

DISSERTATION

submitted to the
Combined Faculties for the Natural Sciences and for Mathematics,
Ruperto Carola University of Heidelberg

for the degree of
Doctor of Natural Sciences

put forward by
Dipl.-Phys. Anikó Udvarhelyi,
born in Budapest, Hungary

Oral examination on 23 July 2014

Molecular mechanism of light activation in the blue-light photoreceptor BLUF

Referees: Prof. Dr. Ilme Schlichting
Prof. Dr. Horst Köppel

Ildinek

Abstract

Photoreceptor proteins are molecular sensors that translate photon energy into biological information. The BLUF (Blue Light using FAD) protein is such a sensor that switches between its dark and light states by means of photoinduced proton-coupled electron transfer (PCET). In this thesis, I present the first detailed and systematic computational study of photoinduced PCET in BLUF using state-of-the-art electronic-structure methods. The photoactivation in BLUF results in the tautomerization and rotation of a conserved glutamine side chain. The computed potential-energy landscapes presented in this thesis reveal the energies of glutamine rotamers and tautomers and serve as a basis to identify the structure of the glutamine side chain in the functional dark and light states of BLUF. To map the pathway connecting the dark and light states on the excited-state potential-energy surface, I established a computational procedure employing multi-configurational multi-reference electronic-structure methods, and built and characterized quantum-mechanical cluster and hybrid quantum-mechanical/molecular-mechanical models. After establishing and benchmarking the computational protocol, I computed several PCET photoreaction pathways. The energy profiles obtained serve as a basis to answer, for the first time, questions related to how PCET is realized in photoactivation, photostability, and redox tuning in BLUF.

Zusammenfassung

Photorezeptor-Proteine sind molekulare Sensoren, die die Photonenergie in biologische Information umwandeln. Das Protein BLUF (Blue Light using FAD) ist solch ein Sensor, der zwischen seinem Dunkel- und Licht-Zustand mittels Protonen-gekoppeltem Elektronentransfer (PCET) schalten kann. In dieser Arbeit wird die erste detaillierte und systematische rechnergestützte Studie über photoinduzierten PCET in BLUF mit state-of-the-art Elektronenstrukturmethoden präsentiert. Die Photoaktivierung von BLUF führt zur Tautomerisierung und Rotation der Seitenkette eines konservierten Glutamins. Die in dieser Arbeit errechneten Potentialenergieprofile zeigen die Energien verschiedener Rotamere sowie Tautomere dieses Glutamins auf und dienen als Grundlage, die Struktur der Glutamin-Seitenkette in den Dunkel- und Licht-Zuständen von BLUF zu identifizieren. Um den Reaktionspfad abzubilden, der den Dunkel- und Licht-Zustand auf der Potentialfläche des elektronisch angeregten Zustandes verbindet, wurde ein rechnergestütztes Verfahren mit Multikonfigurations-Multireferenz-Elektronenstrukturmethoden entwickelt; darüber hinaus wurden quantenmechanische Cluster und quantenmechanisch/molekular-mechanische Hybrid-Modelle erstellt und charakterisiert. Nach dem Benchmarking des rechnergestützten Protokolls wurden mehrere PCET-Photoreaktionspfade berechnet. Mit den erhaltenen Potentialenergieprofilen kann erstmals beantwortet werden, wie PCET bei der Photoaktivierung, Photostabilität und beim Redox-tuning in BLUF wirksam wird.

The work put forward in this thesis was carried out in the Department of Biomolecular Mechanisms at the Max Planck Institute for Medical Research, Heidelberg, Germany, under the supervision of Dr. Tatiana Domratcheva.

I have performed all calculations and carried out all analysis myself, except when noted otherwise, and prepared this thesis myself. All other resources used are specified.

Major parts of this dissertation have already been published or will be published in peer-reviewed journals. In this dissertation, I have used text and figures from these published and planned articles (except for the book chapter), with permission of the respective journals:

- Udvarhelyi, A. & Domratcheva, T. Photoreaction and redox tuning in the BLUF dark and light states. *In preparation.*
- Udvarhelyi, A. & Domratcheva, T. Conformational dynamics of the glutamine side chain in the BLUF W_{in} and W_{out} flavin-binding pockets. *In preparation.*
- Udvarhelyi, A., Olivucci, M. & Domratcheva, T. The role of the macromolecular environment in color and redox tuning: supermolecular cluster versus QM/MM model. *In preparation.*
- Domratcheva, T., Udvarhelyi, A., Moughal Shahi, A. R. (2014) Computational Spectroscopy, Dynamics, and Photochemistry of Photosensory Flavoproteins. Chapter 10 in *Flavins and Flavoproteins – Methods and Protocols*. Eds. S. Weber, E. Schleicher, ISBN 978-1-4939-0451-8. *Springer Protocols*. Humana Press.
- Udvarhelyi, A. & Domratcheva, T. (2013) Glutamine Rotamers in BLUF Photoreceptors: A Mechanistic Reappraisal. *The Journal of Physical Chemistry B* 117, 2888–2897.
- Udvarhelyi, A. & Domratcheva, T. (2011) Photoreaction in BLUF Receptors: Proton-coupled Electron Transfer in the Flavin-Gln-Tyr System. *Photochemistry and Photobiology* 87, 554–563.

Additional published and planned articles, not within the scope of this thesis:

- Udvarhelyi, A., Harding, M. & Domratcheva, T. The principal-orbital complete-active-space (POCAS) approach for uracil photochemistry. *In preparation.*
- Tuna, D., Udvarhelyi, A., Sobolewski, A. L., Domcke, W. & Domratcheva, T. Ab initio study of the photophysics of eumelanin: electronic absorption spectra of isolated and π -stacked oligomers of 5,6-dihydroxyindole. *In preparation.*
- Winkler, A., Udvarhelyi, A., *et al.* (2013) Characterization of Elements Involved in Allosteric Light Regulation of Phosphodiesterase Activity by Comparison of Different Functional BlrP1 States. *Journal of Molecular Biology* 426 (4), 853–868.
- Smeulders, M. J., Barends, T., Pol, A., Scherer, A., Zandvoort, M. H., Udvarhelyi, A., *et al.* (2011) Evolution of a new enzyme for carbon disulphide conversion by an acidothermophilic archaeon. *Nature*. 478, 412–416.
- Lang, M., Jegou, T., Chung, I., Richter, K., Münch, S., Udvarhelyi, A., *et al.* (2010) Three-dimensional organization of promyelocytic leukemia nuclear bodies. *Journal of Cell Science*. 123, 392–400.

Contents

1	Introduction	17
1.1	<i>Blue Light using FAD sensors</i>	17
1.2	<i>Photophysical and photochemical concepts</i>	23
1.3	<i>Electronic-structure methods</i>	25
1.4	<i>Molecular models</i>	27
1.5	<i>Aims of this thesis</i>	30
2	Method benchmarking: POCAS	31
2.1	<i>CASSCF and POCAS</i>	31
2.2	<i>Computational details</i>	34
2.3	<i>Results and discussion</i>	35
2.3.1	Low-lying excited states in lumiflavin	35
2.3.2	POCAS-optimized geometry of lumiflavin	39
2.3.3	Excitation energies in the BLUF cluster	40
2.4	<i>Conclusions</i>	44
3	Cluster models vs. QM/MM models	45
3.1	<i>Cluster and QM/MM models in photoreceptor studies</i>	45
3.2	<i>Models and analysis</i>	47
3.2.1	Cluster model and QM/MM models	47
3.2.2	Analysis of intermolecular interactions	50
3.3	<i>Computational details</i>	52

3.3.1	Cluster models	52
3.3.2	QM/MM calculations	53
3.4	<i>Results</i>	55
3.4.1	Geometry of the flavin-binding site.....	55
3.4.2	Excitation energies – QM calculations.....	57
3.4.3	Excitation energies – QM/MM calculations	60
3.4.4	Electrostatic potential maps and charge-transfer effects.....	61
3.4.5	Excitation-energy shift upon QM/MM embedding	64
3.4.6	Modifications of the MM charges.....	66
3.4.7	Intermolecular interactions and the excitation energy of the ET state	68
3.4.8	Comparison of the full cluster model with the reference QM/MM model	72
3.5	<i>Discussion and conclusions</i>	74
4	Glutamine rotamers	77
4.1	<i>BLUF structures</i>	77
4.2	<i>Computational details</i>	81
4.2.1	W_{in} cluster models and geometry optimization.....	81
4.2.2	Extended cluster models: ext- W_{in} and ext- W_{out}	82
4.2.3	Dihedral-angle energy scans and excitation spectra	83
4.3	<i>Results and Discussion</i>	84
4.3.1	The Q63 rotamer in the W_{in} structure.....	84
4.3.2	Extended cluster and QM/MM models	88
4.3.3	Dynamic aspects: glutamine rotation in the dark state	91
4.3.4	Dihedral-energy scans in the extended models: dark and light states.....	93
4.3.5	Comparison to NMR and MD studies.....	95
4.4	<i>Summary and conclusions</i>	97
5	BLUF photoreaction	99
5.1	<i>Excited-state pathways and BLUF photochemistry</i>	99
5.2	<i>Methods and computational details</i>	102
5.2.1	Cluster and QM/MM models.....	102
5.2.2	Pathway calculations	104
5.3	<i>Results and discussion</i>	106

5.3.1	PET in the dark state	106
5.3.2	Proton transfer pathways	110
5.3.3	PCET pathway without flavin protonation	116
5.3.4	BLUF redox tuning with fluorotyrosine	120
5.3.5	PCET in the light state	123
5.3.6	Electron transfer rate estimates using excited-state energies and Marcus theory	127
5.4	<i>Summary and conclusions</i>	129
6	Conclusions and Outlook	131
	Abbreviations	137
	References	139
	Acknowledgements	157

1 Introduction

Light – through its interaction with biological matter – powers and controls the molecular machinery of life. It serves as an energy source (photosynthesis), provides information about the environment (phototaxis, visual perception, light-dependent metabolism), but may also represent a threat (UV damage). The interaction between light and matter is governed by the laws of quantum-mechanics: photon absorption excites an electron bound in a molecule; the excess energy is temporarily stored in the excited state and may be transferred between molecular species; upon excited-state relaxation, a part of the energy is dissipated. Biological matter converts the photon energy either into vibrational energy or into “chemical” energy by inducing modifications of the chemical structure. One fundamental reaction mechanism that is common in biological photoreactions is proton-coupled electron transfer (PCET). PCET is ubiquitous in biology and occurs both in light-dependent and light-independent processes, for example during respiration, photosynthesis, and cell signaling. The central topic of this thesis is the photoactivation mechanism of the unique bacterial photosensor BLUF, in which, by means of PCET, the photon energy is translated into chemical and vibrational energy, and ultimately into biological information.

1.1 Blue Light using FAD sensors

Light-sensitive proteins, so-called photoreceptors, detect light using photosensitive chromophore molecules. To date, seven classes of photoreceptors are known (van der Horst & Hellingwerf, 2004; Rizzini et al., 2011): rhodopsins (chromophore: retinal), phytochromes (phytochromobilin), xantopsins (p-coumaric acid), the UV sensor UVR8 (14 tryptophan residues) and the three flavin-binding blue-light photoreceptor classes cryptochromes, light-oxygen-voltage sensing (LOV) domains, and blue light using FAD sensors (BLUF) that bind flavin mononucleotide (FMN) or flavin adenosine dinucleotide (FAD). A distinct photochemistry characterizes each photoreceptor

class, i.e. upon photon absorption a chemical reaction occurs that is specific for each class. Rhodopsins for example isomerize the retinal chromophore. The photoactivation mechanism of flavin-binding photoreceptors is rather different and still a subject of debate. The flavin-binding photoreceptors bind flavin in its oxidized form that absorbs light in the blue spectral region, thereby mediating various blue-light responses in plants, bacteria and algae (Losi & Gärtner, 2011, 2012). Detailed knowledge about the photoreceptors' structural and chemical properties is essential to provide a comprehensive mechanistic understanding of their function. However, the size and complexity of the molecular systems involved in sensing light present a great challenge for studies investigating them. For an adequate interpretation of the spectroscopy data and a full understanding of the reaction mechanisms, a detailed and systematic characterization of molecular properties is required for which computational studies are indispensable (Domratcheva et al., 2014). Thus a combination of experimental and computational approaches is required to investigate photoreceptor function. A better understanding of the mechanisms involved in the underlying photochemical processes paves the way to powerful applications in the recently emerging fields of optogenetics (Fenno et al., 2011; Deisseroth, 2011) and synthetic biology (Cheng & Lu, 2012): Photoreceptor proteins can serve as components in rationally-designed photoswitches to control and manipulate targeted biological processes non-invasively with light, even *in vivo*. The ubiquitously available flavin in the cell makes blue light photoreceptors an interesting candidate for the "optogenetic toolkit" (Christie et al., 2012). An impressive *in vivo* example using the LOV photoreceptor in an artificial photoswitch to control the motility of cells was demonstrated in (Wu et al., 2009).

Figure 1.1 shows the chemical structure of lumiflavin (LF), and the naturally occurring flavin analogs riboflavin (RF), FMN, and FAD. Unbound flavins in aqueous solution are found in the oxidized form (Song et al., 2007). The experimental UV-Vis absorption spectrum of the oxidized flavin is shown in Figure 1.2A, revealing two absorption bands centered at around 450 nm (2.76 eV) and 370 nm (3.35 eV); a third, most intense band is found at 275 nm (4.51 eV), which is not of interest in this work. LF, FMN, and FAD possess this characteristic absorption spectrum due to the flavin core ring (isoalloxazine ring). The absorption bands originate from single electronic excitations among the π and π^* frontier molecular orbitals (MO) of flavin (Neiss et al., 2003): the first excited S_1 state is predominantly characterized by a transition from the highest occupied molecular orbital (HOMO) to the lowest unoccupied molecular orbital (LUMO), and the S_2 state by a transition from the HOMO-1 to LUMO, as schematically indicated in Figure 1.2B. The ability of flavins to adopt different redox and protonation states makes them remarkably versatile biological cofactors, exploited by diverse flavin-binding proteins in different biochemically relevant reactions (Miura, 2001). The local environment of the non-covalently bound flavin in the protein fine-tunes the broad reactivity of flavin for a specific reaction. In blue-light photoreceptors the local interactions between the electronically excited flavin and its neighboring residues determine the specific photoreaction.

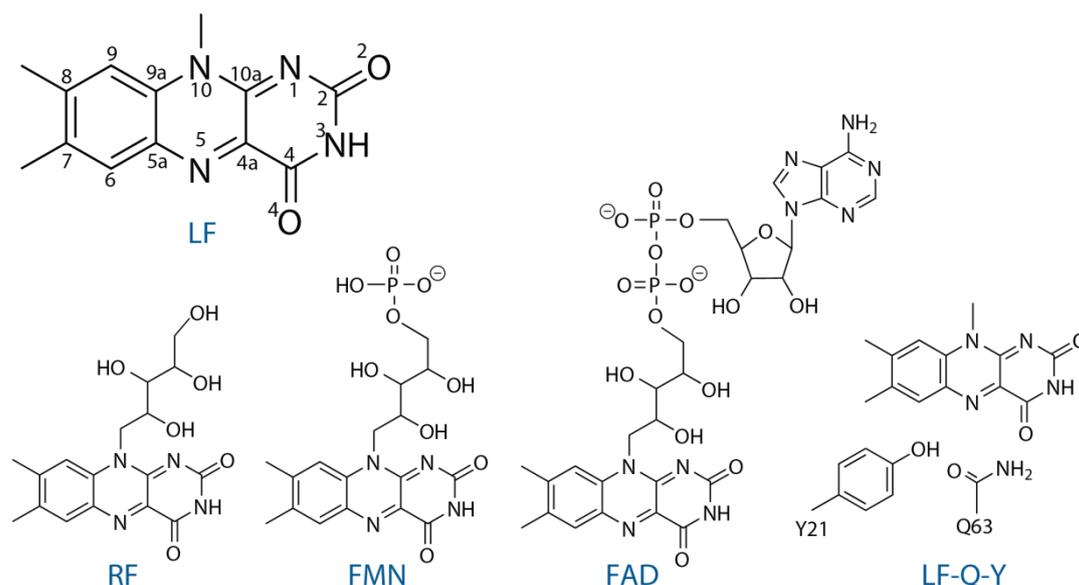


Figure 1.1. Chemical structure of lumiflavin (LF) with conventional numbering of atoms, riboflavin (RF), flavin mononucleotide (FMN), flavin adenine dinucleotide (FAD) and the BLUF lumiflavin-glutamine-tyrosine photoactive triad (LF-Q-Y).

The BLUF protein domain was independently discovered by two groups: it was described as a light-sensory unit of the AppA protein involved in the light-dependent regulation of photosynthetic gene expression in the purple proteobacterium *Rhodobacter sphaeroides* (Masuda & Bauer, 2002); and as a protein domain in photoactivated adenylyl cyclase (PAC) involved in the photoavoidance response in the eukaryote *Englena gracilis* (Iseki et al., 2002). Following these discoveries, several proteins with a similar sequence were identified in many bacterial genomes and the new photoreceptor domain was termed BLUF (Gomelsky & Klug, 2002). BLUF proteins are modular units that can communicate the blue light signal to various output domains (Han et al., 2004), and can be classified to two categories based on the domain architecture: (i) either they serve as a sensory unit that regulates the photoresponse in a multidomain protein or (ii) they appear as a standalone protein that interacts with an effector output protein. Examples from the first category include YcgF from *Escherichia coli* and BlrP1 from *Klebsiella pneumoniae*, two phosphodiesterases that putatively control virulence and biofilm formation (Tschowri et al., 2009; Tyagi et al., 2008; Barends et al., 2009). Examples from the second category include PixD from *Synechocystis* sp. PCC6803 that mediates phototactic response (Okajima et al., 2005); BlsA from *Acinetobacter baumannii* that regulates biofilm formation in a temperature-dependent manner (Mussi et al., 2010); or BlrB from *Rhodobacter sphaeroides* with an unknown effector and unknown function (Jung et al., 2005). So far, from the many BLUF proteins identified in bacterial genomes, the physiological function of only few has been described and only few BLUF domains have been characterized.

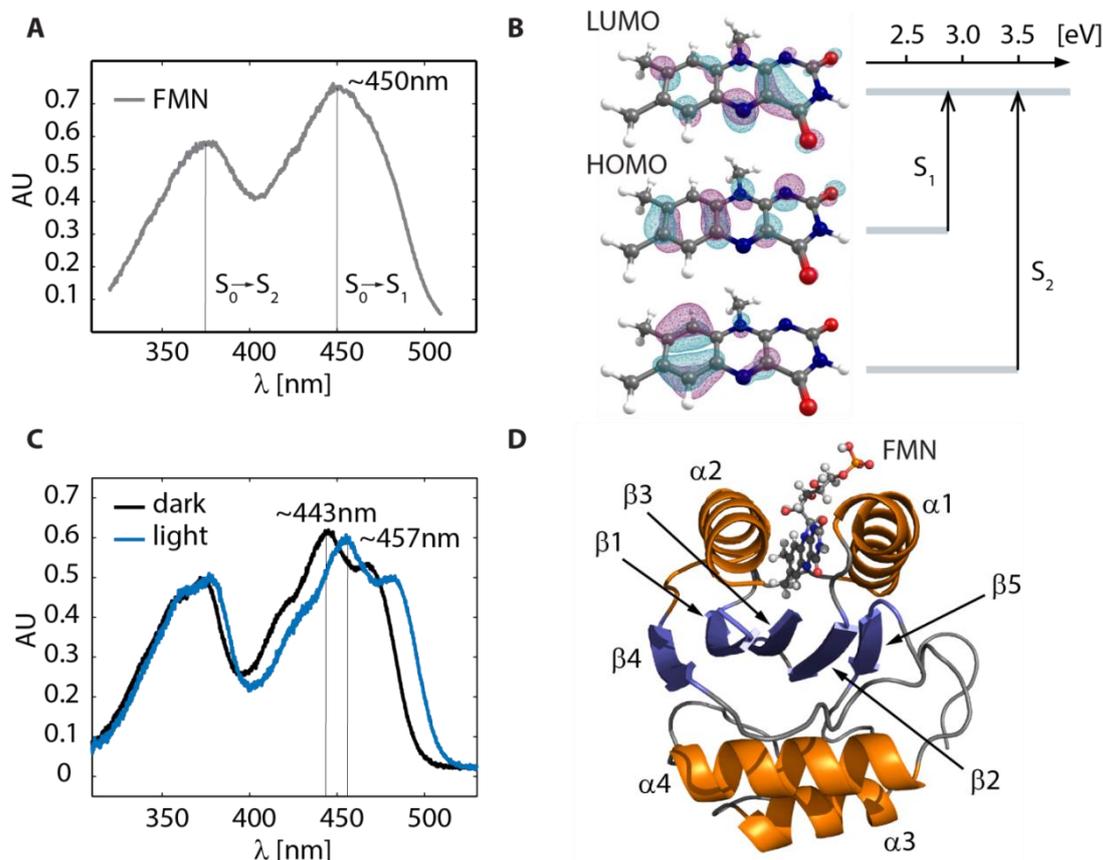


Figure 1.2. (A) Absorption spectrum of FMN in water. The absorption maxima corresponding to the S_1 and S_2 transitions are indicated. (B) Electronic structure and energies of the excited states in oxidized lumiflavin. Electronic excitations among the frontier MOs predominantly contributing to the S_1 and S_2 excited states are indicated. (C) Absorption spectrum of BLUF in the dark and light states. The absorption maximum corresponding to the S_1 transition in the two states is indicated. (D) X-ray crystal structure of BLUF with indicated secondary structure elements in cartoon representation. FMN is shown in balls-and-sticks. The absorption spectra in (A) and (C) were measured in the course of my diploma thesis (Udvarhelyi, 2009).

BLUF photoreceptors non-covalently bind the oxidized FAD as the chromophore. Figure 1.2C shows the flavin absorption spectrum in BLUF that closely resembles that of the free flavin in solution, with the absorption maxima only slightly shifted and the absorption bands vibrationally resolved. Remarkably, after photoexcitation, the flavin absorption is identical to the one in the dark state, with a merely 15-nm red-shifted first flavin band and a few nm red-shifted second band, as shown in Figure 1.2C. This red shift of the flavin absorption is attributed to the formation of the light state (Masuda & Bauer, 2002; Kraft et al., 2003; Masuda et al., 2004; Zirak et al., 2005, 2006; Fukushima et al., 2005), indicating that the FAD chromophore remains in its oxidized form in the light state. Thus, BLUF undergoes perhaps the most peculiar photoreaction among the photoreceptor classes where the chromophore itself does not change chemically. The red-shifted photoproduct is stable for seconds to minutes and spontaneously recovers the dark

state with different kinetics (Zirak et al., 2005, 2006, 2007; Tyagi et al., 2008; Hasegawa et al., 2006; Fukushima et al., 2005; Penzkofer et al., 2011), when the full BLUF photocycle is completed. The photoproduct decays exclusively thermally to recover the dark state, without any light-induced back reaction (Toh et al., 2008). Concomitant with the red-shifted UV-Vis absorption, a 20 cm^{-1} red-shifted IR absorption in the carbonyl-stretching spectral region characterizes the light state (Masuda et al., 2004). The formation or strengthening of hydrogen bonds involving the flavin C4=O4 (for the atom numbering see Figure 1.1) carbonyl group explains these spectral shifts (Masuda et al., 2004; Hasegawa et al., 2004; Unno et al., 2005), thus a photoinduced hydrogen-bonding rearrangement must take place to form the BLUF light state.

To reveal the specific structure of the hydrogen-bonding network around flavin in BLUF, molecular structures are necessary. In the past years, X-ray crystal (Anderson et al., 2005; Jung et al., 2005, 2006; Kita et al., 2005; Yuan et al., 2006; Barends et al., 2009; Winkler et al., 2013) and solution NMR structures (Wu et al., 2008; Wu & Gardner, 2009) of various BLUF domains from several bacteria were determined. These structures revealed a globular ferredoxin-like fold, shown in Figure 1.2D, with a five-stranded beta sheet and two alpha helices ($\alpha 1$ and $\alpha 2$) sandwiching the FMN part of the FAD chromophore. This core structure is nearly identical in all BLUF domains. C-terminal alpha helices ($\alpha 3$ and $\alpha 4$) are found on the other side of the beta sheet, with different conformations and orientations with respect to the beta sheet in various BLUFs. Their role in signal transduction has been proposed in several studies, reviewed in (Zoltowski & Gardner, 2011). To uncover the role of the C-terminus, structural studies of proteins in which the N-terminal BLUF sensor is connected to the C-terminal effector domain are necessary. So far one complex structure was determined: the BlrP1 X-ray structure of the BLUF domain covalently linked to its functional phosphodiesterase EAL effector domain (Barends et al., 2009). More recently, the X-ray structure of AppA BLUF in complex with its non-covalent PpsR effector was also determined (Winkler et al., 2013).

The X-ray and NMR structures revealed that the residues conserved in the BLUF sequences form hydrogen bonds with flavin, most importantly a glutamine residue (in the following referred to as Q53, according to the one-letter amino acid code of glutamine and the residue numbering of the PixD protein). Mutational studies showed that Q53 is essential for the light-switch function of BLUF as its mutation abolishes the formation of the red-shifted light state (Masuda et al., 2007; Stelling et al., 2007; Dragnea et al., 2010). The other hydrogen-bonding partner of Q53 is a conserved tyrosine residue, Y11. Ultrafast spectroscopy studies attributed the quenched flavin fluorescence, which is observed only for tenth of picoseconds, to photoinduced electron transfer (PET) from a nearby residue to the electronically excited flavin (Kraft et al., 2003; Zirak et al., 2005; Gauden et al., 2005; Dragnea et al., 2005). Mutational studies identified the Y11 residue as the electron donor to the excited flavin (Gauden et al., 2007; Bonetti et al., 2009). In addition, H/D exchange experiments revealed a pronounced kinetic isotope effect, suggesting that the photoinduced electron transfer is coupled to proton transfer (Gauden et al., 2006;

Bonetti et al., 2008). The recombination of the radical intermediates on a nanosecond timescale finally yields the red-shifted photoproduct (Gauden et al., 2006). Because the Q53 side chain is located between the electron donor Y11 and the electron acceptor flavin, the photoinduced proton-coupled electron transfer (PCET) reaction is mediated by Q53 and must ultimately result in changes of the Q53 side chain in the photoproduct: a rotation (Gauden et al., 2006; Unno et al., 2006; Jung et al., 2006; Grinstead, Avila-Perez, et al., 2006) or a tautomerization (Domratcheva et al., 2008; Khrenova et al., 2010) of the Q53 side chain were proposed to explain the red-shifted flavin absorption in the light state. The fate of the Q53 side chain during the BLUF photoreaction is one of the central topics of this thesis. Therefore, the relevant entity for the photochemistry in BLUF is not the flavin chromophore alone, but the flavin-glutamine-tyrosine complex, which this thesis refers to as the photoactive triad of BLUF. The minimal model for BLUF photochemistry thus constitutes of the LF-Q-Y complex, as shown in Figure 1.1.

Remarkably, two X-ray crystal structures of the AppA-BLUF domain revealed two conformations of the $\beta 5$ strand that alters the hydrogen-bonding pocket of flavin: in one structure a tryptophan residue points towards flavin (Anderson et al., 2005) and in the other a methionine residue (Jung et al., 2006), apparently even influencing the rotamer conformation of the critical Q53 residue. These two structures with their differently assigned Q53 rotamer conformations were discussed as possible dark and light state structures of BLUF (Gauden et al., 2006; Unno et al., 2006; Grinstead, Avila-Perez, et al., 2006; Stelling et al., 2007; Bonetti et al., 2008, 2009; Mathes, Zhu, et al., 2012; Rieff et al., 2011; Meier et al., 2012; Hsiao et al., 2012), in addition to the alternative mechanism of glutamine tautomerization (Domratcheva et al., 2008; Sadeghian et al., 2008; Khrenova et al., 2010). Chapter 4 gives a detailed account on this debate and the two structures with the possible Q53 rotamer conformations.

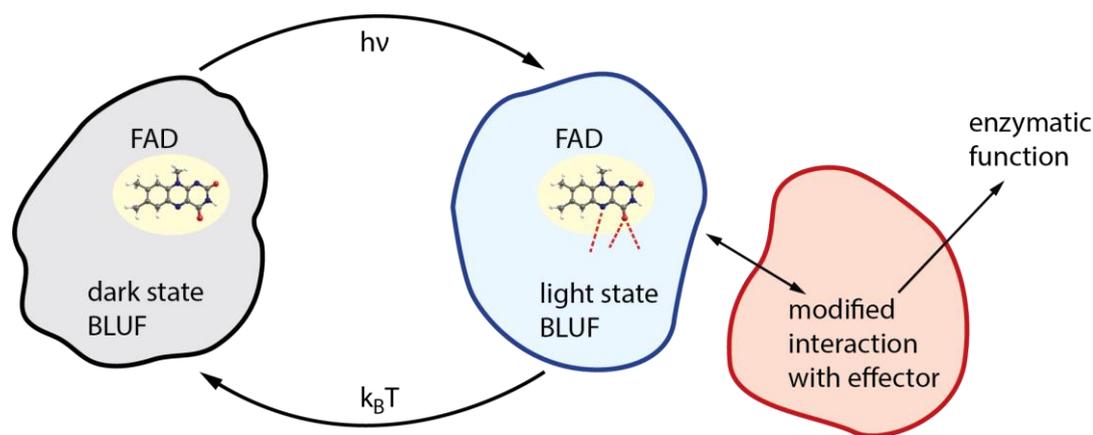


Figure 1.3. Schematic illustration of a BLUF photoswitch. For details see text.

In summary, the structural and spectroscopic studies identify the BLUF photoreceptor essentially as a light-sensitive molecular hydrogen-bonding switch (Kennis & Mathes, 2013).

Figure 1.3 illustrates the photocycle of this switch. Upon light absorption, the flavin chromophore undergoes a complex photoreaction involving the local BLUF protein environment that results in the formation of the red-shifted photoproduct within 1 nanosecond, which constitutes the primary event of photoactivation. The photoproduct contains the oxidized flavin chromophore in an altered hydrogen-bonding environment and initiates the signaling cascade. The altered local flavin-protein interactions, manifested in the switched hydrogen-bonding network around flavin, induce a conformational change of the whole BLUF sensor. The modified conformation influences the sensor-effector domain protein-protein interactions, leading to the enzymatic function of the effector and ultimately to the biological response of the cell.

Usually, the conformational transition of the sensor domain is referred to as a switch from the dark state to the signaling state of the photoreceptor. The timescale for the protein conformational transition is micro- to milliseconds (Henzler-Wildman & Kern, 2007), thus several orders of magnitude longer than the 1 nanosecond required for the formation of the spectroscopically observable photoproduct. This is essential because the information on the ultrafast event of light absorption has to be stored for a rather long time for the subsequent biochemical processes to occur. The photoproduct state is the “end state” of the photoreaction that involves the local environment of flavin, observed via the flavin spectroscopic shifts and is determined by the photoreaction quantum yield. This thesis refers to the red-shifted photoproduct formed within 1 nanosecond as the BLUF “light state” and not to the BLUF protein conformation in the signaling state. Recently, first studies appeared that address the light-induced conformational changes in BLUF proteins, for example in (Udvarhelyi, 2009; Brust et al., 2013; Winkler et al., 2014), yet it is difficult to gain atomic-level insight into the protein structural changes underlying signal transduction.

1.2 Photophysical and photochemical concepts

This section introduces the fundamental physical concepts to describe and understand processes relevant to light-absorbing molecules which undergo a photoreaction. The concept of the Born-Oppenheimer (BO) potential-energy surface (PES) provides the “landscape” for molecular reactions. Because electrons are much lighter than nuclei, their motion is much faster than the vibrational motion of nuclei, i.e. the relevant time scales are well-separated. Therefore, the adiabatic BO approximation considers the electrons instantaneously following the heavy nuclei (Born & Oppenheimer, 1927). The eigenvalues of the electronic Schrödinger equation at given nuclear configurations yield the BO-PES. A large molecule with many vibrational degrees of freedom obviously possesses a complex PES. The minima on this surface represent equilibrium

molecular structures and the saddle points transition states, mediating conformational changes or chemical reactions. The BO-PES concept is applicable for each electronic state. In photophysical and photochemical processes at least two PESs (excited and ground state) are involved and their characterization requires the determination of the two surfaces including the region where the surfaces are degenerate (i.e. the surface crossing seam). However, photochemistry is not limited to two electronic states. A prominent example is PET where two excited states (one of the excited chromophore and an electron-transfer state) and the ground state are involved. As this thesis will show, the BLUF problem requires the description of at least three PESs.

In their seminal paper, Neumann and Wigner showed that in diatomic molecules two energy surfaces of the same symmetry or spin multiplicity cannot cross because these molecules have only one internal degree of freedom (von Neumann & Wigner, 1929). Using their derivation, Teller argued that in contrast to diatomic systems, in polyatomic molecules a degeneracy may occur (Teller, 1937). The surface crossing was termed conical intersection because of the double-cone shape of the surfaces in the appropriate two-dimensional subspace of the nuclear coordinates (Teller, 1937; Robb et al., 1995; Klessinger, 1995; Bernardi et al., 1996; Domcke et al., 2011). The BO approximation breaks down at the conical intersection because non-adiabatic couplings between the two states become infinite. There is a high probability of non-adiabatic transitions between the degenerate electronic states, which provides an efficient and instantaneous channel for radiationless dissipation of the excitation energy (Teller, 1937; Robb et al., 1995; Bernardi et al., 1996; Domcke et al., 2011). Conical intersections were found to be ubiquitous in polyatomic molecules (Robb et al., 1995; Bernardi et al., 1996; Yarkony, 1996; Domcke & Yarkony, 2012), and play a key role in the radiationless decay of excited states. A thorough theory of conical intersections is given, for example, in (Domcke et al., 2011).

Figure 1.4 sketches some of the relevant processes occurring in light-absorbing molecules involving only singlet states (of interest in the BLUF problem). According to the Franck-Condon principle, photon absorption is a “vertical” process because it occurs much faster than the vibrational motion of nuclei. The molecule may return to the electronic ground state via the emission of a photon (radiative deactivation, fluorescence) or via internal conversion through a conical intersection (radiationless deactivation) or via an intersystem crossing (not considered here). In general, transitions that interconvert excited states with each other or excited and ground states without changing the molecular geometry significantly, are called photophysical processes. In general, transitions that yield chemically different molecular structures or different reactive intermediates are called photochemical processes. Often the radiationless deactivation does involve a significant geometry deformation of the molecule to reach a conical intersection. At the conical intersection (with the significantly altered geometry) the photoreaction may enter either (i) a photostable pathway, along which the reactant is regenerated (aborted photochemical reaction) or (ii) a photoactive pathway, along which a chemically distinct species, the photoproduct is formed (photochemical reaction). The topology of the potential energy surfaces

in the vicinity of the conical intersection determines to a significant extent which pathway will be preferred at the conical intersection. The second key determinant is the dynamics: the velocity components in the two directions of the double-cone coordinate subspace determine the branching among the various possible pathways at the conical intersection (Klessinger, 1995).

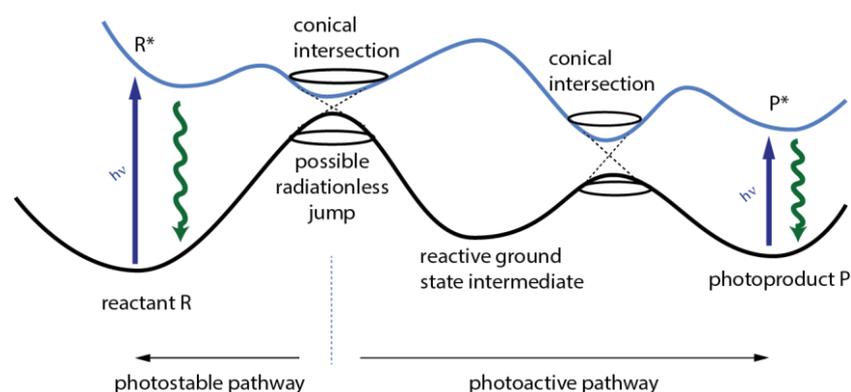


Figure 1.4. Schematic representation of photophysical and photochemical processes relevant in this work. For details see text.

The accurate determination of the electronic PES and the characterization of conical intersections is a prerequisite to understand photostability and photoactivation in biomolecules. State-of-the-art *ab initio* electronic-structure methods provide the means and tools to do so, however, the technical aspects of exploring potential energy surfaces and an overview of electronic-structure methods cannot be given here and the reader is referred to the literature (Helgaker et al., 2000; Cramer, 2004; Serrano-Andrés & Merchán, 2005; Olivucci, 2005; Jensen, 2011). More recently, computational methods of treating excited states and exploring potential energy surfaces in rather large molecules were reviewed in (González et al., 2012) and of characterizing conical intersections in (Schapiro et al., 2011). The next section briefly introduces the electronic-structure methods used in this thesis.

1.3 Electronic-structure methods

The exact form of the Schrödinger equation for a polyatomic system describing all its chemical properties is known, however, too complex to be solved. Therefore, as Paul Dirac called for in 1929, approximate methods are needed to apply the quantum-mechanical laws to explain the main features of molecular systems. Since then, many methods have been developed and the field of quantum chemistry was born. This section briefly outlines the methods pertaining to the excited-state calculations of rather large molecules like flavins.

The computed UV-Vis and IR spectra of flavin are in very good agreement with experiments, which shows the merit of the quantum-chemistry methods and raises confidence for the predictive capabilities of the calculations to answer experimentally unresolvable questions. Very good agreement in this context means that the calculations reproduce spectral signatures and explain the origin of spectral shifts. This is the case in the calculations of flavins where essentially most methods yield similar results. The computational aspects of computing excited states in flavins were recently reviewed in (Domratcheva et al., 2014).

To compute absorption and emission spectra, first the coordinates of the molecule are optimized in the ground and excited electronic state with a given method. At the optimized geometries, the vertical excitation energies are usually computed with a more accurate, higher-level method. This procedure is denoted `property-method//geometry-method` in the following. The most popular computational method for ground-state properties is based on density functional theory (DFT) because of its reasonable accuracy and computational efficiency (Hohenberg & Kohn, 1964). In DFT the electronic energy is expressed as a functional of the electron density. The exchange-correlation energy term in the Kohn-Sham formulation has to be approximated (Kohn & Sham, 1965). The semi-empirical hybrid exchange-correlation functional B3LYP (Lee et al., 1988; Becke, 1993) is most widely used in the literature, also employed in the calculations presented in Chapter 4. Over the years many density functionals have been developed, recently reviewed, for example, in (Sousa et al., 2007; Grimme, 2011; Peverati & Truhlar, 2014). Alternative methods for ground-state calculations are the Møller-Plesset perturbation theory of second order (MP2) method (Pople et al., 1976) or the approximate second-order coupled-cluster with singles and doubles (CC2) method (Christiansen et al., 1995), also used for geometry optimization in Chapter 2. CC2 and MP2 feature similar accuracy and in contrast to DFT, they are *ab initio* wave-function based methods.

Similar to DFT for ground-state properties, the most popular choice for computing excited-state properties in large molecules, is the time-dependent TD-DFT method (Runge & Gross, 1984; Marques & Gross, 2004; Casida & Huix-Rotllant, 2012). TD-DFT is based on linear response theory where the excitation energies are obtained without solving the complete Kohn-Sham equations explicitly. The major limitation of the TD-DFT method is a rather poor description of charge-transfer states (Tozer et al., 1999) and a gross underestimation of charge-transfer excitation energies (Dreuw et al., 2003). For a qualitatively correct description of the excited states, the configuration interaction singles (CIS) wave function can be used (Foresman et al., 1992). Besides a HF reference wave function, it includes a linear combination of all possible singly-excited Slater determinants. The configuration coefficients are variationally optimized and characterize each state. The CIS method is rather fast and applicable to large systems, yet vertical excitation energies are severely overestimated. Including also doubly-excited determinants in the expansion, yielding the corresponding CISD wave function, usually results in little improvements (Foresman et al., 1992). A real improvement to the CIS and CISD methods are the so-called

spin-component scaled (SCS) and the scaled opposite-spin (SOS) variants, SCS-CIS(D) and SOS-CIS(D) (Rhee & Head-Gordon, 2007). SOS-CIS for example was shown to provide excitation energies in good agreement with experiment (Khrenova et al., 2010). Computationally more costly, yet usually more accurate excitation energies are obtained with the response-theory based CC2 method, also used in Chapter 2. It is common to optimize the geometry in the excited state with TD-DFT and subsequently compute the excitation energies with CC2, denoted as a CC2//TD-DFT protocol, used for example in (Sadeghian & Schütz, 2007; Sadeghian et al., 2008, 2010). A further method to obtain excited-state properties is the (second-order) algebraic-diagrammatic construction method, ADC(2), based on a polarization-propagator approach that is devised for the treatment of valence excitations in rather large molecules (Trofimov & Schirmer, 1995).

