we find a transition with significant oscillator strength and a rotational gradient about the phenyl-amine bond (analogous cases for smaller systems are discussed in the supplementary material[44]); and (iii) for 5TLF/S_8, we wish to explore the effects of in-plane asymmetric excited-state motion (no rotational motion). Here, these three molecules only serve as examples for the merits and drawbacks of the methodology and of different coordinate systems. The computed spectra are shown in Figures 3 (AFC) and 4 (VFC), while a direct comparison of the stick spectra is provided in Figure 5. The computational details are given in Tables I and II, respectively.

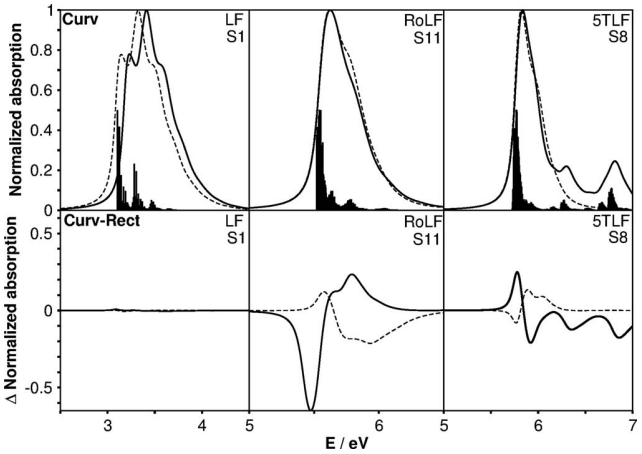


FIG. 3. TI-AFC spectra (solid lines and sticks) of lumiflavin (LF), roseoflavin (RoLF), and 5-thiaflavin (5TLF), based on the PES minima for the given state. Dashed lines: IMDHO approximation. Top row: Curvilinear mode representation spectra. Bottom row: difference spectra (curvilinear-rectilinear). Note that the curvilinear non-IMDHO spectrum of 5TLF/S_8 has very small yield (<0.1), which leads to artificial peaks above 6 eV. Stick spectra in arbitrary units.

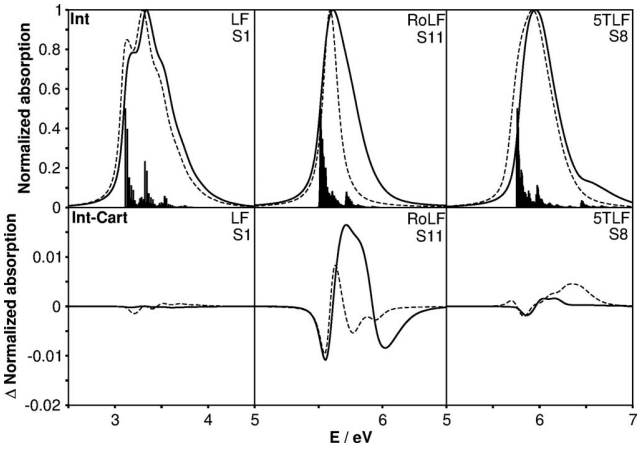


FIG. 4. TI-VFC spectra (solid lines and sticks) of lumiflavin (LF), roseoflavin (RoLF), and 5-thiaflavin (5TLF) for each given state. Dashed lines: IMDHO approximation. Top row: internal coordinate spectra, bottom row: difference spectra (internal-Cartesian). Note that the IMDHO spectra for RoLF/S_{11} have negligible yield and are only given for completeness. Stick spectra in arbitrary units.

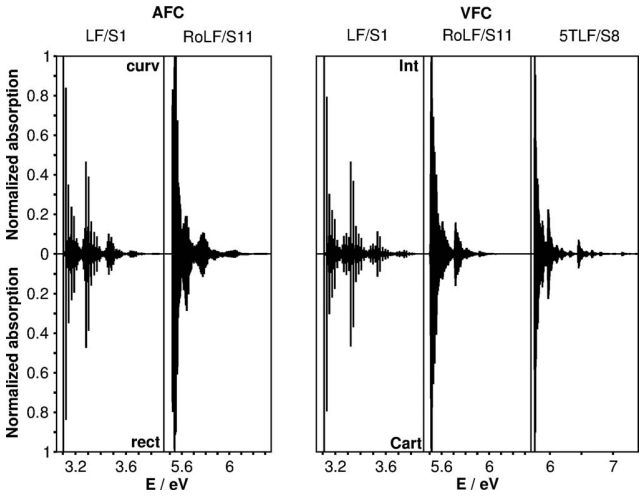


FIG. 5. Comparison of the AFC (VFC) stick spectra between curvilinear and rectilinear mode representations (internal and Cartesian coordinates). No comparison for 5TLF AFC due to low intensity in the curvilinear case.

The computation of the S_1 AFC spectrum of LF has been reported previously for B3LYP.[54] Our CAM-B3LYP LF/S_1 spectra were generated with high spectral yield, both for the AFC and VFC approaches. Judging from the common Cartesian RMSD value of 0.02 Å, the EM structure and the PES minimum are equally close to the FC point, and the reorganization energies are also very similar (−0.26 and −0.27 eV, respectively). Hence, we obtain almost complete and identical spectral convergence (0.987 for AFC curvilinear, 0.937 for all other combinations). Both coordinate representations perform equally well, and the corresponding stick spectra are virtually identical (see Figure 5). Some differences can be found between AFC and VFC calculations, especially in the region between 3.2 and 3.5 eV, which however do not change the shape of the convoluted spectrum much. These differences are independent of the coordinate set, and they can thus be attributed to the changes in the Hessian between the EM and the PES minimum or to the omission of imaginary modes in the VFC treatment (see the supplementary material[44] for a more detailed analysis of this specific case). Generally speaking, most findings for LF/S_1 are analogous to those for ethylene and *at*-HT, which also have C_s symmetry (like LF). We conclude that even for complex molecules, the existence of a symmetry plane significantly simplifies the calculation of FC spectra.

The changes in geometry and energy along the excited-state PES are generally small for the investigated flavin states (e.g., maximum energy shift of −0.33 eV for 5TLF). The AFC spectrum of RoLF/S_{11} is very sensitive to the chosen representation, with internal coordinates generally producing better spectral yield (0.534 vs. 0.766, or 0.832 vs. 0.920 for IMDHO, Table II), despite having a smaller Duschinsky determinant. This causes some significant changes in the convoluted spectra: The difference spectra (Figure 3, lower panels) show that the spectral changes range up to about 65% of the normalized intensity, effectively leading to a shifted onset and the gain/loss of a shoulder at 5.7 eV, as can also be seen in the stick spectrum (see Figure 5). We also note that the character of the state appears to change between the FC point and the PES minimum, as the oscillator strength decreases from 0.11 to zero upon optimization. At the resulting geometry, the excited state actually exhibits more charge transfer character (data not shown). This emphasizes one advantage of the VFC treatment, namely, that it preserves the electronic character of the FC point in the calculation of the spectrum.[34]

In the VFC spectra of RoLF/S_{11}, the internal coordinate description improves the spectral yield (0.684 vs. 0.770), just as in the case of the AFC spectra (0.534 vs. 0.766). This underlines the benefits gained from an internal coordinate/curvilinear description for asymmetric molecules with rotatable groups; a curvilinear description can in general be expected to produce smaller shift vectors and thus higher spectral yield (see the supplementary material[44] for a more detailed discussion on the shift vectors). The difference in the stick or convoluted VFC spectra is not as significant as in the AFC case, but is still present as a small rise in the intensity of a band at about 5.6 eV. The quantitative comparison in Table I shows this effect more clearly than the visual inspection of Figure 4 or Figure 5 (see the spectral yields above). Note that the use of IMDHO results in a drastic drop in the spectral yield (Table I), which can be associated to relatively high shifts of low-frequency modes included in the IMDHO case (data not shown).

5TLF/S_8 provides us with an asymmetric system without relevant rotatable groups. The AFC spectrum of 5TLF/S_8 could only be obtained with a spectral yield of 0.482 in the rectilinear representation, and it appeared only rather weak when using curvilinear normal modes and Duschinsky rotation. The IMDHO approach leads to a higher spectral yield

TABLE II. Details of the PES minima and AFC spectra. f is the oscillator strength at FC point or minimum structure. All spectra computed using CAM-B3LYP and 6-31G(d) basis. RMSD values between excited state minimum and FC point. Cartesian RMSD (Å) after structure alignment; internal RMSD in terms of bonds (Å)/angles (deg)/dihedrals (deg) from a redundant set of internal coordinates. C is spectral yield (ratio obtained/expected FC integral).

Molecule	State	f (FC/min)	ΔE(min-FC)	$RMSD^{cart}$	$RMSD^{int}$	Rep.	Det($\boldsymbol{J}$)	C (IMDHO)
Riboflavin	S_1	0.29/0.24	−0.27 eV	0.02	0.02/1.26/0.14	Rect	1.00	0.937 (0.999)
						Curv	0.944	0.987 (0.999)
Roseoflavin	S_{11}	0.11/0.00	−0.19 eV	0.10	0.01/1.10/4.01	Rect	0.996	0.534 (0.832)
						Curv	0.659	0.766 (0.920)
5-thiaflavin	S_8	0.29/0.58	−0.33 eV	0.12	0.01/0.89/5.03	Rect	0.991	0.482 (0.967)
						Curv	0.275	0.079 (0.945)

(Table II). This suggests that the low spectral yield is associated with a low Duschinsky determinant, at least for curvilinear coordinates. When using the VFC scheme, these problems are partially solved: The spectral yield for 5TLF/S_8 is about 0.7, with slight variations depending on whether Duschinsky rotation is employed and whether we use a Cartesian or internal coordinate representation.

For interpretation purposes, it is important to know that the curvilinear aspects are not as strong for 5TLF as for RoLF, since internal rotations are not important for 5TLF. The excited-state motion can be roughly characterized as a change in the bending angle (C6–S5–N9–C4, see Figure 1 for numbering) in the flavin plane, from -140° to -152° upon going from the FC point to the PES minimum (data not shown). This in-plane motion of 5TLF is asymmetric, but does not involve any major rotational degrees of freedom. Hence, the choice of coordinate system matters, but not as much as for RoLF/S_{11}.

Finally, we compare the computed flavin spectra to experimental data. Figure 6 presents the observed He droplet LF/S_1 spectrum[55] and the computed VFC spectrum in internal coordinates. The experimental LF/S_1 data are available only in the given energy range, which limits the comparisons with the computed spectrum. On the other hand, solvent effects should be negligible in an ultracold He droplet, which facilitates comparisons with the results from our gas-phase calculations. The dominant peaks in the observed spectrum are well reproduced in terms of relative intensity, although they are blue-shifted by about 0.5 eV. The energy difference between the two dominant peaks at 2.67/3.11 eV and 2.81/3.32 eV (experiment/theory) is larger in the computed spectrum (0.14 vs. 0.21 eV), which may partly be due to the neglect of anharmonicity in the excited-state modes. The computation also predicts the large peak directly following the 0-0 peak, but overestimates its relative intensity by almost 60%. Compared with the previous computational study on LF/S_1,[54] it seems that B3LYP is giving better energetic peak positions than CAM-B3LYP, even though CAM-B3LYP reproduces the characteristics of the stick spectra well.

Experimental spectra are available for all three flavin compounds in solution. Experimentally, the LF/S_1 peak is found at 2.5 eV in benzene[56] (i.e., at significantly lower energy than in a He droplet, see above). The strong S_{11} excitation of RoLF is observed in benzene at ∼250 nm (about 5 eV);[56] our computed gas-phase spectra are blue-shifted by about 0.75 eV. Finally, 5TLF/S_8 in water (acidic environment; neutral species) shows a strong peak at 255 nm;[57] we again find a blue shift of about 1 eV in our gas-phase results. While some of this blue shift appears to be due to solvent effects (as indicated by the experimental results for LF/S_1), we believe that the major part of the deviations originates from shortcomings of the chosen CAM-B3LYP density functional. Given the rather high amount of exact Hartree-Fock exchange in CAM-B3LYP, the blue shift in the computed flavin spectra is actually not surprising and as expected.

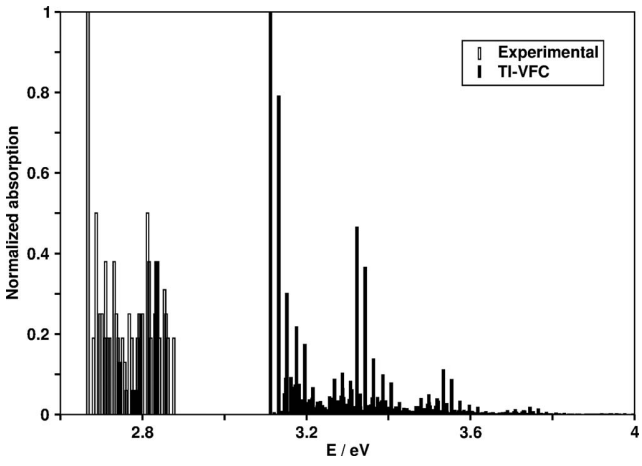


FIG. 6. Stick representation of the experimental He droplet spectrum (left) and the computed internal-coordinate VFC spectrum (right) for LF/S1. Experimental data from Vdovin, Slenczka, and Dick.[55]

SUMMARY AND CONCLUSIONS

In this article, we have examined the influence of the coordinate representation (rectilinear Cartesian vs. internal curvilinear) on the computation of optical FC spectra. For simple cases (no rotatable groups, plane of symmetry), we find that the computed AFC and VFC spectra do not depend on the choice of the coordinate system. However, when moving towards large biological chromophores, the use of Cartesian coordinates tends to be less advantageous (e.g., in terms of the spectral yield). The choice of the coordinate system may sometimes also affect the actual convoluted spectrum (e.g., with regard to the presence or absence of shoulders) in the AFC case. On the other hand, the VFC spectra are essentially the same when computed in internal or Cartesian coordinates. In our implementation, taking ethylene as an example, the differences between VFC/Cart and VFC/Int are less than 2% in relative spectral intensity, which can be attributed to slight differences in the Hessian eigenvalues (see the supplementary material,[44] p. 2). Contrary to the AFC case, internal coordinates provide no distinct advantage in VFC calculations.

If AFC spectra are at all obtainable, they are of similar quality as the VFC spectra in the examples studied presently. While this may not be true in general, there is apparently sometimes no real need to embark on the time-consuming and often error-prone process of excited-state optimizations for the purpose of computing FC spectra. Indeed, for large systems where the excited-state minimum may be far away from the FC point, we consider the VFC approach to be actually superior to AFC, as it guarantees the inclusion of the FC region in the harmonically approximated excited-state PES. In our slightly simplified VFC variant, we omit the imaginary modes of the excited state and their ground-state counterparts. This is a simplification taken to reduce overall computational effort compared to the original VFC method. The resulting spectra are found to be similar to the available experimental data and the results from previous VFC calculations,[29] at least in the present test case of LF/S1. Further studies with more comparisons to experimental spectra are needed to confirm this observation. In the considered cases, the error due to the omission of imaginary modes tends to become smaller as the system size increases, as the ratio of imaginary modes to the total number of modes decreases. This, however, needs

to be checked in future studies on a case-by-case basis. An important criterion in this respect is the actual value of the shift vector elements, as discussed in the supplementary material,[44] since large shift vector elements tend to have a larger influence on the spectral shape.

All our present computations were done at 0 K. This helps avoid certain technical issues such as spectral convergence problems that may appear at ambient temperatures. Proper comparisons with experiment should of course be done at the temperature, at which the spectrum was taken. Temperature effects will be addressed in future work, especially with regard to the influence of temperature on individual low-frequency modes and on the proposed omission of imaginary-frequency modes in our VFC approach.

Finally, the present data indicate that the presence of a symmetry plane in the molecule tends to diminish the differences between the Cartesian and internal coordinate representations, possibly for the following reason: If out-of-plane motions in C_s symmetric molecules are suppressed (e.g., through compensation of the corresponding forces acting in opposite directions), the in-plane motions can be described similarly well in both representations (i.e., through in-plane Cartesian or internal bond length displacements). Furthermore, it is obvious from Eqs. (16)–(18) that all differences between coordinate representations will vanish when using the IMDHO approximation. Hence, neglecting the effects of Duschinsky rotation will ensure that the choice of coordinate representation is not relevant. However, when the effects of Duschinsky rotation are important, as for several cases in this article, we suggest the use of internal coordinates in AFC calculations, because the results will normally be of equal or better quality than those from the Cartesian approach.

ACKNOWLEDGMENTS

This work was partially funded by the Volkswagenstiftung (Grant No. I/83915 to W.T.). The authors are grateful to Professor Jeffrey Reimers for providing the DUSHIN source code. J.P.G. is also grateful to Dr. Mario Barbatti, Dr. Dominik Kröner, and Dr. Taras Petrenko for many helpful discussions. We thank the reviewers for their constructive comments.

[1] M. Grätzel, Nature (London) **414**(6861), 338 (2001).
[2] A. Mishra and P. Bäuerle, Angew. Chem., Int. Ed. **51**(9), 2020 (2012).
[3] A. Möglich, X. J. Yang, R. A. Ayers, and K. Moffat, in *Annual Review of Plant Biology*, edited by S. Merchant, W. R. Briggs, and D. Ort (Annual Reviews, Palo Alto, 2010), Vol. 61, p. 21.
[4] M. E. Casida, C. Jamorski, K. C. Casida, and D. R. Salahub, J. Chem. Phys. **108**(11), 4439 (1998).
[5] S. Grimme and M. Waletzke, J. Chem. Phys. **111**, 5645 (1999).
[6] S. Hirata and M. Head-Gordon, Chem. Phys. Lett. **314**(3–4), 291 (1999).
[7] A. Dreuw, J. Phys. Chem. A **110**(13), 4592 (2006).
[8] C. M. Marian and N. Gilka, J. Chem. Theory Comput. **4**(9), 1501 (2008).
[9] M. R. Silva-Junior and W. Thiel, J. Chem. Theory Comput. **6**(5), 1546 (2010).
[10] L. S. Cederbaum and W. Domcke, J. Chem. Phys. **64**(2), 603 (1976).
[11] W. Domcke and L. S. Cederbaum, J. Chem. Phys. **64**(2), 612 (1976).
[12] W. Domcke, L. S. Cederbaum, H. Köppel, and W. Von Niessen, Mol. Phys. **34**(6), 1759 (1977).
[13] H. Köppel, W. Domcke, and L. S. Cederbaum, Adv. Chem. Phys. **57**, 59 (1984).
[14] L. Seidner, G. Stock, A. L. Sobolewski, and W. Domcke, J. Chem. Phys. **96**(7), 5298 (1992).
[15] R. Crespo-Otero and M. Barbatti, Theor. Chem. Acc. **131**(6), 1237 (2012).
[16] M. Dierksen and S. Grimme, J. Chem. Phys. **122**(24), 244101 (2005).
[17] F. Santoro, R. Improta, A. Lami, J. Bloino, and V. Barone, J. Chem. Phys. **126**(8), 084509 (2007).
[18] F. Santoro, A. Lami, R. Improta, and V. Barone, J. Chem. Phys. **126**(18), 184102 (2007).
[19] F. Santoro, A. Lami, R. Improta, J. Bloino, and V. Barone, J. Chem. Phys. **128**(22), 224311 (2008).
[20] F. Santoro and V. Barone, Int. J. Quantum Chem. **110**(2), 476 (2010).
[21] J. Huh and R. Berger, Faraday Discuss. **150**, 363 (2011).
[22] D. J. Tannor and E. J. Heller, J. Phys. Chem. **77**(1), 202 (1982).
[23] F. Duschinsky, Acta Physicochim. URSS **7**(4), 551 (1937).
[24] S. Banerjee, D. Kröner, and P. Saalfrank, J. Chem. Phys. **137**(22), 22A534 (2012).
[25] K. C. Thompson, M. J. T. Jordan, and M. A. Collins, J. Chem. Phys. **108**(20), 8302 (1998).
[26] R. P. A. Bettens, M. A. Collins, M. J. T. Jordan, and D. H. Zhang, J. Chem. Phys. **112**(23), 10162 (2000).
[27] D. H. Zhang, M. H. Yang, M. A. Collins, and S. Y. Lee, Proc. Natl. Acad. Sci. U.S.A. **99**(18), 11579 (2002).
[28] M. A. Collins, Theor. Chem. Acc. **108**(6), 313 (2002).
[29] A. Hazra, H. H. Chang, and M. Nooijen, J. Chem. Phys. **121**(5), 2125 (2004).
[30] G. I. Csonka, A. Ruzsinszky, and J. P. Perdew, J. Phys. Chem. A **109**(30), 6779 (2005).
[31] S. J. Kolmann and M. J. T. Jordan, J. Chem. Phys. **132**(5), 054105 (2010).
[32] F. J. A. Ferrer and F. Santoro, Phys. Chem. Chem. Phys. **14**(39), 13549 (2012).
[33] M. Rätsep, Z. L. Cai, J. R. Reimers, and A. Freiberg, J. Chem. Phys. **134**(2), 024506 (2011).
[34] F. Neese, T. Petrenko, D. Ganyushin, and G. Olbrich, Coord. Chem. Rev. **251**(3–4), 288 (2007).
[35] M. J. Frisch, G. W. Trucks, H. B. Schlegel *et al.*, Gaussian 09, Revision B.01, Gaussian, Inc., Wallingford, CT, 2010.
[36] C. Lee, W. Yang, and R. G. Parr, Phys. Rev. B **37**, 785 (1988).
[37] T. Yanai, D. Tew, and N. Handy, Chem. Phys. Lett. **393**, 51 (2004).
[38] P. C. Hariharan and J. A. Pople, Theor. Chim. Acta **28**(3), 213 (1973).
[39] F. Jorgensen, Int. J. Quantum Chem. **14**(1), 55 (1978).
[40] E. A. Coutsias, C. Seok, and K. A. Dill, J. Comput. Chem. **25**(15), 1849 (2004).
[41] J. R. Reimers, J. Chem. Phys. **115**(20), 9103 (2001).
[42] J. C. Decius, P. C. Cross, and E. B. Wilson, Jr., *Molecular Vibrations* (McGraw-Hill, New York, 1955).
[43] C. Y. Peng, P. Y. Ayala, H. B. Schlegel, and M. J. Frisch, J. Comput. Chem. **17**(1), 49 (1996).
[44] See supplementary material at http://dx.doi.org/10.1063/1.4844055 for additional results, tests, and comparisons.
[45] T. Petrenko and F. Neese, J. Chem. Phys. **127**, 164319 (2007).
[46] S. Salzmann, V. Martinez-Junza, B. Zorn, S. E. Braslavsky, M. Mansurova, C. M. Marian, and W. Gärtner, J. Chem. Phys. **113**(33), 9365 (2009).
[47] T. Petrenko and F. Neese, J. Chem. Phys. **137**(23), 234107 (2012).
[48] H. F. Davis and Y. T. Lee, J. Chem. Phys. **100**(1), 30 (1996).
[49] T. E. Sharp and H. M. Rosenstock, J. Chem. Phys. **41**, 3453 (1964).
[50] R. Borrelli and A. Peluso, J. Chem. Phys. **125**(19), 194308 (2006).
[51] A. D. Becke, J. Chem. Phys. **98**, 5648 (1993).
[52] J. R. Platt, H. B. Klevens, and W. C. Price, J. Chem. Phys. **17**(5), 466 (1949).
[53] G. F. Woods and L. H. Schwartzman, J. Am. Chem. Soc. **70**(10), 3394 (1948).
[54] B. Klaumünzer, D. Kröner, and P. Saalfrank, J. Phys. Chem. B **114**(33), 10826 (2010).
[55] A. Vdovin, A. Slenczka, and B. Dick, Chem. Phys. **422**, 195 (2013).
[56] P. Zirak, A. Penzkofer, T. Mathes, and P. Hegemann, Chem. Phys. **358**(1–2), 111 (2009).
[57] M. Janda and P. Hemmerich, Angew. Chem., Int. Ed. Engl. **15**(7), 443 (1976).

APPENDIX F

Assessment of Franck-Condon Methods for Computing Vibrationally Broadened UV/vis Absorption Spectra of Flavin Derivatives: Riboflavin, Roseoflavin and 5-thioflavin

Bora Karasulu, Jan P. Götze and Walter Thiel

J. Chem. Theo. Comp., **2014**, 10 (12), pp 5549-5566.

JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

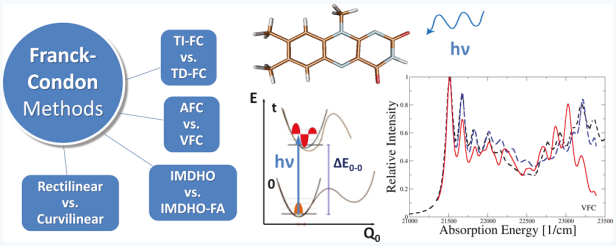
Assessment of Franck−Condon Methods for Computing Vibrationally Broadened UV−vis Absorption Spectra of Flavin Derivatives: Riboflavin, Roseoflavin, and 5-Thioflavin

Bora Karasulu, Jan Philipp Götze, and Walter Thiel*

Max-Planck-Institut für Kohlenforschung, Kaiser-Wilhelm-Platz 1, 45470, Mülheim an der Ruhr, Germany

Supporting Information

ABSTRACT: We address the performance of the vertical and adiabatic Franck−Condon (VFC/AFC) approaches combined with time-independent or time-dependent (TI/TD) formalisms in simulating the one-photon absorption spectra of three flavin compounds with distinct structural features. Calculations were done in the gas phase and in two solvents (water, benzene) for which experimental reference measurements are available. We utilized the independent mode displaced harmonic oscillator model without or with frequency alteration (IMDHO/IMDHO-FA) and also accounted for Duschinsky mixing effects. In the initial validation on the first excited singlet state of riboflavin, the range-separated functionals, CAM-B3LYP and ωB97xD, showed the best performance, but B3LYP also gave a good compromise between peak positions and spectral topology. Large basis sets were not mandatory to obtain high-quality spectra for the selected systems. The presence of a symmetry plane facilitated the computation of vibrationally broadened spectra, since different FC variants yield similar results and the harmonic approximation holds rather well. Compared with the AFC approach, the VFC approach performed equally well or even better for all three flavins while offering several advantages, such as avoiding error-prone geometry optimization procedures on excited-state surfaces. We also explored the advantages of curvilinear displacements and of a Duschinsky treatment for the AFC spectra in cases when a rotatable group is present on the chromophore. Taken together, our findings indicate that the combination of the VFC approach with the TD formalism and the IMDHO-FA model offers the best overall performance.



1. INTRODUCTION

Computational spectroscopy of chromophores is a powerful tool for interpreting the experimental UV−vis spectra in terms of underlying structural and vibrational properties. Hitherto, simulations of absorption and emission spectra of large molecules are often performed by comparing the energy differences between electronic states at a fixed geometry (vertical excitations) with the band maxima.[1] This approach disregards the important information provided by the spectral shape, which can be accessed by considering the vibrational progressions in the Franck−Condon (FC) region.[2,3] Based on the FC principle,[4−6] both time-dependent (TD)[7−15] and time-independent (TI)[16−24] treatments have been devised to account for vibrational progressions in optical spectra, each having their own advantages and limitations.[1] In the TD scheme,[7] the spectrum is computed as a Fourier transform of the autocorrelation function for wave packet propagation on the potential energy surface (PES) of the electronic state of interest. As the TD approach does not require calculation of the molecular eigenstates on the target surface, it is fast and can be directly used in anharmonic and nonadiabatic systems.[1] On the contrary, the TI scheme involves costly FC integral calculations using recursion formulas,[24] but it is more suitable for high-resolution spectra and their assignment, as individual contributions from each vibrational level are considered explicitly. The high cost of the TI approach can be reduced by use of efficient prescreening models.[18,19,22,25] Both TI and TD treatments can account for Herzberg−Teller effects (i.e., changes in transition dipole moments)[16,26] and can handle the problematic cases of anharmonicity and nonadiabaticity using special techniques.[1] At their computational limit, both frameworks are expected to yield identical spectra.[1] Apart from FC-based methods, other approaches, e.g., normal mode sampling using the Wigner distribution, are also available.[27,28]

1.1. Variants within the FC Approach. Both TD and TI approaches require the analysis of the potential energy surface of the ground and electronically excited states (GS and ES, respectively) of interest. In the widely used adiabatic FC (AFC) method, the vibrational characterization of the relevant states is done around the minimum geometry of the corresponding PES.[29] For large molecules, however, characterization of the ES PES is computationally demanding and sometimes unfeasible. Besides, following the gradients on an ES PES to locate minima is prone to root flipping and changes in state character due to crossing with other states. This may lead to a minimum far away from the FC region, which may no longer be a good basis for the representation of the FC region. To deal with this

Received: September 12, 2014
Published: November 3, 2014

   dx.doi.org/10.1021/ct500830a | J. Chem. Theory Comput. 2014, 10, 5549−5566

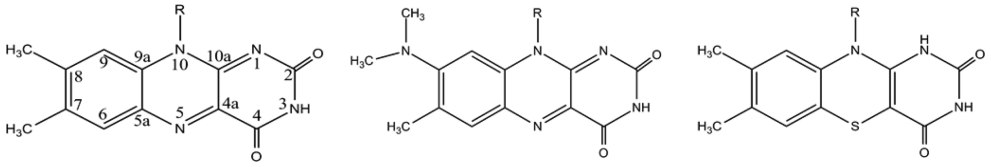


Figure 1. Schematic representation of the models used for three flavin derivatives: LF, RoLF, and 5TLF. The side chain consists of a methyl group (R = CH_3). LF has C_s symmetry, while RoLF and 5TLF have C_1 symmetry.

problem, the vertical FC (VFC) approach was introduced,[15,29−31] in which an artificial extrapolated minimum (EM) generated by a Newton−Raphson step from the FC point is utilized instead of the real minimum (RM) of the ES PES. In general, VFC overcomes problems related to gradient following and obeys the harmonic approximation by definition; however, it does not properly represent the vibrational character at the ES minimum. Since VFC is based on the second derivatives of the FC point, the vibrational analysis around an EM will likely lead to modes with imaginary frequencies. In the literature, three ways have been proposed to handle such modes: (a) simply excluding them from further analysis,[31] (b) replacing them with a real frequency (or its module),[29] or (c) performing costly one-dimensional PES scans to obtain a vibrational Hamiltonian to be used in spectrum calculation.[30]

Another noteworthy aspect is the choice of coordinates to represent the PES of interest and the corresponding Hessian. To date, Cartesian coordinates (describing rectilinear displacements) have been mostly utilized.[1,21,32] As an alternative, within the original framework of Wilson et al.,[33] Reimers introduced a generalized internal coordinate system (using curvilinear displacements) for generating adimensional shifts between two sets of normal modes, reorganization energies and, concomitantly, FC factors;[34] this scheme has been revisited by Borrelli and Peluso.[21] Curvilinear displacements have been shown to outperform rectilinear ones in generating AFC spectra for nonsymmetrical molecules that undergo large torsional displacements upon photoexcitation.[21,31,32,34−36] On the other hand, curvilinear coordinates do not offer any significant advantages over rectilinear ones when used within the VFC framework, because both descriptions are equivalent after the transformation to an orthonormal space.[31,37] However, for numerical reasons, an internal coordinate VFC spectrum may be slightly different from the corresponding Cartesian one.

The overlap between the vibrational wave functions of the GS and the target ES is decisive for the resolution and quality of the simulated FC spectrum. Application of the Duschinsky transformation[38] is expected to increase the overlap via alignment of two normal spaces (GS and target ES) by constructing the ES vibrational modes as a linear combination of GS modes. This becomes more crucial when substantial geometrical changes occur in the chromophore after photoexcitation. Duschinsky mode mixing has been implemented both for the TI-FC[16,34,39,40] and the TD-FC approach.[14,41]

1.2. Flavin Derivatives As Model Systems. Flavin derivatives have been used as active site probes for various types of enzymes since they display different spectral, chemical, and/or mechanistic properties than the enzymes with a native flavin.[42,43] In particular, flavin derivatives with red-shifted absorption have been studied to check whether the cellular processes induced by flavin photodynamics can be carried out at lower energies.[44] Among the myriad of known flavin derivates, riboflavin, roseoflavin, and 5-thioflavin (5TF) were chosen for this study due to their interesting structural properties, the availability of experimental UV−vis spectra, and their biological importance.