In addition to correctly estimating excited states, the treatment of photochemical problems requires the description of bond-breaking processes, biradicals and conical intersections, all cases where the merits of the above-mentioned methods is intrinsically limited. In these cases the use of a multi-determinant reference wave function is required, for example as in the multi-reference CI (MRCI) approach. MRCI//TD-DFT calculations were employed for example in (Salzmann et al., 2008, 2009). If also the MOs are variationally optimized in addition to the configuration coefficients, the so-called multi-configurational self-consistent field (MCSCF) procedure is carried out. Unlike single-reference methods, the MCSCF method yields a qualitatively correct description of practically any chemical electronic-structure problem. Chapter 2 gives examples for MCSCF wave functions and introduces an approach to apply it to the treatment of the BLUF photoreaction.

1.4 Molecular models

The molecular models considered in this thesis are based on the available X-ray crystal structures of the AppA BLUF domain, as well as on the BlrB and PixD proteins (Jung et al., 2006, 2005; Yuan et al., 2006). Several models were built in the course of this project, thus each chapter defines the models considered separately. As explained above, the minimal model to study BLUF photochemistry includes the photoactive triad. In the simplest case a so-called cluster model is constructed (Siegbahn & Himo, 2011) based on the coordinates of the lumiflavin part of the chromophore and the Y11 and Q53 side chains: their coordinates are “taken out” of the protein structure and capped with hydrogen atoms, exemplified in Figure 1.5A. The geometrical constraints of the missing protein environment are introduced through the coordinate locking scheme. Interactions with the local environment in some cases are considered by including respective side chains into the cluster. As a first step for example, the asparagine side chain,

which restrains the position of flavin by means of two hydrogen bonds, may be included to mimic the specific arrangement of the chromophore-binding pocket. These hydrogen bonds are important as they may also influence the redox properties of flavin. Further hydrogen-bonding partners were added to the cluster in the calculations presented in this thesis, as exemplified in Figure 1.5B. This procedure makes the calculations substantially more expensive. However, the obvious advantage is that all interactions are treated at the same (high) level of theory with the chosen quantum-chemical method, which is why this thesis refers to the local environment of the photoactive triad in such cluster models as the QM environment. In all figures in this thesis oxygen atoms are colored red, nitrogen atoms blue, hydrogen atoms white and sulfur yellow; carbon atoms are depicted in gray, green or orange, to distinguish certain parts of the model.

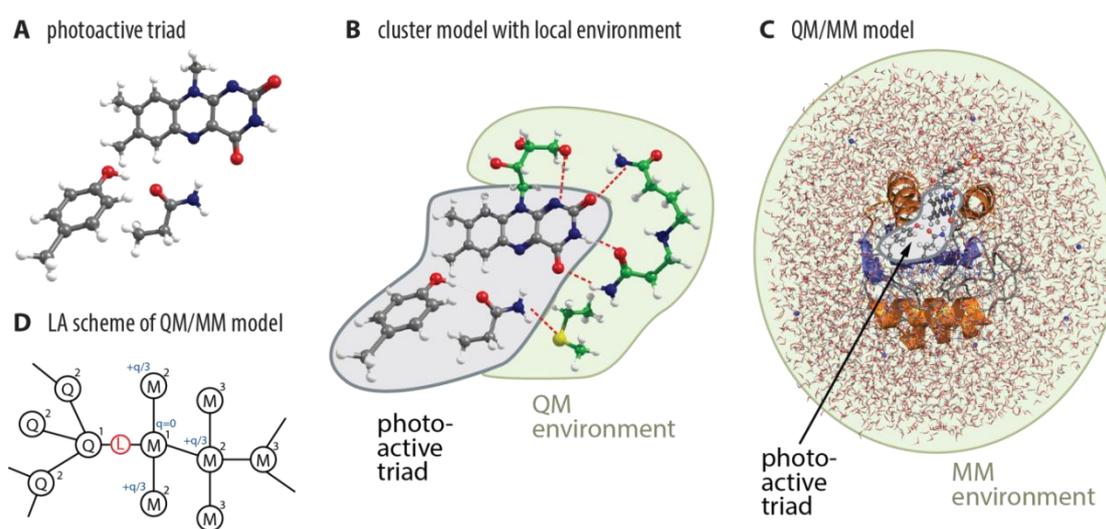


Figure 1.5. (A) Minimal model of BLUF: the photoactive triad cluster model comprising LF, Y21 and Q63. (B) Larger cluster model that includes side chains of the local environment of the photoactive triad, treated at the same level of theory as the photoactive triad. (C) Hybrid QM/MM model of BLUF in a water droplet, with the photoactive triad in the QM subsystem and the rest of the protein together with the water and ion atoms in the MM subsystem. (D) QM/MM LA-scheme (inspired by (Senn & Thiel, 2009)). Q labels the QM atoms, M labels MM atoms, L labels the link atom at the QM/MM boundary. The partial charges illustrate the redefinition of the MM atomic charges of the MM2 atoms.

Because the protein and solvent environment in tuning spectral properties and controlling the photodynamics of biological chromophores usually plays an important role, this thesis also considers so-called hybrid QM/MM models, recently reviewed in (Lin & Truhlar, 2007; Senn & Thiel, 2009; Ananikov et al., 2010). The idea has been introduced by the 2013 Nobel Prize Laureates Warshel and Levitt in the study of an enzymatic reaction (Warshel & Levitt, 1976). Warshel and Levitt suggested to treat the biochemically active part of the protein (in this case the photoactive triad) at the QM level (QM subsystem), and the rest of the protein at a classical force-field-based molecular-mechanics (MM) level (MM subsystem), as indicated in Figure 1.5C.

It is generally assumed that the QM part can describe the reactive chemistry accurately, taking into consideration the interactions with the whole environment in an adequate way. Properties like, for example, protonation energies of the protein active site cannot be reproduced if the charge distribution of the surrounding protein is neglected (Warshel & Levitt, 1976; Kamerlin et al., 2009). With the QM/MM approach, the simulation of several hundred thousand atoms becomes possible. The coupling of the QM and MM subsystems is introduced by QM/MM interaction terms in the QM Hamiltonian and by the QM and MM descriptions of the boundary region.

The energy function of the QM/MM approach is given in Equation (1.1). The total energy of the system is determined by the QM Hamiltonian of the QM subsystem, the MM force field function of the MM environment and an explicit QM/MM coupling term describing the interactions between the QM and MM subsystems:

$$H = H_{QM} + H_{MM} + H_{QM/MM} \quad (1.1)$$

The QM Hamiltonian is determined by the quantum-chemical method used. The MM energy function represents the classical force-field terms: bond stretching, angle bending, dihedral torsions, improper torsions (the simple bonded terms), as well as the Lennard-Jones-type van-der-Waals contribution and the Coulomb interactions between fixed point charges. The explicit coupling term $H_{QM/MM}$ defines the given QM/MM method, where the electrostatic-interaction energy between the two subsystems is the most important term. This work uses the so-called electrostatic potential fitting (ESPF) scheme, that simply uses monopoles (Ferre & Angyan, 2002). This scheme will be described and discussed in detail in Chapter 3. More advanced QM/MM schemes, which are still not common, account also for the polarization of the MM charges by the QM subsystem (Monari et al., 2013) or include QM/MM interactions of the higher order electrostatic moments and additional explicit terms, for example, for dispersion or charge transfer (Nemukhin et al., 2003; Leverentz et al., 2011; Gordon et al., 2013).

Splitting the system into a QM and MM subsystem requires an adequate treatment of the covalent bonds across the QM/MM boundary. This work uses the most common, so-called link-atom (LA) approach (Singh & Kollman, 1986), explained in Figure 1.5D. This approach introduces an additional hydrogen atom into the QM subsystem that covalently bounds to the carbon atom at the boundary and thus caps the dangling bond. Usually, a special treatment of the boundary MM atom is required, for example its charge is reset to zero in order to avoid overpolarization of the QM subsystem close to the boundary. The hydrogen LA does not take part in the QM and MM interactions and is “invisible” to the MM subsystem.

1.5 Aims of this thesis

The central aim of this thesis is to map the potential-energy surface characterizing the BLUF photoreaction. Of particular interest are reaction pathways underlying the photoactivation of the dark state, the photostability of the light state and the redox effects observed in BLUF, controversially interpreted in the literature. The prerequisite for mapping the potential-energy surface is an adequate molecular model. The work presented in this thesis formulates the methodology, exploiting state-of-the-art multi-configurational multi-reference electronic-structure methods, and establishes an adequate computational protocol to achieve this goal.

This thesis is organized as follows. In Chapter 2, I explain the computational approach devised to specifically treat a large molecular system featuring the BLUF photoactive triad (flavin, glutamine, and tyrosine) with high-level multi-reference electronic-structure methods. In Chapter 3, I present benchmark calculations that compare the cluster and QM/MM approaches for the computations of the BLUF excited-state energies and of the intermolecular interaction effects. In Chapter 4, I focus on the available BLUF X-ray and NMR structures and analyze in detail the possible glutamine rotamers and present glutamine dihedral-energy potentials defining the dynamical properties of the flavin-binding pocket. In Chapter 5, I present the results of the photoinduced PCET pathway calculations in several models of BLUF and discuss the emerging mechanistic picture of BLUF photoactivation, photostability, and redox tuning. A summary and outlook are provided in Chapter 6.

2 Method benchmarking: POCAS

The PT2//CASSCF method is widely used in studies of photophysical and photochemical reactions in organic molecules. In this chapter, I present a simple and physically intuitive way of constructing a CASSCF wave function in order to enable calculations on large systems such as the BLUF photoactive triad. First, a CASSCF wave function accounting for only static electron correlation is constructed; the MOs that provide this description of the excited states of interest are called principal orbitals. Including them in the active space yields the uniquely defined POCAS (principal-orbital complete-active-space) wave function. The dynamic electron correlation is included through the state-of-the-art multi-reference perturbation-theory methods, XMCQDPT2 and MS-CASPT2. POCAS is a computationally efficient alternative to the conventional way of choosing the active space in CASSCF calculations, according to which all MOs (or as many as possible) of the π and π^* subspace are included. The POCAS approach provides a balanced description of the excited states of interest, even if they have different electronic character.

After a general introduction to CASSCF, I present the computational details of the two benchmark systems lumiflavin and the BLUF photoactive triad. Next, I compare the lumiflavin excited-state energies and the optimized S_0 and S_1 geometries computed with various active spaces and also with other quantum-chemistry methods. I also benchmark the excited-state energies in the BLUF photoactive triad, by comparing the CASSCF results with those of other methods. Section 2.3.3 in this chapter contains one figure (Figure 2.5) and some text from (Udvarhelyi & Domratcheva, 2011).

2.1 CASSCF and POCAS

Roos and coworkers pioneered the complete-active-space self-consistent field (CASSCF) method (Roos, 1972; Roos et al., 1980; Roos, 1987), that marks a cornerstone in applying the MCSCF

procedure to chemical problems. The CASSCF method reduces the problem of selecting important configuration state functions (CSF) for the MCSCF procedure to selecting important molecular orbitals for the so-called complete active space (CAS): The CASSCF wave function is the linear combination of all possible CSFs that can be constructed by distributing all active electrons among the active MOs. Thus a CASSCF calculation corresponds to performing a full configuration interaction calculation in the chosen active space. The remaining occupied and virtual MOs stay doubly occupied and empty, respectively, however, are also included in the orbital optimization during the SCF iterations. When treating several states in the CASSCF calculation, the energy optimization should be carried out in a state-averaged manner to ensure the orthogonality of the different wave functions (Werner & Meyer, 1981). Usually equal weights on the ground and excited state wave functions are used.

By construction, the CASSCF wave function accounts for quasi-degenerate correlation effects, which are often referred to as static or non-dynamic correlation. The CASSCF ansatz describes the correlation between two electrons at a large separation in space (long-range effects arising from mixing of degenerate states); e.g. a full valence CASSCF wave function will always dissociate a molecule correctly. However, the CASSCF wave function cannot account for short-range correlation effects because it cannot describe the electron cusp owing to the instantaneous repulsion of electrons at short distances, which is often referred to as dynamic correlation. The definition and partitioning of the correlation energy into static and dynamic parts is not well-defined, nevertheless it is an important and often useful concept in quantum chemistry. A given active space is going to be usually realistic if dynamic correlation remains constant along a reaction coordinate, leading to reasonable CASSCF-optimized geometries. At those geometries, dynamic correlation can be accounted for by multi-reference (multi-state) perturbation theory with the CASSCF wave function as a reference (Schmidt & Gordon, 1998). Because of the computational cost, one only computes the perturbation-theory energy correction until second order (MR-PT2 methods). There exist different ways to realize the multi-reference perturbative energy correction, like CASPT2 (Andersson et al., 1992) or NEVPT2 (Angeli et al., 2001), or those multi-state methods which allow for state interactions as well, like MS-CASPT2 (Finley et al., 1998) or XMCQDPT2 (Granovsky, 2011). This work employs CASPT2, MS-CASPT2, and XMCQDPT2. In the studies of potential energy surfaces, where many geometries of the molecule of interest are relevant, it is likely that so-called intruder states will deteriorate a region of the PES and result in unphysical energies. As a standard solution to this problem, the energy denominator around poles is shifted so that far from the poles the effect of the shift is insignificant. Throughout this work, the energy-denominator shift technique was used in the XMCQDPT2 calculations with a recommended edshft value of 0.02 hartree (Witek et al., 2002), and, likewise, in the CASPT2 calculations with an imaginary shift value of 0.2 hartree (Forsberg & Malmqvist, 1997). In contrast to the MOLCAS recommendation (www.molcas.org), the so-

called ipsea shift (Ghigo et al., 2004), an empirical parameter, was not used in the calculations of this thesis.

In recent decades, the success of computing spectroscopic properties and excited-state relaxation pathways of small molecules with the CASPT2//CASSCF approach made it a standard, though computationally rather expensive method. Recent instructive examples are given in the works of Olivucci and coworkers (Altoè et al., 2010; Gozem et al., 2012; Bernini et al., 2013; Rinaldi et al., 2014) or Domcke and coworkers (Schultz et al., 2004; Sobolewski et al., 2005; Frutos et al., 2007; Domcke & Sobolewski, 2013; Sobolewski & Domcke, 2006). The decision on the MOs comprising the active space represents the challenge in carrying out PT2//CASSCF calculations. There only exist rules of thumb but not a rigorous theory to guide the selection. Usually, the “chemically intuitive” full π electronic system with chemically important lone pairs is required to be included in the active space in order to be considered “balanced” (Veryazov et al., 2011). However, for larger molecules this recommendation becomes rather impractical since computational cost prohibits considering more than 14 electrons in 14 MOs. Organic molecules of the size of lumiflavin possess an extended π -bonding system and several lone pairs. Usually a “top-down” approach serves as a solution to construct a computationally affordable active space, where one tries to reduce the “full π plus lone pairs” active space by excluding some MOs. One possibility is usually to exclude lone pairs on oxygen and nitrogen atoms (n MOs), in case one is not interested in the corresponding $n\pi^*$ excited states. If this still results in an active space too large, one can try to exclude also the lone pairs that are part of the π system (p MOs) and check the resulting energies and occupation numbers of natural MOs. Occupied MOs with occupation numbers close to two and unoccupied MOs with occupation numbers close to zero are considered to be excludable from the active space (Veryazov et al., 2011). Usually, one tries to keep the number of active space occupied and virtual MOs approximately the same.

For large molecules this top-down approach often leads to the so-called “intermediate” active spaces. However, these might not represent the best choice for computing excited states and excited state relaxation pathways. For example, Martínez and coworkers already presented several examples where a rather small active space with a “chemically unintuitive” choice of the MOs (not selected based on the “top-down” approach) yields better energies for valence excited states than the full π CAS: An active space with only two-electrons-in-two MOs of the GFP chromophore yielded reasonable potential energy surfaces and allowed for extensive dynamics calculations in (Olsen et al., 2010); another example is the photochemistry of malonaldehyde where the authors argued that the four-electrons-in-four-MOs active space was not inferior to a large active space in locating and characterizing conical intersections in (Coe & Martínez, 2006); moreover, in (Levine & Martínez, 2009) the authors show that the full- π active space in butadiene even predicts a qualitatively incorrect PES. In the case of large molecules the full π -valence active space results in large active spaces that partially also account for dynamic correlation – already by

the sheer number of CSFs. In contrast, the smaller active space represents a more balanced treatment of correlation effects in only describing the static correlation.

In (Domratcheva et al., 2006), the idea of using a small, or “minimal” active space in the excited-state computations of flavins was introduced: the active space should only comprise the MOs needed for the description of the excitations of interest. This is especially simple for single excited states like the HOMO-LUMO excitations. Often the low-lying excited states of interest are already known or, in case not, they can be easily identified by methods like CISD, CC2, or TD-DFT, which do not require any selection of the active MOs. In this thesis, the most important MOs from which and to which the electron excitation occurs, are called principal orbitals (POs). It is proposed that only those principal orbitals should be included in the active space, thus yielding the principal-orbital complete-active-space (POCAS) wave function. The POCAS wave function can be an advantage when the aim is to treat excited states of different chemical character in a balanced way because, by construction, POCAS accounts for static correlation and multi-reference perturbation theory recovers the dynamic correlation for all states in a balanced way. Here, CASSCF with the POCAS selection of the active space is tested for geometry optimization to elucidate photochemical pathways. Besides this work, the POCAS approach was also successfully employed in the studies of other flavoproteins, like the 6,4-photolyase (Domratcheva, 2011; Moughal Shahi & Domratcheva, 2013) or cryptochrome (Solov'yov et al., 2012, 2014). This chapter contains the first systematic comparison of excitation energies with various active spaces and demonstrates the merit of the POCAS approach.

2.2 Computational details

First, lumiflavin is considered at the DFT-B3LYP optimized geometry, with the DZV(P) basis set with polarization functions on the “heavy” atoms. First, the excitation energies were computed by a number of single-reference methods, CIS, CISD, ADC(2), CC2, and TD-B3LYP to identify the structure of low-lying single-electronic excitations. The TZVP basis set was used; the computations were carried out with the quantum-chemistry software TURBOMOLE (ver. 6.3, TURBOMOLE GmbH, Karlsruhe, Germany). Next, PT2-CASSCF calculations with the DZV(P) basis set were carried out, at the same B3LYP/DZV(P) geometry. The RHF MOs (or in some cases CASSCF MOs) served as starting MOs for the CASSCF wave function optimization. All CASSCF wave functions were optimized using state averaging with equal weights for each state. The CASSCF wave function averaged over r number of states with n number of active electrons in m active MOs is denoted as $\text{CASSCF}(n,m)r$. As PT2 theories XMCQDPT2, CASPT2, and MS-CASPT2 were employed. $\text{PT2}(n,m)r$ denotes the energies computed with either MS-CASPT2 or XMCQDPT2, using the $\text{CASSCF}(n,m)r$ zero-order wave function. The

XMCQDPT2-CASSCF calculations were carried out with the quantum-chemistry package Firefly (Firefly ver. 7.1, 8; Granovsky) which is partially based on the GAMESS US source code (Schmidt et al., 1993); whereas the (MS-)CASPT2-CASSCF calculations were carried out with the MOLCAS package version 7.6 (Aquilante et al., 2010).

To evaluate the effect of the active-space composition on the equilibrium geometry, the ground state of the lumiflavin molecule was optimized with different active spaces, including the POCAS and the conventional way of choosing the MOs, with the DZV(P) basis set with Firefly. The LF geometry was also optimized with different methods using different software: CC2/cc-pVDZ and BHLYP/def-TZVP with Turbomole; and RHF/DZV(P), MP2/DZV(P), and B3LYP/DZV(P) with Firefly. The geometry optimization was performed without imposing symmetry constraints. In addition, it was also possible to optimize the LF geometry with Cs symmetry constraints at the XMCQDPT2(2,2)2/cc-pVDZ level with numerical gradients with the newest version 8.0 of the Firefly code. In this case, the recommended extra-tight convergence thresholds, a numerical gradient order of 4, and an edshft value of 0.02 hartree were applied. The S_1 excited-state geometry of LF was optimized with CC2/cc-pVDZ and TD-BHLYP/def-TZVP in Turbomole; and with CASSCF/DZV(P) using different active spaces with Firefly.

The supermolecular cluster model LF-Q-Y-N consisting of the photoactive triad and the N45 residue was built based on the AppA X-ray crystal structure (PDB code 2IYG) and optimized with B3LYP/DZV(P). To mimic the restraints of the protein, the geometry optimization was carried out with the terminal carbon atom in each molecular fragment fixed at the crystallographic positions. The absorption spectrum was computed with CIS/DZV(P) and XMCQDPT2-CASSCF/DZV(P) using different active spaces with the Firefly package.

2.3 Results and discussion

2.3.1 Low-lying excited states in lumiflavin

The π system of lumiflavin consists of seven double bonds and two nitrogen lone pairs on the N3 and N10 atoms (for the LF atom numbering see Figure 1.1), denoted p_N . There are further lone pairs of sigma symmetry on the N1 and N5 atoms, denoted n_N , as well as two lone pairs on the O2 and O4 atoms each, denoted n_O . The active space of the all these MOs, comprises 30 electrons in 22 MOs, denoted CAS(30,22). With simple single-reference methods the principal orbitals describing the excited states of interest were identified. In Figure 2.1, the following excited states are indicated: the lowest-lying singlet excited state is a $\pi\pi^*$ transition from the HOMO to the LUMO, denoted as S_1 ; above it lie two $n\pi^*$ states, involving n_N and n_O MOs, denoted as S_n ; and another $\pi\pi^*$ state, corresponding to a transition from the flavin HOMO-1 to

the LUMO, denoted as S_2 . The $n\pi^*$ states show considerable mixing of various transitions and the structure of the excitations differs substantially in the various quantum-chemical treatments (Salzmann et al., 2008; Neiss et al., 2003). The ADC(2), CC2 and TD-B3LYP excitation energies are close, whereas the CIS and CISD excitation energies are significantly blue-shifted in comparison. The energies are in good agreement with the MRCI//B3LYP/TZVP calculations from (Salzmann et al., 2008). The S_1 and S_2 excitation energy estimates are slightly blue-shifted compared to the experimental FMN absorption maxima in water (Figure 1.2A). Importantly, all methods preserve the order of the states and reveal the principal orbitals involved in the excitations of interest, so that even CIS may be used to identify them. TD-DFT is well-suited to compute the flavin S_1 excitation energy, which will be used in Chapter 4.

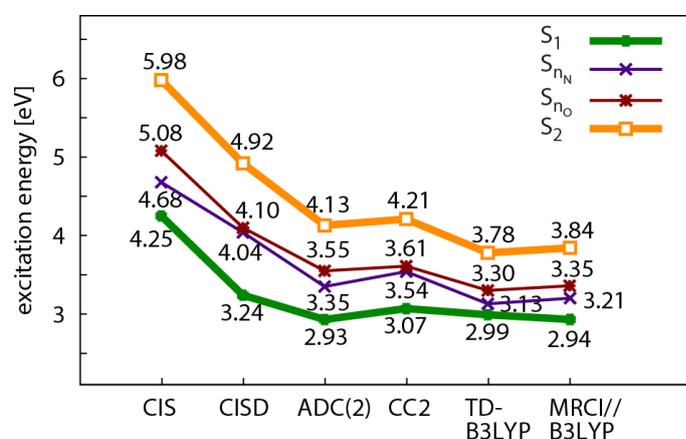


Figure 2.1. Excitation energies of the low-lying states of lumiflavin computed with different methods. The MRCI//B3LYP values are taken from (Salzmann et al., 2008).

In the top-down approach, neglecting the non- π symmetry lone pairs reduces the active space to the full π -valence CAS(18,16), which is at the computationally possible limits. Excluding all lone pairs gives the (14,14) active space, and excluding one π^* MO, the (14,13) active space, used, for example, in (Climent et al., 2006). To consider also $n\pi^*$ excited states, the n MO of interest is added to the (14,13) active space. In the POCAS approach, the selection of the MOs for the active space follows a different logic: The active space is built on the basis of the identified (minimum number of) principal orbitals, according to the excited states of interest. Thus, the flavin S_1 state is computed with a (2,2)2 active space, including the flavin HOMO and LUMO (denoted as (2,2)2- S_1); the S_n state can be computed also with a (2,2)2 active space, but including the n_N MO and the LUMO (denoted as (2,2)2- S_n). The S_1 and S_2 states are considered with a (4,3)3 active space, including the flavin HOMO-1, HOMO, and LUMO orbitals (denoted as (4,3)3- S_1S_2). The (8,5)5 active space includes two lone pairs (one n_N and one n_O), the HOMO-1, HOMO, and the LUMO orbital. Figure 2.2 compares the computed XMCQDPT2-CASSCF energies of the S_1 , S_2 and S_n excited states in the POCAS series (left panel) to the (MS-)CASPT2-CASSCF energies in the series with the conventional way of reducing the active space (right panel). The conventional series was started from the (18,16) active space including the full π system of LF and two p_N lone pairs. The MOs comprising the respective active space in Figure

2.2 are explicitly indicated. The MOs and corresponding occupation numbers of the (18,16) active space are given in Figure 2.3.

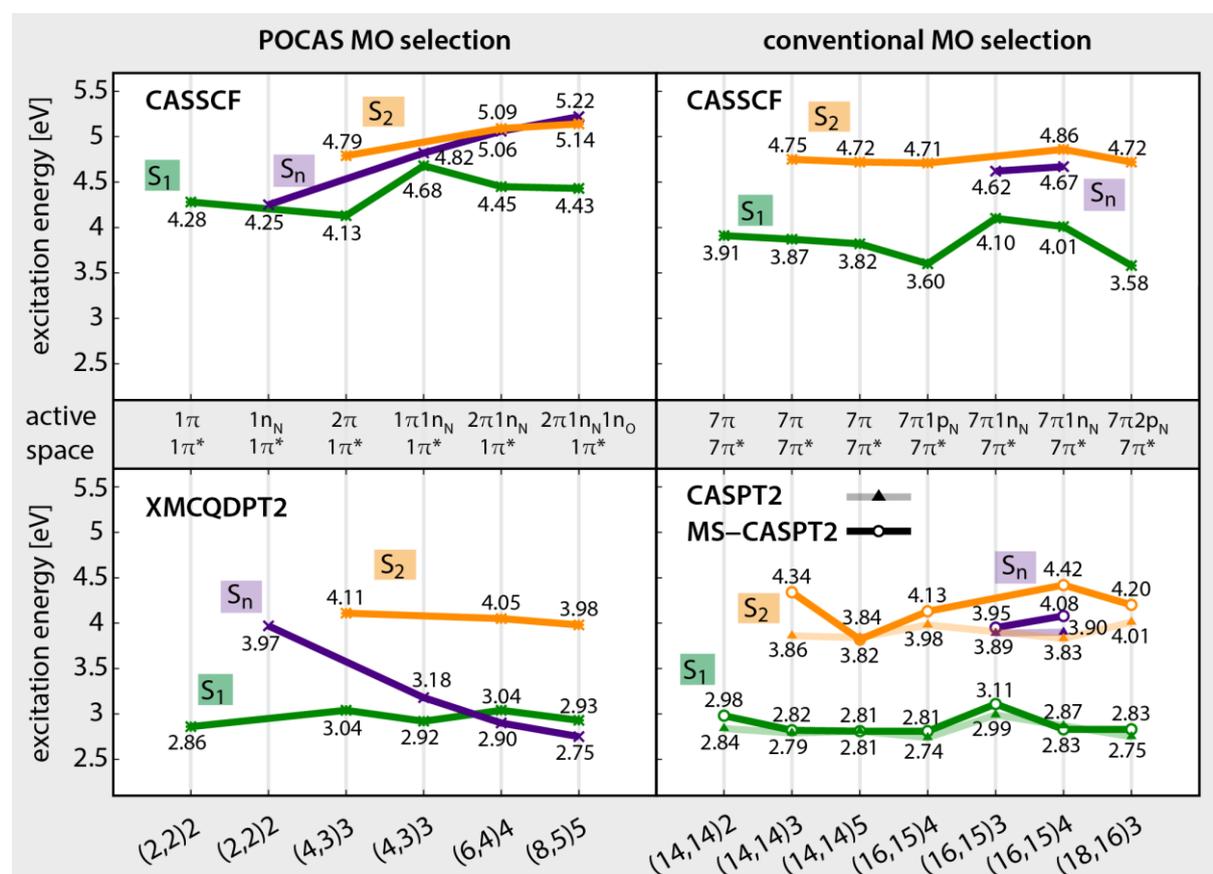


Figure 2.2. CASSCF and PT2 excitation energies of lumiflavin computed with the POCAS and conventional way of constructing the active space. The molecular orbitals comprising the active space are indicated explicitly.

At the CASSCF level, the variation in the excitation energies upon a change in the active space is considerable, both with the POCAS and the conventional way of selecting the active-space orbitals. Notably, there is a substantial variation among the results obtained with a (16,15)4 active space, depending on whether a p or an n_N MO is included in the active space. Furthermore, the S_1 excitation energy is lowered significantly, by 0.29 eV, when the (14,14)3 active space is extended to (18,16)3, although the two p_N MOs additionally included in the CAS(18,16)3 calculation have occupation numbers 2.00 and 1.98 and thus would be considered not necessary to treat in the active space, according to the guidelines given in (Veryazov et al., 2011). The CASSCF S_1 excitation energies obtained with the POCAS way of selecting the active-space orbitals are all higher than those obtained with the conventional selection; however, the variation among the POCAS excitation energies with a changing number of states is not larger than that of the full- π CASSCF results.

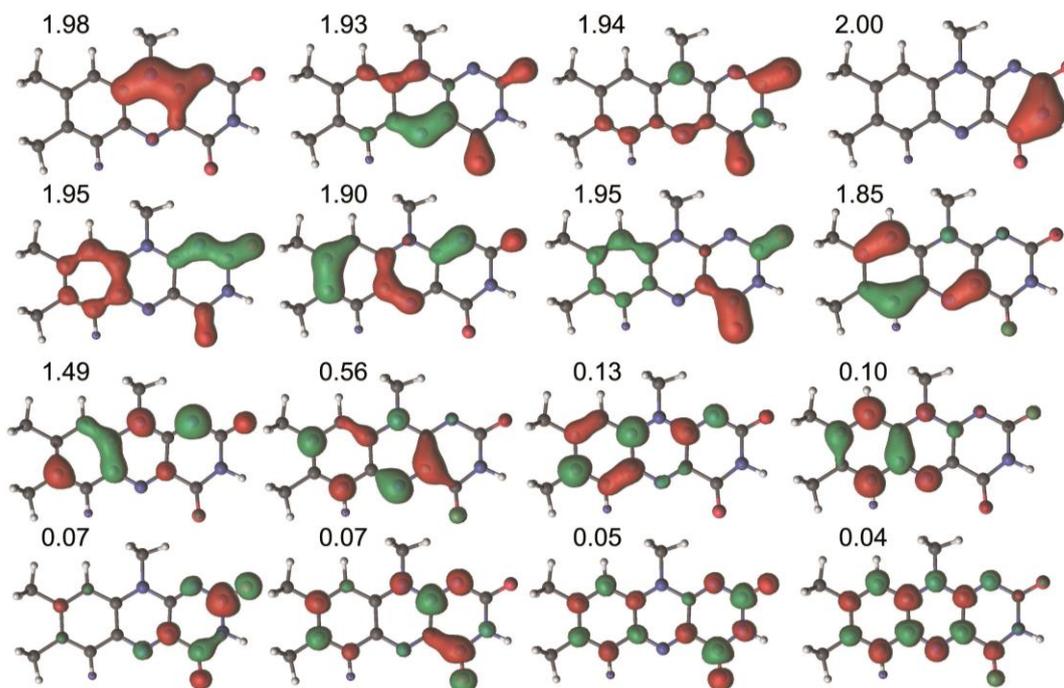


Figure 2.3. Lumiflavin CASSCF-optimized MOs and occupation numbers of the (18,16) active space. The MOs were visualized with the molcas gridviewer.

Accounting for the dynamic correlation with the PT2 method stabilizes the excited state with respect to the ground state, so that the excitation energies are reduced. A comparison of the excitation energies computed with the different PT2 methods (XMCQDPT2 versus CASPT2 and MS-CASPT2) will be given in Chapter 3. Here, it is noted that MS-CASPT2 predicts higher excitation energies of the three considered states and shows substantially more variation than CASPT2. Importantly, the computed CASPT2 and MS-CASPT2 energies depend significantly on the number of states considered. Regarding the results with the POCAS way of choosing the active space MOs, at the XMCQDPT2 level, the S_1 and S_2 excitation energies show little variations. The number of states to be considered determines the active space in the POCAS selection. It remains to be clarified why the XMCQDPT2 correction stabilizes the S_n state so drastically compared to the S_1 and S_2 $\pi\pi^*$ states; the S_n state computed with the active spaces of (6,4)4 and (8,5)5 lies even below the S_1 state. The energies presented in Figure 2.2 suggest that the XMCQDPT2 stabilization is higher, the larger the respective S_n state CASSCF excitation energy is. In the documentation of Firefly (Granovsky, 2013) a list of criteria is given that judge whether the states and the active space provide a good zero-order description in the XMCQDPT2 calculation. For all POCAS calculations, the analysis of the XMCQDPT2 results revealed that indeed (i) the off-diagonal elements of the second-order effective Hamiltonian matrix are small and do not exceed 0.008 hartree; (ii) the eigenvalues of the non-symmetric effective Hamiltonian are all real and close to the eigenvalues of the symmetric effective Hamiltonian; and (iii) the

overlap norm matrix of non-orthogonal eigenvectors contains off-diagonal values of less than 0.05 hartree. Therefore, it is concluded that the XMCQDPT2-CASSCF calculations with the proposed principal-orbital active space provide a good description of the states of interest.

2.3.2 POCAS-optimized geometry of lumiflavin

To tackle photochemical problems with the POCAS way of constructing the active space, benchmarking of the optimized geometries is also relevant. Here, the optimized ground-state and S_1 geometries of lumiflavin are compared. Two comparisons are of interest: the change of the geometry in going from the ground-state to the excited-state minimum and the comparison of the equilibrium bond distances (of the geometry change) obtained with different methods. Figure 2.4 summarizes the results. For flavin photochemistry, the bond distances d_9 and d_{10} are of special interest (see Figure 2.4A for the bond-distance definition). Five different active spaces are compared: (2,2)2- S_1 , (6,4)4 from the previous section, an “intermediate” active space (6,6)2, a “close-to-full π - π^* ” active space (12,12)2, and (14,13) (the latter results are taken from (Climent et al., 2006)). There is no experimental data on the LF structure. The graph in Figure 2.4B uses the structural data from (Wang & Fritchie Jnr, 1973), in which the LF bond distances were estimated based on several X-ray crystal structures of LF-related compounds, like 10-methylisoalloxazine because LF itself could not be crystallized. These estimated LF bond lengths are referred to as experimental data in the following.

Overall, geometry optimization with the listed methods yields similar ground state structures. The CC2, B3LYP, MP2, and XMCQDPT2 methods are close in their bond length predictions; overall these methods give longer bond lengths compared to the experimental ones. Interestingly, there are more variations among the CASSCF-optimized geometries: (i) independent of the active space, the double bonds in ring II and III are significantly shorter than with the other methods or in the experimental structure; (ii) the geometries optimized with CASSCF(2,2)2 and CASSCF(6,4)4 are very close to those optimized with RHF and B3LYP; (iii) notably, the CASSCF(12,12)2- and CASSCF(14,13)-structures deviate more from the reference experimental data than the POCAS geometries; and (iv) the largest deviations are obtained with the intermediate active space (6,6)2.

The geometry changes between the optimized S_1 and ground state minima reveal how the excited flavin relaxes on the excited-state potential-energy surface. All methods yield a longer N5-C4a (d_{10}) and a shorter C4a-C10a (d_9) bond distance in the S_1 minimum compared to the CS ground state minimum, which are the most pronounced geometry changes upon flavin S_1 relaxation and in accord with the electronic structure of the S_1 HOMO-LUMO transition. These changes were also reported in the literature (Salzmann et al., 2008; Climent et al., 2006; Wolf et al., 2008). Overall, CC2 essentially predicts the same geometry changes as B3LYP. The largest

changes in d_9 and d_{10} are predicted by CASSCF(12,12)2 and CASSCF(14,13) and the smallest changes by CASSCF(2,2)2 and CASSCF(6,4)4. The CASSCF results are consistent in ring III but show different, even opposing, geometry changes in ring I when the POCAS and large active-space results are compared. Importantly, the geometry changes computed with CASSCF(2,2)2 and CASSCF(6,4)4 are practically indistinguishable. This property of POCAS-selected active spaces will be exploited during the pathway calculations in Chapter 5.

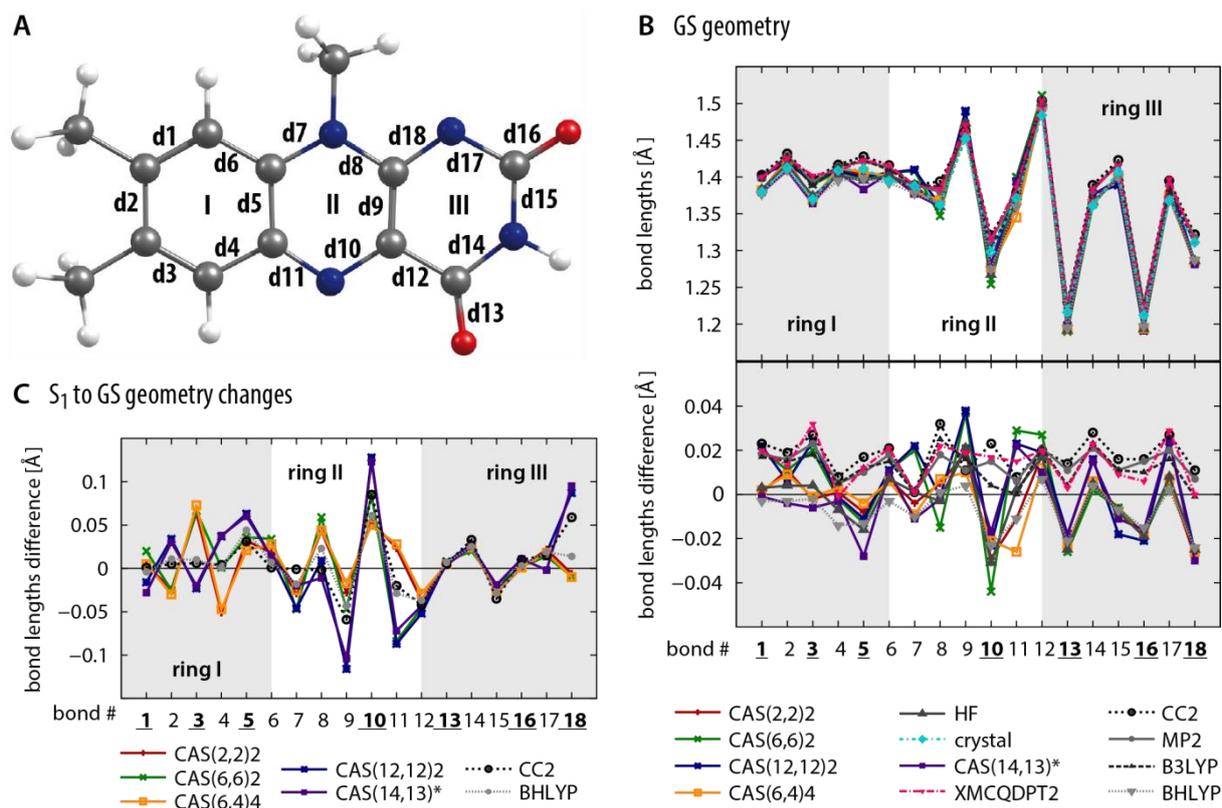


Figure 2.4. Equilibrium geometries of lumiflavin. (A) Labelling of rings and bond lengths in LF. (B) Bond lengths of LF optimized with different methods. The experimental data on the LF bond lengths are taken from (Wang & Fritchie Jr, 1973) (top). Relative distances, i.e. bond lengths differences between the optimized geometries and the experimental reference structure (bottom). (C) Bond-lengths differences between the S_1 and GS minimum-energy geometries. The double bonds are highlighted on the plots by bold and underlined numbers. The geometry data from the CASSCF optimization with the (14,13) active space are taken from (Climent et al., 2006).