Riboflavin (RF, also known as vitamin B_2, lactoflavin) has an isoalloxazine moiety identical to the one present in flavin protein cofactors, flavin mononucleotide (FMN) and flavin adenine dinucleotide (FAD), which are present in the blue-light receptor domains of the cryptochrome, BLUF (blue-light sensors using FAD), and LOV (light oxygen voltage) families.[43] Roseoflavin (RoF) is a modified version of RF, with the methyl group at the isoalloxazine-C8 position replaced by a dimethylamino (DMA) group. This structural modification causes strong changes in the absorption and emission spectra[45] and in the biochemical properties. Whereas RF (or FMN/FAD) acts as protein cofactor, RoF is an antagonist and thus acts as an antibiotic against Gram-positive bacteria. Contrary to RF, RoF was also shown to undergo internal charge transfer (ICT) upon excitation, which is associated with the torsion of the DMA group.[45]

Given the biological importance and unusual photodynamics of these compounds, the optical properties of RF and RoF in various solvents and protein environments were subject to a number of theoretical[46−57] and experimental studies.[43,45,47,56,58−63] Some of the theoretical studies[49,53,56] addressed the vibrationally resolved (TI-FC) absorption spectra of lumiflavin (LF, truncated RF) in vacuum and various solvents, while most others focused on the vertical excitation energies of the respective chromophore (in comparison to measured band maxima). The previous theoretical analyses concentrated on the first electronically ES ($S_0 \rightarrow S_1$: $\pi_H \rightarrow \pi_L^*$), while the experimental spectrum consists of at least three separate bands.[45]

5TF has been subject to only a few experimental studies,[64,65] even though it has interesting photophysical properties because of the structural change in the isoalloxazine moiety (due to thio substitution at the N5 position of RF). This modification leads to a hypsochromic shift in the absorption to the far-violet region as compared to native RF.[65] 5TF is an intriguing but computationally demanding test system that is considered complementary to RF and RoF in this study.

To date, a number of comparative investigations[1,17,21,29,32,34,35,41,66] have been conducted on molecules of different size and in different environments to assess the merits of different approaches for generating vibrationally broadened spectra within the FC framework. The current study considers several low-lying electronic states to cover the complete one-photon absorption spectra of the chosen chromophores that have been measured experimentally. Our aim is to provide a thorough comparison of the performance of TD-FC vs TI-FC approaches in combination with the VFC or AFC scheme using

Cartesian/rectilinear or internal/curvilinear coordinates with or without accounting for Duschinsky rotation effects. Using time-dependent density functional theory (TD-DFT) for the electronic structure calculations, we present a comprehensive comparison of the spectra computed with the aforementioned methods with the available experimental results. We provide results for three flavin derivatives in different media to provide a basis for future analyses of vibrationally broadened spectra of other biologically relevant chromophores.

2. COMPUTATIONAL DETAILS

2.1. Electronic Structure Calculations. Since the ribose side chain is not taking part in flavin photophysics,[53,67] truncated models of the three flavin derivatives were used in computations: RF was modeled as lumiflavin, LF; RoF as roseolumiflavin, RoLF; and 5TF as 5-thiolumiflavin, 5TLF (see Figure 1).

All GS and ES geometry optimizations and the subsequent Hessian calculations were performed using TD-DFT[68−70] with the Gaussian 09 software.[71] For benchmarking purposes, four commonly used functionals, BP86,[72,73] B3LYP,[70] CAM-B3LYP,[74] and ωB97xD,[75] were combined with a Gaussian-type (TZVP)[76] and several Pople-type basis sets[77] (STO-3G, 6-31G(d), 6-31G++G(d,p), and 6-311++G(d,p)). No symmetry constraints were applied during geometry optimizations for any of the chromophores (thus allowing for side-chain torsions).

Single-reference DFT methods treat dynamic electron correlation efficiently,[78] whereas multireference configuration interaction (MRCI) provides a more accurate description of singlet states with double (or higher) excitation character and is capable of treating static correlation effects. To cover such cases, we applied the combined DFT/MRCI method[79] because of its good performance in the computation of vertical excitation energies of flavins (typical errors of 0.2 eV)[55] and in benchmarks for a large set of valence-excited states of organic molecules with $\pi\pi^*$ and $n\pi^*$ character.[78] At the GS geometry optimized at a given DFT level, we used Turbomole 6.3[80] to generate Kohn−Sham (KS) molecular orbitals (MOs) from single-point BHLYP/TZVP calculations. These MOs were utilized in the subsequent DFT/MRCI treatment, in which the lowest 20 roots were determined using standard parameters.[79] The resulting vertical excitation energies and transition dipole moments of the electronic excited states were corrected as follows (in a scheme similar to that employed by Jacquemin and co-workers[81]): all vibrationally resolved peak positions were shifted by the difference between the TD-DFT and DFT/MRCI vertical excitation energies, and the transition dipoles were adopted from DFT/MRCI.

Solvent effects were treated implicitly by the conductor-like polarizable continuum model (CPCM)[82,83] for water and benzene. In DFT/MRCI calculations, when required, solvent effects were taken into account via KS orbitals generated using the COSMO model[84] as implemented in Turbomole 6.3. For COSMO, the dielectric constants ε = 78.00 (water) and ε = 2.27 (benzene) were adopted.

2.2. Calculation of Vibrationally Resolved Spectra. Vibrationally broadened absorption spectra were simulated at near-zero (0.01 K) and finite (298 K) temperatures by the TD-FC and TI-FC methods, using ORCA (ORCA_ASA)[85] and a modified version of FCClasses.[20,86,87] The independent mode displaced harmonic oscillator model without/with frequency alteration (IMDHO/IMDHO-FA)[7] was utilized in the TD approach. We also adapted these models for the TI scheme to allow for an unbiased comparison. For VFC spectra, we used a slightly modified version of the IMDHO model,[31] with the VFC/IMDHO and AFC/IMDHO peak positions being corrected by the IMDHO-FA zero-point vibrational energy (ZPVE) difference between the GS and the target ES. This was done to allow for direct comparison of the spectral shapes from IMDHO to those from IMDHO-FA, for which this ZPVE correction is implicit. Vibrational frequencies, normal mode displacements, and structural properties were extracted from Gaussian 09 outputs and then used in the calculation of adimensional shifts and the related Duschinsky rotation. For AFC spectra, the shifts and Duschinsky rotation were obtained in either rectilinear or curvilinear coordinates with the DUSHIN program,[34] whereas for VFC spectra, they were calculated using a homemade Python script that implements the Newton−Raphson method as outlined previously.[31] By taking the same vectors for both TI and TD calculations, we introduce an implicit Duschinsky rotation of the shift vectors for the ORCA TD spectra as well, thus allowing for more direct comparisons between the TI (FCClasses) and TD (ORCA) approaches. The latter normally lacks Duschinsky rotation (see below). Note that we further employ an implicit Duschinsky rotation in our ORCA TD+VFC spectra for addressing the problem of imaginary modes by using the Duschinsky matrix elements as a selection criterion for mode deletion (see below). For generating TI spectra in FCClasses, we computed up to 10^7 FC integrals considering 60 quanta for C1 and C2 classes, unless stated otherwise. A minimum weight of 0.20 in the Boltzmann distribution was adopted for 298 K spectra.

We automated the procedure for creating FC spectra starting from Gaussian 09 output files for the GS and each ES. To this end, we compiled another Python script that interfaces all required software programs, allows the use of both correction schemes described above, and deals with imaginary modes in case of VFC. Capabilities and the detailed workflow of this Python script are given in Figure S1. In this way, any of the variants discussed above for the TI-FC or TD-FC approaches can be used (e.g., extrapolated or real ES minima, Cartesian/rectilinear, or internal/curvilinear coordinates, with or without application of the Duschinsky transformation). However, the explicit use of Duschinsky rotation in TD-FC spectra is not possible because it is not available in ORCA_ASA (while being available in other implementations of TD-FC).[14,41]

2.3. Quantitative Criteria to Assess Simulated Spectra. To measure the overlap between a computed and an experimental spectrum quantitatively, we applied the following procedure. The stick spectra were convoluted using Lorentzians (see Supporting Information (SI), Part A for an assessment of Lorentzian vs Gaussian broadening). At each convolution point the overlap of intensities in the two convoluted spectra was determined. The overall overlap (O) value was computed via

$$O = 1 - \frac{\sum_i^{N_{\mathrm{conv}}} \frac{|I_{1i} - I_{2i}|}{|I_{1i}| + |I_{2i}|}}{N_{\mathrm{conv}}} \quad (1)$$

where N_{conv} is the number of convolution steps, and I_1 and I_2 are the relative intensities in spectrum 1 and 2 at a given step. By definition, the highest possible value of O is 1 (in case of full overlap); O strongly depends on the choice of half-width at half-maximum (HWHM) used for broadening the spectra (see SI, part A). When comparing two stick spectra (see Section 3.1), we use a representative O value averaged over a selection of HWHM values for determining the degree of overlap. When

dx.doi.org/10.1021/ct500830a | J. Chem. Theory Comput. 2014, 10, 5549−5566

the experimental reference spectrum is already convoluted (see Section 3.2), we present a single O value that is computed using the HWHM value (403 cm^{-1}) for broadening that gives the best overlap between the two convoluted spectra.

3. RESULTS AND DISCUSSION

3.1. Benchmarking. Among the three flavin models investigated in this study, LF is most suitable for an elaborate benchmark. It has a symmetry plane (C_s) and does not exhibit large geometrical changes due to bends and torsions upon photoexcitation. The relaxed ES geometries are close to the FC point, with typical root-mean-square deviations (RMSD) of 0.01−0.03 Å from the GS geometry, which validates the harmonic approximation and enables the use of the AFC and VFC approaches in the TI and TD framework without any significant loss of accuracy. More importantly, there are recent low-temperature measurements of absorption and emission spectra for the lowest-lying optically bright state of LF in superfluid nano-He droplets with detailed band assignments.[56] The vibrational progression peaks[56] are not perturbed by finite-temperature and solvent effects and thus constitute an excellent reference for benchmarking different density functionals, basis sets, and FC-based computational approaches.

3.1.1. Functionals. It is well-known that the calculated spectral shapes and transition energies strongly depend on the chosen density functional (see ref 88 and references therein for benchmarks of different systems). Therefore, we start with a benchmark of functionals and compare the results to available experiments. We chose the range-separated hybrids CAM-B3LYP and ωB97xD in view of successful applications in similar systems (anthraquinones)[88] and the pure BP86 functional because of its ability to give realistic fundamental frequencies.[89] B3LYP was included as a popular general-purpose hybrid functional that has previously been used for investigating RF spectra.[49,53,55,90] The chosen functionals cover a series in terms of Hartree−Fock exchange, namely none (BP86), fixed (B3LYP), or range-dependent (CAM-B3LYP and ωB97xD). The convoluted TI-AFC absorption spectra of the first bright state of LF (LF/S_1) are plotted in Figure 2 along with the experimental spectrum,[56] and the corresponding stick spectra are shown in Figure S2. For a quantitative assessment, some technical details regarding the TI calculations are compiled in Table 1. The reference experimental spectrum is convoluted from the measured absorption peaks and their relative intensities, which are renormalized to account for the missing 0−0 peak (see ref 56, Table 1). As we are mainly interested in spectral progression, shape, and envelope width, the first peaks were aligned by equally shifting all peaks with respect to the reference 0−0 band. This scheme will be called 0−0-alignment (or 0−0-aligned) hereafter. It provides more reliable O values by eliminating any deviations due to 0−0 peak shifts. Correcting the 0−0 peak position in case of missing experimental reference data will be discussed later in this section.

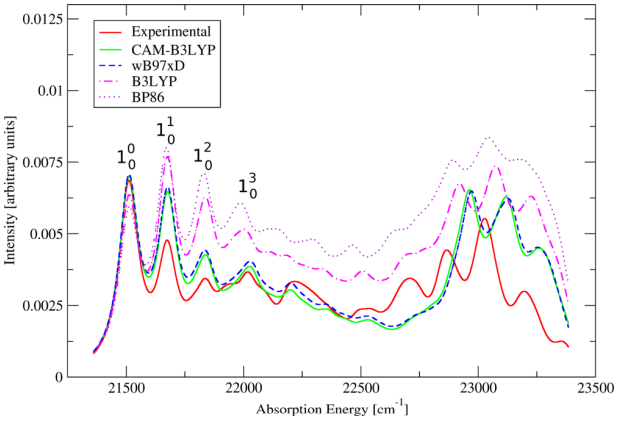


Figure 2. Comparison of LF/S_1 spectra in vacuum obtained from different density functionals with the 6-31G(d) basis set. TI-AFC spectra were computed at 0.01 K using the eclipsed conformation and Cartesian coordinates to generate the adimensional shifts and Duschinsky rotation. HWHM of 50 cm^{-1} for broadening. Spectra are 0−0-aligned. Some peaks are labeled for further discussion in the text.

Visual inspection of the spectra reveals the higher overall accuracy of the range-separated hybrid functionals CAM-B3LYP and ωB97xD, which are of almost equal quality in predicting the relative positions of the progression peaks, the band shape, and the envelope width (O difference of 0.2−6%, see Table 1). Similar high performance of CAM-B3LYP and ωB97xD was also found for anthraquinones.[88] In the low-energy domain (21000−22500 cm^{-1}), CAM-B3LYP and ωB97xD predict peak positions accurately, whereas the relative heights do not match the measured ones, e.g., the 1_0^1, 1_0^2, and 1_0^3 peaks are overestimated. The discrepancy in relative heights is likely due to the different widths of the origin band (1_0^0) and of other progression bands in the measured spectrum. For instance, the 1_0^0 and 1_0^1 peaks are expected to be of similar height, since they have almost equal FC factors (evident from the area under each peak in the measured spectrum), as also indicated by Vdovin et al.[56]

The mismatch in shape and position between the predicted and measured peaks is more pronounced in the high-energy region (22500−23500 cm^{-1}), which can be mainly ascribed to errors in the predicted fundamental frequencies. In SI, Part B, we give a detailed analysis of these errors for the CAM-B3LYP fundamental frequencies and the composition of the vibrational progressions; compared with experiment, CAM-B3LYP overestimates the frequencies in the high-energy region. The use of different scale factors for modes with different character (e.g., stretch, bend, and torsion), as done in previous studies (see ref 32 and references therein), would be one possible solution to this problem but would also introduce some arbitrariness into the comparisons. The common approach of frequency scaling using only a single scaling factor (see ref 91 and references therein) apparently leads to small, unsystematic changes (O values changing by <1%, see Table S1). Therefore, we only present spectra without frequency scaling, unless stated otherwise.

On the other hand, BP86 and B3LYP are giving a qualitatively different picture in terms of spectral shape (Figure 2, Table 1). BP86 assigns the 1_0^1 excitation peak to be the maximum and generally overestimates the relative heights of progression peaks. In the intermediate region (21900−22600 cm^{-1}), BP86 locates the progression peaks inaccurately, but it is superior to the other functionals with regard to the peak positions and topology (except relative intensities) in the high-energy region. The latter finding is contrary to expectation since a broad benchmark on functionals[92] has shown a rather limited accuracy of pure functionals for band shape predictions. In the present case, there may be some fortuitous error cancelation arising from an underestimation of the harmonic

Table 1. Summary of Results From the TI-FC Calculations with Different Density Functionals and the 6-31G(d) Basis: Spectral Yields (C), Reorganization Energies on the GS and ES Surfaces (λ, eV), and Overlap of Intensities (O, see eq 1)[a]

	VFC			
	BP86	B3LYP	CAM-B3LYP	ωB97xD
C	0.632	0.835	0.853	0.856
$\lambda_{GS}/\lambda_{ES}$	0.34/0.26	0.27/0.25	0.25/0.26	0.25/0.26
O	0.632 ± 0.026	0.752 ± 0.038	0.868 ± 0.058	0.860 ± 0.058
	BP86	B3LYP	CAM-B3LYP	ωB97xD
	AFC - Eclipsed Conformation			
$RMSD^{Cart}$	0.02840	0.02691	0.02270	0.02386
$RMSD^{Bonds}$	0.02085	0.01970	0.01960	0.01933
$RMSD^{Ang}$	1.50094	1.40140	1.26334	1.26639
$RMSD^{Dihed}$	0.16899	0.15020	0.13524	0.12386
C_{rect}	0.816	0.860	0.947	0.925
C_{curv}	0.816	0.857	0.947	0.926
Rect: $\lambda_{GS}/\lambda_{ES}$	0.31/0.25	0.29/0.24	0.28/0.26	0.27/0.26
Curv: $\lambda_{GS}/\lambda_{ES}$	0.31/0.25	0.29/0.24	0.28/0.26	0.27/0.26
O_{rect}	0.706 ± 0.034	0.782 ± 0.036	0.866 ± 0.066	0.864 ± 0.066
O_{curv}	0.706 ± 0.034	0.778 ± 0.036	0.866 ± 0.066	0.864 ± 0.066
	AFC - Staggered Conformation[b]			
$RMSD^{Cart}$	0.18601	0.18480	0.18356	0.18385
$RMSD^{Bonds}$	0.02005	0.03061	0.01923	0.01903
$RMSD^{Ang}$	1.58204	2.02544	1.36701	1.38088
$RMSD^{Dihed}$	17.56959	17.56787	17.56899	17.56883
C_{rect}	0.000	no spectrum	0.000	0.000
C_{curv}	0.016	0.055	0.170	0.091
Rect: $\lambda_{GS}/\lambda_{ES}$	12.00/11.77	12.86/12.42	12.60/12.40	12.62/12.38
Curv: $\lambda_{GS}/\lambda_{ES}$	0.50/0.15	0.94/0.35	0.48/0.16	0.43/0.12
O_{rect}	0.156 ± 0.020	no spectrum	0.156 ± 0.020	0.092 ± 0.003
O_{curv}	0.106 ± 0.030	0.154 ± 0.032	0.202 ± 0.024	0.140 ± 0.015

[a]For the eclipsed and staggered AFC minimum conformations, results are given both for rectilinear and curvilinear coordinates as well as RMSD values (RMSD values between GS and ES minima covering 31 atoms, 33 bond lengths, 56 angles, and 70 dihedral angles, respectively) in terms of Cartesian coordinates ($RMSD^{Cart}$, Å), bond lengths ($RMSD^{Bonds}$, Å), angles ($RMSD^{Ang}$, degrees), and dihedral angles ($RMSD^{Dihed}$, degrees). Duschinsky treatment is included. See text for further details. [b]A small imaginary frequency was found for the staggered conformation and excluded from spectrum calculation analogous to the VFC scheme.

frequencies by BP86, as discussed in depth in previous work (ref 89 and references therein). While probably useful in a large-scale approach due to its speed (when combined with the RI approximation), it seems that for LF/S_1, BP86 gives partly good answers for wrong reasons.