2.3.3 Excitation energies in the BLUF cluster

In BLUF the excited states of interest are the light-absorbing S_1 and S_2 states of flavin and the optically dark electron-transfer state that corresponds to a $\pi\pi^*$ electronic transition from the tyrosine HOMO to the flavin LUMO. The CIS method was used to compute the excitation

spectrum at the B3LYP-optimized geometry to identify the low-lying states in the BLUF cluster model, shown in Figure 2.5. The ET state is the eighth excited state in the CIS spectrum. Below it in energy lie three flavin $\pi\pi^*$, two flavin $n\pi^*$, and two tyrosine $\pi\pi^*$ excited states. As expected, the CIS method substantially overestimates the excitation energies, however, reveals the principal orbitals to be included in the active space to compute the excited states of interest. For each state, dominated by a single-electronic excitation, a pair of MOs (occupied and virtual in the reference RHF electronic configuration) is included in the active space. As several states represent an excitation to the flavin LUMO, the POCAS active spaces contain less virtual than occupied MOs. The largest POCAS active space considered here is (8,6)7, describing the flavin S_1 , S_n , S_2 , the tyrosine S_1 , and the tyrosine-flavin ET state (the sixth excited state has major contributions from double excitations).

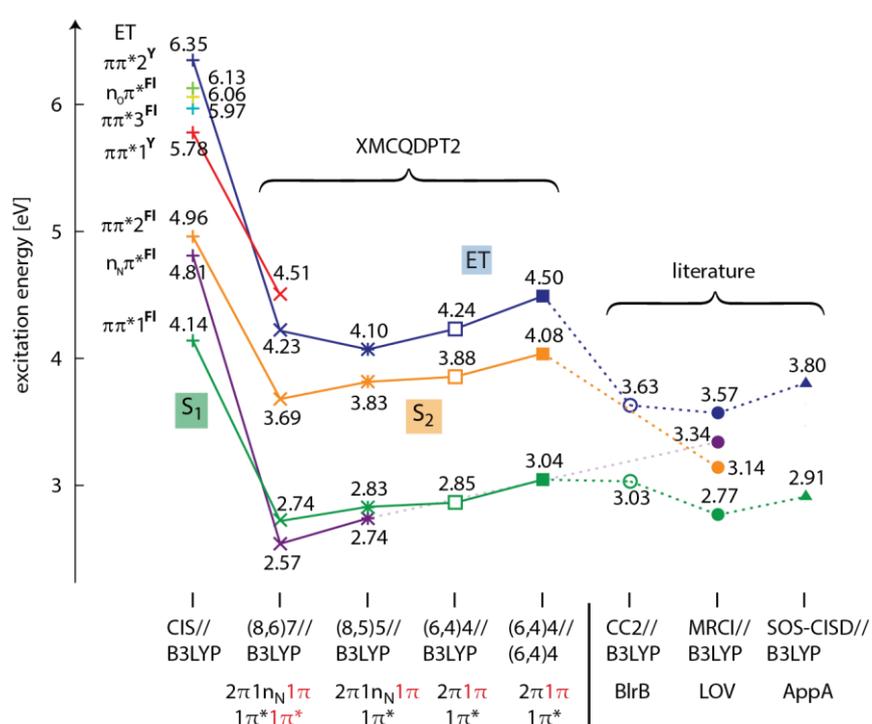


Figure 2.5. Electronic excitation energies in various BLUF models. The composition of the active space is indicated explicitly, black and red letters correspond to flavin and tyrosine MOs, respectively. The main (6,4) active space is given twice, first at the DFT-optimized geometry and second at the CASSCF(6,4)4-optimized geometry. The CC2 values are taken from (Sadeghian et al., 2008), the MRCI values from (Salzmann et al., 2009), and the SOS-CISD values from (Khrenova et al., 2010). Figure was modified from (Udvarhelyi & Domratcheva, 2011).

As Figure 2.5 demonstrates, at the XMCQDPT2 level the states reorder with respect to the CIS result: as in the case of the LF molecule, the S_n state lies below S_1 in energy and the ET state is lower than the tyrosine locally-excited S_1 state. Excluding the doubly excited and tyrosine S_1 state leads to the (8,5)5 active space, with similar excitation energies of the respective states as

with (8,6)7. Considering only the flavin S_1 , S_2 , and the ET state yields the (6,4)4 active space, used in Chapter 5 in the computations of the photoinduced electron-transfer pathway. The (6,4)4 active-space MOs and the electronic configurations corresponding to the four states of interest are depicted in Figure 2.6. CASSCF(6,4)4 was also used to optimize the ground-state geometry. At the CASSCF(6,4)4-optimized geometry the excitation energies are consistently blue-shifted with respect to the CASSCF spectrum calculated at the B3LYP-optimized geometry.

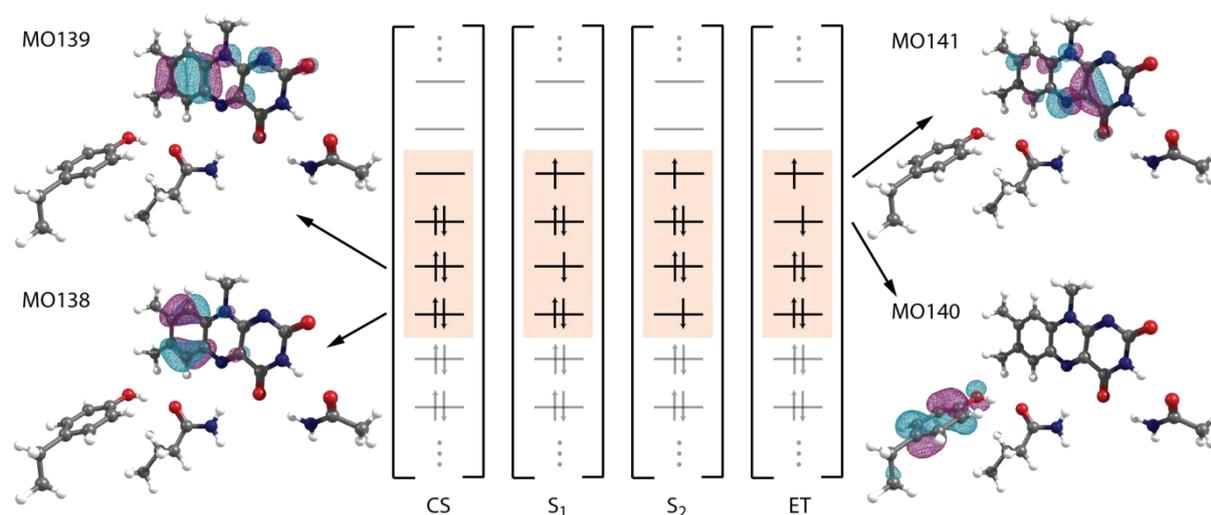


Figure 2.6. POCAS electronic configurations and frontier MOs of the (6,4)4 active space in the BLUF cluster model. The active space is highlighted in pale orange.

A systematic comparison of the XMCQDPT2-CASSCF excitation energies computed with various active spaces constructed in the bottom-up POCAS way and the top-down conventional way is given in Figure 2.7. Here, the top-down conventional way can only be used with “intermediate-sized” active spaces because including the whole π - π^* valence system on both the flavin and tyrosine is computationally too demanding. The BLUF results are consistent with the LF results in the previous sections, therefore the same conclusions can be drawn. The most important points are: The XMCQDPT2 correction lowers the CASSCF excitation energies considerably and makes them less dependent on the active space. The S_n excitation energy varies drastically among the active spaces used. The results computed with active spaces constructed with the conventional way depend significantly on the number of states considered. The ET excitation energy is above 4 eV in all cases.

Figure 2.5 also compares the POCAS results to excitation energies reported for flavoproteins in the literature: BlrB-BLUF (Sadeghian et al., 2008), AppA-BLUF (Khrenova et al., 2010), and a LOV-domain protein (Salzmann et al., 2009), where the ET state involves a cysteine residue instead of tyrosine. Compared to these literature results, the POCAS results for the S_1 state are in good agreement, whereas the excitation energies of the ET state are overestimated. It should be

2.4 Conclusions

This chapter formulated the POCAS approach, a practical way of choosing the active MOs for CASSCF, especially in the case of large molecules and many electronic states of interest in photochemical-pathway calculations. In the POCAS approach the principal MOs are defined by the electronic states of interest and enable the description of static correlation. Excitation-energy calculations on the isolated lumiflavin molecule showed that the POCAS approach in conjunction with single-point XMCQDPT2 or MS-CASPT2 calculations provides good estimates for the S_1 and S_2 $\pi\pi^*$ excited states. Likewise, the POCAS-optimized ground- and excited-state geometries of lumiflavin were found to be consistent with those optimized with other electronic-structure methods. Importantly, the major geometry changes between the ground state FC and the S_1 minima were fully reproduced. Notably, two active spaces chosen according to the POCAS approach (considering different number of states) yielded practically identical optimized geometries in the chosen state. Moreover, the PT2//POCAS excitation energies show very minor variations with respect to the number of electronic states included in the computation. These results raise confidence that the PT2//CASSCF method with the POCAS approach to select active MOs is well-suited to treat photochemical problems of the BLUF flavoprotein that was used in the extensive photoreaction calculations in Chapter 5.

3 Cluster models vs. QM/MM models

In this chapter, I compare the excitation energies computed in a QM cluster and a hybrid QM/MM model to elucidate the origin of the excitation energy shifts caused by intermolecular interactions. The flavin S_1 , S_2 , and the tyrosine-flavin electron transfer (ET) excited states are computed with the PT2//CASSCF method. To determine how intermolecular interactions and their description (QM or QM/MM) affect the excitation energies, I introduced variations in the QM cluster model composition and also modified the QM and MM subsystems of the QM/MM model. The excitation energies computed with the alternative models significantly differ in magnitude, and their shifts also originate from different physical effects that was established by a detailed energy-decomposition analysis. The results show significant changes of the ET excitation energy depending on the donor-acceptor distance, electrostatic interactions as well as wave-function polarization and QM charge-transfer interactions. I analyzed the contributions of specific residue interactions to compare the predictions of the cluster and QM/MM models and to identify the most suitable model for the photochemical pathway calculations, presented in Chapter 5.

3.1 Cluster and QM/MM models in photoreceptor studies

As introduced in Chapter 1, two main approaches to build models of photoreceptors are conceivable: the supermolecular cluster and the QM/MM approach. To date, representatives of all major photoreceptor-protein classes have been studied to some extent with the QM/MM approach (Sekharan & Morokuma, 2011; Sekharan et al., 2012, 2013; Pal et al., 2013; Gozem et al., 2012; Melaccio et al., 2012; Valsson et al., 2013; Strambi et al., 2010; Polli et al., 2010; Rajamani et al., 2011; Salzmann et al., 2009; Sadeghian et al., 2008; Solov'yov et al., 2012; Isborn et al., 2012; Frähmcke et al., 2010; Groenhof et al., 2008; Falklöf & Durbeej, 2013). In the computations of various rhodopsin photoreceptors, the leading role of the electrostatic

environment in color tuning (Melaccio et al., 2012; Valsson et al., 2013; Wanko, Hoffmann, Frauenheim, et al., 2008) and photosensitivity (Gozem et al., 2012) of the retinal chromophore has been demonstrated. Recently, polarization effects were ascribed to an important role in rhodopsin (Valsson et al., 2013; Wanko, Hoffmann, Frauenheim, et al., 2008; Wanko, Hoffmann, Frähmcke, et al., 2008), but also in BLUF (Rieff et al., 2011). However, the general role of polarization effects in the photoreceptor properties and their treatment (extending the QM subsystem or using polarizable MM force field) has not been established yet. Most of these studies model the chromophore in the QM subsystem and the apoprotein in the MM subsystem. For example, including the conserved side chains interacting with the bilin chromophore in the QM subsystem of a bacteriophytochrome QM/MM model, resulted only in a marginal shifts of the computed excitation energies and was considered to be not worthwhile in the view of the significantly increased computational cost (Falklöf & Durbeej, 2013).

In practice, with a given high-level QM method, a cluster model of a larger size than the size of the QM subsystem in QM/MM calculations can be computed. Therefore, with the cluster model the interactions of the chromophore with its local environment can be treated at the QM level. The cluster-model approach has especially been productive in the studies of metal-containing proteins, where it was demonstrated that including the nearest neighbor residues of the active site already converges the activation energies (Siegbahn & Himo, 2011). These studies also emphasized the decisive role of the QM description in getting reliable energies (Siegbahn & Himo, 2011). In the case of photoreceptor proteins, studies employing cluster models were carried out for example in (Gromov et al., 2007; Domratcheva et al., 2008; Khrenova et al., 2010, 2011; Solov'yov et al., 2012). Martínez and coworkers pioneer excited-state calculations with cluster models as large as the entire protein (Isborn et al., 2011). Thus, excited-state calculations in cluster models with increasing size become feasible. It is expected that from the comparison of computed excited-state properties with cluster models to those obtained with QM/MM models further insights into the nature of intermolecular interactions underlying light reception will emerge.

A comparative analysis between cluster and QM/MM models is especially interesting in the case of the BLUF photoreceptor, because its photosensory function depends not only on the excited states of flavin, but also on the tyrosine-flavin ET state. This chapter analyzes how the local and more distant protein environment affects the BLUF excitation spectrum and how reliable different approaches are in modeling the relevant intermolecular interactions. To this end, a fairly large QM cluster and a QM/MM model are used to represent the environment of the photoactive triad in BLUF and to compute the excitation spectrum with the PT2//CASSCF method. The two models were optimized with two different programs: The QM subsystem was embedded into the MM environment by means of the “electrostatic potential fitting” (ESPF) operator (Ferre & Angyan, 2002) with the QM/MM coupling scheme between the MOLCAS quantum-chemistry suite (Aquilante et al., 2010) and the Tinker MM program (Ponder &

Richards, 1987); whereas the QM cluster model calculations were carried out with the program Firefly (ver 7.1, 8; Granovsky). The efficient computational code of Firefly allows treating a much larger cluster model than MOLCAS can treat in the QM subsystem of the QM/MM model, however, Firefly unfortunately does not have QM/MM CASSCF energy gradients implemented. Recently, Firefly was used to compute photochemical problems with cluster models of DNA photolyase and a cryptochrome photoreceptor, for example, in (Domratcheva, 2011; Solov'yov et al., 2012; Moughal Shahi & Domratcheva, 2013; Solov'yov et al., 2014). MOLCAS/Tinker is widely employed for QM/MM studies of photoreceptor proteins, recent examples are (Gozem et al., 2012; Melaccio et al., 2012; Bernini et al., 2013; Solov'yov et al., 2012; Rinaldi et al., 2014).

The excited states treated here are the flavin S_1 , S_2 and the tyrosine-flavin ET states. The PT2//CASSCF method is used and the active space is selected according to the POCAS computational scheme explained and benchmarked in Chapter 2. The two computer codes have different PT2 methods implemented: XMCQDPT2 in Firefly and (MS-)CASPT2 in MOLCAS. This chapter also compares the results of the two PT2 methods. The excitation spectrum is computed at a series of models: the composition of the cluster model and of the QM subsystem is varied, whereas the geometry parameters are conserved. Two equilibrium geometries were considered, one obtained with the large cluster model and the other with the QM/MM model. In addition, the effect of manipulating the MM subsystem in the QM/MM model is evaluated. The origin of the computed excitation energy shifts is analyzed, with special attention paid to the ET excited state because of its high relevance to the BLUF photoreaction.

3.2 Models and analysis

3.2.1 Cluster model and QM/MM models

Based on the PDB coordinates of the PixD BLUF protein (Yuan et al., 2006), a QM-cluster model was built, consisting of the riboflavin (RF) molecule and the side chains Q53, Y11, N34, N35, M96, S31, and L44, in the following referred to as RFQY-NNMSL or full cluster model (in total 135 atoms). To mimic the geometric constraints of the protein on these selected residues, specific atoms were fixed during CASSCF geometry optimization (computational details are presented in Section 3.3.1). On the basis of the final optimized geometry of RFQY-NNMSL, a set of models was prepared by successively excluding the molecular fragments of the side chains (in the following referred to as cluster-model series, RF-series). Analogously, a model series with lumiflavin (LF) as the chromophore was also considered (LF-series). All these models are listed in Table 3.1. With the cluster model series the intermolecular interactions between the photoactive triad and its local environment that includes the ribityl chain, and the side chains S31,

N34, N35, L44 and M96 (termed the QM environment α) was analyzed, as schematically shown in Figure 3.1A. For brevity, the residue numbering in all plots and model designations is omitted; in the cases where only one asparagine is indicated, N35 is included in the model and in the cases where two asparagines are included, they are N34 and N35. Note that every fragment in all models of the series has the same coordinates because the geometries of the reduced models were not reoptimized. For each prepared model the excitation energy spectrum was computed with the XMCQDPT2//CASSCF method.

The QM/MM model was built based on the same BLUF protein, PixD, which contains the LF-Q53-Y11 triad (LFQY) in the QM subsystem and the rest of the protein with the solvent in the MM subsystem. After QM/MM geometry optimization with CASSCF/AMBER (details are presented in Section 3.3.2), several variants of this so-called reference QM/MM model were prepared, to analyze the role of intermolecular interactions in spectral tuning. To this end, the QM and MM subsystems were repartitioned such that the QM subsystem includes, in addition to the photoactive triad, one or two residues of the set $\alpha = \{S31, N35, M96\}$ (here, as in the case of the cluster models, the variable α is used to specify the composition of the QM environment of the triad included in the QM subsystem), as indicated in Figure 3.1A.

Table 3.1. List of models prepared based on the optimized cluster and optimized QM/MM model, in the “cluster series” (left) and the “QM/MM series” (right). The number of atoms and basis functions is also given.

cluster model	# of atoms	# of cc-pVDZ basis functions	QM subsystem of QM/MM model	# of atoms	# of DZV(P) basis functions
LF/RF	31/43	345/465	LF	31	309
LF/RFQY (triad)	59/71	615/735	LFQY (triad)	56	515
LF/RFQY-S	70/82	720/840	LFQY-S	62	553
LF/RFQY-M	71/83	729/849	LFQY-M	65	580
LF/RFQY-N	68/80	700/820	LFQY-N	65	585
LF/RFQY-NN	86/98	870/990	LFQY-SM	71	618
LF/RFQY-SN	79/91	805/925	LFQY-SN	71	623
LF/RFQY-MN	80/92	814/934	LFQY-MN	74	650
LF/RFQY-SM	82/94	834/954			
LF/RFQY-SNN	97/109	975/1095			
LF/RFQY-MNN	98/110	984/1104			
LFQY-SMNN	109	1089			
RFQY-SMNNL	135	1319			

A QM/MM model containing LF only in the QM subsystem and the rest of the system in the MM subsystem was also considered. These QM/MM models are referred to as QM/MM-model series in which the index α indicates the composition of the QM subsystem. Side chains $\alpha = \{S31, N35, M96\}$ were considered at their MM-optimized geometry. In addition to the

QM/MM repartitioning, the reference QM/MM model was modified by zeroing the MM charges of S31, N35, and M96 (the same side chains as in the QM environment) to eliminate the electrostatic QM/MM interactions of these residues with the photoactive triad. The effect of changing the MM charges in the vicinity of the electron donor and acceptor by +1 and -1 a.u. in the reference QM/MM model was also tested; for these computations the residues Y66, S30, H75 and the phosphate group of the FMN chromophore were selected. The atomic coordinates of the geometry-optimized reference QM/MM model were used for every modified model without reoptimization.

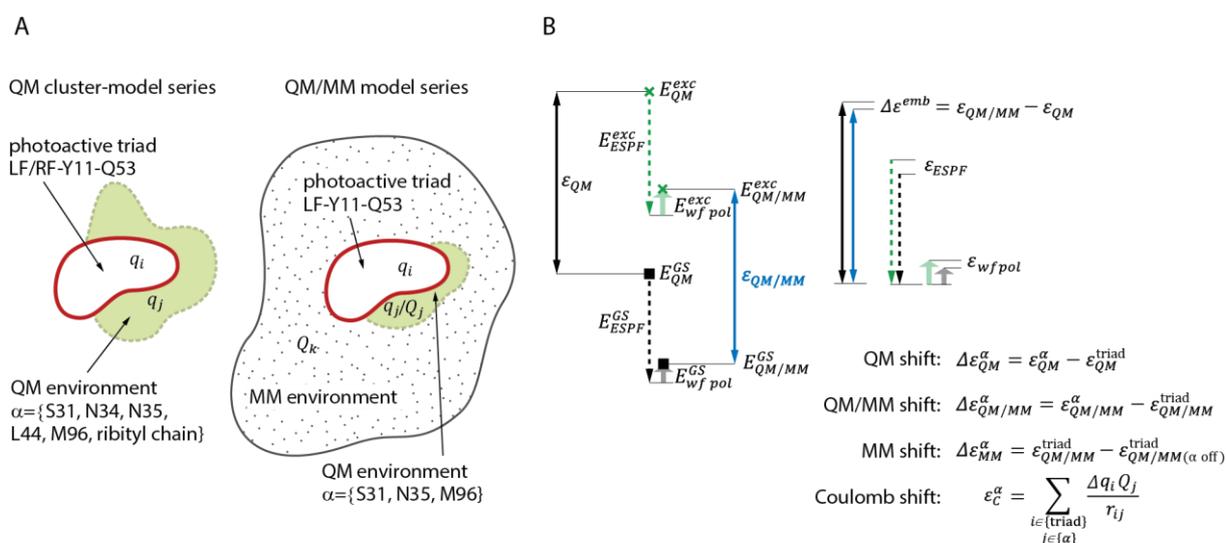


Figure 3.1. BLUF models and the analysis of excitation energies. (A) QM-cluster and QM/MM-model series; (B) Comparison of the intermolecular-interaction energies in the ground and excited states to reveal the role of intermolecular interactions in the excitation-energy shifts. The energy designations and formulas used to evaluate the excitation-energy shifts due to the specific method (QM or QM/MM) and environmental (interacting side chains) effects are explained in the text.

For each QM/MM model the excitation energy spectrum with the (MS)-CASPT2//CASSCF method was computed, as detailed in Section 3.3.2. Additionally, for the QM/MM series with the modified QM subsystems, the excitation spectrum without the QM/MM embedding was also computed, with the (MS)-CASPT2 and XMCQDPT2 methods. From these latter calculations without the MM environment, one can infer both on the effect of the method (i.e. by comparing the XMCQDPT2 results with CASPT2 for the same QM subsystem) and on the effect of the geometry (i.e. by comparing the XMCQDPT2 excitation energies of the QM subsystem without the MM environment with those of the cluster model of the same composition).

3.2.2 Analysis of intermolecular interactions

The environment-induced excitation-energy shift depends on the redistribution of the electronic density upon the transition. To visualize the redistribution, the difference (excited state minus ground state) electrostatic-potential maps of all transitions of interest were computed. Concomitant with the charge redistribution of the electronic excitation, there is a charge transfer between the photoactive triad and its local QM environment. This charge transfer is visualized by the electrostatic-potential maps. As expected, the charge transfer is largest for the transition from the CS ground state to the ET state. To quantify the effects, the differences of the net charges (ET state minus ground state) on the electron acceptor flavin, electron donor tyrosine and on the QM environment of the photoactive triad were computed. To illustrate the effect of the wavefunction polarization by the MM charges, double-difference electrostatic potential maps were computed by subtracting the electronic transition maps of the QM/MM models from its QM subsystem without the MM environment.

In the ESPF QM/MM coupling method (Ferre & Angyan, 2002), the interaction energy between the QM and MM subsystems consists of two components: (i) a classical electrostatic interaction energy term between the electron-density fitted q_i ESPF charges of the QM subsystem atoms and the Q_k MM charges; and (ii) the energy change due to the polarized wave function, $E_{wf\ pol}$, as compared to the QM solution without the MM environment. The ESPF energy is computed as a Coulomb sum

$$E_{ESPF} = \sum_{\substack{i \in \{\text{triad}\} \\ k \in \{\text{MM}\}}} \frac{q_i Q_k}{r_{ik}} \quad (3.1)$$

where r_{ik} is the distance between atom i of the QM subsystem and atom k of the MM subsystem. To determine the polarization energy, two energy calculations for each QM/MM model were carried out: one with and one without the MM environment, to obtain $E_{QM/MM}$ and E_{QM} , respectively. From the QM/MM calculations E_{ESPF} and the energy of the MM subsystem, E_{MM} , are known, so the energy contribution due to the polarized wave function can be computed as

$$E_{wf\ pol} = E_{QM/MM} - E_{QM} - E_{MM} - E_{ESPF} \quad (3.2)$$

The QM/MM excitation energies are given by

$$\begin{aligned} \varepsilon_{QM/MM} &= E_{QM/MM}^{\text{exc}} - E_{QM/MM}^{\text{GS}} \\ &= \varepsilon_{QM} + \varepsilon_{ESPF} + \varepsilon_{wf\ pol} \end{aligned} \quad (3.3)$$

where ε_{QM} , ε_{ESPF} and $\varepsilon_{wf\ pol}$ are the differences of E_{QM} , E_{ESPF} and $E_{wf\ pol}$ in the ground and excited states, respectively. The excitation-energy shift obtained in the QM/MM calculation with respect to the QM calculation, $\Delta\varepsilon^{emb}$, also has two contributions: one due to the QM/MM electrostatic interactions and the other due to the QM-subsystem polarization by the MM environment:

$$\Delta\varepsilon^{emb} = \varepsilon_{QM/MM} - \varepsilon_{QM} = \varepsilon_{ESPF} + \varepsilon_{wf\ pol} \quad (3.4)$$

The excitation-energy shifts upon modification of the QM and MM subsystems are defined as follows. The QM excitation-energy shift obtained upon including the α side chain(s) (α was specified in Section 3.2.1) in the QM environment of the photoactive triad is given by

$$\Delta\varepsilon_{QM}^{\alpha} = \varepsilon_{QM}^{\alpha} - \varepsilon_{QM}^{triad} \quad (3.5)$$

The QM/MM excitation-energy shift caused by zeroing the MM electrostatic charges of the α residue is given by

$$\Delta\varepsilon_{MM}^{\alpha} = \varepsilon_{QM/MM}^{triad} - \varepsilon_{QM/MM(\alpha\ off)}^{triad} = \Delta\varepsilon_{ESPF}^{\alpha\ off} + \Delta\varepsilon_{wf\ pol}^{\alpha\ off} \quad (3.6)$$

The QM/MM excitation energy shift upon moving side chain(s) α from the MM to the QM subsystem is defined as

$$\Delta\varepsilon_{QM/MM}^{\alpha} = \varepsilon_{QM/MM}^{\alpha} - \varepsilon_{QM/MM}^{triad} = \Delta\varepsilon_{QM}^{\alpha} + \Delta\varepsilon_{ESPF}^{\alpha} + \Delta\varepsilon_{wf\ pol}^{\alpha} \quad (3.7)$$

The electrostatic interaction energy contribution C^{α} to the $\Delta\varepsilon_{QM}^{\alpha}$ shift was estimated by computing the respective Coulomb sum between the q_i Mulliken charges of the photoactive triad and the q_j Mulliken charges of the QM environment α as

$$C^{\alpha} = \sum_{\substack{i \in \{\text{triad}\} \\ j \in \{\alpha\}}} \frac{q_i q_j}{r_{ij}} \quad (3.8)$$

The electrostatic interaction energy contribution to the excitation energy shift is

$$\varepsilon_C^{\alpha} = \sum_{\substack{i \in \{\text{triad}\} \\ j \in \{\alpha\}}} \frac{q_i^{\text{exc}} q_j^{\text{exc}}}{r_{ij}} - \sum_{\substack{i \in \{\text{triad}\} \\ j \in \{\alpha\}}} \frac{q_i^{\text{GS}} q_j^{\text{GS}}}{r_{ij}} \quad (3.9)$$

By definition, the model which consists of the photoactive triad only has a Coulomb sum of zero (Equation 3.8); thus $\Delta\varepsilon_C^{\alpha} = \varepsilon_C^{\alpha} - \varepsilon_C^{\text{triad}} = \varepsilon_C^{\alpha}$, and it is referred to as the Coulomb shift,

analogously to the excitation-energy shifts defined above. So the QM shift defined above (Equation 3.5) can be expressed by the electrostatic term ε_C^α and a non-electrostatic term

$$\Delta\varepsilon_{QM}^\alpha = \varepsilon_C^\alpha + \Delta\varepsilon_{non-c}^\alpha \quad (3.10)$$

The Coulomb shift (Equation 3.9) was computed for the QM/MM models and compared with that of the respective QM subsystem computed without the MM environment. For the QM/MM model, either the ESPF charges or the Mulliken charges were used. A graphical presentation of the energy decomposition and a summary of the excitation-energy shifts are given in Figure 3.1B.

3.3 Computational details

3.3.1 Cluster models

The X-ray structure of the PixD BLUF protein (Yuan et al., 2006) (PDB code 2HFN, chain A) was used to generate the QM cluster model introduced above. The geometry was optimized in the ground state with the CASSCF(6,4)4 method, with the coordinates of the selected atoms (the hydroxyl-oxygen atom of S31 and the black-colored carbon atoms indicated in Figure 3.2A) fixed to the values in the PDB model. The cluster model series was prepared as explained above. The LF geometry was obtained from the RF geometry by removing the ribityl fragment and placing a hydrogen atom along the C-C bond at a distance of 1.091 Å. The XMCQDPT2//CASSCF(6,4)4 excitation energies were computed with an edshift value of 0.02 au. The cc-pVDZ basis set was used in the calculations of the QM cluster model series. In the calculations of the isolated LF or RF molecules, only the three MOs localized on flavin were included in the active space and the three states were computed with the CASSCF(4,3)3 method. These calculations give consistent excitation energies with the CASSCF(6,4)4 calculations: in the LFQY cluster model for example, the S_1 and S_2 excitation energies computed with XMCQDPT2//CASSCF(4,3)3 agree within 0.03 eV to those computed with XMCQDPT2//CASSCF(6,4)4.

The electrostatic-potential maps and atomic Mulliken charges were computed from the zero-order QDPT density, generated for each electronic state. The net charge was computed for the isoalloxazine ring of the flavin and for the whole tyrosine fragment in the ground state and ET state. The net charge on the QM environment was computed by excluding all the atoms of the photoactive triad (RFQY in the RF series and LFQY in the LF series) and summing up the charges on the remaining atoms. The electrostatic-potential maps for the full cluster model, computed on a fine grid of approx. $100 \times 100 \times 100$ points, were visualized with Chemcraft. The counterpoise correction (Boys & Bernardi, 1970) using ghost MOs was used to evaluate the basis set superposition error (BSSE). All calculations of the cluster model series were performed with

the quantum chemistry package Firefly (ver. 7.1, 8; Granovsky), which is partially based on the GAMESS US source code (Schmidt et al., 1993).

3.3.2 QM/MM calculations

Based on the X-ray structure of the PixD BLUF protein (Yuan et al., 2006) (PDB code 2HFN, chain A), the initial coordinates of the QM/MM model were prepared with the help of the GROMACS 4.5 package (Prnk et al., 2013). The hydrogen atoms were added to the PDB model assuming standard protonation states of the residues (arginine and lysine residues protonated; aspartate and glutamate residues deprotonated, histidine residues neutral, H75 with Ne-H and the other histidines with N δ -H). In a $75.5 \times 75.5 \times 75.5$ nm³ box, 9055 water molecules and 7 Na⁺ ions solvated and neutralized the protein model. The entire model was subjected to 500 steps of steepest-descent energy minimization with the GROMOS53A6 force field (Oostenbrink et al., 2004) and the SPC water parameters (Berendsen et al., 1981). The energy-minimized model was reduced to contain the protein, the Na⁺ counterions and a water-shell of 15 Å around the protein (4519 water molecules) for the QM/MM calculations with the MOLCAS 7.6/7.8 package (Aquilante et al., 2010) interfaced to the Tinker software (Ponder & Richards, 1987).

The AMBER force field (Cornell et al., 1995) and the TIP3P water model (Jorgensen et al., 1983) were used to describe the MM subsystem. For the flavin mononucleotide (FMN) with the terminal hydrogen phosphate, charge -1 , the RESP charges were derived from a HF/6-31G* calculation using the antechamber package (Wang et al., 2006). The atom types and van-der-Waals parameters of FMN were assigned according to the AMBER atom types of amino acids and nucleobases. The hydrogen link atom (LA) scheme at the QM/MM boundary was used (Singh & Kollman, 1986). The frontier was placed between non-polar sp³ carbon atoms in the FMN cofactor and several amino acid residues. To avoid overpolarization of the QM subsystem, the charge of the boundary MM carbon atom was set to zero and equally redistributed among the three neighboring MM atoms. In addition, the small residual charges of the MM part of the partitioned side chains and FMN were also equally redistributed among the three neighboring MM atoms (charge-redistribution scheme A). The charge-redistribution scheme A was used for the geometry optimization and for the computations with the modified QM part and the zeroed MM charges. In the computations with adding a positive or negative charge to selected residues, only the charge of the boundary MM carbon atom was redistributed among the three neighboring atoms, but the residual non-zero MM charge of the partitioned FMN and side chains was kept (charge-redistribution scheme B). The results of the two charge-redistribution schemes for the reference model with LFQY in the QM subsystem were compared. The respective total QM/MM energies of the four electronic states differ in the second or third digit of an atomic

unit, whereas the largest difference in the excitation energy was found to be 0.11 eV for the ET state, with an ET excitation energy decrease computed with scheme B.

The geometry of the QM/MM model with the LFQY photoactive triad in the QM subsystem (the reference model) was optimized in the ground state with the CASSCF(6,4)4 method, with “microiterations on” (Melaccio et al., 2011) to relax the geometry of the MM subsystem within 4 Å of any QM atom whereas the remaining MM atoms were kept frozen. The excitation energies were computed with the CASPT2 and MS-CASPT2 methods (without applying the ipea shift and with an imaginary shift of 0.2 au). In all calculations with MOLCAS the DZV(P) (d-functions on the heavy atoms) basis set was used. The excitation energies of the QM subsystem without the MM environment were also computed with the (MS)-CASPT2/DZV(P) methods using the MOLCAS program and with the XMCQDPT2/cc-pVDZ method using the Firefly program.

To compute the electrostatic potential from the CASSCF electron densities of the QM subsystems with and without the MM environment, $40 \times 40 \times 40$ grid points were explicitly specified in the MOLCAS input file. The grid points were generated to be similar to the "coarse" grids of the Firefly program. From the MOLCAS output, standard cube files were generated with a perl script for the visualization of the potentials with Chemcraft. For the comparison of the net-charges the same definition of the electron donor and electron acceptor fragments and of the QM environment as in the cluster models were used. For the calculations of the Coulomb shifts, the Mulliken and ESPF charges computed with the CASPT2 method were used.

In Chapter 5, two more QM/MM models will be used in the calculations of the photoreaction pathways. First, a QM/MM model based on the presented model of PixD BLUF but with LF, Q53, Y11, and N35 in the QM subsystem, for which the geometry was reoptimized. This model is referred to as PixD-II QM/MM model in Chapter 5. Second, a QM/MM model based on the X-ray structure of BlrB (Jung et al., 2005) (PDB code 2BYC, chain A) was also prepared, analogously to the PixD model detailed above. The final BlrB QM/MM model contains the 136-amino-acid-long protein, the FMN chromophore with hydrogen phosphate with charge -1 , the photoactive triad (LFQY) in the QM subsystem, and the solvent shell consisting of 4010 water molecules and one Na^+ ion. In both the PixD II and the BlrB QM/MM models the charge redistribution scheme A was used.

3.4 Results

3.4.1 Geometry of the flavin-binding site

The geometry optimization of the QM cluster and QM/MM models yielded two significantly different equilibrium structures of the photoactive triad and its local environment. Whilst the intrinsic geometry parameters of flavin and tyrosine are very similar as the same CASSCF method was employed, the different constraints of the two models resulted in different intermolecular distances. The hydrogen-bonding network of flavin in the two optimized structures of the binding site is shown in Figure 3.2. The equilibrium distances of relevant flavin bond lengths are also indicated. During QM/MM geometry optimization, the MM environment prevents large structural relaxation of the QM subsystem and thus the optimized distances of the hydrogen-bonding network are very similar to those in the initial X-ray structure. In contrast, in the QM-cluster model the geometry relaxation is significant as the fixed terminal atoms provide relatively weak constraints. The equilibrium distances characterizing the inter-fragment interactions increase compared to their magnitudes in the initial crystal-structure geometry, and most of the hydrogen-bond distances are longer in the cluster model as compared to the X-ray structure and the QM/MM model. In the cluster model, the M96 and S31 side-chains move away from their crystallographic positions. As a consequence, in the cluster model the hydrogen bond between the backbone carbonyl of S31 and the backbone nitrogen of N35 is disrupted. The displacement of M96 in the cluster is favorable for the formation of a hydrogen bond between flavin's C4=O4 and the amide group of Q53. Hence the N(Q53)-O4(flavin) distance is shorter in the cluster model than in the QM/MM model. The constraints of the MM subsystem prevent the movement of the M96 side chain away from the photoactive triad, which interferes with the glutamine-flavin interactions. The role of the M96 side chain in modulating the flavin-glutamine interactions has been discussed previously in (Khrenova et al., 2011).