Like BP86, B3LYP predicts the maximum peak to be the 1_0^1 excitation, in contradiction to the experimental findings (where 1_0^0 is the maximum peak). Overall, the spectrum generated by B3LYP overlaps better with experiment than that produced by BP86 (higher O values). In the low-energy region, peaks are located well, but relative intensities are slightly overestimated compared to experiment and to range-separated hybrids. However, all the peaks in the intermediate region are smeared out, preventing a proper analysis of the vibronic structure. B3LYP provides the best accuracy in the high-energy regime in terms of peak positions and intensities compared to the other tested functionals, but it can still not reproduce these features perfectly. Taken together, in spite of the lower quality compared to range-separated hybrids, B3LYP does overall a good job for the gas-phase absorption spectrum of LF/S_1. This is in line with a benchmark on similar systems, namely acenes, which concludes that a functional with 20−30% HF exchange is needed to generate computed spectra of good quality in comparison to experiment.[92] We note that the broadened spectrum here is in perfect agreement with the one presented by Klaumünzer et al.[49] (see Figure S3 for spectra at the B3LYP/TZVP level).

3.1.2. Basis Sets. Based on the above-mentioned findings and other TD-DFT benchmarks,[81,88,93,94] we chose CAM-B3LYP for all following investigations (unless stated otherwise), even though ωB97xD would be equally suitable. As next step, we conducted a benchmark of basis sets, covering STO-3G, 6-31G(d), 6-31++G(d,p), 6-311++G(d,p), and TZVP. The computed convoluted 0−0-aligned spectra are compared to the experimental spectrum in Figure 3, and some technical details regarding the TI-AFC calculations are compiled in Table S2. Overall the choice of basis set is apparently not critical (changes in O values are minute), except that the minimal STO-3G basis completely fails in the high-energy region, as in the case of anthraquinones.[88] In line with the fast basis set convergence for DFT methods,[95] extensions of the basis beyond 6-31G(d) only slightly alter the relative peak heights in the low-energy regime and also correct the peak positions in the high-energy region only to a small extent. The changes in the O value are <2% so that one can readily avoid the high cost of larger basis sets. Therefore, the 6-31G(d) basis was used for the rest of the investigation.

3.1.3. ES Geometry. LF has two C_s symmetry conformers that differ in the relative positions of the two methyl groups at the C7 and C8 positions (eclipsed and staggered, see Figure 4, upper panel). For all four chosen functionals, the eclipsed

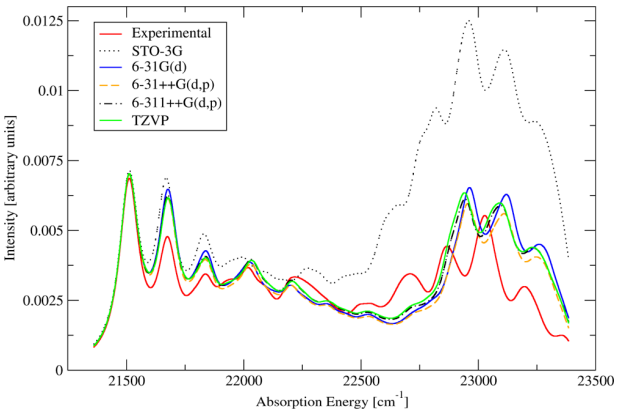


Figure 3. TI-AFC spectra from CAM-B3LYP calculations with different basis sets. Computed spectra are 0−0 aligned. HWHM value of 50 cm^{-1} for convolution of the stick spectra. Spectra for the eclipsed geometry, Duschinsky treatment included.

conformer is the minimum structure for the GS and first ES (S_1), while the staggered one is a transition state (TS) with a single imaginary-frequency mode. CAM-B3LYP predicts a very low barrier connecting two conformers (0.06 eV, see Figure S5), allowing easy access to both. At the CAS(8,8)/6-31G(d,p) level, the staggered conformer is predicted to be the minimum for S_1 and the TS for the GS.[56] The FC factors from the normal modes computed at this level seem to agree well with the experimental vibrational progressions.[56] To check whether this is also the case for DFT, we plotted AFC spectra using the geometries with staggered and eclipsed conformations in Figure 4 (lower panel). We also compiled some technical details for TI-spectra generation in Table 1.

Evidently, the AFC spectra in Figure 4 are almost featureless for the staggered conformation, whereas those for the eclipsed conformation are well structured. Accordingly, we find low overlaps (see *O* values) and very low yields (recovered fraction of the FC integrals, *C* values) for the staggered conformation as opposed to the eclipsed one (Table 1). The main reason for the low yield is the low overlap between the GS and ES vibrational wave functions, which is due to large structural change in going from the eclipsed GS minimum to the staggered S_1 "minimum" (see Table 1, RMSD values) that also leads to very high reorganization energies. Another intriguing feature is the much higher yield when the curvilinear instead of the rectilinear representation is used in the computation of the adimensional shifts and the Duschinsky rotation for the staggered conformation, which reflects the better description of the torsion of the methyl groups in the curvilinear representation.

In cases of substantial geometrical change, the Duschinsky transformation helps to increase the 0−0 overlap and, concomitantly, the yield of the TI-FC spectra. In the specific case of staggered LF/S_1, it was not possible to obtain the TI-FC spectra without the Duschinsky treatment (data not shown). Nonetheless, even with the Duschinsky rotation, the resulting spectra are not accurate; even the one with the highest yield (i.e., CAM-B3LYP, C = 0.170) produces rather diffuse FC factors (see stick spectra, Figure S6) so that the overall shape is featureless (Figure 4) and the overlap is very low (Table 1). Using higher accuracy, with a total of 10^9 FC integrals, in the generation of the TI-AFC spectrum of the staggered conformation increases the yield (C = 0.374) but does not affect the resulting spectra much, as indicated by almost unchanged average *O* values (data not shown). The short-time TD formalism implemented in ORCA does not afford any noticeable improvement to the spectral resolution, and the TD spectra differ from the TI spectra only by a slight red shift.

At the eclipsed geometry, the choice of the FC method (TI vs TD and IMDHO vs IMDHO-FA) and of the coordinate representation does not alter the resulting spectra (Figure 4 and Table 1). This is in line with our expectations, as the change in normal modes upon photoexcitation is subtle because the geometry changes only slightly. All FC methods reproduce the three peaks and the shoulder that are red-shifted with respect to the broad peak for the staggered conformation (lower adiabatic shift). Application of the Duschinsky transformation causes

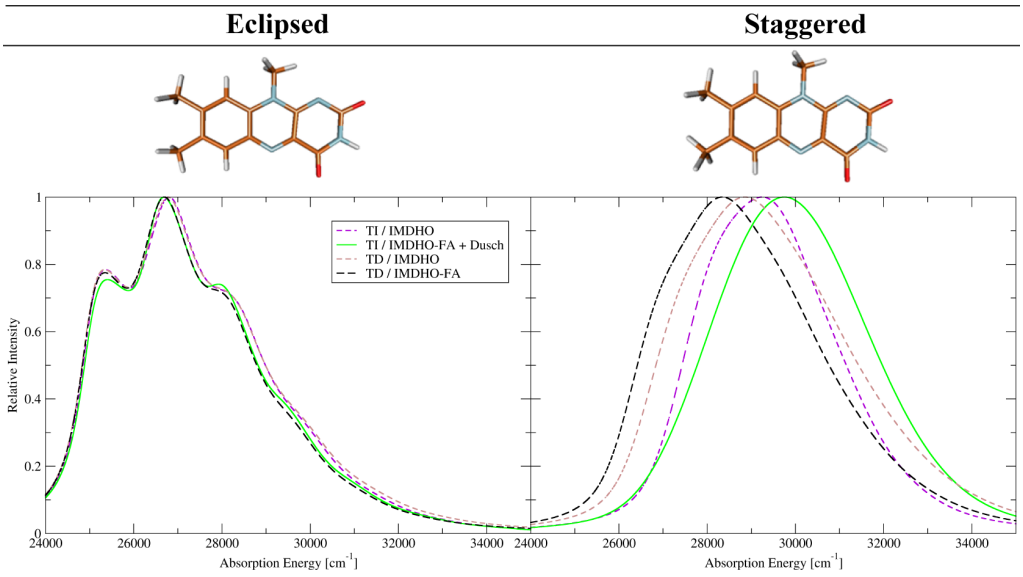


Figure 4. (Top) Stick representations of the eclipsed and staggered conformations of LF. (Bottom) CAM-B3LYP/6-31G(d) AFC spectra of LF/S_1 (relative intensities) using the curvilinear representation at these conformations.

dx.doi.org/10.1021/ct500830a | J. Chem. Theory Comput. 2014, 10, 5549−5566

some changes in the relative heights of the first and third peak, with the tendency to equalizing them.

Evidently, using a "suitable" input geometry (e.g., the eclipsed geometry in the LF/S_1 case) and the corresponding set of vibrational modes is crucial to generate reliable AFC spectra that are comparable to experiment. AFC does not give the correct spectral shape in the case of the staggered LF/S_1 geometry, which is not a minimum on the corresponding PES. The VFC approach may bypass the problem of wrong state character/input geometry by using an EM geometry that is located in close proximity to the FC point in the normal space, despite potentially introducing imaginary frequency modes (see Figure S7 for different possibilities to treat this problem). A more accurate solution to this problem would be the explicit consideration of anharmonic corrections to the vibrational frequencies and energy levels.[30,66,89] However, considering the size of the chromophores studied presently, we opted to discard the imaginary-frequency modes, in order to limit the computational efforts and to avoid introducing additional unphysical peaks.

3.1.4. FC Variants. In order to compare the performance of VFC and AFC, we now consider only the results for the eclipsed conformation. The details of the TI-FC computations for LF/S_1 are given in Table 1 for all employed density functionals. As a specific example, the TI and TD spectra at the CAM-B3LYP/6-31G(d) level are plotted in Figure 5. For the TI spectra, we note that the overall differences in O values (<3%, Table 1) between VFC and AFC are small for all functionals except BP86 (ca. 7%). The similar performance of VFC and AFC can be attributed to the similarity of the geometries at the FC point and the ES minimum and of the corresponding normal mode sets. A closer look at the CAM-B3LYP TI spectra (Figure 5) reveals excellent agreement between VFC and AFC in the low-energy region (21400–22000 cm^{-1}). In the intermediate region (22000–22600 cm^{-1}), VFC provides more accurate peaks in terms of position and relative intensity. In the high-energy regime, the measured spectral shape is almost perfectly reproduced by VFC, albeit hypsochromically shifted by about 200 cm^{-1}, in line with the aforementioned tendency of CAM-B3LYP to overestimate the fundamental frequencies in the higher-energy regime (vide supra, and SI, Part B).

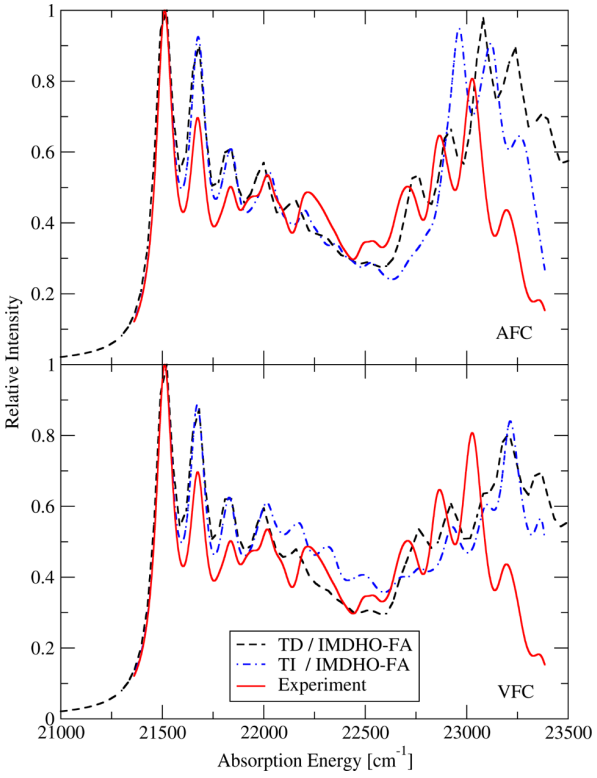


Figure 5. Comparison of 0–0-aligned TD-/TI-AFC spectra (upper panel, eclipsed minimum) and TD-/TI-VFC spectra (lower panel) with the experimental spectrum.[56] The calculations were done at the CAM-B3LYP/6-31G(d) level using the IMDHO-FA model. HWHM = 50 cm^{-1} for Lorentzian broadening.

To enable direct comparisons, TD spectra of LF/S_1 within the AFC and VFC frameworks are also shown in Figure 5. The TI and TD approaches yield almost identical AFC spectra in the low- and medium-energy regions (Figure 5, up to 22600 cm^{-1}). In the high-energy section, TD performs better by predicting the number of peaks and the spectral shape correctly, but it still fails in reproducing the relative heights. The shortcomings in the high-energy region may result from the short-time approximation or the lack of an explicit Duschinsky rotation in the TD formalism implemented in ORCA.[15] The VFC spectra from TI and TD show enhanced agreement in the high-energy regime but larger deviations in the intermediate region.

3.1.5. 0–0 Peak Position. Another important aspect in the calculation of broadened spectra is the positioning of the 0–0 peak; the spectral envelope mainly depends on the adiabatic excitation energy. As seen in Table 2 (visualized in Figure S2), B3LYP predicts a well-positioned 0–0 band for LF/S_1 (21720 cm^{-1}) that is very close to measured band origin (21511 cm^{-1}). The other functionals, especially range-separated hybrids, yield 0–0 peaks off by more than 2500 cm^{-1}. Computed 0–0 peak positions can be refined by including ZPVE corrections[81,88] and by accounting for solvent effects using state-specific solvent corrections.[96−98] In the absence of experimental data for the 0–0 peak, one may introduce further theoretically motivated corrections. Here, we use an approach called MRCI-correction scheme (vide supra), in which the difference between the DFT/MRCI and TD-DFT vertical excitation energies for the desired bright state is added to the TD-DFT adiabatic excitation energy for that state. In the LF/S_1 case, it is actually trivial to select the matching states from the TD-DFT and DFT/MRCI calculations. However, in more complicated cases involving higher excitations, the contributing KS orbitals may need to be analyzed for a correct match.