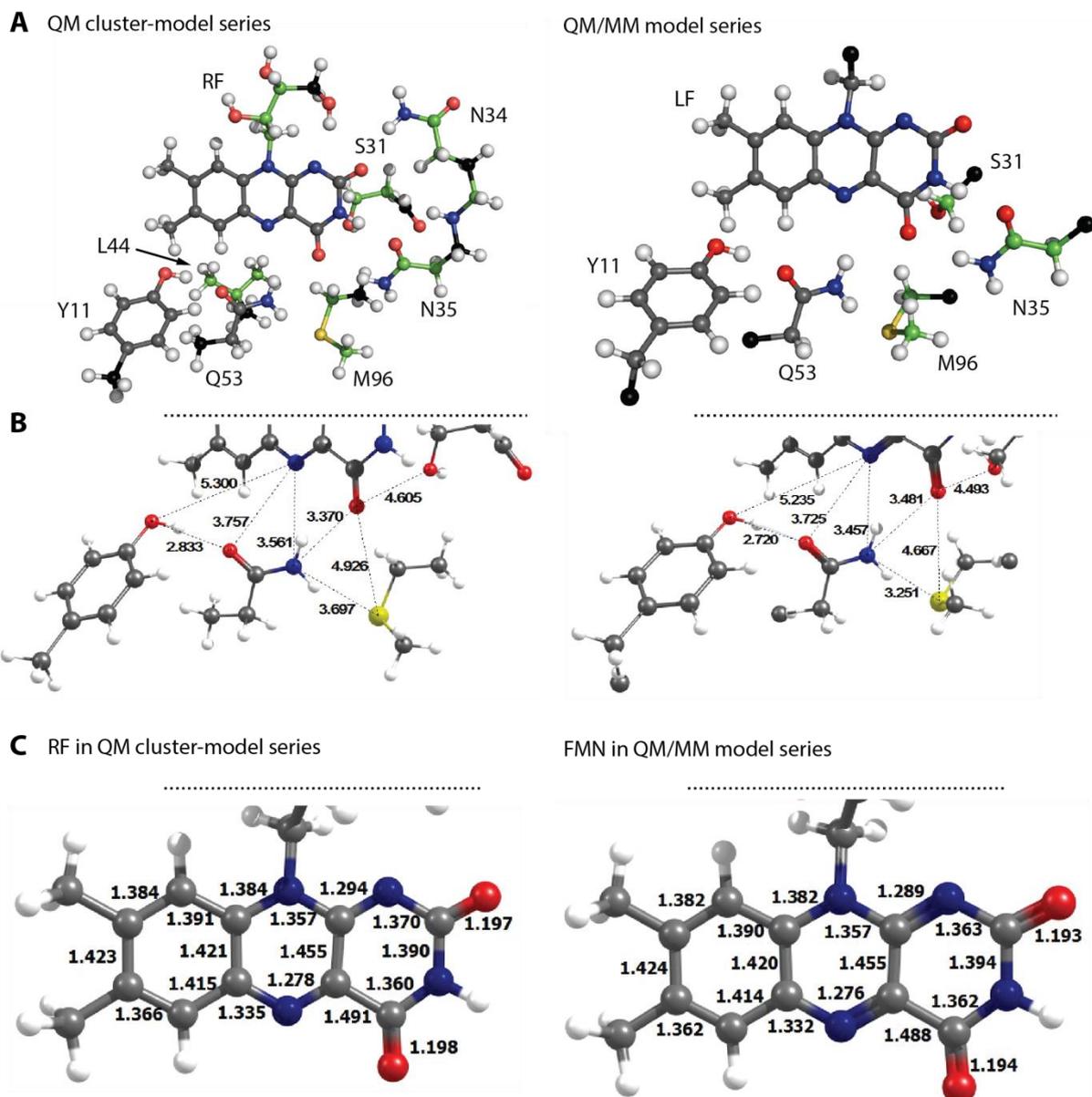


Figure 3.2. Optimized geometries of the BLUF photoactive triad and its local environment. (A) The QM-cluster model (left) and the QM subsystem of the QM/MM model (right). In the QM cluster series, the carbon atoms shown in black and the hydroxyl oxygen of S31 were fixed during geometry optimization. In the QM/MM model, the boundary MM carbon atoms are shown in black. The carbon atoms of the photoactive triad LFQY are shown in grey and those of its QM environment in green. (B) Selected equilibrium distances in Å in the QM-cluster (left) and QM/MM model (right). The N35 side chain is not shown for clarity. (C) Relevant bond lengths in Å of RF in the optimized cluster model (left) and of FMN in the optimized QM/MM model (right).

3.4.2 Excitation energies – QM calculations

Figure 3.3 compares the excitation-energy spectra computed with the CASSCF, XMCQDPT2 and CASPT2 methods at the two optimized geometries. On the abscissa of the plots, the models are arranged according to their decreasing CASSCF ET excitation energy, and this order is conserved throughout this chapter. The order does not correlate with the number of atoms comprising each model (given in Table 3.1); it is rather the specific intermolecular interactions of the included molecular fragments that determine the energy trend. The comparison of the energies at the two geometries shows that the ET excitation energy is 0.3 eV downshifted at the more compact QM/MM-optimized geometry, independent on the computational method. The lower ET excitation energy in the QM/MM models can be attributed to the decreased distances between the electron donor and acceptor as well as to the modifications of the hydrogen bond distances. At the same time, the S_1 and S_2 excitation energies stay unchanged, consistent with the fact that the equilibrium geometry of flavin is very similar in the cluster and QM/MM models.

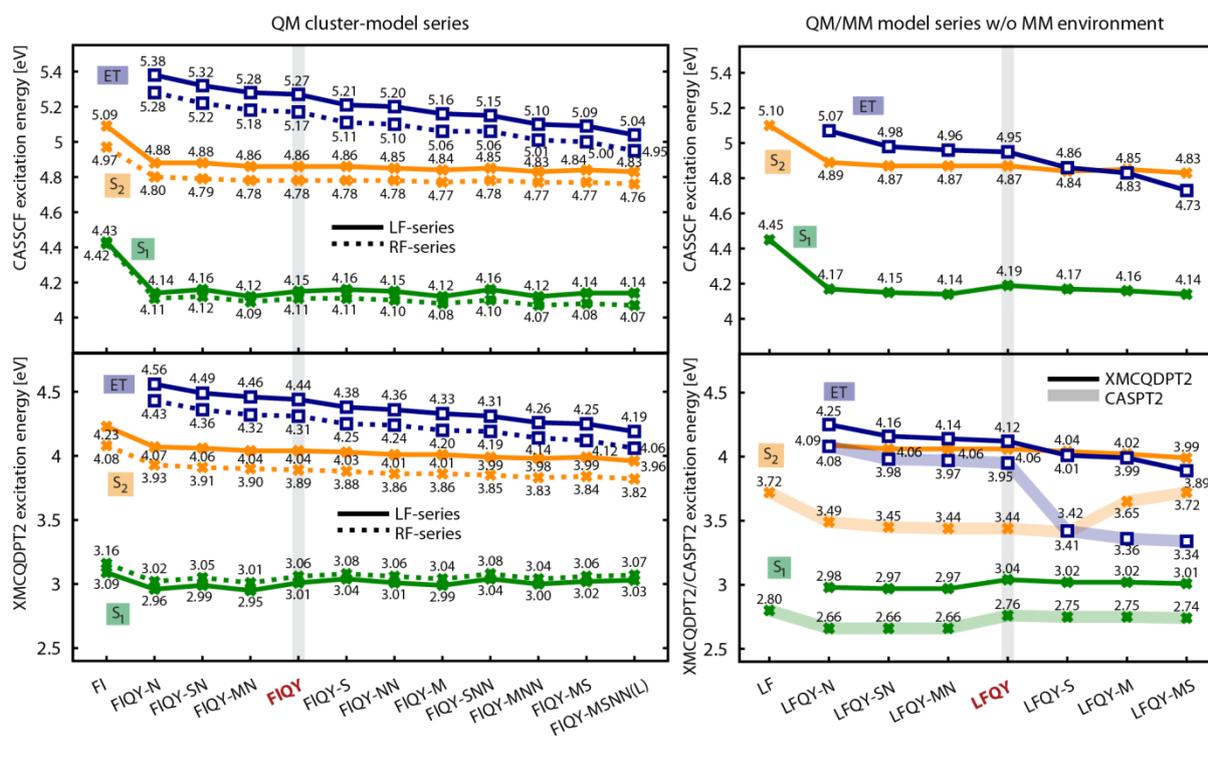


Figure 3.3. CASSCF and XMCQDPT2 excitation energies in the QM-cluster series (left); CASSCF, XMCQDPT2 and CASPT2 energies of the QM subsystem in the QM/MM model series without the MM environment (right). The corresponding MS-CASPT2 energies of the QM subsystems are presented in Figure 3.5. The residue numbering is specified in Section 3.2.1 and in Figure 3.1 but is omitted for brevity in all plots.

The differential dynamic electron correlation (the difference of the CASSCF and PT2 excitation energies) is larger for the S_1 state (1.0-1.27 eV with XMCQDPT2 and 1.40-1.65 eV

with CASPT2) than for the S_2 and ET states (0.84-0.89 eV with XMCQDPT2 and 0.94-1.43 eV with CASPT2). The XMCQDPT2 method gives significantly higher energies than the CASPT2 method, especially for the S_2 state (about 0.5 eV difference). The MS-CASPT2 estimates (shown below, in Figure 3.5) are closer to the XMCQDPT2 ones, but still 0.1-0.3 eV smaller. The XMCQDPT2 S_1 and S_2 energies are in good agreement with the experimental BLUF spectrum, as the estimated spectral maxima are expected to shift to the lower energies upon improving the equilibrium geometry with a method accounting for the dynamic correlation, by extending the basis set and also upon accounting for vibronic interactions. As documented in the literature, all these improvements result in a red-shift of the excitation-energy maximum in flavin (Klaumünzer et al., 2010).

The decrease of the ET energy is larger than of S_2 in the computed series, which leads to the crossing of these two states at the QM/MM-optimized geometry when the S31 and M96 side chains are included in the QM subsystem. Adding both S31 and M96 side chains results in the reordering of the S_2 and ET states. There is a significant decrease of the CASPT2 ET excitation energy upon addition of S31; upon further addition of M96 not only the ET energy significantly decreases but also the S_2 energy significantly increases. At the same time, the respective changes in the CASSCF and XMCQDPT2 energies are much smaller. Table 3.2 and Figure 3.4 present the charge distribution analysis, demonstrating the S_2 /ET state crossing. In the S_2 state, the electron donor Y11 and acceptor LF are neutral, while in the ET state, they become a cation and anion, respectively. In the LFQY-S and LFQY-M complexes, the S_2 and ET states are significantly mixed. At the cluster-model geometry, the energy of the ET state is systematically up-shifted compared to the QM/MM models without the MM environment, and the state crossing is absent.

Table 3.2. Mulliken charge differences between the 3rd and 4th electronic state and the ground state on the LF (full LF fragment) and the Y11 fragment, computed with CASPT2 and XMCQDPT2 at the QM/MM model series geometries (without the MM environment taken into account).

	CASPT2				XMCQDPT2			
	LF		Y11		LF		Y11	
	3rd	4th	3rd	4th	3rd	4th	3rd	4th
LFQY-N	-0.0034	-0.9890	0.0046	0.9943	-0.0045	-0.9956	0.0039	0.9952
LFQY-SN	-0.0050	-0.9840	0.0075	0.9920	-0.0107	-0.9871	0.0110	0.9881
LFQY-MN	-0.0064	-0.9849	0.0080	0.9908	-0.0154	-0.9841	0.0150	0.9841
LFQY	-0.0112	-0.9878	0.0114	0.9880	-0.0286	-0.9716	0.0279	0.9713
LFQY-S	-0.5104	-0.4818	0.5119	0.4845	-0.7098	-0.2883	0.7100	0.2891
LFQY-M	-0.8564	-0.1395	0.8570	0.1405	-0.8824	-0.1174	0.8817	0.1174
LFQY-MS	-0.9873	-0.0064	0.9911	0.0082	-0.9869	-0.0105	0.9880	0.0111

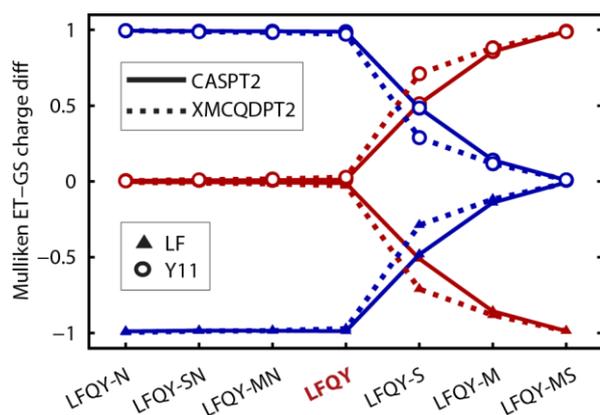


Figure 3.4. Net Mulliken charge differences on the electron donor Y11 and electron acceptor LF between the second (red line) and third (blue line) excited state and the ground state, demonstrating the crossing of the S_2 and ET states.

In the cluster-model series, there is a systematic difference in the excitation energies of the LF- and the RF-containing complexes, which are ascribed to the hydrogen-bonding interactions of the ribityl-chain OH-group with the N1-C2=O2 part of flavin. The energy difference is small (below 0.05 eV) in the case of the S_1 state and it changes the sign when the CASSCF and XMCQDPT2 results are compared. The S_2 energy computed with CASSCF and XMCQDPT2 is lowered by up to 0.08 and 0.15 eV, respectively, when RF is compared to LF. The respective lowering of the ET excitation energy is 0.1 and 0.15 eV. The addition of the side chains to the FlQY triad causes both blue and red shifts of the excitation energies, hence the FlQY model is situated in the middle of the excitation energy plots. The interactions with the side chain of N35, forming two hydrogen bonds with the flavin O4=C4-N3H moiety as a hydrogen donor and acceptor simultaneously, increases the S_2 and ET excitation energies. In contrast, a hydrogen-bond donation of the N34 side chain to the flavin C2=O2 carbonyl group stabilizes the S_2 and ET states. The M96 and S31 side chains do not form hydrogen bonds with flavin, but including them into the complex decreases the energies of the S_2 and ET states. The magnitudes of the respective CASSCF and XMCQDPT2 excitation-energy shifts are consistent with each other.

To aid the comparison of the models of varying size, the basis set superposition error (BSSE) was evaluated, using the counterpoise correction scheme (Boys & Bernardi, 1970). Here, the BSSE is computed as the lowering of the excitation energy upon including ghost MOs on the environment of the triad or the RFQY-SNN model. The results are compiled in Table 3.3. The inclusion of the ghost MOs decreases the excitation energies, especially that of the ET states. Adding the ghost MOs on the SNNML environment to the RFQY triad decreases the ET excitation energy by 0.05 eV. Thus, the BSSE accounts for 0.05 eV (20%) of the total 0.25-eV red shift of the ET excitation energy in the full cluster model RFQY-SNNML as compared to the photoactive triad RFQY model. Adding the ghost MOs of N35 (the side-chain causing the increase of the ET excitation energy) decreases the ET excitation energy by 0.02 eV. Hence, the 0.12-eV blue shift of the RFQY-N cluster with respect to the RFQY model is underestimated by 0.02 eV (15%) without correcting for BSSE. If a larger complex model is considered, such as RFQY-SNN, the BSSE contributes 0.02 eV (15%) to the 0.13-eV red shift of the ET excitation

energy obtained upon extending this model to the full cluster model RFQY-SNNML. Thus, accounting for BSSE increases the blue shift and decreases the red shift of the ET excitation energy. Employing the core-consistent XMCQDPT2' method (Granovsky, 2011), by performing the Firefly calculations with $\text{iropt}=1$ in the LF-series, resulted in a systematic and size-independent blue shift of 0.02-0.03 eV of the ET excitation energies. Thus, the shifts in the excitation energies upon variation of the size and composition of the cluster model is not an artifact of the computational methods, but a manifestation of the intermolecular interactions stabilizing or destabilizing the excited state with respect to the ground state.

Table 3.3. Correction of the XMCQDPT2 excitation energies using the counterpoise correction scheme. a) energies are taken from Figure 3.3; b) BSSEs are given in brackets.

	excitation energy [eV]		
	S_1	S_2	ET
RFQY^{a)}	3.06	3.89	4.31
RFQY with ghost MOs on N35	3.05 (0.01) ^{b)}	3.87 (0.02)	4.29 (0.02)
RFQY with ghost MOs on NNS	3.05 (0.01)	3.87 (0.02)	4.27 (0.04)
RFQY with ghost MOs on SNNML	3.04 (0.02)	3.87 (0.02)	4.26 (0.05)
RFQY-SNN^{a)}	3.08	3.85	4.19
RFQYSNN with ghost MOs on ML	3.07 (0.01)	3.84 (0.01)	4.17(0.02)
RFQYS-NNML^{a)}	3.07	3.82	4.06

3.4.3 Excitation energies – QM/MM calculations

Figure 3.5 shows the excitation energies of the QM/MM models obtained with and without the electrostatic embedding of the QM subsystem into the MM environment. With the CASSCF method a downshift of the S_1 and S_2 excitation energies and a large upshift of the ET-state energy are obtained upon embedding. With the CASPT2 method, only the S_2 state is down-shifted, whereas the S_1 and ET states are up-shifted. The S_2 /ET state crossing is absent in the QM/MM calculations. Upon embedding, the change of the excitation energy with respect to the composition of the QM subsystem is significantly reduced, which is especially apparent for the ET excitation energies. More specifically, without embedding, the ET excitation-energy difference between the most blue-shifted and most red-shifted estimates equals to 0.40 eV (MS-CASPT2 estimates), whereas upon embedding it decreases to 0.10 eV. For the S_1 and S_2 energies, the largest change is found between the QM subsystem consisting only of the LF fragment and the other QM/MM models. Yet again, these shifts are smaller than the respective shifts obtained without the QM/MM embedding.

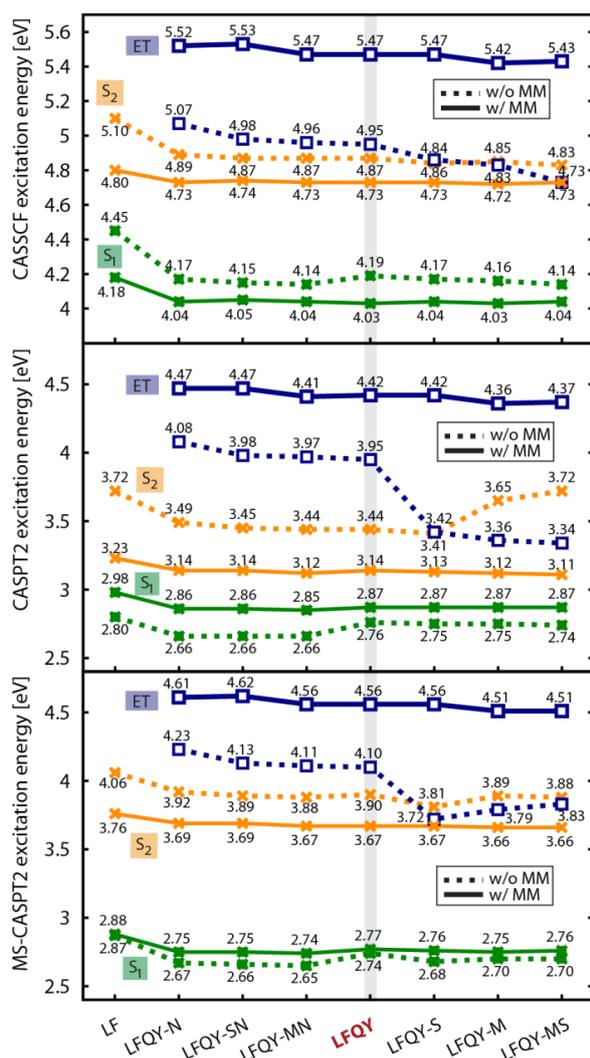


Figure 3.5. The effect of the electrostatic embedding on the excitation energies. To aid the comparison, the dashed lines report the excitation energies of the QM subsystem computed without the MM environment, which are the same CASSCF and CASPT2 energies as in the right panel of Figure 3.3. The QM/MM excitation energies were computed with the charge-redistribution scheme A.

3.4.4 Electrostatic potential maps and charge-transfer effects

Figure 3.6 compares the difference electrostatic-potential maps characterizing the three electronic transitions. Table 3.4 presents the dipole moments in the four electronic states. In the full cluster model, all excited states are more polar than the ground state; the polarity increases in the order S_1 , S_2 , and ET. The excited states of flavin S_1 and S_2 correspond to an intramolecular negative-charge transfer from the phenyl to the uracil ring, consistent with the structure of the MOs involved in the transitions, as discussed in Chapter 2. The ET state corresponds to one electron transfer from the phenol ring of tyrosine to the isoalloxazine system of flavin; thus the transition from the ground state to the ET state is the most polar one considered in this study. For the QM subsystem computed without the MM environment, a very similar picture to that of the full cluster model is found. There are pronounced changes in the charge distribution upon embedding the QM subsystem to the MM environment, demonstrating its polarization. The polarization of the states by the MM environment increases the polarity of all four states but to a different extent. The largest effect is found for the S_1 state that becomes more polar than the S_2

state. The ground-state polarity also significantly increases and becomes comparable to the polarity of the S_2 state. The polarity of the ET state increases to the same extent as that of the ground state. Hence the polarity of the transitions changes: The S_1 transition becomes significantly more polar, the S_2 transition substantially less polar and the polarity of the ET transition does not change as demonstrated by the double-difference maps in Figure 3.6C.

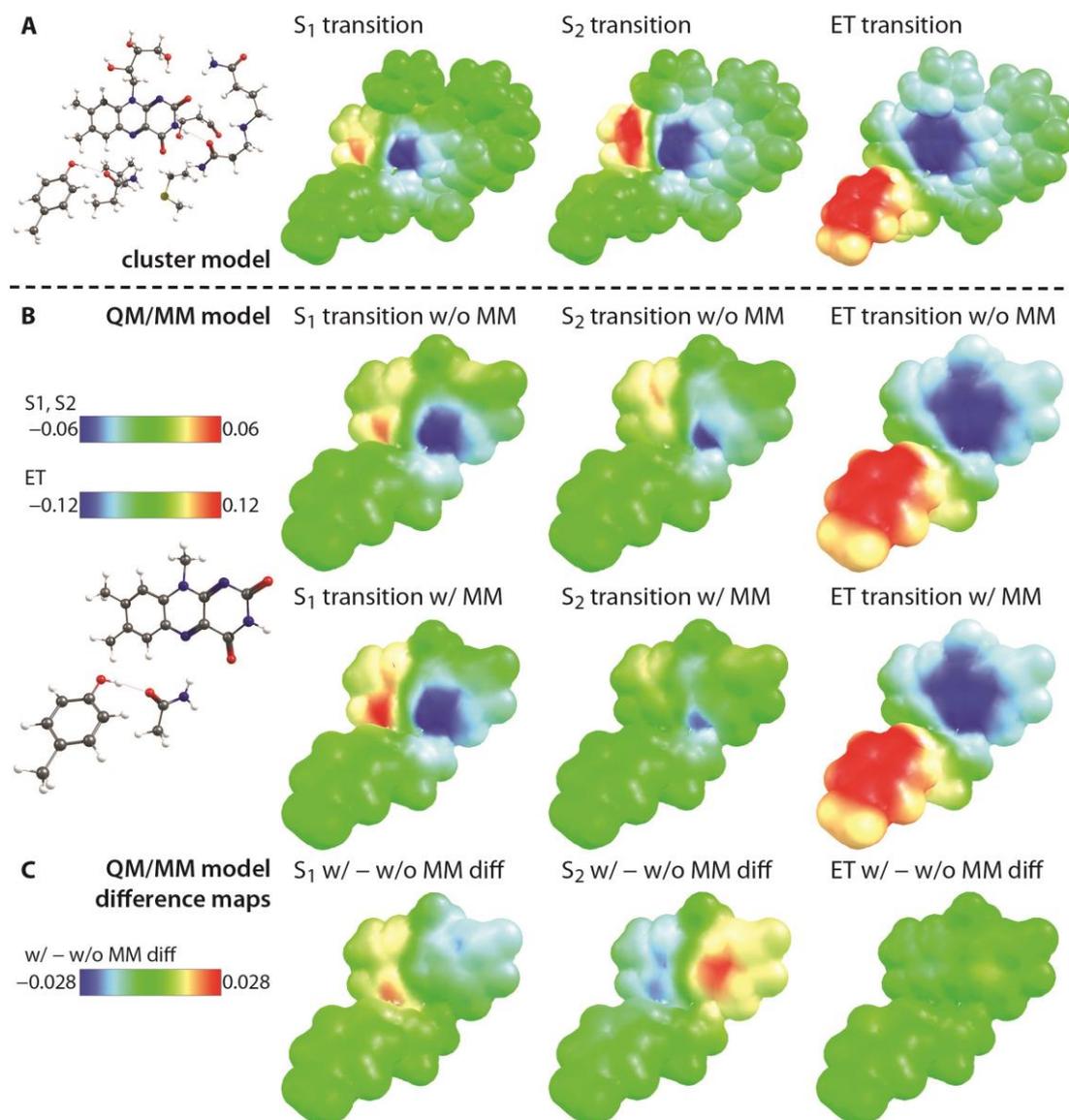


Figure 3.6. Electrostatic potential differences (in atomic units) upon electronic transitions and electrostatic embedding. (A) Molecular model and electrostatic-potential difference maps in the largest cluster model. (B) Molecular model and electrostatic-potential difference maps in the QM/MM models in comparison to the same models without MM environment. The color scale shown is also valid for the maps presented in panel (A). (C) The double difference maps “w/-w/o MM” visualize the polarization of the electrostatic transitions upon including the MM environment. The Chemcraft software was used for the visualization of the maps with the mapped spheres scaled by 1.5 to improve the visibility of the differences.

Table 3.4. Dipole moments (in Debye) for selected models of the cluster series (XMCQDPT2), QM/MM series (XMCQDPT2 and CASSCF without MM) and QM/MM models (CASSCF with MM). The XMCQDPT2 method is abbreviated by XPT2.

QM/MM	GS			S ₁			S ₂			ET		
	XPT2	CASSCF w/o MM	CASSCF w/ MM	XPT2	CASSCF w/o MM	CASSCF w/ MM	XPT2	CASSCF w/o MM	CASSCF w/ MM	XPT2	CASSCF w/o MM	CASSCF w/ MM
LF-Q-Y-N	9.47	10.26	13.65	9.89	11.00	17.25	14.23	11.72	13.64	37.37	37.59	39.58
LF-Q-Y	6.56	7.84	10.46	8.93	10.57	15.48	13.40	10.10	10.99	35.77	36.53	37.66
LF-Q-Y-S	5.88	7.42	10.96	9.10	10.90	16.28	16.60	19.80	11.75	27.67	20.58	37.78
LF-Q-Y-M	6.06	7.00	9.39	8.30	9.85	14.28	12.12	9.11	9.99	30.78	29.94	35.52
LF-Q-Y-MS	5.67	6.79	9.88	8.74	10.40	15.09	12.51	9.03	10.77	34.06	34.40	35.64
Cluster												
LF-Q-Y-N	7.98			8.43			13.01			38.96		
LF-Q-Y	4.92			7.53			12.33			38.47		
LF-Q-Y-MS	5.75			9.37			12.98			35.12		
RF-Q-Y-MSNNL	8.11			11.62			15.31			39.91		

As demonstrated by the maps of the full cluster model in Figure 3.6A, the charge redistribution involves also the environment of the photoactive triad in all three electronic transitions. Especially noticeable is the redistribution of the negative charge into the local environment of the flavin electron acceptor upon the ET transition. To quantify the charge transfer, the net-charge difference (ET state minus ground state) on the electron donor, electron acceptor and the environment of the photoactive triad was computed. The charge-transfer magnitudes correlate with the ET excitation energy, as evidenced by Figure 3.7.

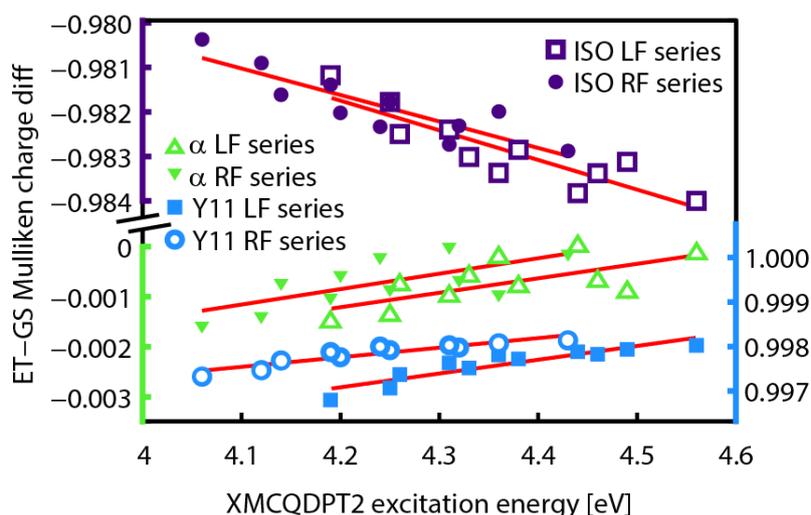


Figure 3.7. Correlation between the charge transfer to the QM environment and the excitation energies of the ET state computed in the cluster model series. For both RF and LF series, the CS ground state to ET charge transfer on the electron acceptor isoalloxazine fragment, the electron donor tyrosine fragment and on the QM environment α is shown. The red lines indicate linear regression, performed with gnuplot.

The more the excitation energy is shifted to the red, the more the negative-charge transfer from tyrosine to flavin deviates from one electron and the larger amount of the negative charge is transferred to the environment. This finding indicates that the charge transfer interactions of the photoactive triad with its environment stabilize the very polar (charge-separated) ET state with respect to the ground state. The absolute magnitude of the charge leakage to the environment is only -0.0057 au. However, even a small amount of charge transfer between interacting molecules has been shown to be associated with significant charge-transfer interaction energies (Mo et al., 2011). In the QM/MM models with the α side chains included in the QM subsystem, the charge-transfer interaction also should play a role. The analogous charge transfer analysis for the QM/MM models was not carried out because of the small number of models and because the S_2 /ET state mixing obviously influencing the atomic charges.

3.4.5 Excitation-energy shift upon QM/MM embedding

Of the two QM/MM interaction energy components E_{ESPF} (Equation 3.1) and $E_{wf\ pol}$ (Equation 3.2), the former represents the stabilization (binding) energy; it is more negative than the QM/MM interaction energy. The $E_{wf\ pol}$ term is approximately one order of magnitude smaller than E_{ESPF} and is positive as it corresponds to the energy of redistributing the equilibrium electron density upon polarization by the MM environment. Although the CASSCF and CASPT2-derived ESPF charges are quite different, the respective E_{ESPF} interaction energies are very similar in magnitude. The $E_{wf\ pol}$ term increases when computed with the CASPT2 method as compared to the CASSCF method. Because the two terms have an opposite sign, the excitation-energy shift upon electrostatic embedding, $\Delta\epsilon^{emb}$ (Equation 3.4), results from the interplay of the stabilization and destabilization of the two electronic states: The blue shift originates from either the increased electrostatic interactions of the ground state compared to the excited state or a larger polarization of the excited state compared to the ground state. The origin of the red shift is a larger electrostatic stabilization of the excited state or a larger wave-function polarization of the ground state. The more polar the transition is, the more significant the ESPF contribution ϵ_{ESPF} becomes. The polarity of the transitions is demonstrated by the difference electrostatic-potential maps in Figure 3.6B. In the case of nonpolar transitions, like the S_2 transition in the QM/MM model, the ϵ_{ESPF} component should be insignificant and thus the wave-function polarization contribution $\epsilon_{wf\ pol}$ determines the excitation-energy shift. The double difference electrostatic-potential maps in Figure 3.6C illustrate the charge redistribution due to wave-function polarization.

The results of the energy decomposition analysis for the CASSCF and CASPT2 excitation-energy shifts upon electrostatic embedding, $\Delta\epsilon^{emb}$, are presented in Figure 3.8. The ϵ_{ESPF} shifts are very similar when computed with the CASSCF and CASPT2 methods, whereas the $\epsilon_{wf\ pol}$ shifts increase significantly when computed with the CASPT2 method for the S_1 and S_2

transitions but are quite similar for the ET transition. The largest $\epsilon_{wf\ pol}$ shifts are obtained for the QM subsystems with the S_2 /ET state crossing; these large shifts are related to the spurious change of the S_2 and ET energies computed with the CASPT2 method. The ϵ_{ESPF} term causes the red-shift of the S_1 energy, does not contribute to the shift of the S_2 energy, and significantly blue-shifts the ET excitation energy. The $\epsilon_{wf\ pol}$ shift is blue for the S_1 transition and thus counteracts the red ϵ_{ESPF} shift. In total, the CASSCF S_1 energy is red-shifted because of the electrostatic interactions, but the CASPT2 S_1 excitation energy is blue-shifted because of wavefunction polarization. The $\epsilon_{wf\ pol}$ energy completely determines the red shift of the S_2 excitation energy. For the S_1 and S_2 transitions, the largest shifts are found when the QM subsystem consists only of the LF chromophore. Interestingly, in the LF model, the ϵ_{ESPF} and $\epsilon_{wf\ pol}$ shifts compensate each other for the S_1 transition, whereas they sum up for the S_2 transition. In the LFQY model, the $\epsilon_{wf\ pol}$ red shift partly counteracts the large ϵ_{ESPF} blue shift for the ET transition.

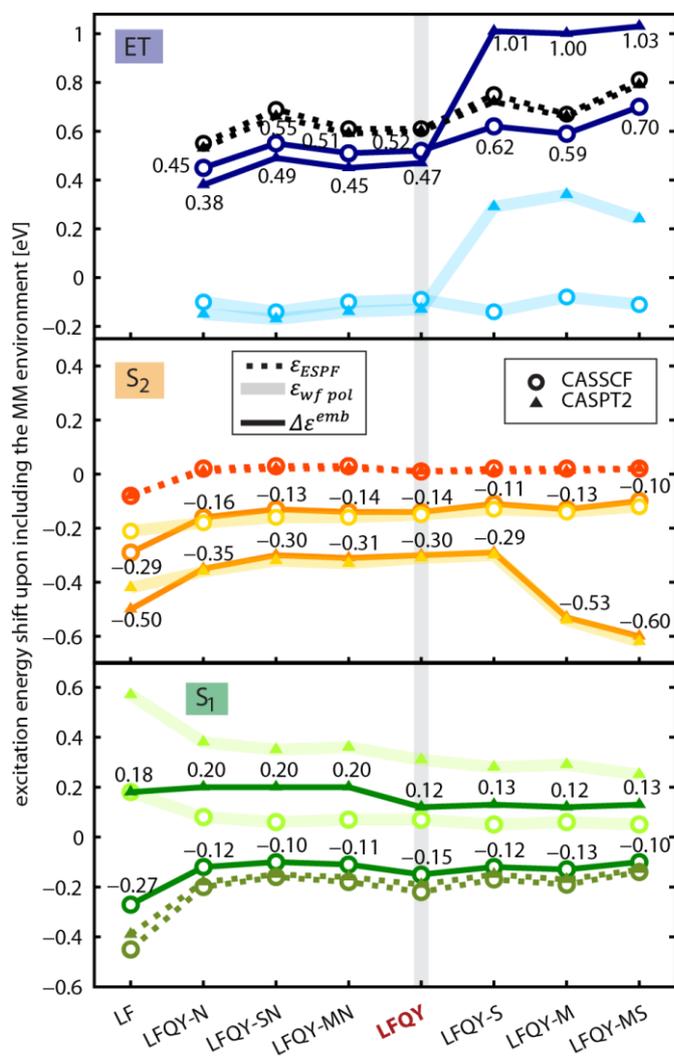


Figure 3.8. Energy decomposition analysis of the QM/MM models. The two contributions, $\epsilon_{wf\ pol}$ and ϵ_{ESPF} to the excitation-energy shifts upon including the interactions with the MM part, $\Delta\epsilon^{emb}$ (Equation 3.4), are shown. Negative and positive values correspond to red and blue shifts, respectively.

3.4.6 Modifications of the MM charges

The modifications of the MM charges of selected side chains resulted in specific changes of the excitation-energy spectrum, especially of the ET excitation energy. Figure 3.9 shows a cartoon of the flavin-binding pocket with highlighted residues for which the MM charges were manipulated. The computed excitation energies after zeroing the MM charges of the S31, N35, or M96 residues along with the energy decomposition analysis is presented in Figure 3.10; the results obtained when modifying the charges on the phosphate group of FMN and on the H75, S31, or Y66 residues are presented in Figure 3.11. Zeroing the MM charges resulted in small changes of the excitation energies, with the largest shift of the ET energy. Zeroing the S31 or N35 charges resulted in a 0.03-0.04 eV (5-6 nm) blue shift of the S_1 excitation energy. Experimental absorption spectra are available for a PixD S31A mutant, which show a 12-nm red shift of the first absorption band, associated with the S_1 transition (Bonetti et al., 2009). Thus, a simple zeroing procedure is not capable of reproducing the experimental trend. Zeroing of the M96 side chain causes only an insignificant S_1 shift, more in line with the experimentally observed unperturbed absorption spectrum of the PixD M96A mutant (Yuan et al., 2011). The interactions of the photoactive triad with N35 result in a blue shift of the ET excitation energy and interactions with either S31 or M96 yield a red shift. The respective ϵ_{ESPF} and $\epsilon_{wf\ pol}$ terms follow the same trend as in the unmodified QM/MM series. The origin of the $\Delta\epsilon^{emb}$ excitation-energy shifts is the same as in the reference QM/MM model, however, because of the small magnitudes of the shifts, neither ϵ_{ESPF} nor $\epsilon_{wf\ pol}$ can be neglected if the excitation energies of the modified models are compared to those of the reference model.

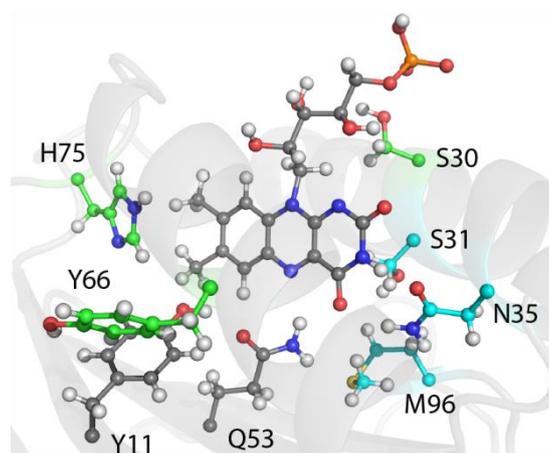


Figure 3.9. Arrangement of the residues with modified MM charges in PixD BLUF. The MM charges of the cyan colored residues were set to zero. The charges on the green-colored residues and the phosphate were changed by +1 or -1 a.u.

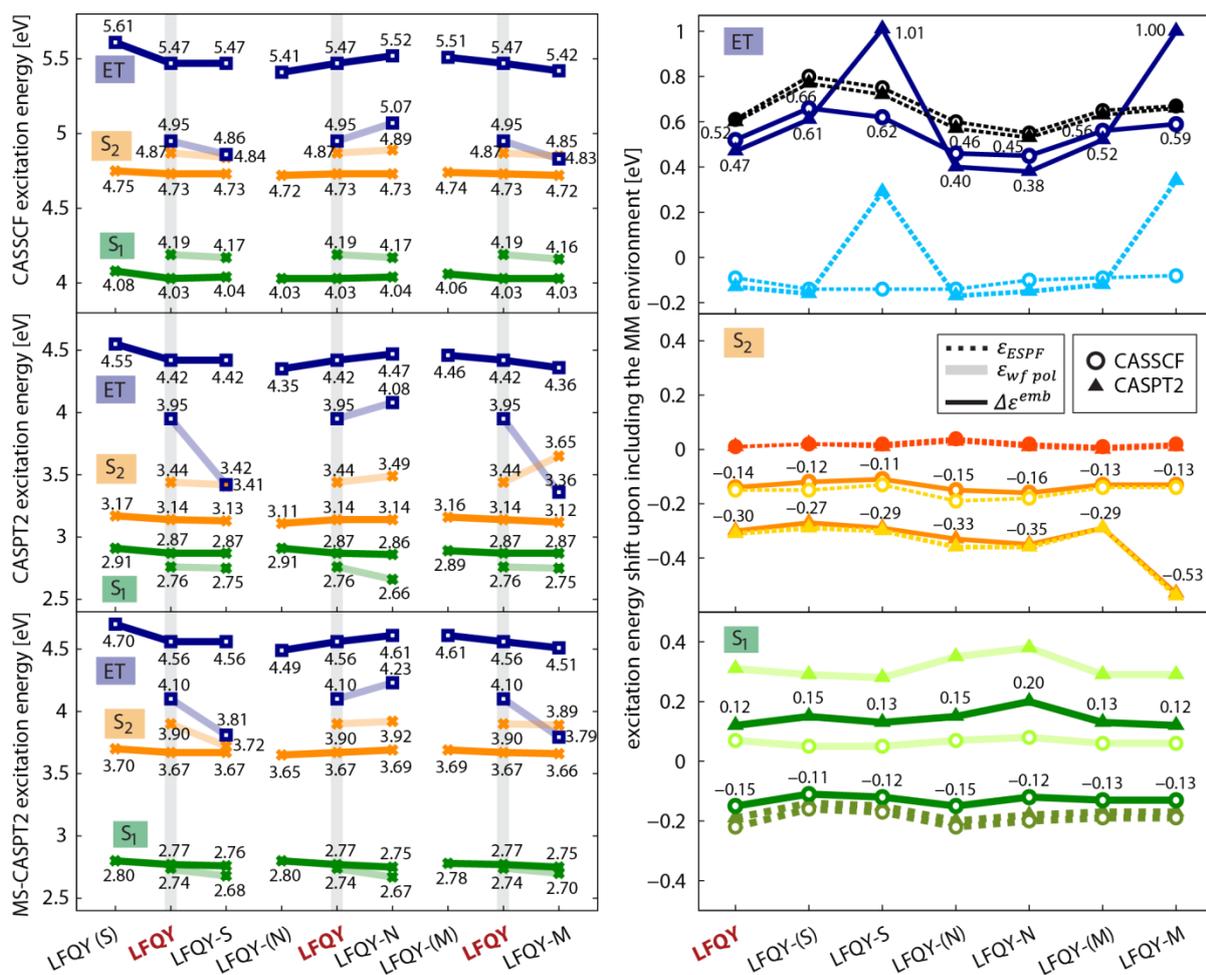


Figure 3.10. CASSCF and (MS-)CASPT2 excitation energies of the S₁, S₂, and ET states in eV with (solid lines) and without (transparent lines) the MM environment. The brackets in (S), (N) and (M) indicate that the corresponding residue's electrostatic interactions were switched off by setting the charges of its atoms to zero (left). Energy decomposition analysis of the excitation energy shifts upon including the MM environment (right). The charge redistribution scheme A and the same color coding are used as in the previous figures.

Upon placing a positive or a negative charge on the H75, S31, and Y66 residues as well as on the phosphate group of FMN, substantial shifts of the ET excitation energy, and small changes of the flavin S₁ and S₂ excitation energies are found, as expected. The energies, along with the shift decomposition, are presented in Figure 3.11. As in all considered models, the changes of the S₁ and ET energies follow the same trend as the ϵ_{ESPF} energies, whereas the changes of the S₂ energy are determined by the polarization term $\epsilon_{wf pol}$. The polarity of the S₁ transition is significantly smaller compared to the ET transition, therefore only small changes of its excitation energy are possible to achieve through the electrostatic interactions with the environment. In contrast, the large polarity of the ET transition underlies its high sensitivity to the electrostatic charges of the environment. The computed changes are consistent with the conjecture that interactions of the electron donor with the negative charge and the acceptor with the positive

charge decrease the ET excitation energy. The reverse is also true: interactions of the donor with the positive charge and of the acceptor with the negative charge increase the ET energy. It is interesting to note that the stabilization of the ET state by placing a negative charge on Y66 results in almost degenerate ET and S_2 states. However, in this case the two states are not mixing at all: the Mulliken charges on the LF fragment in the third and fourth state are -0.9610 and -0.0032 a.u., respectively; and on the Y11 fragment 0.9616 and 0.0036 a.u., respectively, indicating that the QM/MM coupling dominates the QM coupling between the two degenerate states.

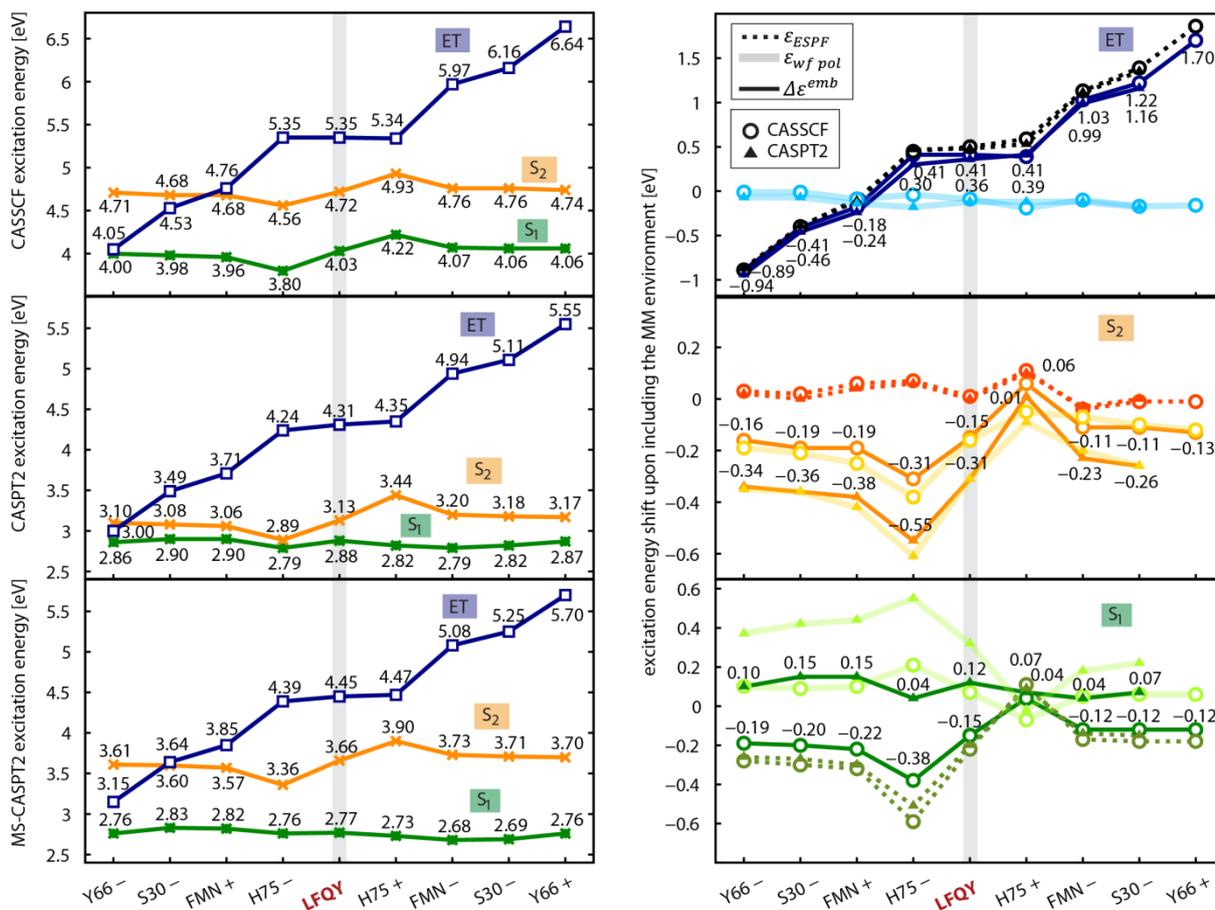


Figure 3.11. CASSCF and (MS-)CASPT2 excitation energies of the S_1 , S_2 , and ET states in eV (left). Energy decomposition analysis of the excitation energy shifts upon including the MM environment (right). The charge redistribution scheme B and the same color coding is used as in the previous figures.

3.4.7 Intermolecular interactions and the excitation energy of the ET state

The above analysis showed that the ET state is most sensitive to environmental effects. These can be both red- and blue-shifting: upon extending the QM model a pronounced red-shift is obtained, whereas by adding the MM environment the ET excitation energy is significantly blue-

shifted. This section further elaborates on the role of electrostatic and non-electrostatic interactions in determining the ET excitation energy. In Figure 3.12, the electrostatic-interaction energy difference in the CS ground state and the ET state, i.e. the Coulomb shift ε_C^α (Equation 3.9), is compared to the XMCQDPT2 or CASPT2 excitation-energy shifts $\Delta\varepsilon_{QM}^\alpha$ (Equation 3.5) and also to the shifts computed with the QM/MM models $\Delta\varepsilon_{QM/MM}^\alpha$ (Equation 3.7), originated from including the side chain(s) α in the QM environment, and to the ESPF component of the latter shift, $\Delta\varepsilon_{ESPF}^\alpha$.

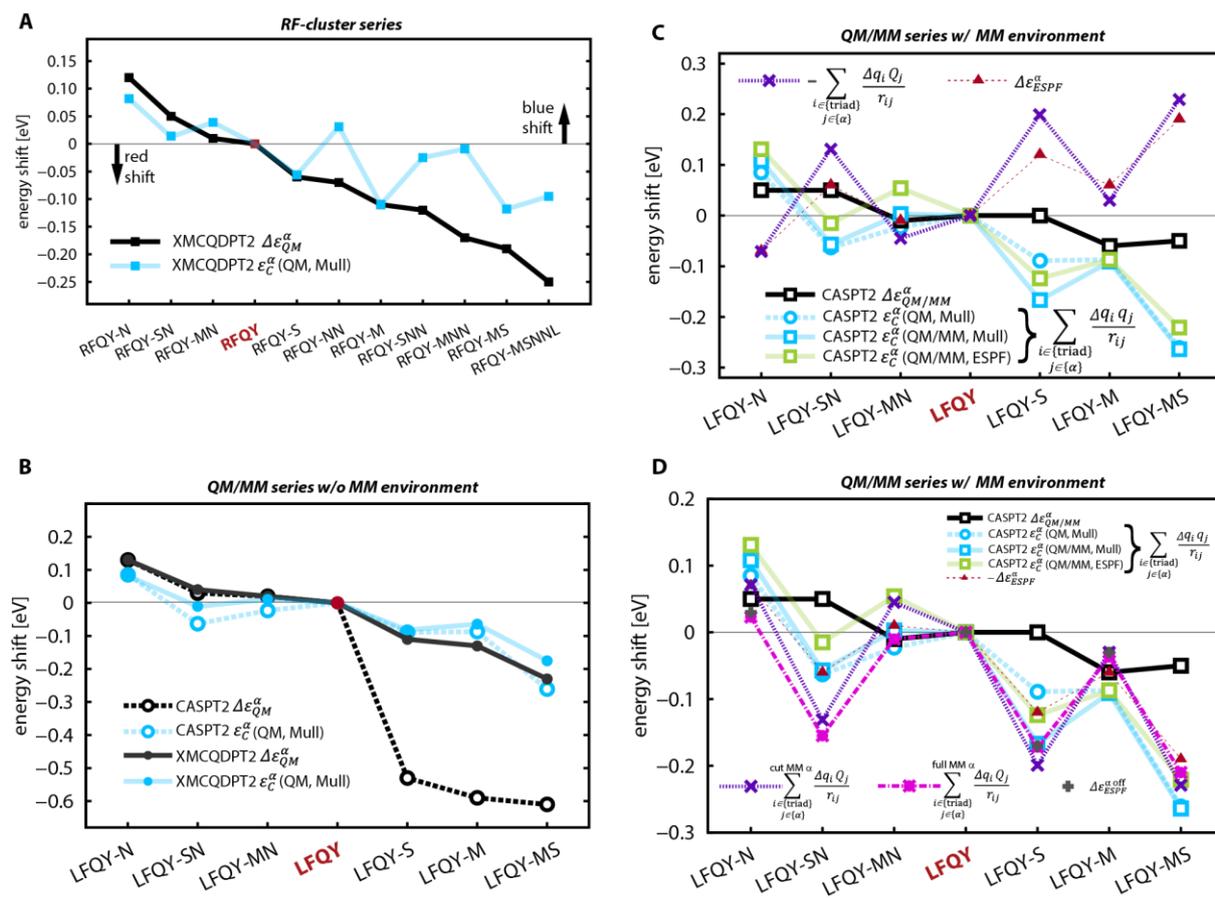


Figure 3.12. Comparison of the ε_C^α , $\Delta\varepsilon_{QM}^\alpha$ and $\Delta\varepsilon_{QM/MM}^\alpha$ shifts of the ET excitation energy obtained upon including the QM environment α . (A) $\Delta\varepsilon_{QM}^\alpha$ and ε_C^α shifts in the RF-cluster model series computed with XMCQDPT2 using Mulliken charges. (B) $\Delta\varepsilon_{QM}^\alpha$ and ε_C^α shifts in the QM subsystems without MM environment, computed with both XMCQDPT2 and CASPT2 using Mulliken charges. (C) $\Delta\varepsilon_{QM/MM}^\alpha$ and ε_C^α shifts in the QM/MM models, computed with CASPT2 using Mulliken and ESPF charges. Additionally, the $\Delta\varepsilon_{ESPF}^\alpha$ shift and a Δq_i -term explained in the text (Equation 3.11) is plotted. (D) Same as in (C) with additional terms Δq_i (Equation 3.13) and $\Delta\varepsilon_{ESPF}^{\alpha, off}$. Note the different energy scale of the plots.

The $\Delta\varepsilon_{QM}^\alpha$ shift includes the contribution of all interactions between the photoactive triad and the particular side chain(s) α (electrostatic, charge transfer, polarization) computed with the

XMCQDPT2 or CASPT2 methods. The deviations between ε_C^α and $\Delta\varepsilon_{QM}^\alpha$, presented in Figure 3.12A and B, indicate the role of interactions other than electrostatic ones in the excitation energy of the ET state. Such interactions are, for instance, polarization and charge-transfer. The charge-transfer analysis in Figure 3.7 indicates that the role of these interactions increases in the red-shifted models and these interactions contribute to the stabilization of the ET state. In the blue-shifted models RFQY-N (Figure 3.12A) and LFQY-N (Figure 3.12B) the electrostatic interactions of the triad with N35 induce a blue-shift, which is further increased by the small shift due to the non-electrostatic interactions. In the model RFQY-NN (Figure 3.12A), including the N35 and N34 side chains results in a small blue shift, but this shift is compensated by the substantial red shift due to the non-electrostatic interactions. In the RFQY-S and RFQY-M models (Figure 3.12A), the Coulomb shift equals to the XMCQDPT2 QM shift, indicating that the non-electrostatic effects are equal in the two states and therefore play no role in the excitation-energy shift. Interestingly, however, when both S31 and M96 side-chains are added in the RFQY-MS model, a significant non-electrostatic red shift appears. Thus, it seems that the contribution of non-electrostatic interactions is increasing with increasing number of molecular fragments included in the environment of the RFYQ triad which stabilizes the ET state (red-shifts the ET excitation energy). In the full cluster model, the shifts due to the electrostatic and non-electrostatic interactions have the same sign (both red) and are about equal in magnitude.

It appears that there is an increasing contributing of the non-electrostatic quantum effects in the LFQY-S and LFQY-M complexes (Figure 3.12B) according to the CASPT2 shifts as compared to the XMCQDPT2 shifts. These models feature the S_2 /ET state crossing (see Figure 3.4) and therefore the polarity of the ET state is reduced by the mixing with the S_2 state, thus decreasing the respective Coulomb sum. Figure 3.12B shows that the XMCQDPT2 and CASPT2 Coulomb shifts are similar, whereas the respective QM shifts deviate significantly in the models with the S_2 /ET state crossing because of the spurious decrease of the CASPT2 excitation energy.

Figure 3.12C compares the Coulomb shifts ε_C^α for the QM/MM series computed from the Mulliken MM-polarized (computed with the MM environment) and not polarized charges (computed without the MM environment). For the considered ET state, these shifts are almost identical except for the LFQY-S model, in which the polarization by the MM environment increases the Coulomb shift because it removes the state crossing and the non-polarized mixed ET/ S_2 transition is less polar than the polarized pure ET transition. The Coulomb shifts computed with the ESPF charges, also presented in Figure 3.12C, are blue-shifted compared to those estimated with the Mulliken charges. The deviation is small, but for a larger number of interacting charges this difference may become more significant and might be contributing to the blue shift of the ε_{ESPF} energies determining the ET excitation energy in the QM/MM models.

The QM/MM excitation-energy shift, caused by moving the α side chain(s) in the QM subsystem, $\Delta\varepsilon_{QM/MM}^\alpha$ (Equation 3.7), plotted in Figure 3.12C consists of three components (the QM contribution, the ESPF contribution and the polarization contribution), as explained in Section 3.2. The electrostatic-energy component $\Delta\varepsilon_{ESPF}^\alpha = \varepsilon_{ESPF}^\alpha - \varepsilon_{ESPF}^{\text{triad}}$, can be written as two Coulomb sums:

$$\begin{aligned}
\Delta\varepsilon_{ESPF}^{\alpha} &= \sum_{\substack{i \in \{\text{triad}, \alpha\} \\ k \in \{MM\}}} \frac{\Delta q_i Q_k}{r_{ik}} - \sum_{\substack{i \in \{\text{triad}\} \\ k \in \{MM, \alpha\}}} \frac{\Delta q_i Q_k}{r_{ik}} \\
&= \sum_{\substack{j \in \{\alpha\} \\ k \in \{MM\}}} \frac{\Delta q_j Q_k}{r_{jk}} - \sum_{\substack{i \in \{\text{triad}\} \\ j \in \{\alpha\}}} \frac{\Delta q_i Q_j}{r_{ij}}
\end{aligned} \tag{3.11}$$

where small q_i and q_j denote the QM charges of the triad and its α environment, respectively; capital Q_j and Q_k denote the MM charges of the α environment and the rest of the protein excluding the α environment and the triad, respectively; $\Delta q = q^{\text{exc}} - q^{\text{GS}}$ is the charge difference upon excitation. The indices i, j , and k are associated with the photoactive triad, α environment, and MM environment atoms, respectively. The term containing Δq_j describes the change of the electrostatic energy related the change in the interactions of the polarized charges of the α environment with the charges of the MM environment. The term containing Δq_i describes the change of the electrostatic interaction because of the excitation in the triad interacting with the classical MM charges of the α environment. The $\Delta\varepsilon_{ESPF}^{\alpha}$ energies together with the Δq_i -term are also plotted in Figure 3.12C. The difference between the two reveals the magnitude of the Δq_j -term, i.e. the effect of including the polarized QM-ESPF charges as opposed to the fixed Amber MM charges of the α environment. The Δq_i -term, describing local interactions, is equal or larger in magnitude than the $\Delta\varepsilon_{ESPF}^{\alpha}$ shift. The magnitude of the Δq_j -term is significant, even though the magnitude of the charge polarization Δq_j is rather small (see below), because the number of the MM atoms in the model of the protein and solvent interacting with the q_j charges is large. To further elaborate on the role of charge-polarization of the photoactive triad, ε_C^{α} (Equation 3.9) is compared to $\Delta\varepsilon_{ESPF}^{\alpha}$. The Coulomb shift ε_C^{α} can be also decomposed into two terms depending on Δq_i and Δq_j :

$$\varepsilon_C^{\alpha} = \sum_{\substack{i \in \{\text{triad}\} \\ j \in \{\alpha\}}} \frac{\Delta q_i q_j^{\text{GS}}}{r_{ij}} + \sum_{\substack{i \in \{\text{triad}\} \\ j \in \{\alpha\}}} \frac{\Delta q_j q_i^{\text{exc}}}{r_{ij}} \tag{3.12}$$

The Mulliken charge analysis shows that the Δq_i -term is the leading term of the Coulomb shift and the Δq_j -term is negligible (does not exceed 0.016 eV) in the considered cases. Figure 3.12C shows that $\Delta\varepsilon_{ESPF}^{\alpha}$ more-or-less compensates ε_C^{α} in the estimates of the $\Delta\varepsilon_{QM/MM}^{\alpha}$ shifts. The magnitude of the polarization component $\Delta\varepsilon_{wf\ pol}^{\alpha}$ can be derived from Figure 3.8: it ranges from -0.05 to 0.01 eV (CASSCF values) or from -0.01 to 0.47 eV (CASPT2 values). Thus, $\Delta\varepsilon_{QM/MM}^{\alpha}$ essentially accounts for the non-electrostatic interactions of the triad with the α environment and thus is roughly comparable to the difference between the ε_C^{α} and $\Delta\varepsilon_{QM}^{\alpha}$ energies in Figure 3.12A and B. Thus the smaller magnitude of the $\Delta\varepsilon_{QM/MM}^{\alpha}$ shifts compared to the

$\Delta\varepsilon_{QM}^\alpha$ shifts are explained by the fact that the electrostatic interaction energies between the photoactive triad and the α side chain(s) are already included in the reference QM/MM model, but not in the cluster model consisting only of the triad.

Figure 3.12D compares the above shifts from Figure 3.12C to the shifts obtained with the zeroed charges. In the excitation energy shifts computed with the zeroed charges of the α environment, $\Delta\varepsilon_{MM}^\alpha$ (Equation 3.6), the energy term depending on the Δq_i and Q_j charges also appears:

$$\Delta\varepsilon_{ESPF}^{\alpha\text{ off}} = \sum_{\substack{i \in \{\text{triad}\} \\ k \in \{\alpha, MM\}}} \frac{\Delta q_i Q_k}{r_{ik}} - \sum_{\substack{i \in \{\text{triad}\} \\ k \in \{MM\}}} \frac{\Delta q_i Q_k}{r_{ik}} = \sum_{\substack{i \in \{\text{triad}\} \\ j \in \{\alpha\}}} \frac{\Delta q_i Q_j}{r_{ij}} \quad (3.13)$$

The obtained expression in (Equation 3.13) looks like the second term in (Equation 3.11), however, they are not completely identical: The $\Delta\varepsilon_{ESPF}^{\alpha\text{ off}}$ energy shift depends on the charges of the entire α residue as defined in the PDB model, whereas the Δq_i term in the $\Delta\varepsilon_{ESPF}^\alpha$ energy shift depends on the charges of the α side chain as included in the QM subsystem (fragments without the back bone atoms). These two Δq_i terms computed for the reference QM/MM model for the full α residue and the α side chain are compared in Figure 3.12D and show small deviations. At the same time, the $\Delta q_i Q_j$ term computed for the full α residue practically equals to the $\Delta\varepsilon_{ESPF}^{\alpha\text{ off}}$ excitation energy shift, which is also shown in Figure 3.12D. The polarization energy shift $\Delta\varepsilon_{wf\text{ pol}}^{\alpha\text{ off}}$, which is the other component of the $\Delta\varepsilon_{MM}^\alpha$ shift, ranges from -0.01 to 0.04 eV (the CASPT2 results) in the three considered zeroed-charge models, and in the case of the N35 residue, it exceeds the $\Delta\varepsilon_{ESPF}^{\alpha\text{ off}}$ shift and makes more than a half of the $\Delta\varepsilon_{MM}^\alpha$ shift. Thus, the excitation-energy shifts computed by manipulating the MM charges are not necessarily dominated by the electrostatic interactions. This conclusion is also supported by the origin of the S_2 excitation energy shifts in the QM/MM models, which are dominated by the polarization of the QM subsystem (by the MM environment), even when the MM charges are manipulated.

3.4.8 Comparison of the full cluster model with the reference QM/MM model

The two alternative models of PixD BLUF, the full cluster model RFQY-NNMSL and the reference QM/MM model, with all differences mentioned in the previous sections, provide overall consistent estimates of the BLUF excitation-energy spectrum. In the cluster model with the XMCQDPT2 method the excitation energies of the S_1 , S_2 , and ET states are 3.07, 3.82, and 4.19 eV, respectively; whereas in the reference QM/MM model with the MS-CASPT2 method, the respective energies are 2.77, 3.67, and 4.56 eV. Compared to the isolated chromophore, both models predict red-shifted S_1 and S_2 excitation energies, as expected for flavin in a polar

environment. The major difference of the two models is the red as opposed to the blue shift of the ET excitation energy upon introducing the environment of the photoactive triad.

Table 3.5 summarizes the effects of the geometry optimization, the PT2 method, the QM environment and the QM/MM embedding in the obtained excitation-energy estimates. The S_1 estimate depends on the choice of the PT2 method and practically does not depend on environmental effects. The environmental shifts are observed only for interactions such as hydrogen bonds of flavin with Q53, N35, and N34. As these hydrogen bonds both red- and blue shift the S_1 excitation energy, in the full cluster model the related effects compensate each other. In contrast, the S_2 excitation energy is rather sensitive to both local and long-ranged effects of the environment. The sensitivity is related to the high polarity of this transition in the cluster model. Interestingly, in the QM/MM model this property is not reproduced because the higher polarization of the ground state with respect to the S_2 state turns this transition into a non-polar one. The red-shift of the S_2 excitation energy in the QM/MM model originated purely from wave-function polarization. The red shift shows no dependence on the size of the QM subsystem, in contrast to the cluster model. The ET state is very sensitive to all factors but the choice of the PT2 method (unless the S_2 /ET state crossing is considered). Remarkably, the energy of the state is sensitive to all types of environmental effects: the electrostatic interactions, the arrangement of the electron donor and the electron acceptor and also the quantum-mechanical effects such as polarization and charge transfer. The latter can only be accounted for by enlarging the QM subsystem. As in the QM/MM model the ET excitation energy depends on the interactions between the atomic charges to a significant extent, the accuracy of these charges is critical for the correct estimation of the ET excitation energy in the protein. The dependence of the ET energies on many factors makes their accurate determination a real challenge. It is noteworthy that the electrostatic interactions of the photoactive triad with the MM environment eliminate the crossing of the ET and S_2 electronic states, indicating that the classical interactions dominate the quantum mechanical ones in the QM/MM models considered here.

Table 3.5. Geometry, PT2 method and QM- or MM-environmental effects on the excitation energies obtained with the full cluster model and reference QM/MM model. Excitation energies are taken from Figure 3.3 and 3.5. All significant energy shifts are marked bold.

effect	energy [eV]		
	S_1	S_2	ET
Cluster-geometry versus QM/MM -geometry Model LF-Q-Y, XMCQDPT2 method	3.01 – 3.04	4.04 – 4.06	4.44 – 4.12
XMCQDPT2 method versus MS-CASPT2 method Model LF-Q-Y at the QM/MM geometry	3.04 – 2.74	4.06 – 3.90	4.12 – 4.10
QM environment: LF-Q-Y cluster versus full cluster, XMCQDPT2 method	3.01 – 3.03	4.04 – 3.96	4.44 – 4.19
MM environment: LF-Q-Y w/o MM environment versus LF-Q-Y with MM environment, MS-CASPT2 method	2.74 – 2.77	3.90 – 3.67	4.10 – 4.56

3.5 Discussion and conclusions

This chapter analyzed the effects of the intermolecular interactions on the excitation energies of the photoactive triad in BLUF by comparing the computed excitation energies in a QM cluster model and in a QM/MM model. The estimates of the protein's role in fine-tuning the excitation energies support mechanistic studies aiming at establishing the mechanism of photoactivation in BLUF. A careful analysis is especially important in the case of the tyrosine-flavin ET state, as the ET state excitation energy is experimentally only indirectly accessible through the determination of the flavin fluorescence quenching rates. Multi-exponential models are needed to describe the measured fluorescence kinetics, making the understanding of the link between the fluorescence life time and the BLUF photoreaction challenging. Therefore, evaluating the effects related to the specific intermolecular interactions on the ET excitation energy is indispensable to complement the experimental detection of fluorescence life times. The flavin S_1 and S_2 excitation energies, however, can be directly derived from the UV/Vis absorption spectra. Experiments observe only minor changes in the UV/Vis spectra of the oxidized flavin in proteins and solutions of different polarity. In polar solvents, there is a small red shift of the S_1 and S_2 bands that is more pronounced for the higher-energy S_2 state. In previous quantum-chemical studies this observation was correlated with the polarity of the S_1 and S_2 states (Salzmann et al., 2008; Hasegawa et al., 2007). Overall, the flavin excitation energies computed with the TD-DFT, CC2, SOS-CIS, MRCI(DFT) and the current XMCQDPT2 and (MS)-CASPT2 estimates are in good agreement with each other and with the experimental excitation band maxima, as already discussed in Chapter 2 and reviewed in detail in (Domratcheva et al., 2014).

This chapter reveals major environmental effects that shift the excitation energy of the S_1 , S_2 and ET excited states. The S_1 state of flavin shows only small variations with respect to the environment. Consistent with this finding, color tuning in BLUF (and in flavoproteins binding an oxidized flavin in general) is not important because the charge separation in the S_1 transition is not significant enough. The method differences play a significant role in the energy of the S_1 state but not in the amount of the environment-induced shifts. The S_2 excitation energy is more sensitive to the environment because of the higher polarity of the S_2 transition. However, surprisingly, the higher polarity is not reproduced by the QM/MM models, which predict the S_2 transition to be rather neutral in contrast to the S_1 and ET transitions. The S_2 state is less important in the photochemistry of BLUF and other flavoproteins (Domratcheva et al., 2014). It was included here for the sake of completeness and because the photochemically relevant ET state is close to S_2 in energy. The ET excited state energy depends on many factors. The geometry conformation of the active site plays an important role as it defines the distance between the electron donor and acceptor. The electrostatic environment, as expected, has a large effect on the energy of the highly polar ET transition that is responsible for the blue-shift of the ET excitation energy in the QM/MM model. Importantly however, the ET excitation energy also depends on the QM interactions of the photoactive triad with its local environment to a substantial extent: these interactions, accounted for only by the cluster model with an extended

QM subsystem, red-shift the ET excitation energy. Therefore, in the estimation of the ET state excitation energy, the QM cluster and QM/MM approaches are rather complementary.

The QM/MM models account for the long-range electrostatic effects. The QM cluster models account only for the local electrostatic effects but also for QM interactions, e.g. charge transfer between the photoactive triad and the local environment. In the cluster models, it is the QM effects that play a substantial role in the stabilization of the ET state along with the electrostatics. The long-ranged electrostatic interactions of the protein with the photoactive triad are missing in the cluster models, which should result in a slow convergence of the excitation energies with the size of the cluster model when a very polar transition is considered. The convergence is slower, the more polar the transition is. Still, a careful choice of the local environment to be included in the cluster to study specific problems may minimize the systematic error due to the missing parts of the protein. On the contrary, the QM/MM model accounts for the long-range electrostatic effects that dominate the computed excitation energy shifts. The presented ESPF-operator based model predicts a blue shift of the ET excitation energy upon including the electrostatic interactions of the protein. However, the amount of the blue shift upon the QM/MM embedding is difficult to validate. The fact that this shift is defined by the rather simple ESPF interactions raises the question whether the ESPF electrostatic model is accurate enough. Moreover, the excitation energy obtained with the QM/MM model also depends on the MM force-field charges; again the dependence of the excitation energies on the magnitude of the atomic charges increases with the increasing polarity of the electronic transition and the number of atomic charges to be considered. However, in the studies of shifts induced by rather local effects, like point mutations, the systematic error related to the MM charges should be alleviated.

How accurate the electrostatic interactions are described by the ESPF scheme is especially critical for the ET excitation energy; an underestimation or overestimation of the electrostatic interaction energy results in artifacts rather than in a more realistic protein-tuned ET excitation energy estimate. This is especially important for photochemistry: The blue shift of the ET-state energy in the QM/MM model compared to the cluster model might be significant, for example, in the estimates of the electron-transfer rates. This chapter showed that for a highly polar state, like the tyrosine-flavin ET excited state in BLUF, the excitation energy depends strongly on the electrostatic description of the protein. When the QM subsystem is enlarged, the ET excitation energy red shifts in the QM/MM models similar to the cluster models, though to a much smaller extent than in the cluster models. The strongly blue-shifting protein ESPF electrostatic interaction terms dominate over the red-shifting local QM interactions (or also the local electrostatic interactions) between the photoactive triad and the α environment. The modification of the MM charges by $+1/-1$ au, though exaggeratedly, but clearly showed that the ET excitation energy is highly sensitive to the MM charges and thus an “unphysical charge” or an “unphysically screened charge” in the QM/MM model results in an unphysical shift in the ET excitation energy. As an example the counterions neutralizing the charge of the protein should be mentioned: the seven Na^+ ions were placed randomly into the model (as is usually done when neutralizing a protein model). However, a more realistic simulation should account for the

presence of both cation and anion counterions, and for their averaged distribution around the negatively and positively charged residues on the protein surface. The importance of the ion distribution and its role in screening the DNA charges for the ET excitation energies in the case of the DNA photolyase was recently pointed out in (Moughal Shahi & Domratcheva, 2013).

The problem of choosing a cluster model or a QM/MM model is related to estimating the systematic errors caused by the treatment of the protein electrostatics. In the cluster model the origin of the excitation-energy shifts is clear and physically well-founded. However, in the QM/MM model used in this study the excitation-energy shifts depend strongly on the ESPF charge model and its rigorous validation would be an additional tedious task to carry out. Therefore, it is found here that large cluster models including the local environment of the photoactive triad are more suitable to address photochemical questions in the BLUF photoreceptor. The highly efficient XMCQDPT2-CASSCF code in Firefly is a further important reason. Chapter 5 will revisit some of the QM/MM issues presented here, in view of the fluorescence quenching mechanism and the computed photoreaction pathways.

4 Glutamine rotamers

In this chapter, I discuss two BLUF crystal structures and identify the glutamine rotamer conformation that is consistent with the experimental electron densities. The analysis presented here solves the long-standing debate on the chemical change underlying the photoactivation in BLUF: The two main hypotheses postulate rotation as opposed to tautomerization of the conserved glutamine residue Q63 (in this chapter the residue numbering of the AppA BLUF protein is used). In this chapter, I propose physical criteria for the dark- and light-state structures as well as for the light-induced photoreaction to evaluate existing models of BLUF. The glutamine rotamer assignment of the crystal structure with the PDB code 1YRX does not satisfy these criteria because after equilibrating the intermolecular forces the glutamine rotamer in 1YRX is incompatible with the experimental density. Therefore, the root of the mechanistic controversy is identified to be the incorrect glutamine rotamer assignment in 1YRX. Furthermore, by performing extensive dihedral-angle energy scans, I show that the glutamine side chain may rotate without light activation in the BLUF dark state. Finally, I demonstrate that the tautomerized glutamine is consistent with the formulated criteria and observations of the BLUF light state. Parts of this chapter have been published in (Udvarhelyi & Domratcheva, 2013); indicated figures and parts of the text are taken from this publication. In addition to the published results, calculations on substantially extended cluster models are also presented that elucidate that specific glutamine rotamers are stabilized in the two crystal-structure conformations of BLUF. On the basis of these preferences, I discuss the two structures as the functional dark and light states.

4.1 BLUF structures

X-ray protein crystallography determines the three-dimensional structure of protein crystals. The obtained electron density (at the usual resolution not better than 1.5 Å) is used to fit the residues

of the protein into, according to a previously known amino-acid sequence. Protons are not “visible” in the density because of the low scattering cross section, and the distinction between the oxygen and nitrogen atoms is usually not possible. Ambiguous side chain rotamers in the case of glutamine, asparagine or histidine side chains are thus distinguished based on the available hydrogen-bonding possibilities. In the case of BLUF, the X-ray structures provide key information on characterizing the hydrogen-bonding network of the flavin-binding pocket. Interestingly, two different crystal structures of the AppA-BLUF domain were independently solved (Anderson et al., 2005; Jung et al., 2006). The structures differ in the rotamer conformation of the Q63 side chain as well as in the fold of the β 5 strand and the position of the conserved W104 residue with respect to flavin as shown in Figure 4.1. In the crystal structure presented in (Anderson et al., 2005) the side chain of W104 points into the flavin-binding pocket (PDB 1YRX), thus it was dubbed W_{in} structure. In the crystal structure presented in (Jung et al., 2006) W104 is solvent exposed and M106 replaces W104 in the flavin pocket (PDB 2IYG), thus it was dubbed W_{out} structure. Critically, the Q63 side chain was placed with different rotamers into the obtained electron density in the two structures. In addition, the orientation of the hydroxyl group of Y21 was also affected. In the following, the Q63 rotamer in the W_{in} and W_{out} structures are referred to by Q63_A and Q63_J, respectively.

Importantly, as Figure 4.1B shows, the electron density defining the arrangement of the Y21, Q63 and N45 conserved residues with respect to the flavin chromophore are virtually identical in the two structures. However, because of the differently assigned glutamine rotamers, two distinct hydrogen-bonding networks around flavin are assumed in the two PDB structures, sketched in Figure 4.1C. The decisive difference for the BLUF mechanistic discussion is that the Q63_J rotamer forms a hydrogen bond to the flavin C4=O4 carbonyl, whereas the Q63_A rotamer does not. This hydrogen bond could in principle explain the red-shifted flavin absorption spectrum in the light state (Unno et al., 2006), thus most experimental (Gauden et al., 2006; Unno et al., 2006; Bonetti et al., 2008; Grinstead, Avila-Perez, et al., 2006; Stelling et al., 2007; Bonetti et al., 2009; Mathes, Zhu, et al., 2012) and some computational (Rieff et al., 2011; Meier et al., 2012; Hsiao et al., 2012) studies consider the W_{in} structure with Q63_A as the dark-state and the W_{out} structure with Q63_J as the light-state model of BLUF.

The assignment of the two X-ray structures to the functional states of BLUF led to a photoreaction hypothesis that explains the light-induced hydrogen-bond switch by glutamine rotation (Gauden et al., 2006; Unno et al., 2006; Jung et al., 2006). It is further assumed that glutamine rotation subsequently leads to a conformational switch of the β 5 strand, where the W104 and M106 residues exchange their position (Jung et al., 2006; Masuda et al., 2005; Dragnea et al., 2009). However, there is no direct experimental evidence that glutamine rotation would indeed be coupled to a W104-M106 conformational switch. Thus, it remains unclear how the hydrogen-bond switch leads to BLUF signaling.

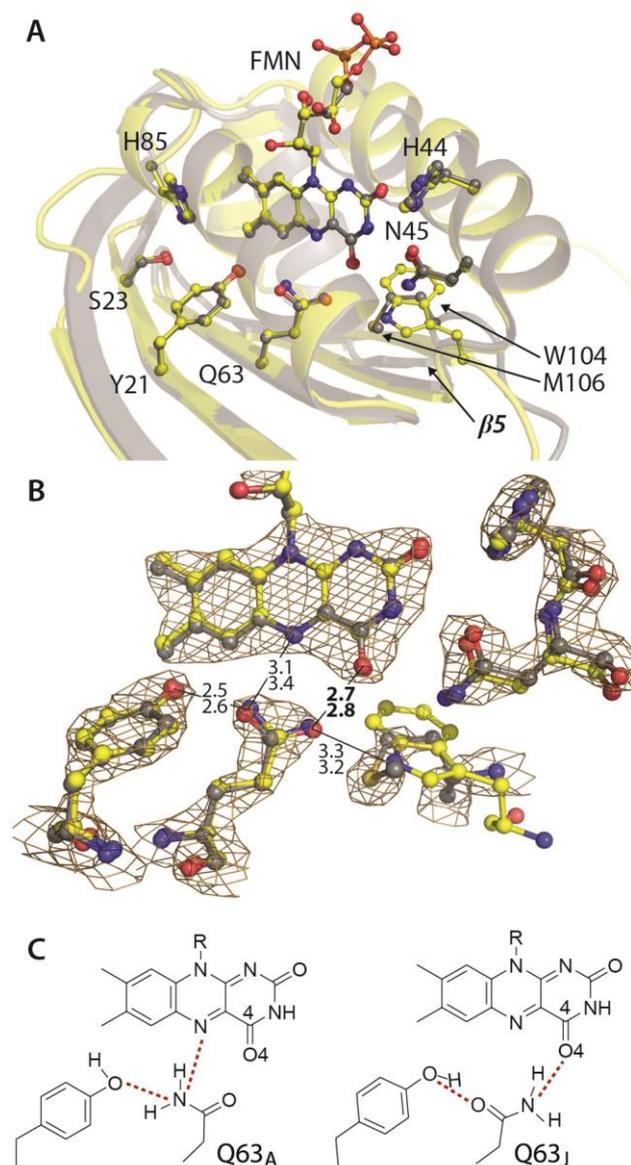


Figure 4.1. Comparison of the AppA-BLUF domain structures, W_{in} PDB structure 1YRX (yellow) and W_{out} PDB structure 2IYG (gray), both chain A. (A) Superposition of the backbone of the two structures in cartoon representation and the side chains of conserved residues around flavin in ball and stick representation. (B) The hydrogen bonding network of the flavin-binding pocket with key distances indicated in Å. The upper and lower values correspond to the 1YRX and 2IYG structures, respectively. The 2IYG experimental electron density contoured at two sigma overlays the two structures. (C) The two hydrogen bonding networks of the two glutamine rotamers Q63_A (W_{in}) and Q63_J (W_{out}). Figure was taken from (Udvarhelyi & Domratcheva, 2013).

The glutamine-rotation mechanism was challenged in several studies (Domratcheva et al., 2008; Obanayama et al., 2008; Khrenova et al., 2010), noting various inconsistencies in the W_{in} X-ray structure containing Q63_A. The major issue concerns the 2.7-Å distance between the flavin carbonyl O4 and the glutamine carbonyl Oe1 atoms being clearly too short to represent an equilibrium distance between the two carbonyl oxygen atoms. The corresponding distance in the W_{out} structure is 2.8 Å, where it describes a hydrogen bond between the flavin O4 and glutamine Ne2 atoms. To explain the short distance in the W_{in} crystal structure, the computational studies in (Domratcheva et al., 2008; Khrenova et al., 2010) suggested a tautomeric glutamine. Taking this into account, glutamine tautomerization was proposed as the photoactivation mechanism in BLUF (Domratcheva et al., 2008; Sadeghian et al., 2008).