As different functionals yield different GS minimum geometries, the DFT/MRCI vertical excitation energies at those minima differ noticeably (see Table 2). As can also be noted from Table 2, the MRCI-corrected 0–0 peak positions (from AFC spectra) for CAM-B3LYP and ωB97xD are in quite good agreement (within 0.03 eV, 250 cm^{-1}) with the measured one (21511 cm^{-1}), whereas the BP86 value deviates strongly. In contrast to the other functionals, the MRCI-corrected 0–0 band for B3LYP is inferior to the original one (20591 cm^{-1} vs 21720 cm^{-1}) that was already aligned well with the measured peak. The DFT/MRCI wave function for the LF/S_1 state has only about 90% single excitation character, implying that the B3LYP 0–0 peak position is likely right for the wrong reason: linear response TD-DFT enforces 100% single-excitation character and should therefore give slightly blue-shifted

dx.doi.org/10.1021/ct500830a | J. Chem. Theory Comput. 2014, 10, 5549–5566

Table 2. Vertical Excitation Energies (E_{vert}, eV) and Oscillator Strengths (f) from DFT and DFT/MRCI Calculations at the GS Minimum of the Corresponding Level of Theory[a]

	DFT			MRCI					
functional	state	f	E_{vert}	state	f	E_{vert}	ΔE_{vert}	0−0 peak w/out corr.	0−0 peak with corr.
BP86	S_2	0.12	2.61	S_1	0.31	2.80	−0.19	18322	19853
B3LYP	S_1	0.19	3.04	S_1	0.32	2.90	0.14	21720	20591
CAM-B3LYP	S_1	0.29	3.46	S_1	0.32	2.99	0.47	25027	21252
ωB97xD	S_1	0.29	3.44	S_1	0.32	2.98	0.46	24958	21256
								experiment[b]	21511

[a]Energy differences (ΔE_{vert}, eV) are used in the MRCI-correction scheme. Positions (in cm^{-1}) of the original and MRCI-corrected 0−0 peaks are given along with the experimental band origin. [b]Taken from ref 56.

excitation energies. This effect is more clearly seen for larger systems, e.g., carotenoids.[98−100]

3.2. Multistate Absorption Spectra of Three Flavins in Different Media. Having provided an assessment of functionals, basis sets, and FC variants for one particular ES (LF/S_1), we now continue our evaluation on a broader scale. In this section, we assess the performance of the different FC variants in reproducing the complete experimental absorption spectra, involving several bright states, of three chromophores, LF, RoLF, and 5-TLF in vacuum, a nonpolar solvent (benzene), and a polar solvent (water). The comparisons will mainly address the vibrationally broadened spectra obtained at the CAM-B3LYP/6-31G(d) level; B3LYP and ωB97xD results are also considered in some specific cases. As the experimental 0−0 peak positions are not known (as opposed to the LF/S_1 case, see Section 3.1), we applied the MRCI-correction scheme (vide supra).

In this section, we also report O values for computed spectra with respect to the corresponding reference experimental spectra (where available). These O values differ from those presented in Section 3.1 in that they are computed using a single HWHM value (403 cm^{-1}). This was necessary since only convoluted experimental spectra are available, rather than stick spectra that can be broadened using different HWHM values. The single HWHM value was chosen to give the highest overlap between any given pair of computed and experimental spectra. To minimize the effect of 0−0 peak shifts, the O values were further optimized by moving the computed spectrum as a whole after convolution of different bands (denoted as "optimized O values" in the captions of related figures).

3.2.1. LF Spectra. Figure 6 shows (a) spectra obtained from Gaussian-broadened DFT/MRCI vertical transition energies, computed at the CAM-B3LYP GS minimum and scaled by their oscillator strengths, and (b) the CAM-B3LYP Lorentzian-broadened FC spectra (absorption cross sections, in cm^2) for the following options: TI-FC or TD-FC approaches; IMDHO or IMDHO-FA models; with or without Duschinsky rotation; rectilinear or curvilinear coordinates; in the gas phase, benzene, and water at 0.01 K. In addition, the absorption cross sections of RF in benzene and water, measured at room temperature by Zirak et al.,[45] are included in Figure 6 for a direct comparison using the data points kindly provided by Prof. Penzkofer. As an alternative representation, overlays of these spectra are shown in Figures S9−S11 separately for each medium. Only transitions to selected low-lying ESs were included in the FC spectra; a full list of these states is given in Table S3 along with the values used for the MRCI-correction scheme.

RF has three different protonation states in aqueous solutions depending on the pH value:[62] fully oxidized/neutral, reduced/ionic, and radical semiquinone. We only considered the neutral form of LF, as the reference experimental spectrum[45] was taken at pH = 7 at which the neutral form is dominant. The latter also applies to RoLF (vide infra). The experimental spectra show three distinct bands in both solvents, which are assigned to four different bright states both at the CAM-B3LYP and DFT/MRCI level. The FC spectra from CAM-B3LYP reveal that the first two observed bands (maxima at ca. 450 nm and ca. 350−360 nm) are dominated by a single electronic excitation, whereas the third band (at ca. 290 nm) is a combination of two separate excitations with intermediate to strong brightness (see Table S3 for state properties). The first band results from the vibrational progression following the $\pi_H \rightarrow \pi_L^*$ transition. It is well resolved in the experimental spectrum in benzene, showing two characteristic peaks (at 425 and 455 nm) and a shoulder (at 490 nm). The common approach of Gaussian broadening the vertical transition "sticks" seems to provide none of these resolved features (Figure 6, "E_{vert} Broadened"), whereas the FC spectra consist of highly structured bands, whose positions and relative heights are predicted quite accurately.

Upon photon absorption the C_s symmetry in LF is preserved, which prevents large geometrical displacements and keeps the chromophore in close vicinity of the FC point. Therefore, the FC principle holds for each electronic state, and the choice of IMDHO/IMDHO-FA has little effect on the computed FC spectra in any environment (O values deviating by less than 1%, see Figure 6). Duschinsky mixing provides a subtle improvement in some cases (e.g., VFC/gas phase and AFC/water) as compared to FC spectra for which the Duschinsky matrix is set to unity; yet, it does not affect the overall topology. Likewise, the AFC spectra of LF generated using rectilinear coordinates are identical to their counterparts from curvilinear coordinates in all three media (see Figures 6 and S9−S11).

As seen in Figure 6, most FC methods reproduce the experimental spectra of RF in benzene and water sufficiently accurately in terms of relative heights, positions, and band shapes. However, in the case of benzene, the relative intensities of the vibronic progressions in the first band are not predicted as well as their positions, with the shoulder being of almost equal height as the central peak, which cannot be remedied by increasing the precision in the TI computation (10^{10} FC integrals); B3LYP yields almost the same topology for the first band (Figure S12). The other two bands in benzene and all three bands in water have less features, and their shapes are predicted more accurately both by CAM-B3LYP and B3LYP.

The MRCI-correction rectifies the CAM-B3LYP band positions and relative heights for the AFC and VFC spectra in water (compared with experiment), whereas in benzene both spectra remain slightly blue-shifted. The former is likely due to a fortuitous error cancellation, since a state-specific solvent

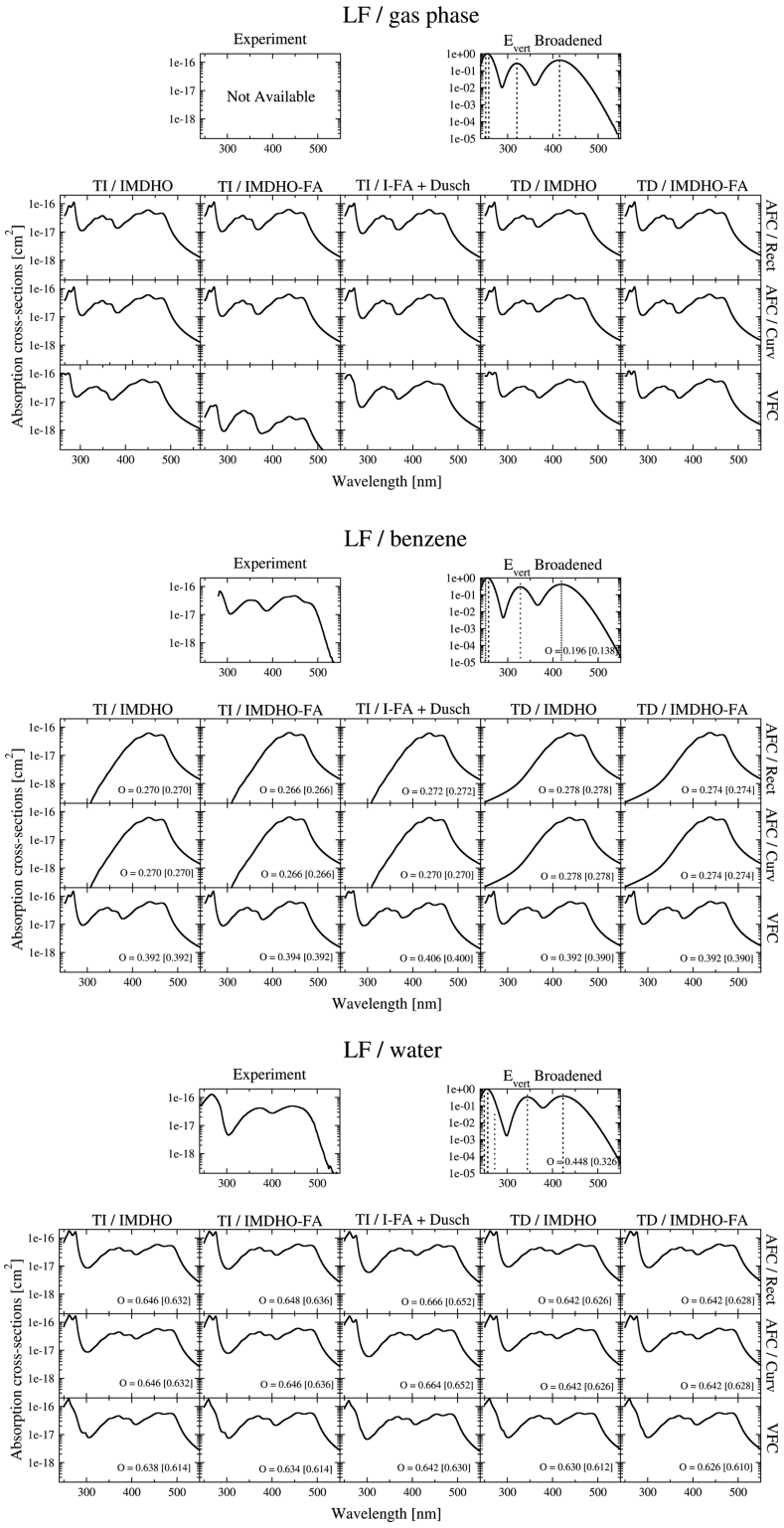


Figure 6. Convoluted vertical absorption spectra ("E_{vert} Broadened") along with TI-AFC, TI-VFC, TD-AFC, and TD-VFC spectra of LF computed using IMDHO or IMDHO-FA models at 0.01 K in the gas phase, benzene, and water; compared to the experimental spectra at 298 K (benzene and water, taken from ref 45). A Duschinsky treatment was applied only for TI/IMDHO-FA spectra. Optimized [original] O values are given for the spectra with an experimental reference. All computed spectra are MRCI-corrected. Spectra based on vertical transition energies were Gaussian broadened with an HWHM of 1500 cm^{-1}; the FC spectra were Lorentzian broadened with an HWHM of 403 cm^{-1}.

correction scheme (as outlined in refs 96−98) is required to account for the fast solvent relaxation following photon absorption. As evident from Table 2, the DFT/MRCI transition energies and oscillator strengths are better reproduced by

dx.doi.org/10.1021/ct500830a | J. Chem. Theory Comput. 2014, 10, 5549−5566

B3LYP than by CAM-B3LYP. The MRCI-corrected B3LYP spectra are slightly less accurate than the original ones, as in the case of the gas-phase LF/S_1 spectra (see Section 3.1). For both benzene and water, the original B3LYP VFC spectra are in very good agreement with the measured spectra (see Figures S16–S17), and they do not seem to benefit from the MRCI-correction scheme. CAM-B3LYP and B3LYP topologies agree well, with B3LYP mimicking the second band better. In summary, B3LYP appears to perform slightly better than CAM-B3LYP in this case.

Having a symmetry element simplifies the vibrational spectrum calculation and removes the discrepancies between various FC methods and coordinate sets.[31] Yet, it does not guarantee to ease the process of locating an RM on the ES PES. Accordingly, for LF in benzene, we could not find the RM for the S_4 and S_9 states with CAM-B3LYP or with B3LYP even though this was possible in the gas phase. The ES minima in the gas phase and in water greatly resemble the corresponding GS minimum (with a very low RMSD of <0.015 Å in each case). Nonetheless, the existence of (almost) degenerate states leads to a stationary point with an imaginary-frequency mode on the surface described by a single-reference method. The VFC scheme produces spectra almost identical to those from AFC (Figure 6, except for benzene), without requiring a costly full optimization.

The vibrationally broadened spectra presented in Figure 6 were prepared at 0.01 K and thus considered only the ground vibrational state in the ground electronic state. As the experimental reference spectra were taken at 298 K, finite-temperature effects should also be probed. However, due to the size of the chosen model systems, the total number of FC integrals needed to simulate TI spectra at finite temperature (T = 298 K) rises very rapidly and often cannot be handled within software limitations. Here, we had to limit the maximum number of integrals to 10^6 (previously 10^7) and to set the minimum weight for inclusion in the Boltzmann distribution to 0.20. This alleviates the total absorption intensity, and to allow for a fair comparison we thus also compiled the TI/IMDHO spectra at 0.01 K using these options. They are shown in Figure 7 along with the TD/IMDHO and the experimental spectra. We could study only IMDHO at 298 K, since this is the only model compatible with a temperature option for the TD formalism implemented in ORCA. The TD spectra at 0.01 and 298 K show no significant discrepancies, since those at 298 K are obtained simply by applying a larger broadening from the original autocorrelation function computed at 0.01 K. In the TI spectra, the higher vibrational states of the ground electronic state are explicitly taken into account, and the TI spectra at 298 K thus differ noticeably from those at 0.01 K (Figure 7). While the shape of the second band in benzene and water is slightly improved in TI spectra at 298 K, the inaccurate topology of the first peak in benzene (as previously discussed) cannot be amended. The latter may be attributed to missing contributions from an H-bonding network, which can only be captured by explicit consideration of solvent molecules. This will be shown to be the case for RoLF (see SI, part C, and Section 3.2.2).

3.2.2. RoLF Spectra. The dimethylamino substitution in RoLF leads to absorption spectra which differ significantly from those of RF, as evident from measured absorption cross sections[45] (Figure 8). Figure 8 also shows (a) spectra obtained from Gaussian-broadened DFT/MRCI vertical transition energies, computed at the CAM-B3LYP GS minimum and

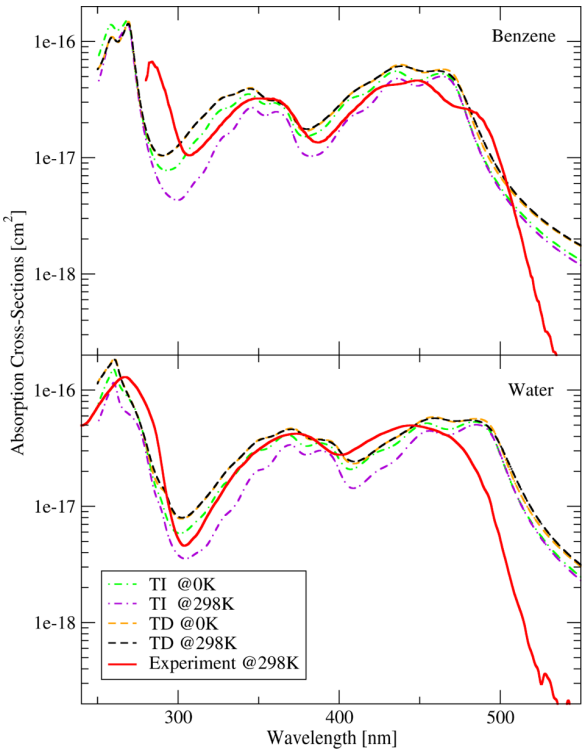


Figure 7. Temperature effects: MRCI-corrected TI/IMDHO-FA (with Duschinsky mixing) and TD/IMDHO VFC spectra of LF computed at T = 0.01 and 298 K in benzene (top) and water (bottom) in comparison to the measured spectra at T = 298 K (taken from ref 45). The accuracy parameter for TI spectra is 10^6, see text.

scaled by their oscillator strengths, and (b) MRCI-corrected CAM-B3LYP Lorentzian-broadened FC spectra, with absorption cross sections in cm^2. The latter were computed with the following options: TI-FC or TD-FC; IMDHO or IMDHO-FA models; with or without Duschinsky rotation; rectilinear or curvilinear coordinates; in the gas phase, benzene, and water at 0.01 K. Overlays of these spectra are given in Figures S18–S20 separately for each medium to provide a different representation. Only transitions to selected low-lying ESs were included in the FC spectra; a full list of these states is given in Table S4 along with the values used for the MRCI-correction scheme. As evident from Figure 8, the AFC spectra of RoLF using rectilinear and curvilinear coordinates agree almost perfectly (almost identical O values) in the gas phase and benzene but not in water. This will be analyzed later in this section.

The measured absorption spectrum of RoLF in benzene consists of two main bands located at 490 and 290 nm, with a shoulder (330–430 nm) hidden in the broad first band. The naive approach of broadening of the vertical excitation sticks ("E_{vert} Broadening") fails almost completely in emulating the overall topology of the measured absorption cross sections (Figure 8, benzene part). The AFC approach can only partly reproduce the experiment (see Figure 8, especially around 290 nm), because some of the bright electronic states that contribute to the measured absorption bands are not included in the AFC spectra since the corresponding RM geometries are not available (see Table S4). The VFC approach provides more complete spectra by accounting for more bright states.

As apparent from Figures 8 and S19, the MRCI-corrected AFC and VFC spectra in benzene are slightly blue-shifted in

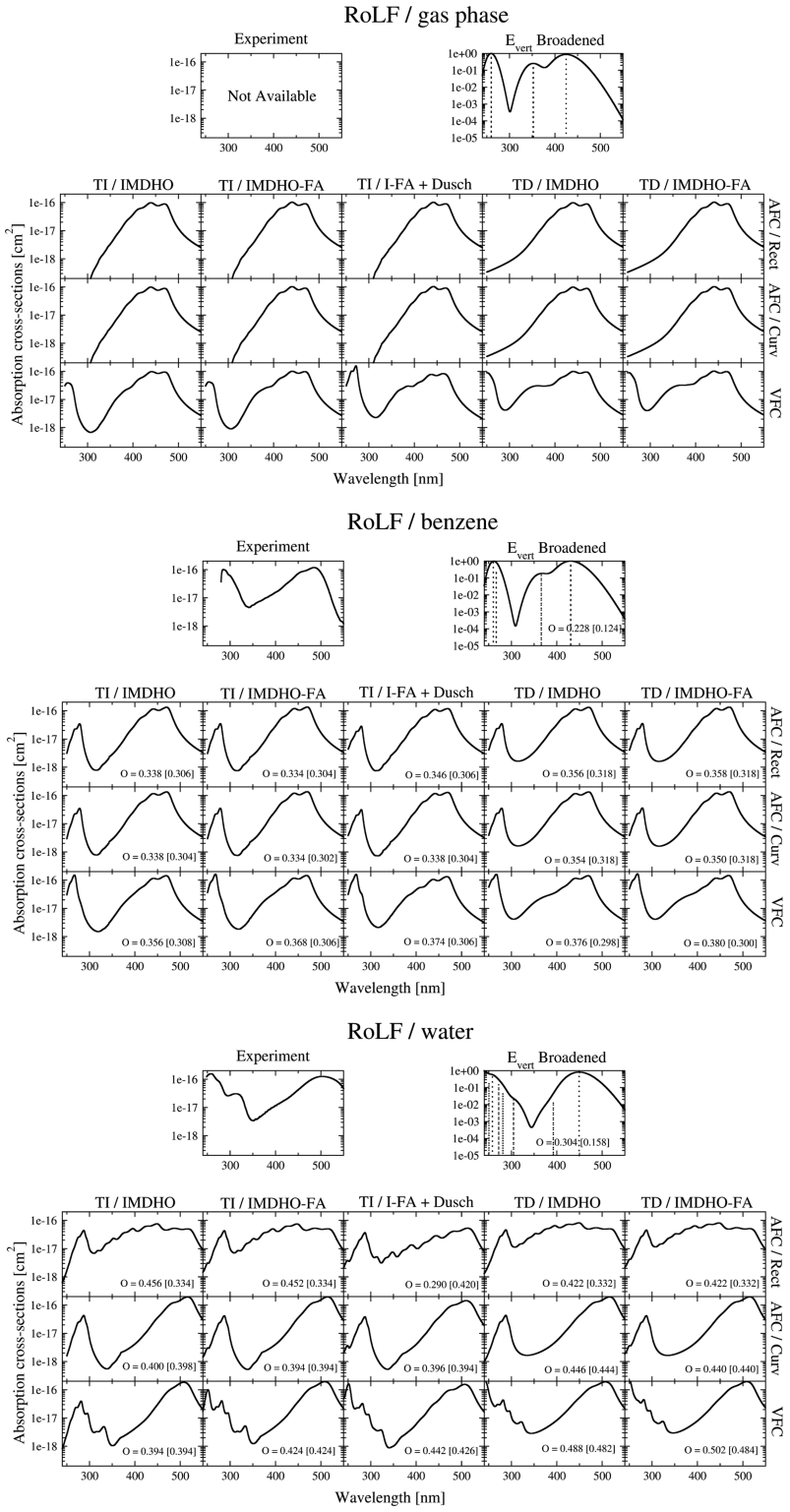


Figure 8. Convoluted vertical absorption spectra ("E_{vert} Broadened") along with TI-AFC, TI-VFC, TD-AFC, and TD-VFC spectra of RoLF computed using IMDHO or IMDHO-FA models at 0.01 K in the gas phase, benzene, and water; compared to the experimentally measured spectra at 298 K (benzene and water, taken from ref 45). A Duschinsky treatment was applied only for the TI/IMDHO-FA spectra. Optimized [original] O values are given for spectra with an experimental reference. All computed FC spectra are MRCI corrected. Spectra based on vertical transition energies were Gaussian broadened with an HWHM of 1500 cm^{-1}; the FC spectra were Lorentzian broadened with an HWHM of 403 cm^{-1}.

comparison to the experiment, like in the case of LF, underlining once again the need for solvent correction schemes.

The benzene spectra are slightly red-shifted compared to the gas-phase spectra (Figure S18) in line with the lower transition

dx.doi.org/10.1021/ct500830a | J. Chem. Theory Comput. 2014, 10, 5549−5566

energies arising from the relaxation of the ESs in the solvent environment.[98] The benzene and gas-phase spectra also differ only slightly in terms of band topology. As opposed to the LF case, the choice of FC method in RoLF affects the resulting spectra to a noticeable extent. As a specific example, the tail part of the first band (300−350 nm) in the gas phase and benzene is different in the TI and TD spectra for both AFC and VFC (see Figures 8, S18, and S19). For the AFC method, the choice of model (IMDHO, IMDHO-FA with/without Duschinsky rotation) does not affect the outcome significantly, whereas for the VFC approach different models yield different spectra.

The experimental absorption spectrum in benzene is best reproduced by the TD-VFC/IMDHO-FA combination, having the highest O value (0.380, see also the VFC spectra in Figures 8 and S19 for a qualitative comparison). In detail, the progression tail of the first band (300−350 nm) as well as the overall shape and relative intensities are computed accurately with this combination, despite the aforementioned minute blue-shift. The TD/IMDHO combination performs almost equally well (O = 0.376). Closer inspection of the TI and TD spectra reveals that the dissimilar prediction of the shoulder of the first band can be associated with the $S_0 \rightarrow S_3$ transition. In detail, the TI formalism gives an intermediate spectral recovery fraction (0.6), which in combination with the moderate oscillator strength of S_3 (f = 0.08) leads to a very low intensity in the corresponding spectral region that is disguised by the strong absorption to the S_1 state. The TD formalism implemented in ORCA recovers the vibronic structure in this region efficiently and reproduces the tail behavior observed experimentally. Hence, the TD approach is apparently especially well suited for cases of low overlap of vibrational wave functions.

In water, RoF displays three absorption bands (Figure 8); the first one is very broad with a maximum at 500 nm (and a little bump at 390 nm), and the other two are relatively narrow being located at 320 and 260 nm (with a shoulder at 280 nm). The broadened vertical spectra reproduce the general structure of the measured spectrum by predicting broad first and third bands of almost equal height (even though the computed first band maximum is off by about 0.28 eV); the second band is computed to be a small bump, whereas it is a stand-alone peak in the measured spectrum.

Calculations at the TD-DFT and DFT/MRCI levels predict the lowest-lying LE state (S_1) to have some CT character in water, leading to a torsion of the DMA group during the relaxation from the FC point to the ES minimum (with a RMSD dihedral of 5.9°). This is in line with the known link between ICT and this torsion[57,63] and with the fact that the ICT is enhanced in a polar solvent like water.[45] The DMA torsion is expected to be better described by curvilinear displacements. This is borne out by the AFC spectra (Figures 8 and S20, water part): the curvilinear representation reproduces the bell-like shape of the first experimental band as well as its high intensity, whereas the rectilinear one fails to do so. By contrast, both representations describe the S_{10} state almost equivalently (band at 280 nm in the simulated spectra), since its RM deviates from the GS minimum only by a small torsion (RMSD of ca. 1.8°).

Considering the curvilinear AFC spectra, both the TI and TD approaches yield the correct bell shape and band maximum. Yet, they differ in the tail part (310−350 nm) where the TI method predicts an unrealistic drop. Both for TI and TD, there is not any conceivable difference with regard to the choice of the IMDHO or IMDHO-FA models (O values 0.400 vs 0.394 and 0.446 vs 0.440, respectively). Duschinsky mixing in the curvilinear AFC calculation slightly improves the shape of the first band (450−550 nm) by alleviating the relative height of its maximum (O values: 0.394 vs 0.398), but it is apparently needed in the rectilinear AFC calculation for a better topology and a higher resemblance to the measured bell shape. However, the O value is higher for the spectrum with Duschinsky treatment (O values: 0.452 vs 0.420), contradicting the latter qualitative comparison. This indicates that the O parameter might be misleading in spectral comparisons in nontrivial cases.

Due to missing RM geometries of higher-lying electronic states, the AFC approach limits the number of bands recovered in the measured full-scale UV−vis spectrum (see Table S4). The VFC approach appears as a remedy for this problem, since normal modes are available for any given state at the FC point and the corresponding EM. In our VFC calculations (see Table S4), we included most of the bright states with intermediate brightness, with the aim of reproducing the three observed absorption bands in water. As seen in Figure 8 (and Figure S20, overlay representation), both TI and TD methods performed well in reproducing the shape of the first band in terms of position, height, and width. Interestingly, the VFC approach yields a first band (with both TI and TD) that is almost identical to that of the curvilinear AFC approach; both agree better with experiment than the spectrum from the rectilinear AFC treatment. This further supports the credibility of the VFC method and the use of curvilinear coordinates (in combination with AFC) for handling controversial cases. Furthermore, we note that the third band in the experimental spectrum is reproduced more accurately by the VFC (with the exception of TI/IMDHO) than by the AFC treatment. This can be attributed to the lacking contributions from the bright S_9 and S_{12} states in AFC due to missing RM structures. Taken all together, TD/IMDHO-FA with VFC is the most successful approach, like in previous cases, according to qualitative and quantitative (highest O = 0.502) comparisons to the measured spectrum.

Even the addition of states with intermediate brightness in the VFC spectra was not enough to recover the second measured band at 320 nm. These states solely contribute small signals but do not convolute into a single separate band that would resemble the experiment. For further analysis, we also prepared VFC spectra with B3LYP and ωB97xD (see Figures S21 and S22). The vibronic structure obtained with these functionals also fails to reproduce the second band, ruling out assignment problems for the missing band due to the choice of density functional.

To scrutinize the finite-temperature effects on the simulated VFC spectra of RoLF, we present the TI and TD spectra using the IMDHO model at 298 K in Figure 9. As for LF, the TD spectra for RoLF at higher temperature differ from the corresponding spectrum at 0.01 K only insignificantly. In the TI formalism, by contrast, accounting for the excitations at higher states abridges the envelope width for the first simulated band in benzene and water. This lowers the intensity for the tail part and thus lessens the similarity to the experiment. The other band in benzene remains almost unaffected, whereas in water the third band differs greatly as the S_9 and S_{12} states cannot be included in spectra, even with Duschinsky mixing, due to the low overlap between GS and ES normal mode sets. More importantly, the second band observed in water remains missing in 298 K spectra. To elaborate on the source of this

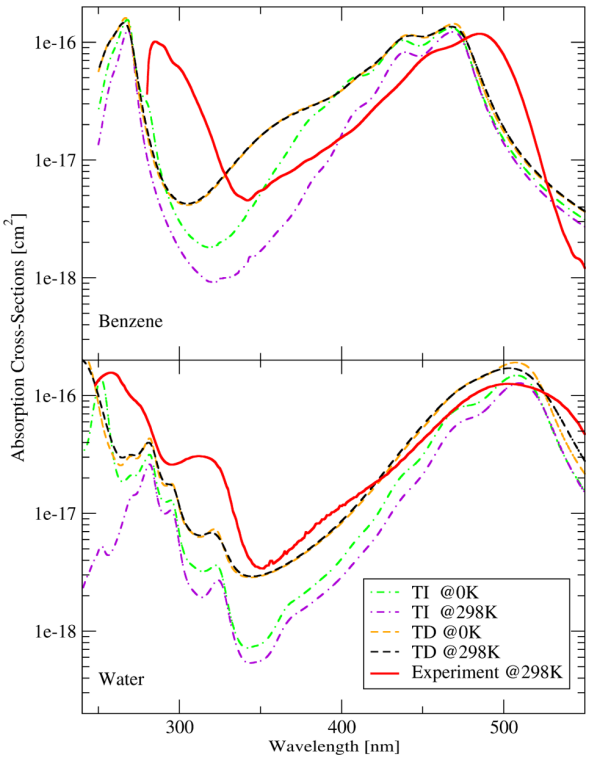


Figure 9. MRCI-corrected TI/IMDHO-FA (with/out Duschinsky mixing) and TD/IMDHO VFC spectra of RoLF computed at T = 0.01 and 298 K in benzene (top) and water (bottom) in comparison to the measured spectrum at T = 298 K (taken from ref 45). Accuracy for TI spectra is 10^6. See text.

missing band, we carried out calculations using a microsolvation scheme, in the spirit of previous studies of assorted flavins.[52,55,57] Test calculations with a shell of explicit water molecules produce a bright band in the correct spectral region (see SI, Part C for detailed discussion). Thus, the missing band can likely be attributed to H-bonding interactions of RoF with surrounding solvent molecules.

3.2.3. 5TLF Spectra. The last chromophore to be analyzed in this work is 5-thioriboflavin (modeled as 5TLF, see Figure 1), which differs from the wild-type LF by thio substitution at the N5 position. The sulfur atom paves the way for breaking the C_s symmetry, which yields a butterfly-like, bent structure as the GS minimum geometry (Figure 10). Upon photon absorption in water, the chromophore gets closer to planarity by opening up the "butterfly wings" (see Figure 10). This pronounced structural change drags the chromophore away from the FC point, putting the harmonic model in danger. In this regard, 5TLF constitutes a challenging case for most spectrum simulation methods.

5-thioriboflavin can adopt two redox states in water: neutral (reduced N3) and anionic (oxidized N3). As it is not clear which form is dominant during the absorption in water at neutral pH,[65] we considered both forms for simulating the spectra of 5TLF. For both, B3LYP and CAM-B3LYP predict a butterfly-like bent structure as the GS geometry (as shown in Figure 10). We could obtain the RM on the S_1 and S_2 surfaces, which are well separated from each other and the GS and do not couple (at both levels of theory), whereas geometry optimization of the other ESs ended up at the S_2 minimum or got stuck close to a crossing seam. We could not obtain AFC spectra for any of these two states, neither with TI nor even with the short-time approximated TD formalism. The latter can be traced to very large shifts between two normal mode sets with very high reorganization energies (larger than 250 eV with curvilinear coordinates, and even higher with Cartesians) that are caused by huge geometrical changes (RMSD > 0.30 Å, as compared to RMSD < 0.03 Å in LF/S_1, for example). Locating the RM of the ESs is difficult also for the anionic form of 5TLF. Hence, VFC is the only option within the FC framework for simulating the absorption spectrum of 5TLF in water.