To model the crystallographic Q63_A rotamer in the W_{in} structure with the short oxygen-oxygen distance as the dark state, other computational studies in (Rieff et al., 2011; Hsiao et al.,

2012) introduced restraints. However, the physical meaning of these restraints stabilizing the Q63_A rotamer is unclear. Another important question is why most crystal structures of BLUF proteins adopt the W_{out} conformation (Jung et al., 2005; Barends et al., 2009; Kita et al., 2005; Yuan et al., 2006), which according to the light-induced glutamine rotation hypothesis represents the light state. Furthermore, the hydrogen-bonding network postulated by the W_{in} structure is inconsistent with the IR study in (Takahashi et al., 2007) that established Y21 to be a hydrogen-bond donor in both the dark and light states. A recent IR study concluded that the hydrogen bond between Y21 and Q63 in the light state is unusually strong: The Y21-OH stretching frequency is 400-600 cm⁻¹ lower than the typical frequency of a hydrogen bond between a phenolic OH and a carbonyl group (Iwata et al., 2011). This finding is in fact inconsistent with the glutamine rotation mechanism that proposes the W_{out} structure as the light state, where the Y21 hydroxyl group forms a “normal” hydrogen bond with the Q63_J carbonyl group. All these observations suggest that the Q63_A rotamer in the W_{in} structure cannot represent the dark and the Q63_J rotamer in the W_{out} structure cannot represent the light state. Finally, it is questionable, from the theoretical viewpoint, how the energy barrier of glutamine rotation in the flavin-binding pocket is raised to be high enough to ensure the photosensitivity of BLUF.

To resolve the inconsistencies listed above, three criteria for physical models of BLUF are proposed here. First, both the dark and light state models must contain an atom arrangement that represents equilibrated molecular forces that correspond to minimum-energy structures. After optimization of the respective molecular models, at least the dark-state model must be consistent with the experimental X-ray electron density. Second, the light-state model must reproduce the 15-nm red-shifted flavin absorption compared to the dark state. Third, the transition from the dark to the light state must have a high energy barrier in the electronic ground state to ensure light sensitivity. All criteria must be satisfied simultaneously. In particular, it is not sufficient for physically meaningful models of BLUF to only reproduce the spectral red shifts.

In the following, using these three criteria, the crystallographic assignment of the Q63 side chain rotamer in the W_{in} structure is re-examined and the Q63_A and Q63_J models as proposed dark and light states are reappraised. This chapter complements previous computational studies addressing the Q63 orientation and its tautomeric forms in both the W_{in} and W_{out} structures (Khrenova et al., 2010, 2011, 2013). First, the W_{in} structure is investigated, in light of the conformation of the Q63 side chain. The optimized model structures are compared directly to the experimental X-ray electron density. Clarifying the Q63 rotamer assignment in the W_{in} structure is of highest importance since it directly influences the interpretation of all spectroscopy data. Second, extensive Q63 dihedral-angle energy scans characterize the energy barriers for glutamine rotation in both the amide and imid forms in both the W_{in} and W_{out} structures. Along these lines, dynamics aspects are discussed.

4.2 Computational details

4.2.1 W_{in} cluster models and geometry optimization

Supramolecular cluster models with a complete quantum-mechanical description of the hydrogen-bonding network around flavin and a substantial part of the beta sheet (parts of the $\beta 1$, $\beta 3$ and $\beta 5$ strands) are used. The AppA BLUF W_{in} model, built based on the PDB coordinates 1YRX chain A, includes lumiflavin, the residues Y21, Q63, N45, L65, W104, and A46 as well as the backbone atoms of residues C20, R22, F62, W64, E66, H105, G103, H44 (AppA numbering); in total 184 atoms, shown in Figure 4.2A. To account for the ambiguous assignment of the hydrogen atom positions in the crystal structures, three forms of the glutamine side chain were assumed: the two rotamers $Q63_A$ and $Q63_J$ as well as the tautomeric form, $Q63_{tau}$. In addition, the $Q63_A$ rotamer in the W_{in} molecule of the PixD BLUF protein was also considered because this structure was determined with the highest resolution among the BLUF X-ray structures (Yuan et al., 2006). The PixD model, built based on the PDB coordinates 2HFN chain D, consists of lumiflavin, residues Y11, Q53, V54, L55, N35, P36, A37, N38, and W94 as well as the backbone atoms of residues I10, S12, L52, E56, G39, N34, S95, V93 and the water molecule WAT1005 (PixD numbering); in total 225 atoms, shown in Figure 4.2B. In the PixD model only the crystallographic $Q53_A$ rotamer was considered. The computer program HyperChem (HyperCube, Inc., Gainesville, FL, USA) was used to add the hydrogen atoms and to optimize their coordinates in the starting geometries with the Amber force field, keeping the PDB coordinates of the heavy atoms fixed. In the following, the standard PDB names for residue atoms are used.

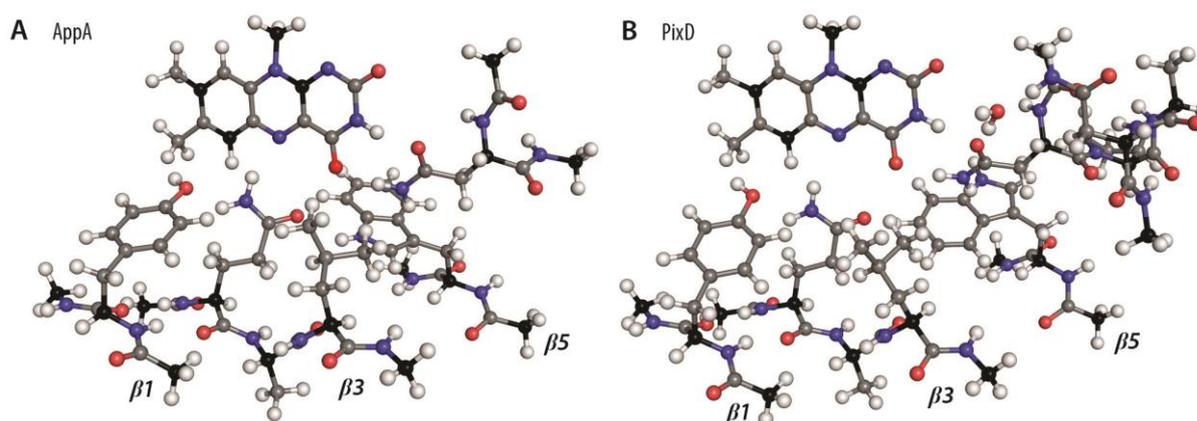


Figure 4.2. Supermolecular cluster models of BLUF. (A) W_{in} AppA model based on the PDB structure 1YRX, chain A. (B) W_{in} PixD model based on the PDB structure 2HFN, chain D. The black-colored atoms (all $C\alpha$ and four flavin atoms) were constrained during geometry optimization. Figure was taken from (Udvarhelyi & Domratcheva, 2013).

To characterize the hydrogen bonding network in BLUF, the geometries of the prepared starting models were optimized with B3LYP/6-31G**. To mimic the constraints due to the protein scaffold, the Cartesian coordinates of the C α carbon atoms, the water oxygen atom in the PixD model and the flavin carbon atoms C6, C8, C1' and C10a were constrained to their crystallographic values (black-colored atoms in Figure 4.2). A three step optimization procedure was used to conserve the crystallographic distance between the glutamine O ϵ 1/Ne2 and flavin O4 atoms as much as possible. First, starting from the X-ray coordinates, the geometry was optimized with constrained PDB coordinates of the glutamine O ϵ 1/Ne2 and flavin O4 atoms to conserve the glutamine-flavin interactions postulated by the PDB structures (opt1). Then, the optimization was continued with the released glutamine and the constrained PDB coordinates of the flavin O4 atom (opt2). Finally, the geometry optimization was completed without the additional flavin–glutamine constraints (opt3). The geometries were optimized to an energy gradient of 0.0001 hartree/bohr in all three steps. The finally obtained models (opt3) are regarded as models for the flavin hydrogen-bonding network in the W_{in} BLUF conformation. The opt3 optimized geometries are compared with the experimental electron densities computed with the ARP/wARP program (Perrakis et al., 1999) using the structure factors of the respective PDB models by performing 25 cycles of “map improvement by atoms update and refinement”. A Perl script was written to evaluate the heavy-atom RMSD values with respect to the crystallographic coordinates along the three-step optimization procedure.

4.2.2 Extended cluster models: ext-W_{in} and ext-W_{out}

In addition to the AppA-W_{in} and PixD-W_{in} models presented above, extended models of both the W_{in} and W_{out} conformations of AppA BLUF were considered. These extended models include also parts of the β 4 strand, and the α 1 and α 2 helices in addition to the residues of the AppA-W_{in} model above. The ext-W_{in} model, built based on the PDB coordinates 1YRX chain A (Anderson et al., 2005), includes lumiflavin and the first part of the ribityl chain, the residues Y21, I37, S41, N45, A46, G52, A53, L54, Q63, L65, I79, and W104 as well as the backbone atoms of residues C20, R22, D36, V38, T40, Q42, A43, H44, T51, F55, F62, W64, E66, H78, Q80, G103, H105, M106; in total 346 atoms, shown in Figure 4.3A. In this model, the glutamine side chain was considered in three forms: Q63_A, Q63_J, and Q63_{tau}. The ext-W_{out} model, built based on the PDB coordinates 2IYG chain A (Jung et al., 2006), includes the same residues and backbone atoms as the ext-W_{in} model, except for the β 5 strand, where the ext-W_{out} model contains M106 and the backbone atoms of residues W104, H105, Q107, and L108. Thus, the ext-W_{out} model consists of 349 atoms in total, shown in Figure 4.3B. In this model, the glutamine side chain was considered in the Q63_J and Q63_{tau} forms.

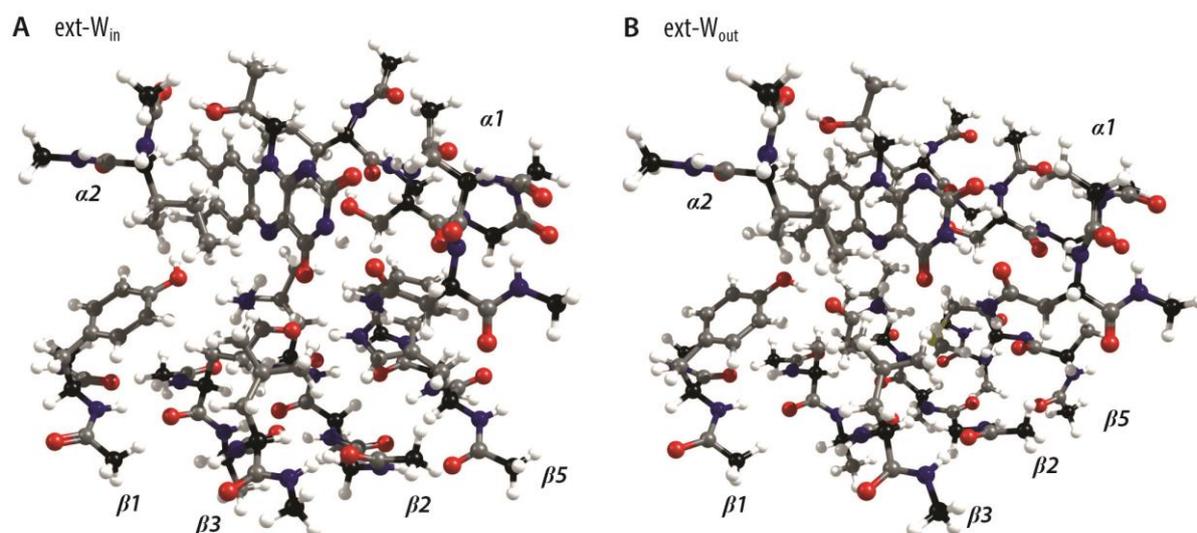


Figure 4.3. Extended supermolecular cluster models of AppA BLUF. (A) W_{in} model based on the PDB structure 1YRX, chain A. (B) W_{out} model based on the PDB structure 2IYG, chain A. The black-colored atoms (all C_{α} and one flavin atom) were constrained during geometry optimization.

Altogether five extended models (three in the W_{in} conformation and two in the W_{out} conformation) were subjected to geometry optimization with B3LYP/cc-pVDZ. The C_{α} atoms and one flavin atom were kept fixed to their crystallographic position, indicated by the black-colored atoms in Figure 4.3, and no additional constraints on the Q63 atoms were imposed. The gradient threshold was selected as 0.0003 hartree/bohr. The finally obtained ext- W_{in} and ext- W_{out} models are the largest equilibrium-geometry models of BLUF optimized at the density-functional level (3430 and 3449 basis functions were used, respectively).

4.2.3 Dihedral-angle energy scans and excitation spectra

To estimate the energy barriers between different glutamine rotamers, relaxed energy scans along the glutamine and glutamine-flavin dihedral angle coordinates were computed. As the scans were computed not in Cartesian but in the delocalized internal coordinates (Granovsky, 2013), protein scaffold constraints were applied as constraints of distances between all pairs of atoms constrained in the previous geometry optimization procedure (black-colored atoms in Figure 4.2 and Figure 4.3). To produce scans every 10 degrees of the chosen dihedral angle, geometry optimization was carried out until an energy gradient of 0.0003 hartree/bohr with a fixed value of the dihedral angle. Complete 360 degree scans in both clock- and counter clockwise directions were obtained and analyzed. The energy scans were computed with the B3LYP method, with the 6-31G** and cc-pVDZ basis sets in the W_{in} models and the extended models (W_{in} and W_{out}), respectively.

At the optimized structures, and at the identified local minima in the dihedral-angle energy scans, the excitation spectrum was computed with the TD-B3LYP/cc-pVDZ method for the ten lowest lying excited states. The HOMO-LUMO transition of the flavin is of interest in this chapter as that corresponds to the absorption maximum around 450 nm (2.76 eV). Among the ten computed states, the state corresponding to the HOMO-LUMO flavin excitation was identified according to the contributions of the respective MOs. This state is also the lowest energy excited state characterized by a considerable oscillator strength. The excitation energy of this state can be considered as a measure for the position of the flavin absorbance maximum. The computer program Firefly (ver 7.1, 8; Granovsky) which is partially based on the GAMESS US source code (Schmidt et al., 1993) was used for all quantum-chemical calculations presented in this chapter.

4.3 Results and Discussion

4.3.1 The Q63 rotamer in the W_{in} structure

The conservative three-step geometry-optimization procedure employed here assures finding the local-energy-minimum structure closest to the initial AppA and PixD W_{in} PDB structures (the final opt3 models). In the optimized models containing Q63_j or Q63_{tau} (i.e. opt3-Q63_j and opt3-Q63_{tau} models), the glutamine side chain is close to its crystallographic position, whereas in the opt3-Q63_A model both Y21 and Q63 side chains move away from the crystal coordinates, as demonstrated in Figure 4.4 and Table 4.1. During geometry optimization, the largest geometry changes occur between the opt1-Q63_A and opt2-Q63_A models. The final heavy-atom RMSD values of the fully optimized Q63_j, Q63_{tau} and Q63_A models are 0.32 Å, 0.29 Å, and 0.53 Å, respectively. The Q63 carbonyl oxygen and Y21 hydroxyl oxygen atoms in Q63_A undergo the largest displacements. The distance between the glutamine Oε1 and flavin O4 atoms increases from 2.7 Å in the PDB structure to 3.7 Å in the optimized AppA model, and from 2.7 Å to 4.1 Å in the PixD model. The Q63_A side chain no longer stays parallel to flavin, but is rather perpendicular. In Figure 4.5 the optimized structures are overlaid with the experimental electron density, which demonstrates how the Q63_A side chain, in contrast to Q63_j or Q63_{tau}, rotates out of the experimental electron density.

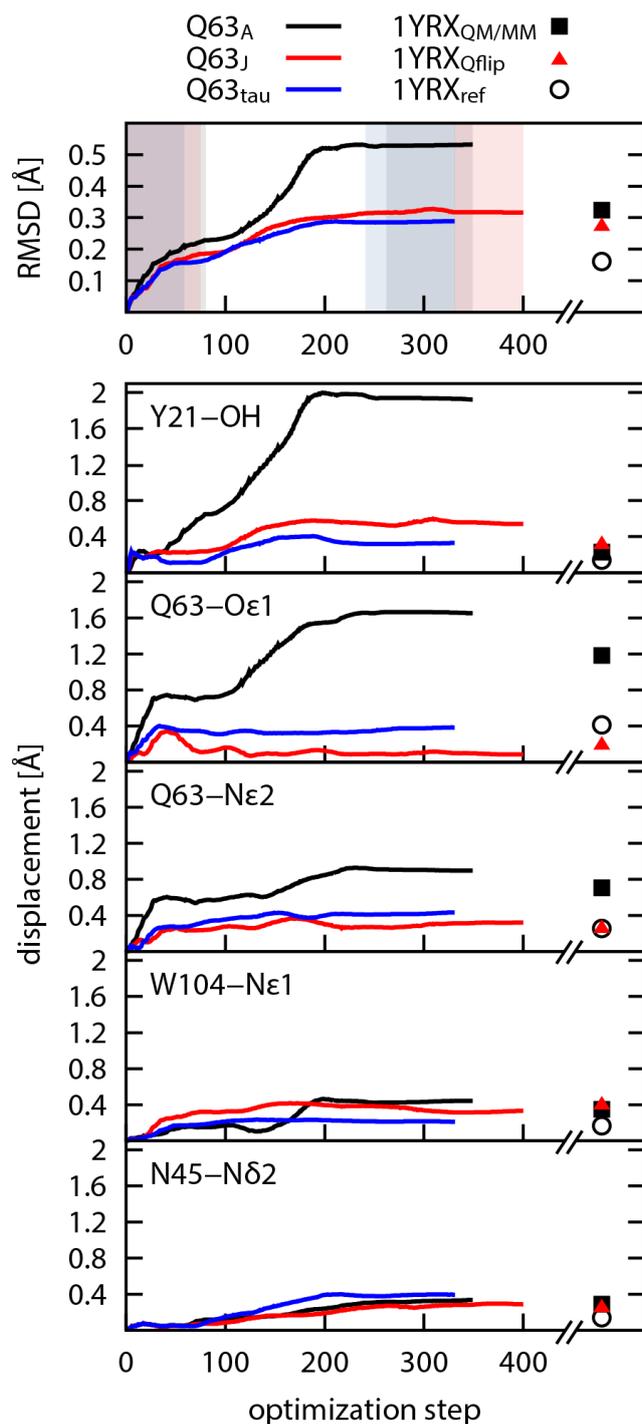


Figure 4.4. Heavy atom RMSD values and atom-specific displacements of the AppA cluster models Q63_A (black), Q63_J (red) and Q63_{tau} (blue) as the function of the geometry optimization step, determined with respect to the starting 1YRX crystal structure. The colored background rectangles in the RMSD plot indicate the three-step optimization procedure of the corresponding models: gray – Q63_A model (opt1 and opt3), light red – Q63_J (opt1 and opt3) and light blue – Q63_{tau} (opt1 and opt3). As a comparison, the corresponding RMSDs and displacement values for the comparable QM/MM models in (Hsiao et al., 2012) are also indicated: 1YRX_{QM/MM} corresponds to the QM/MM-optimized model with the Q63_A rotamer (opt3-Q63_A model), 1YRX_{Qflip} – to the 180 degree rotated glutamine (opt3-Q63_J model) and 1YRX_{ref} – to the “quantum-refined” model (opt1-Q63_A). Figure was taken from (Udvarhelyi & Domratcheva, 2013).

Table 4.1. Four selected distances in Å involving glutamine heteroatoms of the optimized AppA models, along the three-step optimization procedure: opt1 (constrained glutamine O ϵ 1 or N ϵ 2 to flavin O4 distance), opt2 (constrained flavin O4 atom) and opt3 (fully optimized without flavin-Q63 constraints). Table was taken from (Udvarhelyi & Domratcheva, 2013).

Q63_A model	opt1	opt2	opt3
Tyr-OH to Q63-N ϵ 2	3.0	3.1	3.1
Q63-N ϵ 2 to fl-N5	3.2	3.0	3.0
Q63-Oϵ1 to fl-O4	2.7	3.7	3.6
Q63-O ϵ 1 to W104-N ϵ 1	3.1	2.8	2.8
Q63_J model	opt1	opt2	opt3
Tyr-OH to Q63-O ϵ 1	2.8	2.8	2.8
Q63-O ϵ 1 to fl-N5	3.1	3.1	3.1
Q63-Nϵ2 to fl-O4	2.7	2.8	2.8
Q63-N ϵ 2 to W104-N ϵ 1	3.4	3.5	3.4
Q63_{tau} model	opt1	opt2	opt3
Tyr-OH to Q63-N ϵ 2	-	2.8	2.8
Q63-N ϵ 2 to fl-N5	-	3.1	3.1
Q63-Oϵ1 to fl-O4	-	2.8	2.7
Q63-O ϵ 1 to W104-N ϵ 1	-	2.9	2.9

To keep the Q63_A rotamer in the experimental density, i.e. to preserve the unphysical oxygen-oxygen distance that is inconsistent with the interactions of the two carbonyl groups, the position of the two oxygen atoms must be constrained. As a consequence of the distance constraint, the energy of opt1-Q63_A is 9.5 kcal/mol higher than that of opt3-Q63_A, demonstrating the repulsive oxygen-oxygen interactions. In contrast, the energy of the opt1-Q63_J model is almost identical to the opt3-Q63_J model, indicating that the intermolecular forces between Q63 and flavin are equilibrated already in the starting structure. The same is true for Q63_{tau}. Thus, in the fully optimized opt3 models, the Q63_A rotamer is incompatible with the experimental electron density in contrast to the Q63_J and Q63_{tau} models. Therefore either the Q63_J or Q63_{tau} form but not the Q63_A rotamer should be used to model the X-ray density of the W_{in} structure. From Figure 4.5 it must be concluded that Q63_A does not satisfy the first criterion, but Q63_J and Q63_{tau} do. With the Q63_J rotamer, the crystallographic W_{in} and W_{out} hydrogen bonding networks would be the same (differing only in the W104 position) and with the Q63_{tau} form they would differ because of the chemically distinct Q63_{tau}. It is emphasized that in either case the glutamine rotation mechanism cannot be proposed on the basis of the BLUF X-ray structures.

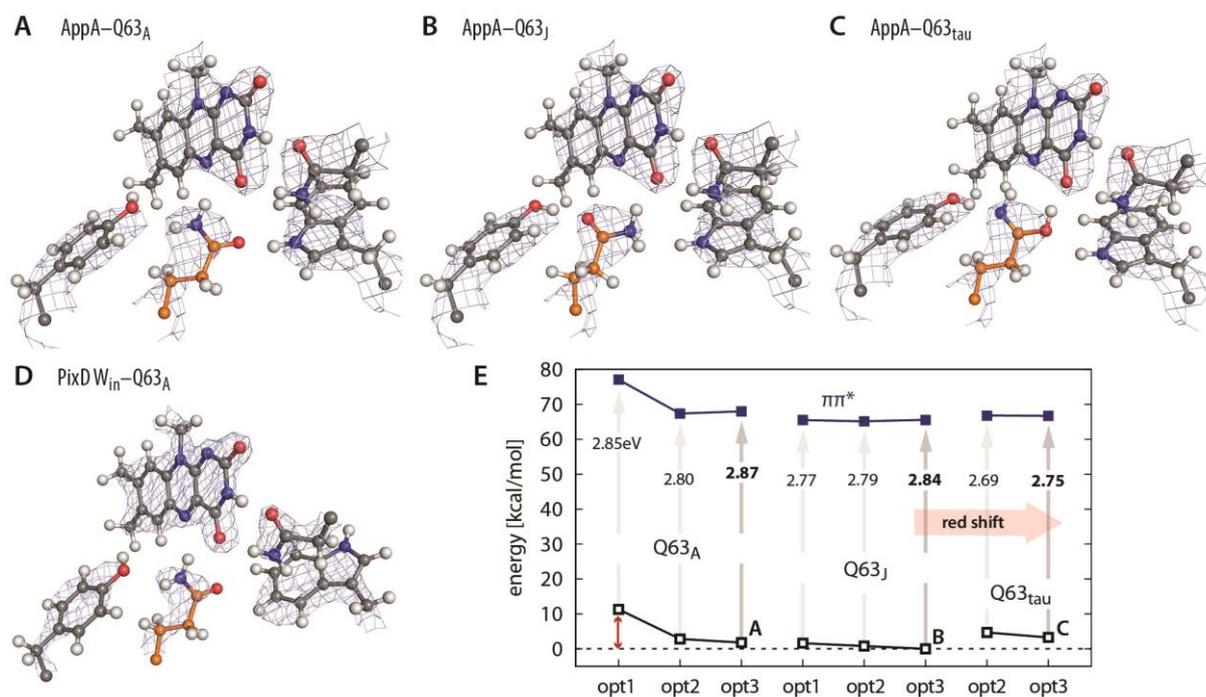


Figure 4.5. (A-C) Part of the optimized W_{in} AppA-BLUF cluster models overlaid with the experimental X-ray electron density of the PDB structure 1YRX, chain A, contoured at two sigma. The carbon atoms of the glutamine side chain are colored orange. The full cluster models are shown in Figure 4.2. (D) W_{in} PixD-BLUF cluster model overlaid with the experimental X-ray electron density of the PDB structure 2HFN, chain D, contoured at three sigma. (E) Relative ground state energies (black empty squares) of the AppA-BLUF models optimized with and without constraints computed with respect to the energy of the fully optimized opt3-Q63_J model, and flavin excitation energies (blue filled squares). The small red arrow indicates the high energy of the crystallographic Q63_A rotamer. Figure was taken from (Udvarhelyi & Domratcheva, 2013).

Figure 4.5E presents the flavin excitation energy of the identified energy-minimum structures in the AppA- W_{in} models with and without constraints. In the case of the constrained opt1-models, the excitation energy in the Q63_A model is significantly blue-shifted with respect to the Q63_J model. After releasing the distance constraint (opt2), the difference in the predicted absorption almost vanishes. In the final opt3 models, a small difference of 0.03 eV (4.5 nm) is obtained. These calculations demonstrate that there is a caveat when working with constrained or restrained models since constraints/restraints influence the absorption maximum and may introduce artifacts. The formulated physical criteria related to the protein photoactivation function help to minimize artifacts of computational models. According to the first criterion, only the fully optimized opt3 Q63_A, Q63_J and Q63_{tau} models, which correspond to the equilibrated structures, should be considered as BLUF models. Only opt3-Q63_J and opt3-Q63_{tau} consistent with the electron density can represent the dark state. The flavin absorption energy in the opt3-Q63_{tau} model is 2.75 eV and 14 nm red-shifted with respect to that of the Q63_J model. Thus only Q63_{tau} reproduces the absorption properties of the light state.

4.3.2 Extended cluster and QM/MM models

The question naturally arises how well the W_{in} cluster models from the previous section can represent BLUF states since parts of the protein and the solvent are missing from the models. It is expected that the observed structural relaxation is well-described by these models since the important geometry changes pertain to Y21 and Q63 because of the local interactions and not because of the extended protein environment. To further elucidate how much the protein environment influences the relaxation of Y21 and Q63, the models were extended to include more residues, referred to as the ext- W_{in} models, as shown in Figure 4.3. In addition, the W_{out} structure was also modeled with an extended environment of Q63 and compared to the results obtained with the ext- W_{in} models.

Figure 4.6 shows the five optimized extended models, overlaid with the respective experimental electron density. Concerning the W_{in} structure, the same conclusions are drawn on Q63, as in the previous section with the smaller W_{in} models: The $Q63_A$ rotamer is incompatible with the experimental electron density, whereas $Q63_j$ and $Q63_{tau}$ can both be used to model the electron density of the 1YRX structure. Comparing the relative energies of the optimized structures, interestingly, one finds that the ext- W_{in} - $Q63_{tau}$ equilibrium structure lies 2.3 kcal/mol below the ext- W_{in} - $Q63_A$ minimum, indicating that the $Q63_A$ rotamer in the BLUF pocket is indeed energetically so unfavorable that even the chemically modified tautomeric form of glutamine is more stabilized by the BLUF environment. Concerning the W_{out} structure, both optimized ext- W_{out} - $Q63_j$ and ext- W_{out} - $Q63_{tau}$ minima are compatible with the experimental electron density and the relevant interatomic distances are close to those in the X-ray structure 2IYG. The ext- W_{out} - $Q63_{tau}$ geometry lies 14.7 kcal/mol higher in energy than that of ext- W_{out} - $Q63_j$. Thus the energy difference between the imid and amide forms of glutamine is significantly increased in the W_{out} structure compared to the W_{in} structure.

The S_1 excitation energies computed in all five extended models are also indicated in Figure 4.6. In the models containing the amide Q63, the S_1 excitation energy varies between 2.91-2.94 eV and containing the imid Q63 it is 2.81 and 2.83 eV. Thus, irrespective of the methionine and tryptophan position, the glutamine chemical structure defines the position of the S_1 transition: the tautomeric glutamine-containing models feature a significantly red-shifted S_1 excitation energy compared to those computed containing the amide glutamine. Thus, similar to the findings in the smaller W_{in} models of the previous section, it is emphasized that there is no significant excitation-energy shift between the equilibrated ext- W_{in} models containing either the $Q63_A$ or $Q63_j$ rotamer. However, the environment (tryptophan or methionine) plays a role in the stabilization of the tautomeric glutamine form relative to the amide form. The different relative energies suggest that the imid Q63 is more favored in the W_{in} environment compared to the W_{out} environment, therefore, W_{in} with $Q63_{tau}$ is suggested here to represent the light state and W_{out} with $Q63_j$ the dark state of BLUF.

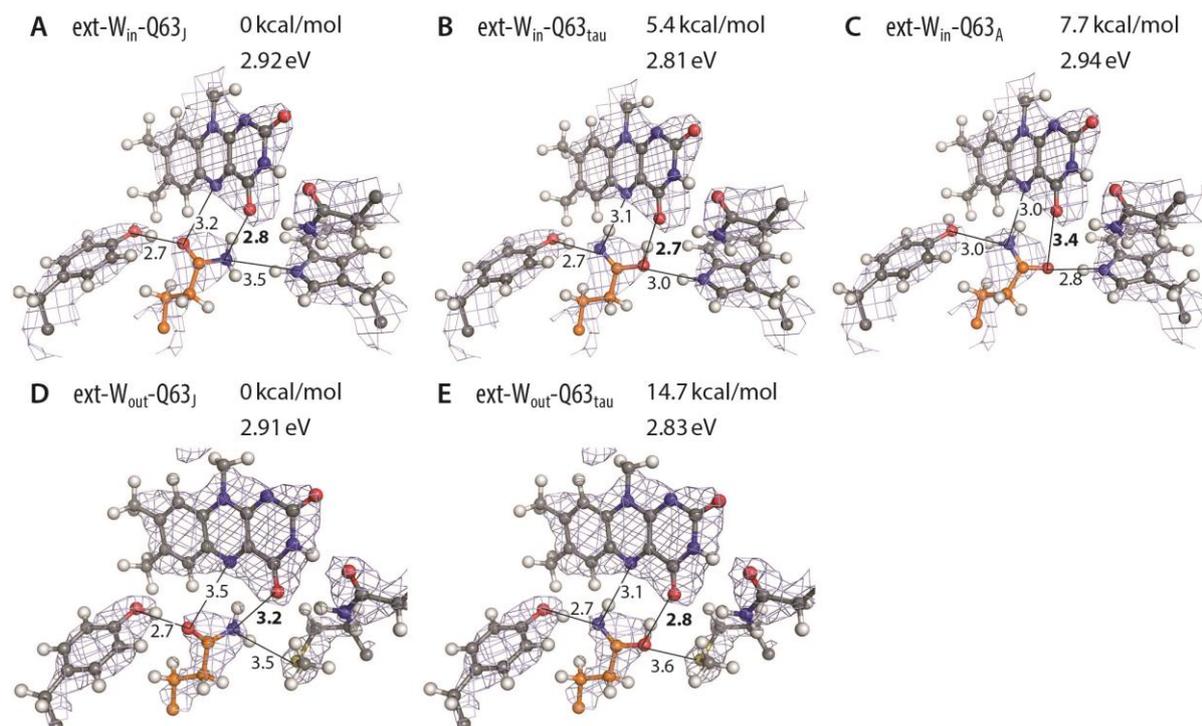


Figure 4.6. Optimized structures of the extended models. (A-C) Part of the optimized ext- W_{in} AppA-BLUF cluster models with the three forms of Q63 overlaid with the experimental X-ray electron density of the PDB structure 1YRX, chain A, contoured at two sigma. The full ext- W_{in} cluster model is shown in Figure 4.3A. (D-E) Part of the optimized ext- W_{out} AppA-BLUF cluster models with the two forms of Q63 overlaid with the experimental X-ray electron density of the PDB structure 2IYG, chain A, contoured at two sigma. The full ext- W_{out} cluster model is shown in Figure 4.3B. Selected interatomic distances are given in Å. The relative energies with respect to the Q63_j-containing model are indicated in kcal/mol. The excitation energies in eV define the position of the flavin absorption maximum.

In the following, the QM/MM study presented in (Hsiao et al., 2012) is discussed in view of the Q63_A rotamer relaxation in the W_{in} structure and compared to the results of the cluster models presented here. Hsiao et al. reported “quantum-refined” AppA BLUF crystal structures obtained by QM/MM calculations with crystallographic restraints and compared these to fully-optimized QM/MM models. The authors considered models with both the Q63_A and Q63_j rotamers and computed their excitation energies. The model with the Q63_A rotamer was found to have a blue-shifted flavin absorption with respect to the 180 degree rotated glutamine rotamer, Q63_j. Following this finding, Hsiao et al. assigned the Q63_A and Q63_j rotamers to the dark and light states respectively, thereby supporting the glutamine rotation as the light-activation mechanism of BLUF. To identify the origin of the apparent contradiction in their work and the presented results, the QM part of three of their QM/MM models is analyzed here according to the formulated three criteria of the BLUF dark and light state models. The three models are (i) the “quantum-refined” 1YRX_{ref} model (comparable to the opt1-Q63_A model); (ii) 1YRX_{QM/MM} structure without restraints (equivalent to the opt3-Q63_A model) and (iii) the 1YRX_{Qflip} structure

with the Q63_J rotamer (equivalent to the opt3-Q63_J model). Figure 4.7 shows the superposition of the QM part of the QM/MM-optimized models by Hsiao et al. with the 1YRX electron density. A picture fully consistent with the results of the cluster models is obtained: The Q63_A rotamer is only compatible with the electron density when the geometry is restrained but not when its geometry is fully equilibrated. The energy difference between the restrained 1YRX_{ref} and fully equilibrated 1YRX_{QM/MM} models is 15 kcal/mol because of the restraints, which is even higher than in the case of the cluster models. The distance after the unrestrained QM/MM optimization increases to 3.6 Å (because of the repulsion between the two carbonyl oxygen atoms), which is very similar to the geometry change in the cluster models. This repulsion by no means represents a “weak hydrogen bond” as the authors refer to it in (Hsiao et al., 2012). Just as in the case of the cluster models, the Q63_A side chain undergoes the most significant geometry change upon relaxation in contrast to its 180 degree rotated counterpart, Q63_J. The latter is compatible with the experimental electron density. Unfortunately, Hsiao et al. did not perform constrained optimization (“quantum-refinement”) of the 1YRX_{Qnip} model containing Q63_J to further confirm the incorrect assignment of the Q63_A orientation in the 1YRX crystal structure.

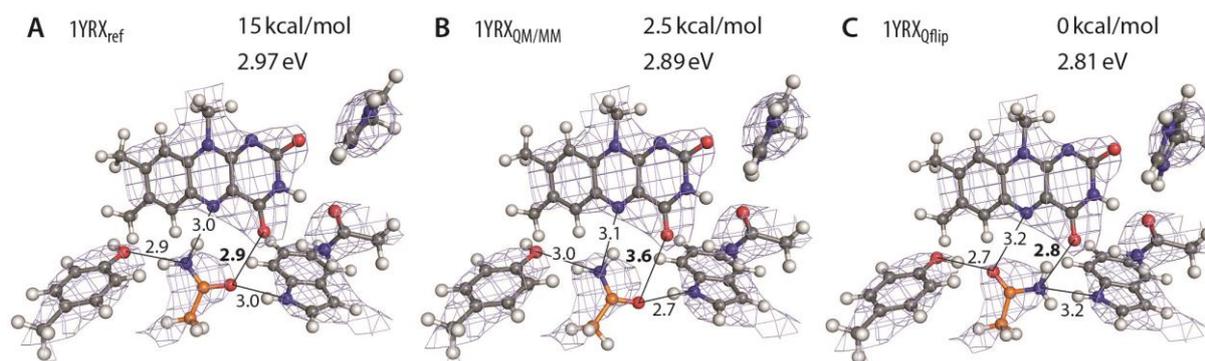


Figure 4.7. The QM part of the QM/MM-optimized geometries by Hsiao et al. (Hsiao et al., 2012) superposed with the arp/wARP-refined electron density of the PDB structure 1YRX, contoured at two sigma. Selected interatomic distances are given in Å. The relative energies with respect to the 1YRX_{Qnip} model are indicated in kcal/mol. The excitation energies in eV define the position of the flavin absorption maximum. Figure was taken from (Udvarhelyi & Domratcheva, 2013).

Hsiao et al. reported on the major structural rearrangement of Q63_A, nonetheless, on the basis of the small overall RMSD values of 0.3 Å (and the blue-shifted flavin absorption) they assigned the restrained optimized 1YRX_{ref} model with Q63_A to the BLUF dark state. However, overall RMSD values are not suitable to judge the agreement of a specific side chain conformation with the experimental electron density. Instead, atom-specific displacements should be looked at, like those displayed in Figure 4.4, from which it is apparent that in both the cluster and the QM/MM models of (Hsiao et al., 2012), Q63_A deviates significantly from the starting PDB structure in contrast to the Q63_J rotamer. Most importantly, the Q63_A-Oε1 atom

displacement in 1YRX_{ref} is larger than the displacement of the Nε2 atom at the same position in 1YRX_{Qflip} (Q63_J model), thus clearly demonstrating that Q63_J is preferred over Q63_A. Unfortunately, Hsiao et al. did not consider a tautomeric form of Q63. They noted however, that the bond distances in the glutamine side chain in all of their models were typical for amides and incompatible with an imid form of glutamine, thus arguing against the latter. However, this result is an obvious consequence of their choice to add hydrogen atoms assuming only the amide glutamine because geometry optimization always results in an energy-minimum structure consistent with the selected protonation pattern. Therefore, the results presented in (Hsiao et al., 2012) cannot be used to evaluate the tautomeric form of Q63 in BLUF.

4.3.3 Dynamic aspects: glutamine rotation in the dark state

As demonstrated above, the energy difference between all fully optimized models containing Q63_J and Q63_A is rather small. To estimate the energy barrier of glutamine rotation, energy scans along different glutamine dihedral angles can be used. Two types of scans were computed: (i) changing the glutamine Cβ-Cγ-Cδ-Oε1 dihedral angle to rotate the amide group and (ii) changing the glutamine Cδ=Oε1 flavin C4=O4 dihedral to control hydrogen-bonded interactions of flavin and glutamine. Both energy scans revealed several minima with different Q63 rotamers that are close in energy and separated by low-energy barriers. The results for the smaller W_{in} models are shown in Figure 4.8.

The Q63 dihedral angle scan starting from the opt3-Q63_A model (minimum I) in both the clockwise (cw) and counterclockwise (ccw) directions revealed two other minima lower in energy than Q63_A, separated by low-energy barriers, shown in Figure 4.8A-C. The ccw scan starting from the opt3-Q63_J model (minimum V) also revealed two more minima, but with higher energy, shown in Figure 4.8D-F. In this case the full 360 degree scans cannot be used to analyze the Q63 rotamers in BLUF as a model artifact occurs in the cw direction after 7 steps: the scan arrives at a minimum where W104 forms a hydrogen bond with the backbone carbonyl of W64. Along the ccw scan, a similar hydrogen bond is formed after 18 steps. Since the hydrogen bond between W104 and the backbone carbonyl most likely cannot form in BLUF (the β2 strand preventing such interactions is missing from this model), the respective minima have to be excluded from further analysis and consistently, Figure 4.8 only shows minima relevant for BLUF. Importantly, none of these minima feature a short distance between the glutamine and flavin carbonyl groups, as in the crystal structure 1YRX. The glutamine Cδ=Oε1 flavin C4=O4 dihedral angle scan starting from the opt3-Q63_A model revealed a low-energy barrier between the Q63_A and Q63_J rotamers, shown in Figure 4.8G-J. In this case again, only parts of the scan can be used because of the W104-backbone hydrogen bond formation. However, the part of the energy scan meaningful for BLUF displays a low-energy path between the Q63_A and Q63_J (minimum IV)

rotamers, which have almost equal energies. Minimum IV differs from the opt3-Q63_j model by the orientation of the Y21 hydroxyl group because the scan was started from the opt3-Q63_A model. The orientation of the Y21 hydroxyl group with respect to Q63 represents another degree of freedom, which is interesting to consider in future studies. In fact, when changing a given dihedral angle, all the dihedral angles of glutamine and not only those of the head group change, indicating the coupling of various torsional degrees of freedom.