The absorption spectrum of 5-thioriboflavin in distilled water has been kindly provided by Prof. Gärtner.[65] It is shown as a plot of relative absorption against absorption wavelength in Figure 11. 5TLF absorbs mainly in the far UV and exhibits a significant hypsochromic shift compared to wild-type LF, likely due to the reduced aromaticity of the 5-thioisoalloxazine moiety compared to normal isoalloxazine. The absorption spectrum is dominated by three narrow bands (225, 255, and 310 nm), accompanied by two broader shoulders (340 and 400 nm). The simple approach of broadening vertical excitations computed at the DFT/MRCI level ("E_{vert} Broadening") yields spectra that resemble experiment reasonably well. The choice of geometry (from either B3LYP or CAM-B3LYP) has little, if any, qualitative effect on the resulting broadened vertical spectrum. The computed spectra of the neutral and anionic forms of 5TLF have several characteristic features, which resemble experiment in some regions but not as a whole. The anionic form gives the highest match with experiment (O values of 0.634 and 0.646). Computed spectra are also shown for a 1:1 mixture of the two forms in water, which is likely not present in reality but would explain the loss of some features with regard to the measured spectrum.

The VFC spectra of anionic and neutral 5TLF in water were simulated at 0.01 K using the TI and TD formalisms at the B3LYP and CAM-B3LYP levels. Since 5TLF is a challenging case, we computed all ESs in the relevant range (S_1–S_9) to avoid missing any relevant contribution to the spectral shape. The corresponding MRCI-corrected FC spectra are plotted in Figure 11 (see also Figures S23 and S24 for overlay representations of these spectra). State properties and technical details are given for both forms in Tables S5 and S6, respectively. The original B3LYP and CAM-B3LYP FC spectra (without the MRCI-correction) are presented in Figures S25

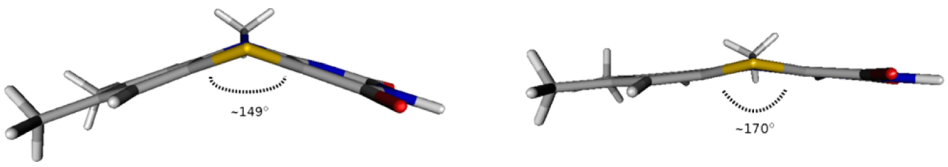


Figure 10. Side view of 5TLF in stick representation, neutral form: (left) S_0 and (right) S_2 geometries at the CAM-B3LYP/6-31G(d) level.

dx.doi.org/10.1021/ct500830a | J. Chem. Theory Comput. 2014, 10, 5549–5566

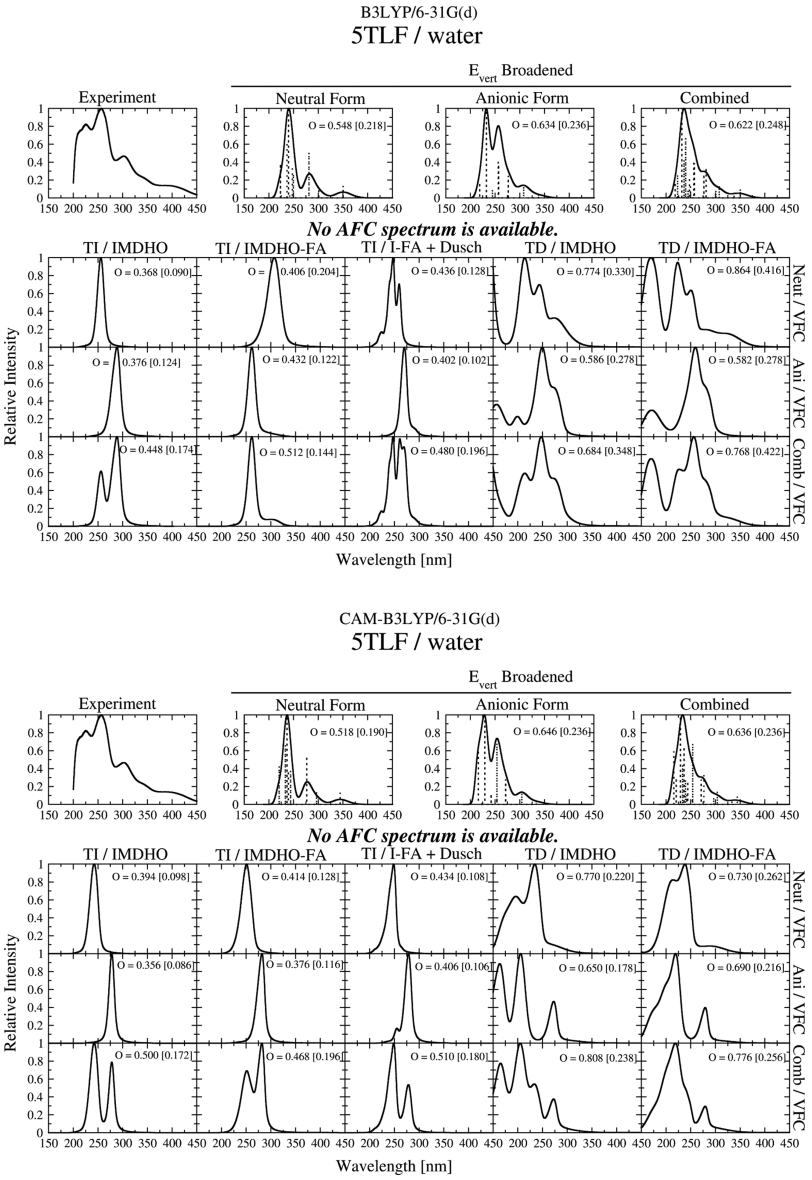


Figure 11. Convoluted vertical absorption spectra ("E_{vert} Broadened") along with TI-VFC and TD-VFC spectra of 5TLF (neutral and anionic forms and their 1:1 mixture) computed using IMDHO or IMDHO-FA models at 0.01 K in water compared to experiment at 298 K, taken from ref 65. A Duschinsky treatment was applied only for TI/IMDHO-FA spectra. Optimized [original] *O* values are also given for computed spectra. All computed spectra are MRCI-corrected. Spectra based on vertical transition energies were Gaussian broadened with an HWHM of 1500 cm^{-1}; the FC spectra were Lorentzian broadened with an HWHM of 403 cm^{-1}.

and S26 for comparison. As can be seen from Figures 11 and S23−S26, the TI spectra are almost featureless and show less resemblance to experiment than the TD spectra (as endorsed by up to 2-fold higher *O* values). This is mainly due to missing contributions from most of the ESs, because of very low overlap between the normal mode sets of GS and target ES. By contrast, the TD formalism with the short-time approximation did produce spectra with all states included, thanks to the low reorganization energies (data not shown) at the EM geometries used in the VFC scheme.

It is further evident from Figure 11 that the two density functionals yield quite different spectra (as apparent from one-to-one comparisons of each FC combination), whereas they were generally in good agreement in the case of LF and RoLF. For both functionals, the computed spectra of the neutral and anionic forms have distinctive features similar to the broadened vertical spectra. The spectra for a 1:1 mixture of the two forms resemble experiment better for CAM-B3LYP but not for B3LYP. The use of different TD models (IMDHO or IMDHO-FA) leads to noticeable differences for both functionals and 5TLF forms (unlike in the case of RF and RoF, vide supra). One would expect that IMDHO-FA provides a more accurate description than IMDHO, because it also accounts for the ES modes. However, none of the applied FC methods can reproduce the topology of the reference 5TLF spectra as accurately as in the case of LF and RoLF. This was to be expected, since the harmonic model holds less well for 5TLF as explained before.

A detailed inspection of the original and the MRCI-corrected TD/IMDHO-FA spectra (Figures 11 and S23−S26) clearly

suggests that B3LYP can reproduce the positions of the main band maxima more accurately than CAM-B3LYP, which yields hypsochromic shifts. More specifically, the B3LYP TD/IMDHO-FA spectrum of the 1:1 mixture of neutral and anionic 5TLF gives the best qualitative match with experiment in terms of spectral shape, although there exist spectra with a better quantitative match (higher O values, Figure 11). It reproduces the band maxima at 225 and 255 nm and their relative heights very well. However, the experimental band at 310 nm is calculated by B3LYP to be hypsochromatically shifted, giving a shoulder to the main peak (255 nm). This band is predicted as a separate peak by CAM-B3LYP (see TD/IMDHO-FA spectrum of the 1:1 mixture), which however greatly underestimates the height of the peak at 225 nm relative to the main peak at 255 nm. Another significant feature of the B3LYP TD/IMDHO-FA spectra is the additional peak at 170 nm that is present in both 5TLF forms. This extra peak is not found with CAM-B3LYP and might thus be an artifact; the experimental spectrum does not resolve this issue since it does not cover the region below 200 nm. We note that the water spectra may suffer from the crude solvent correction, which may not cover all solvent effects, especially for the anionic form.

The two shoulders observed in the low-energy region (340 and 400 nm) are not simulated accurately by any functional or FC method. As documented in Table S5, for the neutral form of 5TLF, B3LYP, CAM-B3LYP, and DFT/MRCI predict the S_1 state to have little oscillator strength ($0.03 < f < 0.06$) and to lie close to the relevant spectral region, with computed vertical transition energies ranging from 3.47 to 3.89 eV. Accordingly, the two shoulders in the low-energy region may possibly arise from non-Condon effects (e.g., intensity borrowing) and thus require an explicit Herzberg−Teller treatment, which is beyond the scope of the current study. Given the rather poor performance of FC-based methods in predicting the absorption spectrum of 5TLF, it will be intriguing to see how well other more advanced methods can handle this challenging case in the future.

4. SUMMARY AND CONCLUSIONS

In this contribution, we carried out an in-depth assessment of various FC-based methods in simulating optical absorption spectra of medium-sized chromophores of biological interest using a time-dependent density functional approach combined with a polarizable continuum model. To this end, we compared the measured spectra with the TI-FC and TD-FC spectra generated within the AFC and VFC frameworks using either the IMDHO or IMDHO-FA model (as outlined in ref 15). Effects of Duschinsky mixing could only be considered for the TI/IMDHO-FA combination due to limitations of the software used for the TD calculations. As test systems, we picked three flavin derivatives (using truncated models: LF, RoLF, and 5TLF) for which experimental spectra are available. The three test molecules differ in their symmetry. While LF has C_s symmetry, the DMA substitution in RoLF removes the symmetry plane, and the DMA group rotates while the isoalloxazine moiety still stays planar. The thio substitution in 5TLF causes a complete loss of planarity, leading to a butterfly-like minimum-energy structure that partly flattens out after photoexcitation.

First, we presented a benchmark for a particular electronically ES (LF/S_1) in the gas phase, the vibrational structure and fundamental frequencies of which had been measured in He nanodroplets at 0.3 K.[56] This experimental reference spectrum allowed for a fair evaluation of different density functionals, basis sets, and FC variants (in the absence of environmental or temperature effects). Qualitative and quantitative comparisons of TI-AFC spectra showed that the choice of basis set affects the resulting spectra only very slightly. The use of extended basis sets turned out to be unnecessary for the studied flavin derivatives, but one should of course refrain from using basis sets that are too small. The quality of the vibrationally resolved spectra was found to depend strongly on the choice of functional. The vibrational modes obtained with the range-separated functionals, CAM-B3LYP and ωB97xD, give almost identical absorption stick-band sets with the highest match to the measured spectra. Although they provide a realistic spectral topology, both functionals fail at predicting the vertical transition energies accurately so that the band envelopes are blue-shifted because of the high amount of HF exchange. As a remedy for this problem, we use a MRCI-correction scheme with the aim of correcting the 0−0 transition using a higher-level theory (DFT/MRCI). This scheme also aims at improving the relative absorption intensities by using transition dipoles from DFT/MRCI when multistate absorptions are considered. B3LYP and BP86 perform less well compared to range-separated hybrid functionals but still give satisfactory spectral shapes. In particular, B3LYP provides a good compromise between 0−0 peak position and topology, at least for the model systems covered here. Our findings are overall in good agreement with those for similar chromophores (anthraquinones).[88]

The detailed analysis for the LF/S_1 state emphasized the need for using accurate normal modes, created at reliable geometries, in generating AFC spectra. Considering the two possible conformers of LF, the measured vibrational structure could be accurately reproduced only for the eclipsed one, but not for the staggered one where featureless AFC spectra with very low match to experiment were obtained. The VFC approach yielded high-quality spectra for LF/S_1 in all cases. In this regard, the VFC approach appears superior to AFC, since it is not prone to errors arising from failures of the harmonic approximation; the artificial minimum used as geometry input in VFC will likely be in the immediate FC region where photon absorption occurs.

For a broader assessment, we extended our investigation to full-range absorption spectra of the chosen flavin derivatives in different media, also accounting for excitations to higher electronic states. For all chromophores, the simple approach of computing broadened vertical spectra roughly reproduced the experimental spectra but often gave only featureless bands and rather inaccurate band maxima. Obviously a more realistic modeling of the vibronic progressions is needed to capture the fine details in the observed spectra.

For LF, the computed spectra were generally quite similar for different FC variants, owing to the molecular rigidity of LC that supports the validity of the harmonic approximation. However, C_s symmetry does not always ensure the applicability of the AFC approach, because it may be impossible to locate the needed ES minimum especially for higher-lying states (e.g., LF in benzene). By contrast, the VFC scheme does not require geometry optimization and yields spectra of the same (or often higher) quality as AFC (even if imaginary-frequency modes are ignored).

In RoLF, the DMA substituent breaks the overall C_s symmetry, and the choice of FC method starts to make a difference. In one case (RoLF in benzene), only the

combination of the TD method and the VFC approach was capable of reproducing the slow decrease in the tail part of the first experimental band. The computed AFC spectra for RoLF in water were greatly affected by the choice of coordinate system for the Hessian, Duschinsky mixing, and shift vectors. While Cartesian coordinates failed, the spectral shape of the RoLF/S_1 band could be well reproduced by using internal coordinates (affording a better description of the curvilinear displacements during DMA twists), with TD performing better than TI schemes. The VFC approach, on the other hand, does not depend much on the choice of coordinates,[31,37] and it also reproduced the observed spectral shape for the RoLF/S_1 excitation very well.

For 5TLF, the VFC approach appears to be the only working option within the FC framework, since most of the ES minima needed in the AFC treatment could not be located properly for this asymmetric molecule. Even in the VFC scheme, the extrapolated minima are far away from the FC point in normal mode space, which causes low overlap with the GS wave function and high shifts so that reliable TI spectra cannot be obtained. The short-time approximation in the TD formalism in combination with the IMDHO-FA model seems to overcome this problem, since the main features of the measured absorption spectrum can be reproduced reasonably well (especially when assuming a 1:1 mixture of the neutral and anionic form), although the agreement is less good than in the simpler cases of LF and RoLF.

We also considered finite-temperature effects on the simulated spectra. In the applied TD formalism, this effect is taken into consideration by using a larger broadening factor, which led only to small changes when comparing the spectra computed at 0.01 and 298 K. By contrast, the TI formalism explicitly accounts for the excitations from the higher vibronic levels of the ground electronic state, which can noticeably alter the envelope width and shape of individual bands, as shown for LF and RoLF. However, especially for unsymmetrical molecules, the TI treatment can quickly become unpractical at higher temperatures because of the quickly rising number of integrals needed for recovering the TI spectrum.

■ ASSOCIATED CONTENT

Supporting Information

Comparison of Gaussian and Lorentzian broadening, assignments of predicted vibration modes to experimentally observed fundamental modes of LF/S_0 and LF/S_1, explicit treatment of H-bond effects via a microsolvation scheme, vibrational frequencies of LF/S_1, 1D PES scan for the interconversion between staggered and eclipsed conformations of LF, detailed technical data on TI spectrum calculations, MRCI-correction values and properties of the studied electronic states, simulated spectra in overlay representation, and simulated spectra not included in the main text. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: thiel@kofo.mpg.de.

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was partially funded by the Volkswagenstiftung (Grant No. I/83915 to W.T.). It was performed in the context of the Cluster of Excellence RESOLV (EXC 1069) funded by the Deutsche Forschungsgemeinschaft. The authors are grateful to Prof. Jeffrey R. Reimers for providing the DUSHIN source code and to Prof. Wolfgang Gärtner for absorption spectra of 5-thioriboflavin. B.K. acknowledges Prof. Alfons Penzkofer for providing the absorption cross section data of RF and RoF, and Prof. Bernhard Dick for providing the absorption spectrum of lumiflavin in He droplets before publication of the work.

■ ABBREVIATIONS

5TF, 5-thioflavin; 5TLF, 5-thiolumiflavin; AFC, adiabatic Franck−Condon; CPCM, conductor-like polarizable continuum model; DFT, density functional theory; DMA, dimethylamine; EM, extrapolated minimum; ES, excited state; FAD, flavin adenine dinucleotide; FC, Franck−Condon; FMN, flavin mononucleotide; GS, ground state; HT, Herzberg−Teller; HWHM, half-width at half-maximum; ICT, internal/intramolecular charge transfer; IMDHO, independent mode displaced harmonic oscillator; IMDHO-FA, independent mode displaced harmonic oscillator with frequency alternation; KS, Kohn−Sham; LF, lumiflavin; LOV, light oxygen voltage; MRCI, multireference configuration interaction; PES, potential energy surface; RF, riboflavin; RI, resolution of the identity; RM, real minimum; RMSD, root-mean-square deviation; RoF, roseoflavin; RoLF, roseolumiflavin; TD, time-dependent; TI, time-independent; VFC, vertical Franck−Condon; ZPVE, zero-point vibrational energy

■ REFERENCES

(1) Improta, R.; Lami, A.; Barone, V.; Santoro, F. *Int. J. Quantum Chem.* **2010**, *110*, 624−636.
(2) Domcke, W.; Cederbaum, L. *Chem. Phys. Lett.* **1975**, *31*, 582−587.
(3) Cederbaum, L. S. *J. Chem. Phys.* **1976**, *64*, 603.
(4) Franck, J. *Trans. Faraday Soc.* **1925**, *21*, 536−542.
(5) Condon, E. *Phys. Rev.* **1926**, *28*, 1182−1201.
(6) Condon, E. *Phys. Rev.* **1928**, *32*, 852−872.
(7) (a) Tannor, D.; Heller, E. J. *J. Chem. Phys.* **1982**, *77*, 202−218. (b) Heller, E. J. *Acc. Chem. Res.* **1981**, *14*, 368−375. (c) Gauss, J.; Heller, E. J. *Comput. Phys. Commun.* **1991**, *63*, 375−388. (d) Heller, E. J. *Acc. Chem. Res.* **2006**, *39*, 127−134. (e) Reimers, J. R.; Zheng-Li, C.; Kobayashi, R.; Rätsep, M.; Freiberg, A.; Krausz, E. *Sci. Rep.* **2013**, *3*, 2761.
(8) Bloino, J.; Biczysko, M.; Santoro, F.; Barone, V. *J. Chem. Theory Comput.* **2010**, *6*, 1256−1274.
(9) Barone, V.; Bloino, J.; Biczysko, M.; Santoro, F. *J. Chem. Theory Comput.* **2009**, *5*, 540−554.
(10) Peng, Q.; Niu, Y.; Deng, C.; Shuai, Z. *Chem. Phys.* **2010**, *370*, 215−222.
(11) Tang, J.; Lee, M. T.; Lin, S. H. *J. Chem. Phys.* **2003**, *119*, 7188.
(12) Beck, M. H.; Jäckle, A.; Worth, G. A.; Meyer, H.-D. *Phys. Rep.* **2000**, *324*, 1−105.
(13) De Groot, M.; Buma, W. J. *Chem. Phys. Lett.* **2007**, *435*, 224−229.
(14) Tatchen, J.; Pollak, E. *J. Chem. Phys.* **2008**, *128*, 164303.
(15) Petrenko, T.; Neese, F. *J. Chem. Phys.* **2007**, *127*, 164319.
(16) Santoro, F.; Lami, A.; Improta, R.; Bloino, J.; Barone, V. *J. Chem. Phys.* **2008**, *128*, 224311.
(17) Borrelli, R.; Peluso, A. *J. Chem. Phys.* **2008**, *128*, 044303.

(18) Jankowiak, H.-C.; Stuber, J. L.; Berger, R. *J. Chem. Phys.* **2007**, *127*, 234101.
(19) Santoro, F.; Lami, A.; Improta, R.; Barone, V. *J. Chem. Phys.* **2007**, *126*, 184102.
(20) Santoro, F.; Improta, R.; Lami, A.; Bloino, J.; Barone, V. *J. Chem. Phys.* **2007**, *126*, 084509.
(21) Borrelli, R.; Peluso, A. *J. Chem. Phys.* **2006**, *125*, 194308.
(22) Dierksen, M.; Grimme, S. *J. Chem. Phys.* **2005**, *122*, 244101.
(23) Dierksen, M.; Grimme, S. *J. Chem. Phys.* **2004**, *120*, 3544−3554.
(24) Doktorov, E. V. *J. Mol. Spectrosc.* **1977**, *8*, 507−326.
(25) Hazra, A.; Nooijen, M. *Int. J. Quantum Chem.* **2003**, *95*, 643−657.
(26) Ma, H.; Liu, J.; Liang, W. *J. Chem. Theory Comput.* **2012**, *8*, 4474−4482.
(27) Crespo-Otero, R.; Barbatti, M. *Theor. Chem. Acc.* **2012**, *131*, 1237.
(28) Nooijen, M. *Int. J. Quantum Chem.* **2006**, *106*, 2489−2510.
(29) Avila Ferrer, F. J.; Santoro, F. *Phys. Chem. Chem. Phys.* **2012**, *14*, 13549−13563.
(30) (a) Hazra, A.; Chang, H. H.; Nooijen, M. *J. Chem. Phys.* **2004**, *121*, 2125−2136. (b) Thompson, K. C.; Jordan, M. J. T.; Collins, M. A. *J. Chem. Phys.* **1998**, *108*, 8302−8316. (c) Bettens, R. P. A.; Collins, M. A.; Jordan, M. J. T.; Zhang, D. H. *J. Chem. Phys.* **2000**, *112*, 10162−10172. (d) Kollmann, S. J.; Jordan, M. J. T. *J. Chem. Phys.* **2010**, *132*, 054105.
(31) Götze, J.; Karasulu, B.; Thiel, W. *J. Chem. Phys.* **2013**, *139*, 234108.
(32) Rätsep, M.; Cai, Z.-L.; Reimers, J. R.; Freiberg, A. *J. Chem. Phys.* **2011**, *134*, 024506.
(33) Wilson, E. B.; Decius, J. C.; Cross, C. C. *Molecular vibrations: the theory of infrared and Raman vibrational spectra*; McGraw-Hill: New York, 1955; pp 54−77.
(34) Reimers, J. R. *J. Chem. Phys.* **2001**, *115*, 9103.
(35) Capobianco, A.; Borrelli, R.; Noce, C.; Peluso, A. *Theor. Chem. Acc.* **2012**, *131*, 1181.
(36) Borrelli, R.; Peluso, A. *J. Chem. Phys.* **2008**, *128*, 044303.
(37) Cerezo, J.; Zúñiga, J.; Requena, A.; Avila Ferrer, F. J.; Santoro, F. *J. Chem. Theory Comput.* **2013**, *9*, 4947−4958.
(38) Duschinsky, F. *Acta Physicochim. URSS* **1937**, *7*, 551.
(39) Sando, G. M.; Spears, K. G. *J. Phys. Chem. A* **2001**, *105*, 5326−5333.
(40) Özkan, İ. *J. Mol. Spectrosc.* **1990**, *62*, 147−162.
(41) Banerjee, S.; Kröner, D.; Saalfrank, P. *J. Chem. Phys.* **2012**, *137*, 22A534.
(42) Ghisla, S.; Massey, V. *Biochem. J.* **1986**, *239*, 1−12.
(43) Mathes, T.; Vogl, C.; Stolz, J.; Hegemann, P. *J. Mol. Biol.* **2009**, *385*, 1511−1518.
(44) Marian, C. M.; Setsuko, N.; Rai-Constapel, V.; Karasulu, B.; Thiel, W. *J. Phys. Chem. B* **2014**, *118*, 1743−1753.
(45) Zirak, P.; Penzkofer, A.; Mathes, T.; Hegemann, P. *Chem. Phys.* **2009**, *358*, 111−122.
(46) Shiga, K.; Nishina, Y.; Ohmine, I.; Horiike, K.; Kasai, S.; Matsui, K.; Watari, H.; Yamano, T. *J. Biochem.* **1980**, *87*, 281−287.
(47) Song, P.; Walker, E.; Vierstra, R.; Poff, K. L. *Photochem. Photobiol.* **1980**, *32*, 393−398.
(48) Choe, Y. Y.-K.; Nagase, S.; Nishimoto, K. *J. Comput. Chem.* **2007**, *28*, 727−739.
(49) Klaumünzer, B.; Kröner, D.; Saalfrank, P. *J. Phys. Chem. B* **2010**, *114*, 10826−10834.
(50) Merz, T.; Sadeghian, K.; Schütz, M. *Phys. Chem. Chem. Phys.* **2011**, *13*, 14775−14783.
(51) Sadeghian, K.; Bocola, M.; Schütz, M. *Phys. Chem. Chem. Phys.* **2010**, *12*, 8840−8846.
(52) Sadeghian, K.; Schütz, M. *J. Am. Chem. Soc.* **2007**, *129*, 4068−4074.
(53) Salzmann, S.; Martinez-Junza, V.; Zorn, B.; Braslavsky, S. E.; Mansurova, M.; Marian, C. M.; Gärtner, W. *J. Phys. Chem. A* **2009**, *113*, 9365−9375.
(54) Neiss, C.; Saalfrank, P.; Parac, M.; Grimme, S. *J. Phys. Chem. A* **2003**, *107*, 140−147.
(55) Salzmann, S.; Tatchen, J.; Marian, C. M. *J. Photochem. Photobiol., A* **2008**, *198*, 221−231.
(56) Vdovin, A.; Slenczka, A.; Dick, B. *Chem. Phys.* **2013**, *422*, 195−203.
(57) Karasulu, B.; Thiel, W. *J. Phys. Chem. B* **2014**, ASAP article, DOI:10.1021/jp506101x.
(58) Tyagi, A.; Penzkofer, A.; Mathes, T.; Hegemann, P. *J. Photochem. Photobiol., B* **2010**, *101*, 76−88.
(59) Matsui, K.; Kasai, S. In *Chemistry and Biochemistry of Flavoenzymes*; F. Müller, Ed.; CRC Press: Boca Raton FL, USA, 1991; pp 105−120.
(60) Sikorska, E.; Khmelinskii, I. V.; Prukała, W.; Williams, L.; Patel, M.; Worrall, D. R.; Bourdelande, J. L.; Koput, J.; Sikorski, M. *Society* **2004**, 1501−1508.
(61) Tyagi, A.; Zirak, P.; Penzkofer, A.; Mathes, T.; Hegemann, P.; Mack, M.; Ghisla, S. *Chem. Phys.* **2009**, *364*, 19−30.
(62) Drössler, P.; Holzer, W.; Penzkofer, A.; Hegemann, P. *Chem. Phys.* **2002**, *282*, 429−439.
(63) Zirak, P.; Penzkofer, A.; Mathes, T.; Hegemann, P. *J. Photochem. Photobiol., B* **2009**, *97*, 61−70.
(64) Fenner, H.; Grauert, R. W.; Tessendorf, L. *Arch. Pharm. (Weinheim, Ger.)* **1981**, *314*, 874−879.
(65) Neubert, D. *Chemical Synthesis of 5-Thiariboflavin as a Potential Chromophore in Blue Light-Sensitive Photoreceptors*; M.Sc. Thesis; Hochschule Niederrhein, 2011; pp 24−31.
(66) Peluso, A.; Borrelli, R.; Capobianco, A. *J. Phys. Chem. A* **2009**, *113*, 14831−14837.
(67) Bowd, A.; Byrom, P.; Hudson, J.; Turnbull, J. *Photochem. Photobiol.* **1968**, *8*, 1−10.
(68) Parr, R.; Yang, W. *Density-functional theory of atoms and molecules*; Oxford University Press: Oxford, 1994; Vol. *16*; pp 142−215.
(69) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, *98*, 11623−11627.
(70) Becke, A. *J. Chem. Phys.* **1993**, *98*, 5648−5652.
(71) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*, Revision A.02; Gaussian, Inc.: Wallingford, CT, 2009.
(72) Perdew, J. *Phys. Rev. B* **1986**, 8822−8824.
(73) Becke, A. *Phys. Rev. A* **1988**, *38*, 3098−3100.
(74) Yanai, T.; Tew, D. P.; Handy, N. C. *Chem. Phys. Lett.* **2004**, *393*, 51−57.
(75) Chai, J.-D.; Head-Gordon, M. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6615−6620.
(76) Schäfer, A.; Huber, C.; Ahlrichs, R. *J. Chem. Phys.* **1994**, *100*, 5829−5835.
(77) Francl, M.; Pietro, W.; Hehre, W.; Binkley, J.; Gordon, M.; DeFrees, D.; Pople, J. *J. Chem. Phys.* **1982**, *77*, 3654−3665.
(78) Silva-Junior, M. R.; Schreiber, M.; Sauer, S. P. A.; Thiel, W. *J. Chem. Phys.* **2008**, *129*, 104103.
(79) Grimme, S.; Waletzke, M. *J. Chem. Phys.* **1999**, *111*, 5645−5655.
(80) Ahlrichs, R.; Bär, M.; Baron, H.-P.; Bauern- schmitt, S.; Böcker, S.; Ehrig, M.; Eichkorn, K.; Elliott, S.; Furche, F.; Haase, F.; Häser, M.; Horn, H.; Huber, C.; Huniar, U.; Kattannek, M.; Kölmel, C.; Kollwitz,

M.; May, K.; Ochsenfeld, C.; Öhm, H.; Schäfer, A.; Schneider, U.; Treutler, O.; von Arnim, M.; Weigend, F.; Weis, P.; Weiss, H. *TURBOMOLE* v. 6.3; TURBOMOLE GmbH: Karlsruhe, Germany, 2011.
(81) Jacquemin, D.; Adamo, C. *Int. J. Quantum Chem.* **2012**, *112*, 2135−2141.
(82) Cossi, M.; Rega, N.; Scalmani, G.; Barone, V. *J. Comput. Chem.* **2003**, *24*, 669−681.
(83) Barone, V.; Cossi, M. *J. Phys. Chem. A* **1998**, *102*, 1995−2001.
(84) Schäfer, A.; Klamt, A.; Sattel, D.; Lohrenz, J. C. W.; Eckert, F. *Phys. Chem. Chem. Phys.* **2000**, *2*, 2187−2193.
(85) Neese, F.; Petrenko, T.; Ganyushin, D.; Olbrich, G. *Coord. Chem. Rev.* **2007**, *251*, 288−327.
(86) Santoro, F.; Barone, V. *Int. J. Quantum Chem.* **2010**, *110*, 476−486.
(87) Santoro, F.; Cappelli, C.; Barone, V. *J. Chem. Theory Comput.* **2011**, *7*, 1824−1839.
(88) Jacquemin, D.; Brémond, E.; Planchat, A.; Ciofini, I.; Adamo, C. *J. Chem. Theory Comput.* **2011**, *7*, 1882−1892.
(89) Neugebauer, J.; Hess, B. A. *J. Chem. Phys.* **2003**, *118*, 7215.
(90) Salzmann, S.; Silva-Junior, M. R.; Thiel, W.; Marian, C. M. *J. Phys. Chem. B* **2009**, *113*, 15610−15618.
(91) Alecu, I. M.; Zheng, J.; Zhao, Y.; Truhlar, D. G. *J. Chem. Theory Comput.* **2010**, *6*, 1−10.
(92) Dierksen, M.; Grimme, S. *J. Phys. Chem. A* **2004**, *108*, 10225−10237.
(93) Peach, M. J. G.; Benfield, P.; Helgaker, T.; Tozer, D. J. *J. Chem. Phys.* **2008**, *128*, 044118.
(94) Stendardo, E.; Ferrer, F. A. *J. Chem. Theory Comput.* **2012**, *8*, 4483−4493.
(95) Christensen, K. A.; Jensen, F. *Chem. Phys. Lett.* **2000**, *317*, 400−403.
(96) Improta, R.; Barone, V.; Scalmani, G.; Frisch, M. J. *J. Chem. Phys.* **2006**, *125*, 054103.
(97) Improta, R.; Scalmani, G.; Frisch, M. J.; Barone, V. *J. Chem. Phys.* **2007**, *127*, 074504.
(98) Götze, J. P.; Thiel, W. *Chem. Phys.* **2013**, *415*, 247−255.
(99) Kleinschmidt, M.; Marian, C. M.; Waletzke, M.; Grimme, S. *J. Chem. Phys.* **2009**, *130*, 044708.
(100) Starcke, J. H.; Wormit, M.; Schirmer, J.; Dreuw, A. *Chem. Phys.* **2006**, *329*, 39−49.