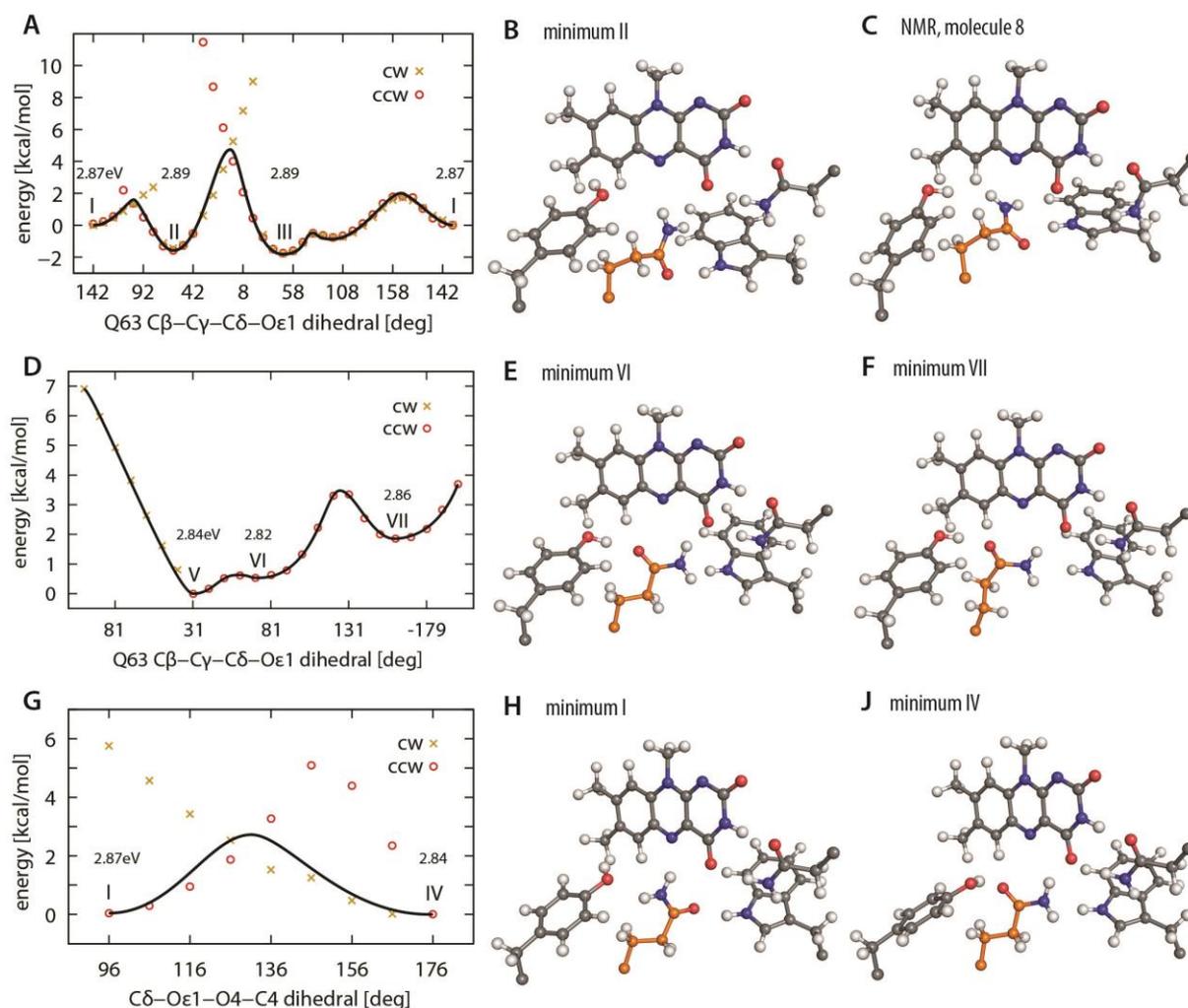


Figure 4.8. (A), (D) and (G) energy scans of the Q63 dihedral angle and the Q63-flavin dihedral angle, starting from the Q63_A model (minimum I) or the Q63_j model (minimum V) in clockwise (cw) and counter-clockwise (ccw) directions in 10 degree steps. The black lines represent the dihedral angle potential energy that connects respective minima. The S_1 excitation energies for the minimum-energy structures are indicated in eV. (B) and (C) The structure of the flavin binding pocket at the energy minimum II and of molecule 8 of the NMR ensemble (PDB 2BUN), respectively. (E), (F), (H) and (J) The structure of the flavin binding pocket at the indicated energy minima. Figure was taken from (Udvarhelyi & Domratcheva, 2013).

Figure 4.8 also indicates the flavin S_1 excitation energy at the identified local minima. The differences of 0.02–0.04 eV (3–6 nm) are small. All of these minima together should be regarded as the dark state model of BLUF. This picture is consistent with the results of a recent MD study (Khrenova et al., 2010), that identified four conformational substates with different Q63 rotamers in the BLUF W_{out} structure with the flavin excitation energy varying by 0.03 eV. In contrast, the excitation energy of the opt3-Q63_{tau} model is significantly red-shifted by 0.09–0.12 eV compared to the amide forms. The energy barrier of glutamine tautomerization is certainly much higher than the energy of thermal fluctuations because it requires the rearrangement of covalent bonds. In contrast to the rotation energy barrier, the tautomerization energy barrier ensures photosensitivity. Rotations of both the amide Q63 and imid Q63_{tau} side chains are implicated in protein dynamics of the dark and light states, respectively. The next section elaborates further on them in the extended models.

4.3.4 Dihedral-energy scans in the extended models: dark and light states

The dihedral-energy scans were also carried out in the extended models, both in the W_{in} and W_{out} conformations and with Q63 in both the amide and imid forms. In the extended models, the additional $\beta 2$ strand prevented the formation of the unphysical hydrogen bond between W104 and the backbone carbonyl atom encountered above. However, also in the extended models, parts of the energy scans had to be disregarded because of major conformational changes involving the Y21 residue that are unlikely to occur in the protein as parts of the protein environment around Y21 is missing in these extended models (e.g. $\beta 4$). Nevertheless, by merging the suitable parts of the energy scans in the cw and ccw directions, full 360-degree energy scans and dihedral-angle potential energy curves were obtained revealing several local minima, as presented in Figure 4.9. Figure 4.9 shows the results from the energy scans of changing the glutamine $C\delta=O\epsilon 1$ flavin $C4=O4$ dihedral angle. The energy scans of changing the glutamine $C\beta-C\gamma-C\delta-O\epsilon 1$ dihedral angle revealed very similar minima as those shown in Figure 4.8, and are therefore not discussed further.

Consistent with the findings in the previous section, the energy scan starting from the ext- W_{in} -Q63_A model (minimum VIII) revealed two local minima (IX and X), substantially stabilized in energy and separated by only a few kcal/mol energy barrier from minimum VIII. Minimum X, containing the rotated glutamine Q63_J, is 5.7 kcal/mol lower in energy than minimum VIII containing Q63_A. In contrast, the energy scan starting from the ext- W_{out} -Q63_J model (minimum XI) revealed a high-energy barrier for glutamine rotation, which is attributed to the relatively strong hydrogen bond between the Y21 hydroxyl and Q63 carbonyl oxygen atoms. In the energy scan starting from the ext- W_{in} -Q63_J model (minimum XII) the energy barrier for glutamine rotation is not that high and a Q63_A-type local minimum is identified (minimum XIII), 5

kcal/mol higher in energy than the structure with Q63_J. This finding is fully consistent with the results in the energy scan starting from minimum VIII. Apparently, in the W_{in} structure the amide Q63 is more likely to rotate inside the flavin-binding pocket than in the W_{out} structure. This characteristic could be relevant for the thermal recovery of the dark-state, when the protein, following the recovery of the amide state of glutamine, returns from the W_{in} light state to the W_{out} dark state.

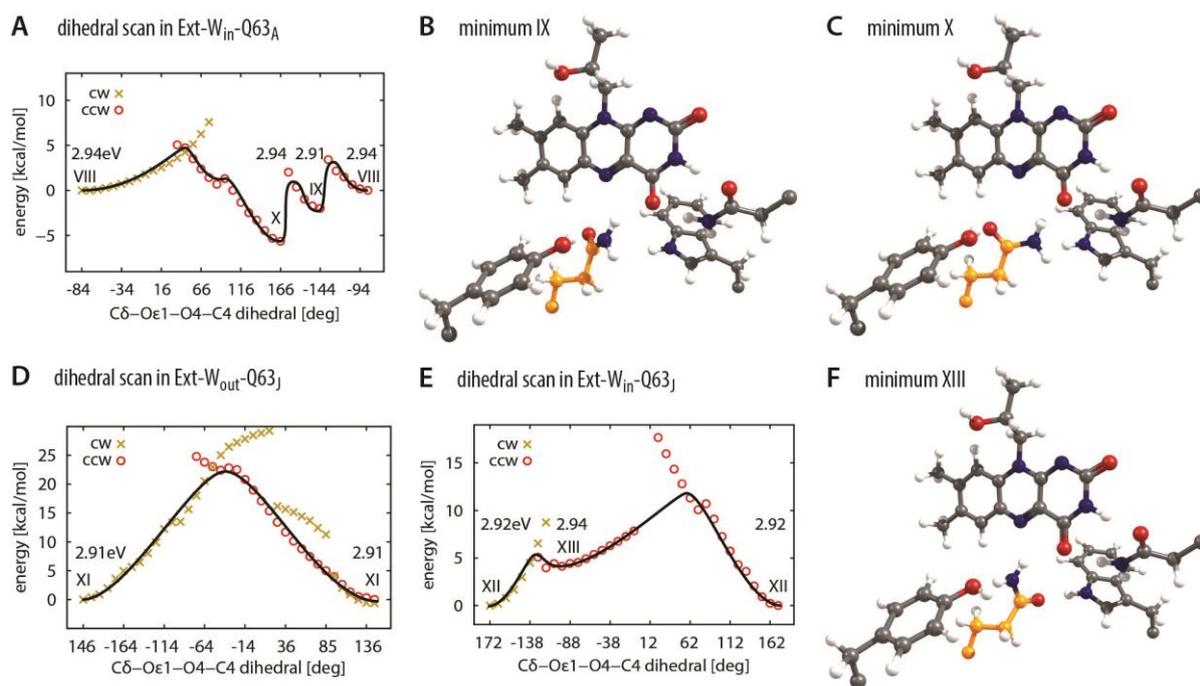


Figure 4.9. (A), (D) and (E) Full 360 degree energy scans of the Q63-flavin dihedral angle, starting from the ext-W_{in}-Q63_A (minimum VIII) or the ext-W_{out}-Q63_J (minimum XI) or the ext-W_{in}-Q63_J (minimum XII) models in clockwise (cw) and counter-clockwise (ccw) directions in 10 degree steps. The black lines represent the dihedral angle potential energy that connects respective minima. Note the different energy scale of the plots. The S₁ excitation energies for the minimum-energy structures are indicated in eV. (B), (C) and (F) The structure of the flavin binding pocket at the indicated energy minima. The rest of the models is not shown for clarity. The structure of the flavin binding pocket at the minima VIII, XI, and XII is shown in Figure 4.6.

The dihedral-angle energy scans starting from the ext-W_{out}-Q63_{tau} model (minimum XIV) and from the ext-W_{in}-Q63_{tau} model (minimum XV) are presented in Figure 4.10. The energy barrier for the rotation of the tautomeric glutamine is around 10 kcal/mol in the two cases; a little above 10 kcal/mol in the W_{out} structure and a little below 10 kcal/mol in the W_{in} structure. In the W_{in} structure, a well-defined local minimum is identified (minimum XVI), 4.4 kcal/mol above minimum XV. In contrast, in the W_{out} structure, such a local minimum was not found. Thus, the rotation of the tautomeric glutamine to adopt the orientation of minimum XVI is strongly

promoted in the W_{out} structure. This finding is consistent with the assumption that upon the formation of the tautomeric glutamine during the photoreaction in the W_{out} dark-state structure, the tautomeric glutamine rotates to adopt its more stable rotamer, which is further stabilized by the formation of the W_{in} light-state structure. In contrast, a tautomeric glutamine in the rotamer form of minimum XVI is proposed as the light-state model in (Sadeghian et al., 2008, 2010). On the basis of the presented energy scans however, it can be concluded that the rotated $Q63_{\text{tau}}$ form as in minimum XV is energetically more favorable, supporting the light-state model suggested in (Domratcheva et al., 2008; Khrenova et al., 2010). In Figure 4.9 and Figure 4.10 the S_1 excitation energies are also indicated at the equilibrium geometries. In the models containing the amide glutamine, the excitation energy varies between 2.91 and 2.94 eV, whereas in the models containing the imid glutamine, it varies between 2.79 and 2.83 eV. These findings are consistent with the results on the smaller W_{in} cluster models presented above. As was argued in the previous section, only a glutamine undergoing the amide-imid transition is consistent with the observed red-shifted flavin absorption in the BLUF light state.

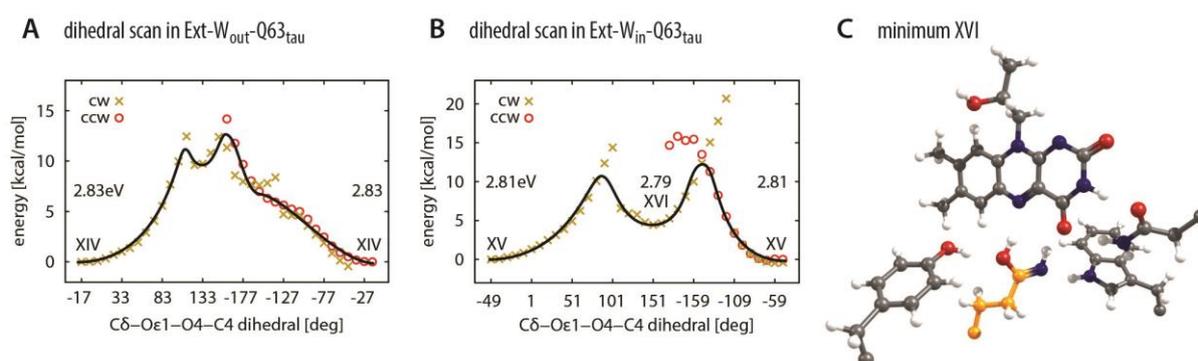


Figure 4.10. (A) and (B) Full 360 degree energy scans of the $Q63$ -flavin dihedral angle, starting from the ext- W_{out} - $Q63_{\text{tau}}$ (minimum XIV) or the ext- W_{in} - $Q63_{\text{tau}}$ (minimum XV) models in clockwise (cw) and counter-clockwise (ccw) directions in 10 degree steps. The black lines represent the dihedral angle potential energy that connects respective minima. Note the different energy scale of the plots. The S_1 excitation energies for the minimum-energy structures are indicated in eV. (C) The structure of the flavin-binding pocket at minimum XVI. The rest of the model is not shown for clarity. The structure of the flavin binding pocket at the minima XIV and XV is shown in Figure 4.6.

4.3.5 Comparison to NMR and MD studies

The results of these calculations suggest a rather dynamic flavin-binding pocket in the dark state with several glutamine rotamers present, especially in the W_{in} BLUF conformation, which all are consistent with the experimental absorption maximum of the dark state at about 450 nm. The proposed structural disorder is in agreement with the NMR study of the AppA-BLUF domain that revealed intermediate exchange broadening of the $Q63$ signals and suggested several $Q63$

rotamers in the W_{in} conformation of the dark state structure (Grinstead, Hsu, et al., 2006). Importantly, none of the 20 NMR structures of the W_{in} conformation contains the 1YRX crystallographic Q63_A rotamer. The hydrogen bonding network around flavin in all NMR structures is distinct from the W_{in} crystal structure 1YRX. The Y21-OH group points towards Q63 in all NMR molecules and not away from it as it is assumed in the crystal structure. Therefore, the NMR solution structure cannot support the crystallographic Q63_A rotamer, as other studies implied (Grinstead, Avila-Perez, et al., 2006; Rieff et al., 2011; Meier et al., 2012; Hsiao et al., 2012). Interestingly, one of the 20 NMR structures resembles the low-energy minimum II shown in Figure 4.8 B and C. Further NMR studies on other BLUF proteins revealed a rather complex picture of protein dynamics. For instance, BlrB BLUF appeared highly dynamic (Wu et al., 2008), similar to AppA BLUF, whereas BlrP1 BLUF (W_{out} conformation) turned out to be more stable (Wu & Gardner, 2009). Whether these differences are implicated in the light-sensing properties of BLUF is still unknown and must be elucidated in upcoming studies.

In the literature one only finds limited molecular dynamics (MD) analysis of the flavin-binding pocket of BLUF (Obanayama et al., 2008; Rieff et al., 2011; Meier et al., 2012; Khrenova et al., 2010) and no studies on the energy barriers separating conformational substates. In (Obanayama et al., 2008) a rotation of the Q63_A side chain in the W_{in} structure was observed during the course of the MD trajectory, yielding the Q63_J rotamer as the stable conformation in the W_{in} structure. Consistent with the results and argumentation presented here, the authors explained the Q63 side-chain rotation to be promoted by the electrostatic repulsion between the carbonyl oxygen atoms of Q63 and flavin. However, in contrast to the presented results, Obanayama et al. argued that the transition from the solvent-exposed to the protein-buried conformation of W104 is sufficient to induce the spectral red shift between the dark and light states, notably without any change in the Q63_J rotamer conformation. In (Rieff et al., 2011) the dynamics of the W_{in} and W_{out} conformations using crystal and solution NMR structures of several BLUF proteins was simulated. The observed movements of the residues destabilizing the hydrogen bonds in the flavin-binding pocket were ascribed to the lack of electrostatic polarizability in the force field employed, but not to the non-equilibrated W_{in} starting structure, containing the wrong conformation of the glutamine side chain. To maintain the crystallographic positions of the atoms, Rieff et al. introduced harmonic restraining potentials. The stability of the protein was judged on the basis of average atom-specific RMSD values larger than the thermally expected deviation in the restraining potential for each atom. In another MD study comparing the W_{in} and W_{out} structures (Meier et al., 2012) a different definition of protein stability was used: the existence of the characteristic secondary structure motifs throughout the MD simulation and an average overall RMSD of maximum 3.5 Å led the authors to conclude that both the W_{in} and W_{out} X-ray structures are “reasonably stable” and that the W_{in} structure is preferred in solution. However, the hydrogen-bond statistics presented in this study reveal a substantial deviation of

the simulated structures, in particular of the flavin-binding pocket, from the starting crystal structures throughout the trajectories. It is emphasized that atom-specific RMSD values used here and in (Rieff et al., 2011) are more instructive to judge the stability of a hydrogen-bonding network than overall RMSD of the whole protein.

The presented dihedral energy scans – though computed on cluster models excluding parts of the protein and the solvent – are an expensive yet highly valuable way to characterize protein dynamics through the estimation of energy barriers separating neighboring local minima. Undoubtedly, the DFT method describes hydrogen bonds better than any molecular-mechanical force field. In this respect, the equilibration of the structures and the description of assessable side-chain rotamers are expected to be more accurate than those derived from the MD trajectories. However, for larger scale calculations this method is too expensive and can serve only as a complement for MD studies like, for example, those presented in (Khrenova et al., 2010, 2011). In the future, a systematic study of the various dihedral degrees of freedom of the hydrogen-bonding network residues and a comparative study of various BLUF proteins would be of high interest. Furthermore, the analysis of protein-protein interactions most likely influencing the conformational dynamics of BLUF is needed, as most BLUFs are part of a multi-domain protein. Finally, large-scale conformational transitions, like the W_{out} – W_{in} (or vice versa) transition and the refolding of the $\beta 5$ strand, will require the use of enhanced conformational sampling and free-energy techniques that deform the potential-energy surface along a pre-defined collective coordinate.

4.4 Summary and conclusions

In this chapter, the Q63 rotamer assignment in the W_{in} X-ray structure of AppA BLUF was investigated and shown to be wrong. The wrong assignment has been fuelling the debate on the BLUF light-activation mechanism for several years. The interpretation of the structural and computational findings in the context of the three criteria formulated here help to overcome the debate: agreement with the X-ray electron density, flavin S_1 spectral shifts and the light-sensitivity energy-barrier requirement. The analysis of rather large cluster models describing the environment of the flavin-binding pocket demonstrates that from the three possible forms of the conserved glutamine only Q63_J and Q63_{tau} but not Q63_A are compatible with the experimental electron density of the W_{in} structure. Thus the Q63_A rotamer, originally proposed in the 1YRX X-ray structure, cannot serve as an experimental basis for the BLUF dark-state model. It is noted that the NMR structures in fact do not support the original Q63_A assignment in the W_{in} crystal structure because they contain different Y21 and Q63 rotamers.

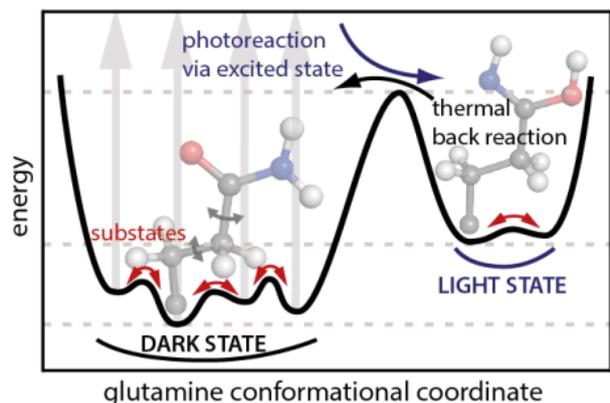


Figure 4.11. The potential energy surface of Q63 rotamers and tautomers in BLUF. Figure was taken from (Udvarhelyi & Domratcheva, 2013).

Figure 4.11 summarizes the main conclusions of this chapter. In the dark state of BLUF the Q63 side chain (and possibly that of other residues) may adopt several conformational substates. The dihedral energy scans in both the W_{in} and W_{out} structures revealed that the energy barriers between these substates, in particular in the W_{in} structure, is low. The computed red shift of flavin absorption between the geometry-optimized $Q63_A$ and $Q63_j$ models is insignificant, irrespective of the tryptophan orientation. In contrast, the significant spectral shift computed between the amide Q63 and imid Q63-containing models is consistent with the experimental red shift. The conclusions drawn from the results in the cluster models are not altered to any extent if QM/MM models are considered. The extended W_{in} and W_{out} cluster models discussed in this chapter represent the hitherto largest BLUF models optimized at the DFT level providing insights into the dynamic nature of the two BLUF protein conformations. Their relative energies allow speculating that the W_{out} conformation with $Q63_j$ and the W_{in} conformation with $Q63_{tau}$ correspond to the dark and light states, respectively. The dihedral-angle energy scans in both structures with both amide and imid Q63 forms support this conclusion. Furthermore, it is concluded from the light-sensitivity criterion that glutamine rotation does not require photoactivation and is unlikely to trigger the BLUF light response. Instead, light-induced glutamine tautomerization is consistent with the formulated light-sensitivity criteria and thus proposed as the light-activation mechanism in BLUF.

5 BLUF photoreaction

Having obtained a good understanding of the method and model limitations in the previous chapters, in this chapter, I present the results of the photoinduced proton-coupled electron transfer pathway calculations of the BLUF photoreaction mechanism. I have computed the hitherto most complete pathways that describe the photoactivation of the BLUF dark state, the photostability of the BLUF light state and the effect of the chemical (redox) modification of the conserved tyrosine side chain. From the computed pathway energies, I estimate the electron-transfer rate that determines the flavin fluorescence life time using Marcus theory. The results of the calculations consistently reproduce experimental trends. The results on the dark-state photoactivation mechanism in a small cluster model were previously published in (Udvarhelyi & Domratcheva, 2011). In my more recent studies of the BLUF photoreaction, I employed a significantly larger cluster model and also hybrid QM/MM models. I further refined the protocol of the pathway calculations to obtain more accurate estimates of the energy barriers. This chapter contains the results from the most recent calculations.

5.1 Excited-state pathways and BLUF photochemistry

The knowledge of the pathway along which the excited-state chromophore relaxes back to the ground state potential energy surface is required to determine the photoreaction mechanism. Along the relaxation pathway crossings of excited-excited and excited-ground state potential energy surfaces govern the formation of the photoproduct. Molecular geometry optimization in the excited state is needed to find the relaxation pathway. For a molecular complex of the size of the BLUF photoactive triad, consisting of flavin, glutamine, and tyrosine, these calculations are very rare and bear many challenges. In fact, in the literature one finds detailed photochemical calculations only for the isolated flavin molecules embedded in different environments but

without considering intermolecular electron transfer (Climent et al., 2006; Salzmann et al., 2008; Klaumünzer et al., 2010).

Photoreactions of the flavin molecule start with the population of the S_1 excited state through blue light absorption. Subsequent S_1 -state relaxation may take several routes, the most typical of which are fluorescence, triplet formation or photoinduced electron transfer from a nearby residue inside the protein. The latter two are computationally described by locating state-crossing geometries between the S_1 and another excited state, a triplet or an intermolecular electron-transfer (ET) state. The location of the state crossings is the primary task in the characterization of photoreaction pathways. The different electronic character of the two crossing excited states or of an excited state and the ground state ensures that the crossing geometry can be located by following the energy gradient of the higher-energy state because the energy decrease of the higher-lying state is concomitant with the energy increase of the lower-lying state, finally leading to the state crossing.

The triplet formation in the isolated flavin molecule (or flavin in a polar environment) was described using different computational techniques. In (Salzmann et al., 2008), a linear interpolation path connects two minima of different electronic structure (S_1 and T_n) that contains the crossing of the two electronic states. Alternatively, a minimum energy path, used in (Climent et al., 2006), initiated from the FC point in the S_1 state on the excited-state potential-energy surface leads to a state crossing between the S_1 and T_n states. An intersystem crossing from S_1 to the triplet state plays no role in BLUF photochemistry (Bonetti et al., 2009). Instead, an intermolecular electron transfer from the conserved tyrosine Y11 residue to the excited flavin is responsible for the fluorescence quenching of the S_1 state. The observed multiple S_1 decay times point to a complex process (Gauden et al., 2006). Electron transfer in the excited state corresponds to the S_1 /ET state crossing, which in the case of the BLUF photoreceptor is characterized in detail in the present work for the first time, using multi-configurational quantum-chemistry methods. Previously, Sadeghian et al. attempted to locate the S_1 /ET state crossing with the combination of TD-DFT geometry optimization in the ET state and subsequent CC2 single-point energy calculations (denoted CC2//TD-DFT) (Sadeghian & Schütz, 2007; Sadeghian et al., 2008, 2010). As already discussed in Chapter 2, TD-DFT severely underestimates the energy of the ET state and gives an incorrect S_1 and ET state ordering in BLUF. The TD-DFT-computed ET state is the lowest-lying excited state in the computations of Sadeghian et al., which they used to perform excited-state geometry optimization to find the ET-state energy minimum. Then, along the interpolation path connecting the FC point and the ET minimum, Sadeghian et al. recomputed the energies of the ground and excited states with the CC2 method and obtained correctly-ordered states and estimated the energy of the ET/ S_1 crossing. However, both the CC2 and TD-DFT methods are unsuitable to locate and characterize the state crossing of an excited state with the ground state; therefore the computed CC2//TD-DFT pathway of BLUF remains incomplete.

In BLUF photochemistry, electron transfer is followed by proton transfer (Gauden et al., 2006). Proton-coupled electron transfer (PCET) leads to the ET/CS state crossing, mediating the system's decay to the ground state. Only the multi-reference PT2//CASSCF methods are suitable to describe complete reaction pathways consisting of the photoinduced electron transfer (PET) part with the ET/S₁ and the proton transfer (PT) part with the ET/CS electronic state crossings. Therefore, the multi-reference PT2//CASSCF methods, despite the impressive development of the single-reference excited-state methods in the recent years, remain the method of choice for the computation of intermolecular electron and proton transfer processes in excited organic molecules. For example, in (Sinicropi et al., 2000) the intermolecular PCET process between amine and electronically excited azoalkane was computed with the PT2//CASSCF approach. Domcke and coworkers studied the photoinduced electron and proton transfer reactions with the PT2//CASSCF approach and demonstrated that photoinduced PCET provides an efficient deactivation pathway for an excited pyridine molecule in the hydrogen-bonded pyridine-pyrrole system (Frutos et al., 2007) and proposed that PCET explains DNA photostability (Schultz et al., 2004; Sobolewski et al., 2005).

The BLUF dark-state photoreaction (i.e. the formation of the BLUF light state) is proposed to involve sequential electron and proton transfer reactions, leading to a neutral biradical intermediate (Gauden et al., 2006). Then, a second PCET reaction results in the recombination of the biradical to form the photoproduct characterized by the 15-nm red-shifted flavin absorption. The formation of the photoproduct via the two PCET reactions takes place in less than a nanosecond (Gauden et al., 2006). The photoproduct (i.e. the light state) contains the oxidized flavin in an altered hydrogen-bonding environment compared to that in the dark state. In particular, the hydrogen bond to flavin C4=O4 becomes stronger as evidenced by its well-pronounced downshift in the IR region (Masuda et al., 2004). Moreover, it was found that Y11 forms an "unusually strong" hydrogen bond with Q53 in the light state (Iwata et al., 2011). Several studies proposed that the spectral red shifts observed in the light state are signatures of stronger intermolecular interactions of flavin with its hydrogen-bonded environment (Mathes, Zhu, et al., 2012; Toh et al., 2008; Shibata et al., 2009; Lukacs et al., 2011; Yuan et al., 2011). These studies also showed that the flavin fluorescence life time is significantly shortened and the heterogeneity of the excited-state decay is reduced in the light state compared to the dark state. It was suggested that similar to the dark-state photoreaction, the light state also undergoes a PCET process upon photoexcitation, yet with a distinct photodynamics consistent with a concerted proton and electron transfer reaction, in contrast to the sequential reaction of the dark state (Mathes, Zhu, et al., 2012). Upon photoactivation of the light state, a neutral biradical intermediate highly similar to the one formed during the dark-state photoreaction was identified and the photoreaction was completed in about a picosecond. However, in contrast to the dark state, the light state is highly photostable because the internal conversion leads back to the light-state Franck-Condon (FC) point (Toh et al., 2008). The light state thermally decays to the dark

state within seconds to minutes (spanning three decades), depending on the BLUF domain (Masuda et al., 2004; Zirak et al., 2006, 2005; Fukushima et al., 2005; Rajagopal et al., 2004).

Notably, the biradical intermediate could only be resolved in the experiments with the PixD protein (Gauden et al., 2006), but not with AppA (Gauden et al., 2007; Stelling et al., 2007), BlrB (Mathes et al., 2011) or BlsA (Brust et al., 2014). Recently, on the grounds of the missing radical intermediates, the Meech group proposed that BLUF photoactivation does not follow the PCET pathway (Lukacs et al., 2014). Instead, they proposed a “neutral pathway” along which the formation of the S_1 flavin induces the tautomerization of the glutamine side chain. In contrast, the Kennis group demonstrated that the dynamics of both photoactivation and the dark-state recovery are linked to redox properties (Mathes, Stokkum, et al., 2012), which the authors interpreted according to the PCET mechanism. The observed redox tuning was achieved by mutations of residues around flavin and by the chemical modifications of the tyrosine electron donor (Mathes, Stokkum, et al., 2012). For example, when the tyrosine electron donor was replaced by a fluorotyrosine analog, a substantially slower fluorescence quenching was observed in PixD (Mathes, Stokkum, et al., 2012) and AppA (Lukacs et al., 2014).

Here the PT2//CASSCF approach is applied to find the BLUF photoreaction pathway of the dark and light states to reveal the mechanisms responsible for photoactivation and photostability. In addition, the redox-modulating effect of Y11 fluorination is computed and compared to the measured ET rates in fluorotyrosine-containing BLUFs. Geometry optimization is carried out in the excited state with the CASSCF method in various BLUF models containing the photoactive triad and its environment in the protein. At the equilibrium geometries, the energies are recomputed with the XMCQDPT2 and (MS-)CASPT2 methods to include dynamic electron correlation. As demonstrated in Chapter 2, the POCAS selection of active space orbitals provides a chemically sensible and computationally efficient choice that allows mapping the complete photoreaction pathway in BLUF (consisting of the PET and PT parts). The ET/ S_1 and ET/CS state crossings are located along the pathways and the ET and PT energy barriers are estimated. The results obtained in the dark-state, fluorinated, and light-state models are consistent with experimental observations. The hitherto most complete model of BLUF photochemistry is presented.

5.2 Methods and computational details

5.2.1 Cluster and QM/MM models

The “small cluster model” representative of any BLUF protein (introduced in Chapter 2) as well as the “large cluster model” and a QM/MM model based on the PixD X-ray structure

(introduced in Chapter 3) were used for the electron and proton transfer pathway calculations in this chapter. This chapter refers to the QM/MM model as PixD I. In addition, a second QM/MM model, denoted PixD II, including the N35 side chain into the QM subsystem is also considered. Furthermore, this chapter also considers a QM/MM model based on the 2BYC PDB coordinates (chain A) of the BlrB BLUF protein, prepared analogously to the PixD QM/MM model (for details see Chapter 3).

To model the effect of tyrosine fluorination, cluster and QM/MM models including 3-fluorotyrosine and 2-fluorotyrosine were used. For both fluorotyrosine analogs two conformations resulting from the relative orientation of tyrosine to glutamine were considered, indicated in Figure 5.1. All four fluorotyrosine analogs were used in the small cluster and the PixD-I QM/MM model PET pathway calculations, whereas only the two conformations of 3-fluorotyrosine were considered in the large cluster model. Finally, the photoreaction of the BLUF light state, that contains the tautomeric glutamine, was also computed in the small cluster model (built based on the respective crystal structure, as described in Chapter 2), and in the PixD I QM/MM model (built based on the structure obtained at the end of the dark-state photoreaction calculations).

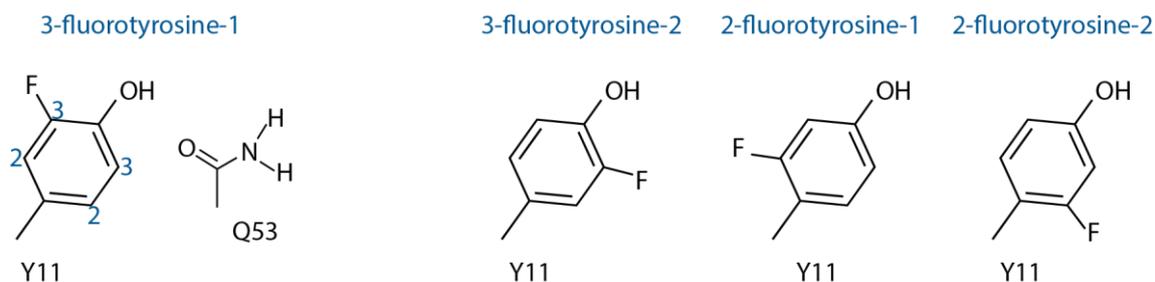


Figure 5.1. Two orientations of the 3-fluorotyrosine and 2-fluorotyrosine analogs with the fluorine atom either pointing away from the Q53 side chain (orientation 1) or towards it (orientation 2).

Similar to the previous chapters, geometry optimization was carried out with the state-averaged CASSCF method. In the cluster models the distances between pairs of the terminal atoms of each molecular fragment were kept constant. In the QM/MM models the “microiterations on” (Melaccio et al., 2011) scheme was used to optimize the geometry of the QM subsystem and to relax the geometry of the MM subsystem within 4 Å of any QM atom, whereas the remaining MM atoms were kept frozen. The excitation energies at the CASSCF geometries were computed with the XMCQDPT2/cc-pVDZ (with an edshift value of 0.02 au) method with the Firefly program and the (MS-)CASPT2/DZV(P) (without applying the ipewa shift and with an imaginary shift of 0.2 au) method in the cluster and QM/MM models, respectively.

5.2.2 Pathway calculations

The photoactivation pathways computed in this chapter consist of two parts, as mentioned in Section 5.1, the PET part starting from the FC geometry and leading to the zwitterionic biradical minimum, and the PT part describing proton transfer and neutralization of the biradical charge. Geometry optimization was performed at the CASSCF level of theory and at the optimized geometries single-point energy calculations were performed at the XMCQDPT2 level in the cluster models and at the (MS-)CASPT2 level in the QM/MM models. The POCAS way of choosing the active-space MOs was employed as described in Chapter 2. For brevity, energy-CAS//optimization-CAS denotes the active spaces used for the CASSCF geometry optimization and the subsequent PT2 energy calculations in this chapter. In the dark-state photoinduced electron-transfer reaction calculations, the (6,4)4//((6,4)4 active space was used, describing the closed-shell, the flavin S_1 and S_2 states and the tyrosine-flavin ET state. At the FC geometry, the ET state is the fourth state, whereas at the radical-pair geometry reached after electron transfer from tyrosine to flavin, the ET state becomes the second state. In the subsequent PT reactions starting from the radical-pair (biradical) minimum, only the CS and ET states were considered and thus the (2,2)2//((2,2)2 active space was employed. As shown for the isolated LF molecule in Section 2.3.2, changing the number of states in the POCAS calculations yields virtually identical optimized structures. This is also the case for the BLUF cluster. Figure 5.2 demonstrates the structural similarities of the biradical minimum, optimized with (2,2)2 and (6,4)4 in the small cluster model. In the light-state PET pathway calculations also the (6,4)4//((6,4)4 active space, whereas in the light-state PT pathway calculations the (2,2)2//((2,2)2 active space was employed. It should also be noted here that along the reaction pathway, the POCAS active space is stable and not even the occupation numbers of the considered MOs change, which is an important prerequisite of a balanced description of the states along the pathway.

The absolute energies of the states obtained from the POCAS calculations with the different number of states differ significantly, however, the respective XMCQDPT2 excitation energies are very close to each other. At the biradical minimum for example, the excitation energy of the ET state computed with (2,2)2 and (6,4)4 differs by less than 5%. This good agreement was used to “merge” the PET part of the pathway computed with the (6,4)4 active space to the PT part of the pathway computed with the (2,2)2 active space by shifting the (2,2)2-computed energies (CS and ET state) such that the ET state obtained with (2,2)2 matches that computed with (6,4)4. The fact that the POCAS excitation energies show a negligible dependence on the number of states included in the calculation allows for an easy comparison of the relative energies of various species occurring along the photochemical pathways studied here.

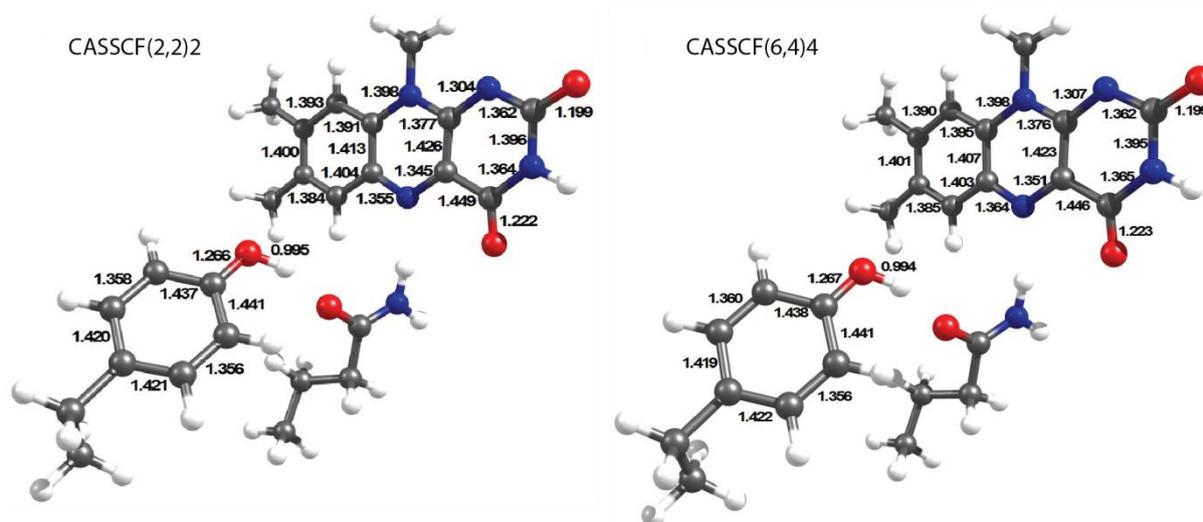


Figure 5.2. CASSCF(2,2)2/cc-pVDZ and CASSCF(6,4)4/cc-pVDZ-optimized geometries of the biradical minimum in the small cluster model (N35 is not shown for clarity). Relevant bond lengths are indicated in Å.

In the fluorotyrosine-containing models there are several cases in the geometry-optimization and in the subsequent single-point calculations of changing the number of states in the POCAS calculations. In the fluorotyrosine-containing small cluster and PixD-I QM/MM models the FC geometry was optimized with CASSCF(4,3)3 (with the S_0 , S_1 , and S_2 states) but the single-point energies were recomputed either with the (6,4)4 active space, i.e. (6,4)4// (4,3)3 for the small cluster 3-fluorotyrosine-1, 2-fluorotyrosine-1, 2-fluorotyrosine-3 and the PixD-I 3-fluorotyrosine-2 models or with (8,5)5// (4,3)3 for the small cluster 3-fluorotyrosine-2 model. In that case the (8,5)5 active space was used for the single-point energy calculation because the flavin S_n state had to be included into the active space to compute the ET state. The biradical minima in the all four fluorotyrosine PixD-I QM/MM models were computed with (6,4)4// (2,2)2. In the large cluster models containing fluorotyrosine models the whole pathway was computed with (6,4)4// (6,4)4.

The PT part of the pathway calculations was carried out for the dark state models (with the native tyrosine) in the small cluster, large cluster and PixD-I QM/MM models. To determine the transition state geometry and the energy barrier of the proton transfer reactions, relaxed energy scans along hydrogen bonds, denoted by donor-proton...acceptor (D-H...A), were performed. Along the energy scan, the H...A distance was successively reduced by 0.1 Å along the vector connecting the D and A atoms, and at each step the geometry was fully optimized with the coordinates of the D-H...A atoms kept fixed. The standard gradient thresholds (0.0001 hartree/bohr) were used except for the calculations of the large cluster model, where a threshold of 0.0005 hartree/bohr was applied. The geometry with the highest energy along the evaluated scan is considered as the PT transition state. The difference between the ET-state energies at this structure and at the D-H...A minimum gives the PT energy barrier. At the end of the energy

scan, geometry optimization without freezing the D...H-A atomic coordinates leads to the structure with the proton transferred to the acceptor atom.

The reaction coordinate of the ET and PT pathways was determined from the analysis of the obtained energy minima by finding a suitable linear combination of the bond and hydrogen bond distances, which are specified in Figure 5.3. Because of the different chemical structure and orientation of Q53, the bonds and hydrogen bonds comprising the reaction coordinate are distinct in the dark and light states. The gradient threshold for geometry optimization in the small cluster model was 0.0001 hartree/bohr and in the large cluster model and all fluorotyrosine-containing cluster models was 0.0003 hartree/bohr. Test calculations revealed no significant effect on the geometries with these thresholds.

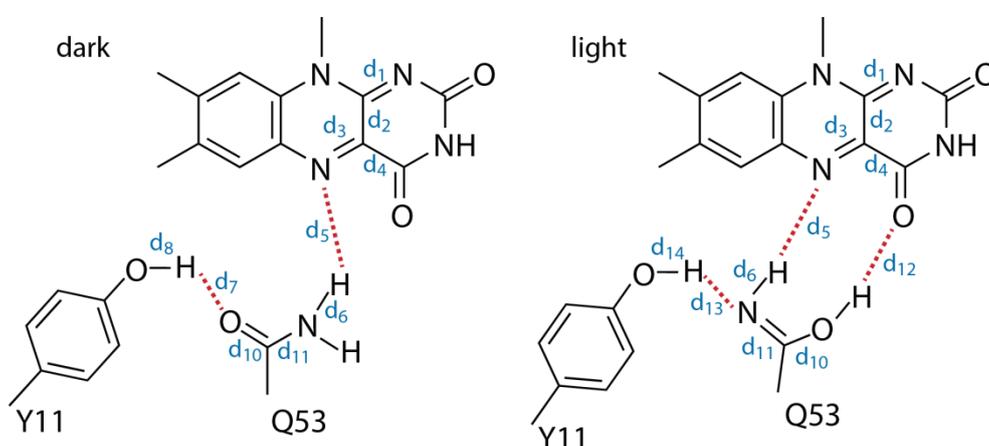


Figure 5.3. Definition of interatomic distances used to describe the ET and PT reaction coordinates in the dark-state model containing the amide and in the light-state model containing the imid Q53 side chain.

5.3 Results and discussion

5.3.1 PET in the dark state

The geometry in the closed-shell ground state was optimized to reach the Franck-Condon minimum, denoted as geometry A in the following. Table 5.1 lists the vertical S_1 and ET state excitation energies in the five BLUF models employed to study the PET pathway. At the FC point, the ET state lies always above the S_1 state in energy. As the following PET pathways will show, the S_1 -ET gap is relevant for quenching of the flavin fluorescence by photoinduced electron transfer. The lowest ET state excitation energy and smallest S_1 -ET gap is found in the large cluster and the BlrB QM/MM models, and the largest in the PixD QM/MM models. The energy diagrams presented in Figure 5.4 illustrate the PET reaction energies in dark-state BLUF.

Both CASSCF and PT2 pathways are shown for the five models. In the following, the construction of these pathways is detailed and the implications discussed.

Geometry optimization in the S_1 state, the first excited state, leads to the S_1 minimum from which fluorescence is observed, denoted as geometry B. The relaxation from A to B decreases the energy of the S_1 state by ≈ 12 kcal/mol in the cluster models and less, by ≈ 8 kcal/mol, in the QM/MM models (here, and below XMCQDPT2 and MS-CASPT2 values are given for the cluster and QM/MM models, respectively). At geometry B, the vertical S_1 excitation energy gives an estimate for the emission maximum, also listed in Table 5.1. The obtained vertical S_1 energies at B and the Stokes shifts (computed as an energy difference between the vertical S_1 excitation energy at A and B) are consistent in all models and are in reasonable agreement with experiment (Zirak et al., 2006, 2007; Weigel et al., 2011; Fujiyoshi et al., 2011). It is noteworthy that the Stokes shift in the PixD II QM/MM model, which has the QM subsystem of the same composition as the small cluster model, is identical to the one obtained in the small cluster model.

Table 5.1. The vertical S_1 and ET excitation energies at the FC minimum A, along with the S_1 -ET energy gap are listed together with the vertical S_1 excitation energy (emission) at the optimized S_1 minimum B and the resulting Stokes shift. XMCQDPT2 and MS-CASPT2 values are given for the cluster and QM/MM models, respectively.

	absorption [eV]	ET [eV]	S_1 -ET gap [eV]	emission [eV]	Stokes shift [nm]
small	3.03	4.58	1.55	2.37	114
large	3.07	4.06	1.00	2.34	125
PixD I	2.77	4.45	1.68	2.12	137
PixD II	2.76	4.63	1.87	2.20	114
BlrB	2.87	3.85	0.98	2.30	108

Geometry optimization starting from the FC point in the ET state, the third excited state, first leads to the ET/ S_2 state crossing. The geometry optimization in the second excited state after the crossing follows the energy gradient of the ET state until the ET/ S_1 state crossing is reached, denoted as geometry C. Here again, after the crossing the geometry optimization is continued in the lower, first excited state which becomes the ET state, leading to a zwitterionic biradical minimum, denoted as geometry D. Structure D contains the flavin radical anion and the tyrosine radical cation. Upon relaxation from C to D, the biradical is stabilized by 5.6 kcal/mol in the large cluster model and by 8 kcal/mol in the BlrB model, given by the ET-state energy decrease from C to D. The biradical stabilization is insignificant in the small cluster and both PixD QM/MM models. The energy difference between the ET state at geometry D and the S_1 state at geometry B is one of the determinants of the ET rate, which will be discussed in Section 5.3.6. The lowest-energy geometry D is found in the large cluster and the BlrB QM/MM model, at a similar energy as the S_1 state at geometry B. The ET state at geometry D is still rather high in energy; its further stabilization takes place by proton-transfer reactions, which neutralize the

zwitterionic biradical. The PT steps concomitantly increase the energy of the closed-shell ground state ultimately leading to the ET/CS state crossing, discussed in the next section.

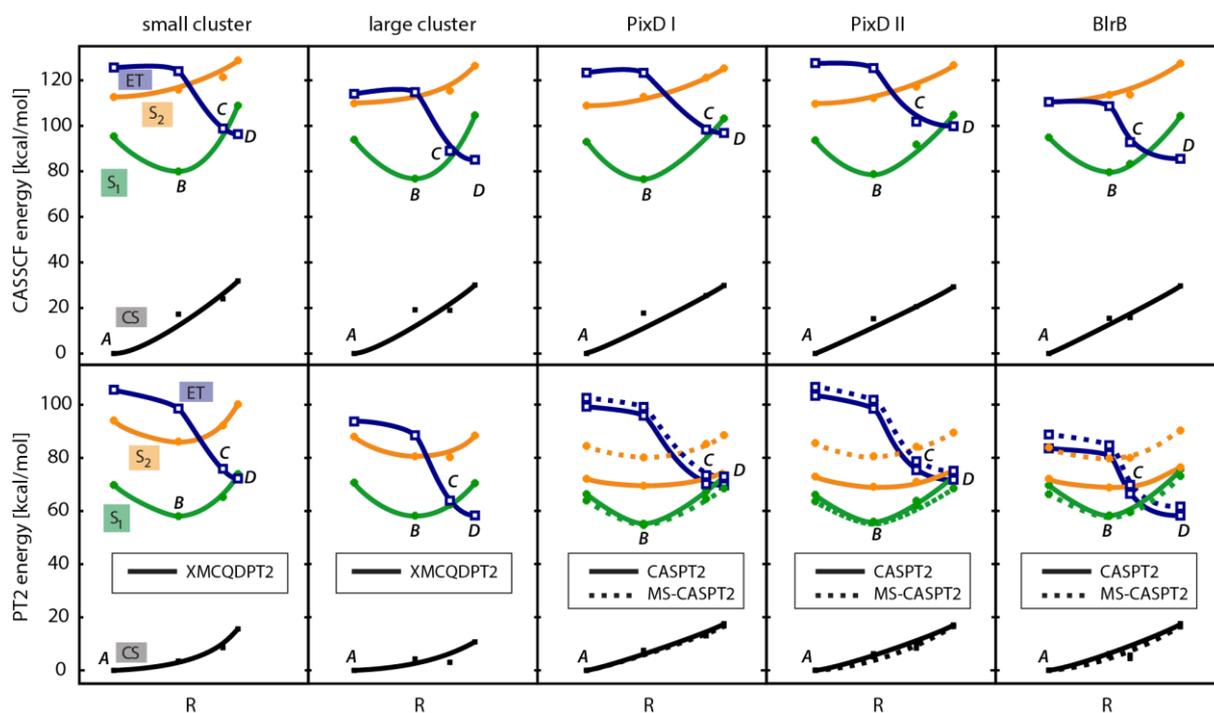


Figure 5.4. Photoinduced electron transfer pathway in various BLUF cluster and QM/MM models at the CASSCF and PT2 level of theory. Optimized geometries A-D and energies were obtained as described in the text, with $(6,4)4//((6,4)4)$ calculations in all cases, except for geometry D in the PixD-II model, where $(6,4)4//((2,2)2)$ was employed.

An analysis of the changes in the bond lengths helps to describe the photoreaction coordinate connecting the located geometries A-D and to construct the energy diagrams in Figure 5.4 for different models in a consistent way. Table 5.2 lists the bond lengths that change most during geometry optimization of the S_1 and ET states. The structural relaxation from geometry A to B involves a lengthening of the flavin N5-C4a bond (d_3) by approx. 0.06 \AA and a shortening of the C4a-C10a (d_2) and C4a-C4 (d_4) bonds to a lesser extent. The lengthening of d_3 is larger in the cluster models, whereas the shortening of d_4 is more pronounced in the QM/MM models. The structural changes are consistent with the flavin HOMO-LUMO electron density redistribution (see Chapters 2 and 3). As a result of the S_1 transition, the negative charge on the flavin N5 atom increases, that shortens the flavin-glutamine hydrogen bond (d_5) and increases the N-H (d_6) bond lengths of the glutamine amide group. Even the Q53 C=O (d_{10}) and C-N (d_{11}) bonds are affected to a small extent; the lengthening of d_{10} and the shortening of d_{11} are more pronounced in the cluster than in the QM/MM models.

Table 5.2. Selected bond lengths of the optimized structures along the PET pathway in the five models given in Å. For the definition of the distances see Figure 5.3.

small	d₁	d₂	d₃	d₄	d₆	d₈	d₁₀	d₁₁
geo A	1.278	1.465	1.273	1.492	1.004	0.956	1.214	1.335
geo B	1.275	1.435	1.337	1.453	1.009	0.959	1.222	1.326
geo C	1.301	1.424	1.349	1.447	1.019	0.981	1.229	1.313
geo D	1.307	1.423	1.351	1.446	1.020	0.994	1.235	1.307
large	d₁	d₂	d₃	d₄	d₆	d₈	d₁₀	d₁₁
geo A	1.294	1.455	1.278	1.491	1.003	0.953	1.214	1.336
geo B	1.292	1.426	1.340	1.455	1.007	0.956	1.221	1.329
geo C	1.314	1.417	1.351	1.449	1.012	0.970	1.227	1.318
geo D	1.325	1.412	1.356	1.445	1.015	0.988	1.233	1.310
PixD I	d₁	d₂	d₃	d₄	d₆	d₈	d₁₀	d₁₁
geo A	1.289	1.455	1.276	1.488	1.006	0.965	1.209	1.334
geo B	1.286	1.428	1.337	1.452	1.012	0.966	1.214	1.330
geo C	1.315	1.418	1.353	1.446	1.022	1.010	1.230	1.311
geo D	1.319	1.415	1.356	1.444	1.024	1.029	1.235	1.306
PixD II	d₁	d₂	d₃	d₄	d₆	d₈	d₁₀	d₁₁
geo A	1.289	1.454	1.275	1.489	1.007	0.964	1.209	1.334
geo B	1.286	1.425	1.333	1.445	1.011	0.965	1.212	1.330
geo C	1.306	1.422	1.342	1.447	1.021	0.997	1.226	1.313
geo D	1.316	1.417	1.347	1.446	1.028	1.037	1.237	1.303
BlrB	d₁	d₂	d₃	d₄	d₆	d₈	d₁₀	d₁₁
geo A	1.298	1.449	1.276	1.487	1.005	0.964	1.208	1.337
geo B	1.297	1.416	1.333	1.445	1.010	0.965	1.212	1.333
geo C	1.310	1.416	1.335	1.448	1.014	0.981	1.219	1.322
geo D	1.322	1.412	1.342	1.446	1.023	1.035	1.234	1.304

At the ET/S₁ state-crossing geometry C, the bond lengths d₃, d₆, d₈, and d₁₀ further increase and d₂, d₄, and d₁₁ further decrease compared to minimum B. From B to C a significant lengthening of the d₁ bond length also occurs. In addition, the hydrogen bonds between Y11 and Q53 as well as between Q53 and flavin N5 shorten, following the charge separation between the negatively charged flavin and the positively charged tyrosine radicals. The bond distance changes in the Y11 phenol ring (not shown in Table 5.2) are consistent with removing an electron from the tyrosine HOMO. Structural relaxation from the state-crossing geometry C to minimum D involves further increase/decrease of the respective bond lengths, however, altogether to a smaller extent. Thus, the ET state minimum D lies rather close in geometry to the ET/S₁ state-crossing geometry C, except for the BlrB QM/MM model, where the structural relaxation from C to D is even more pronounced than that from B to C, presumably because of the large energy stabilization of the ET state.

The observed structural changes upon the relaxation of the S_1 and ET states characterize the reaction coordinate (R) for the photoinduced electron transfer from Y11 to flavin in BLUF. The relaxation of the photoactive triad monotonically progresses along the same bond-stretching collective coordinate. Two set of coordinates are relevant: First, the flavin bond-stretching mode $-d_2+d_3-d_4$ related to the flavin LUMO, which is the same singly-occupied MO in both the S_1 and ET states. Second, the hydrogen-bond mode $+d_6+d_8+d_{10}-d_{11}$ points to the important role of glutamine in mediating the photoreaction in BLUF, already at this early stage. Therefore, in Figure 5.4 the PET reaction coordinate in BLUF is described by a linear combination of the discussed bond lengths: $R=d_1-d_2+d_3-d_4+d_6+d_8+d_{10}-d_{11}$. It is emphasized that the order of geometries A to D along this coordinate is revealed by the bond distance changes during the independent optimization of the S_1 and ET states without any restraints (except for the terminal atom constraints).

The energy diagrams in all models are consistent but show some variations. The important difference is the barrier of electron transfer predicted by the models, determined here by the energy at the crossing geometry C. The large cluster and the BlrB QM/MM models predict the smallest ET barrier and thus shortest flavin fluorescence life time. These are also the models with the smallest S_1 -ET gap at the FC geometry. Notably, the S_1 -ET gap correlates with the barrier height. In fact, the presented QM/MM models cannot explain the experimentally observed approximately twice-long fluorescence life time in BlrB (Zirak et al., 2006), compared to PixD (Zirak et al., 2007). The BlrB QM/MM model predicts a substantial ET-state stabilization by the protein environment, in drastic contrast to the PixD QM/MM model. In BlrB a (positively charged) arginine side chain is located close to the flavin C2=O2 carbonyl group, whereas in PixD a polar but neutral asparagine residue is located at that position. Consistent with the calculations in Chapter 3, the positive charge on the arginine close to flavin may explain the large stabilization of the ET state in the BlrB QM/MM model.

5.3.2 Proton transfer pathways

Proton-transfer pathways, along which the ET state is further stabilized, were computed using the small cluster, the large cluster and the PixD-I QM/MM models. To reach the neutral biradical intermediate, geometry optimizations were carried out considering only the CS and ET states, starting at the zwitterionic biradical minimum D. At D, the ET state is the first excited state. In all models, geometry optimization of geometry D with CASSCF(2,2)2 resulted in minimum D', virtually identical to minimum D, as shown in Section 5.2.2. Proton transfer from the electron donor to the electron acceptor stabilizes the zwitterionic biradical through neutralization; a process referred to as "the proton follows the electron" along the hydrogen bond connecting the donor and acceptor (Sobolewski et al., 2005). As a consequence, the ET state energy decreases

and concomitantly the CS state energy increases. In BLUF, hydrogen-bonding between the Y11 electron donor and flavin electron acceptor is established via Q53. Therefore, starting from structure D', there are two hydrogen bonds and two PT reactions considered here, which may occur in different order.

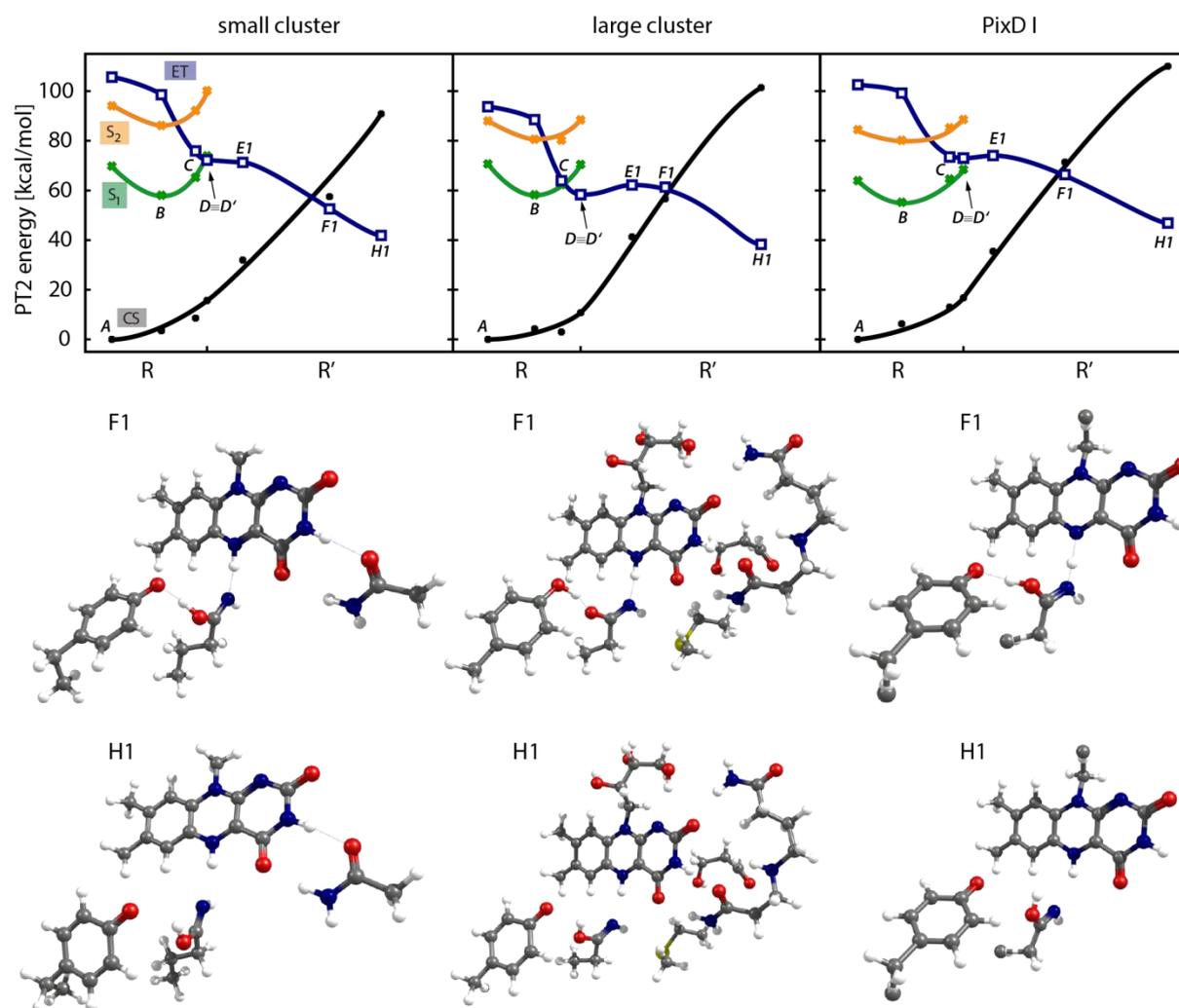


Figure 5.5. Full XMCQDPT2 and MS-CASPT2 energy profiles of the ET/PT-I pathway and structures F1 and H1 in the three models. The L44 side chain (see Figure 3.2A) in the large cluster model is not shown for clarity. The energies at geometries A-D were obtained with the (6,4)4//((6,4)4 calculations and at geometries D'-H1 with (2,2)2//((2,2)2, as described in Section 5.2.2.

First, the PT reaction mediated by the hydrogen bond between Q53-N ϵ 2 and flavin N5, is discussed, denoted as PT-I. PT-I corresponds to flavin protonation. The computed energy diagram for PT-I is presented in Figure 5.5 along with the preceding PET pathway described in the previous section. At the identified transition state, denoted as geometry E1, the proton is “half way” between the glutamine N ϵ 2 and flavin N5 atoms. After the barrier, unrestrained

geometry optimization in the ET state yields the ET/CS state crossing geometry, denoted as geometry F1. Starting from the F1 crossing point, there are two relaxation pathways in the ground state: the CS pathway leading back to the Franck-Condon minimum A and the radical-pair ET-state pathway. The latter was computed by continuing geometry optimization from F1 in the ground state, finally leading to the neutral biradical minimum, denoted as H1. Structure H1 contains the neutral N5-protonated flavin radical, the neutral deprotonated Y11 radical, and a tautomeric E-imidic Q53 in the three models. The tautomeric Q53 is rotated approximately 90° with respect to the neutral flavin radical. The hydrogen bond between the E-imidic Q53 and the neutral Y11 radical is substantially weakened and even broken in the PixD-I model. Deprotonation of the Y11 radical occurs practically concomitant to the protonation of the flavin radical, therefore only a single PT barrier (E1) was located along this pathway. The E1 barrier of flavin protonation in the ET state is rather small and slightly changes from one model to the other: 0.9 kcal/mol (small cluster), 3.9 kcal/mol (large cluster), and 1 kcal/mol (PixD I QM/MM). With respect to D', the energy of the ET state at the neutral biradical H1, after the two PT reactions are completed, is lowered by 30 kcal/mol (small cluster), 20 kcal/mol (large cluster), and 26 kcal/mol (PixD-I QM/MM). Thus the two PT reactions substantially stabilize the ET state.

As Table 5.3 demonstrates, the structural relaxation from D' to H1 mainly concerns the bonds involved in the glutamine-flavin (described by d_5 and d_6) and involved in the tyrosine-glutamine (described by d_7 and d_8) hydrogen bonds, as well as the Q53 d_{10} and d_{11} bonds. Proton transfer from the tyrosine hydroxyl to the Q53-O ϵ 1 oxygen atom occurs concomitantly to the proton transfer from Q53-N ϵ 2 to flavin N5, however, the three models differ in their structures at the ET/CS state crossing geometry F1. In the small cluster model, both protons are transferred and the Y11 and flavin radicals are both neutralized at F1 in a concerted fashion. In the large cluster model, F1 is reached upon flavin protonation but before tyrosine deprotonation. Thus, F1 contains a deprotonated anionic glutamine that explains the higher energy barrier of the PT process compared to the other models and the high energy of the ET state at F1. In the PixD-I QM/MM model the spontaneous deprotonation of the Y11 radical precedes the protonation of the flavin radical so that F1 is reached before the flavin radical is protonated. A deprotonation of Q53-N ϵ 2 and a concomitant protonation of Q53-O ϵ 1 result in a substantial increase and decrease in the d_{10} and d_{11} bonds, respectively. From the observed structural changes, the following PT reaction coordinate R' can be identified: The hydrogen bond distance changes are described by the obvious $-d_5+d_6-d_7+d_8$ coordinate and the changes in the bond lengths of Q53 C δ -O ϵ 1 and C δ -N ϵ 2 by the $+d_{10}-d_{11}$ tautomerization coordinate. Thus, the reaction coordinate $R'=0.1 \times (-d_5+d_6-d_7+d_8)+d_{10}-d_{11}$ is proposed for the PT process in BLUF. The weighting coefficient 0.1 is arbitrarily chosen to emphasize the changes in the tautomerization coordinate.

Table 5.3. Selected interatomic distances (defined in Figure 5.3) in Å along the pathway PT-I.

small	d₅	d₆	d₇	d₈	d₁₀	d₁₁
geo D'	1.992	1.021	1.551	0.995	1.236	1.306
geo E1	1.392	1.171	1.455	1.017	1.252	1.297
geo F1	1.058	1.788	1.016	1.492	1.311	1.271
geo H1	1.008	2.274	0.954	2.199	1.328	1.257
large	d₅	d₆	d₇	d₈	d₁₀	d₁₁
geo D'	2.102	1.015	1.599	0.988	1.233	1.310
geo E1	1.302	1.277	1.422	1.030	1.255	1.299
geo F1	1.063	1.742	1.365	1.053	1.269	1.288
geo H1	1.013	2.131	0.952	2.322	1.327	1.258
PixD I	d₅	d₆	d₇	d₈	d₁₀	d₁₁
geo D'	2.152	1.024	1.443	1.030	1.235	1.306
geo E1	1.585	1.183	1.389	1.048	1.244	1.302
geo F1	1.482	1.205	1.001	1.599	1.325	1.264
geo H1	1.011	2.304	0.958	3.028	1.330	1.251

Next, the PT reaction mediated by the hydrogen bond between Q53 and Y11 is discussed, denoted as PT-II and shown in Figure 5.6. PT-II corresponds to Y11 deprotonation. Analogously to PT-I, in the large cluster and the PixD-I models, the relaxed proton transfer (Y11-OH to Q53-O ϵ 1) energy scan renders the transition state E2, the ET/CS state crossing F2, and the neutral biradical minimum H2. The same energy scan in the small cluster model resulted in a different PT pathway, discussed in the Section 5.3.3. Deprotonation of the Y11 radical requires only a small energy at the CASSCF level (transition state geometry E2); at the PT2 level the transition state disappears because at E2 the ET-state energy is even lowered by 1.5 kcal/mol (large cluster) and 1.9 kcal/mol (PixD I) compared to that at D'. This finding indicates a completely barrierless relaxation towards the ET/CS state crossing geometry F2. Continuing geometry optimization in the ground state after F2 leads back to the Franck-Condon minimum on the CS path. To stay on the ET path in the ground state, proton transfer from Q53-N ϵ 2 to flavin N5 is needed.

In the PixD-I model, the F2 crossing is “delayed” significantly. It can only be located by performing the second PT energy scan (corresponding to flavin protonation). At the beginning of this scan, the ET/CS crossing is reached, denoted as F2'. The transition state for the Q53-N ϵ 2 to flavin N5 proton transfer, denoted as G2, is only a transition state at the CASSCF level, but at the MS-CASPT2 level its energy lies below that of F2 and F2'. Thus the downhill energy path from D' to H2 in the PixD-I model indicates that both proton transfer reactions occur spontaneously and in a concerted fashion. In the large cluster model, F2 contains the protonated Q53 and the second proton transfer to flavin N5 stabilizes the ET state and yields the neutral biradical H2. The biradical intermediate H2 contains the neutral Y11 and flavin radicals and the

tautomeric E-imidic Q53. The latter still forms a strong hydrogen bond to Y11 in the PixD-I model, whereas in the large cluster model this hydrogen bond is broken and the tautomeric Q53 is rotated approximately 90° with respect to flavin. With respect to D', the energy of the ET state at H2 is lowered by 19 kcal/mol (large cluster) and 23 kcal/mol (PixD I). Because of the obvious similarity to the PT-I pathway, the same reaction coordinate for the PT-II pathway, $R'=0.1 \times (-d_5+d_6-d_7+d_8)+d_{10}-d_{11}$ including the hydrogen bond and glutamine tautomerization coordinates, is proposed. Table 5.4 and Figure 5.6 summarize and illustrate the results.

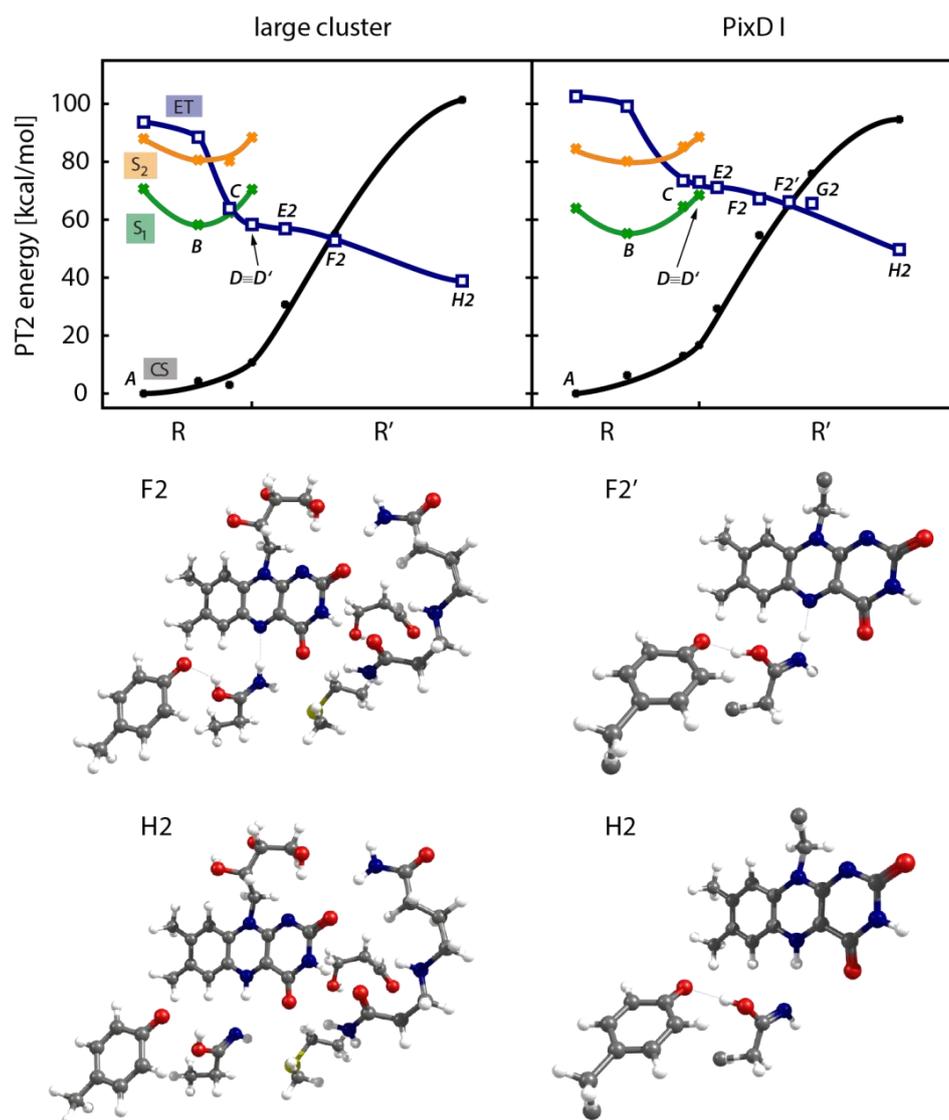


Figure 5.6. Full XMCQDPT2 and MS-CASPT2 energy profiles of the ET/PT-II pathway and structures F2/F2' and H2 in the large cluster and PixD-I models. The L44 side chain in the large cluster model is not shown for clarity. The energies at geometries A-D were obtained with the (6,4)4// (6,4)4 calculations and at geometries D'-H2 with (2,2)2// (2,2)2, as described in Section 5.2.2.

Table 5.4. Selected interatomic distances (defined in Figure 5.3) in Å along the pathway PT-II.

large	d_5	d_6	d_7	d_8	d_{10}	d_{11}
geo D'	2.102	1.015	1.599	0.988	1.233	1.310
geo E2	2.066	1.020	1.199	1.155	1.248	1.297
geo F2	2.031	1.025	0.972	1.825	1.275	1.285
geo H2	1.013	2.145	0.953	2.292	1.327	1.258
PixD I	d_5	d_6	d_7	d_8	d_{10}	d_{11}
geo D'	2.152	1.024	1.443	1.030	1.235	1.306
geo E2	2.128	1.029	1.250	1.122	1.245	1.299
geo F2	2.144	1.037	0.995	1.552	1.275	1.284
geo F2'	1.644	1.118	0.987	1.588	1.288	1.280
geo G2	1.444	1.251	0.977	1.652	1.302	1.273
geo H2	1.015	2.230	0.966	1.856	1.347	1.247

Figure 5.5 and Figure 5.6 depict the results from the combined ET and PT pathway calculations. The relaxation of the respective first excited state starting from the Franck-Condon minimum reveals a PCET downhill energy path until the biradical intermediates H1 and H2. Both PT pathways essentially predict two concerted and barrierless proton transfer reactions that result in the formation of the very similar intermediates H1 and H2, containing the tautomeric glutamine. The energy of the ET state at H2 is only 0.6 kcal/mol (large cluster) and 2.8 kcal/mol (PixD I) above that at H1. The energy of the ET state both at H1 and H2 is higher in the PixD-I model than in the cluster models with respect to the Franck-Condon minimum A. The S_1 minimum B is separated by the energy barrier E1 or D from the biradical minima H1 and H2, respectively. The different energy barriers in the different models suggest significantly different life times of the flavin fluorescence. The energy barrier is negligible in the large cluster model, thus suggesting a completely downhill energy path (especially the ET/PT-II pathway) in full accordance with the experiments that show no kinetic isotope effect upon H/D exchange on the flavin fluorescence life time (Gauden et al., 2006). Remarkably, a rather clear barrier (E1 or D') exists in the PixD-I model QM/MM. As Chapter 3 discussed, the PixD protein environment rather destabilizes the ET state and as found here, the barrier in the PixD-I model is even the highest among all models.

Radical recombination from H1 and H2 in the ground state either leads to the photoproduct minimum or back to the FC minimum. Radical recombination is not considered here, however, with the presented approach its calculation is straightforward. A further rotation of the tautomeric Q53 is likely (Khrenova et al., 2011, 2013). A second ET reaction coupled to two PT reactions with a low energy barrier in the ground state leads to the photoproduct minimum containing the tautomeric and rotated glutamine, which forms specific hydrogen bonds with the oxidized flavin explaining the red-shifted flavin absorption. A radical recombination pathway was computed in (Khrenova et al., 2013) with the UDFt method, along with energy barriers for the

rotation of the various imidic forms of Q53 in BLUF. Khrenova et al. proposes the same rotated, tautomeric Q53 in the BLUF light state as this work.

A reaction pathway leading to the formation of the tautomeric Q53 that is overall similar to the PT-II pathway presented here was previously computed and analyzed in (Sadeghian et al., 2008) for the BlrB protein. The same two PT reactions (first from Y11 to Q53 and then from Q53 to flavin N5) were computed with CC2//TD-DFT in the ET state, leading to a similar biradical intermediate as H1 and H2 shown here, but with a Z-imidic Q53. Next, the radical recombination was postulated to take place via the concerted protonation of the Q53-O ϵ 1 atom and the Y11 radical, yielding the final photoproduct with a tautomeric, but non-rotated Q53. The authors explicitly ruled out the rotation of the tautomeric Q53 based on the computed high energy barrier of the rotation. The presented work and the calculations in (Khrenova et al., 2011, 2013) challenge this conclusion. The next section demonstrates that the rotation of the tautomeric Q53 in the ET state in fact may be even barrierless. Moreover, the dihedral-energy scans presented in Chapter 4 and the cited works by Khrenova et al. further support that the flavin-binding complex representing the BLUF light state is more likely to contain the rotated Q53 tautomer than the non-rotated one.

The PT-I and PT-II results presented here are significantly different from those discussed earlier in (Udvarhelyi & Domratcheva, 2011), especially for the case of PT-I. The unrelaxed energy scans in our earlier study suggested that the PT-I pathway has a substantial energy barrier. However, this energy barrier vanishes when the energy scans are performed with a concomitant relaxation of the other degrees of freedom, as presented here. How similar/different the PT-I and PT-II pathways indeed are, i.e. how fast and in which order the two PT reactions occur, might also be determined by protein dynamics that further studies should elucidate.

5.3.3 PCET pathway without flavin protonation

In the small cluster model, the proton-transfer energy scan Y11-OH to Q53-O ϵ 1 yielded a PCET pathway, denoted as PT-III and shown in Figure 5.7. PT-III describes a new pathway, which does not involve the protonation of the flavin radical. Deprotonating the Y11 radical and protonating Q53 occurs with a small energy barrier denoted as E3 at the CASSCF level and without an energy barrier at the XMCQDPT2 level of theory. The ET/CS state crossing F3 is reached after the E3 transition state. At F3 the Y11-Q53 hydrogen bond is broken. From F3, geometry optimization was continued in the ground state (corresponds to the ET state) and surprisingly, the observed structural relaxation involves a spontaneous rotation of the protonated glutamine. As a result, a hydrogen bond forms between Q53-O ϵ 1 and the flavin radical anion O4 atom. During the rotation, the hydrogen bond between Q53-N ϵ 2 and the flavin radical anion N5 atom remains intact. Thus, the formation of a salt bridge between the flavin anion and the

glutamine cation is a mechanism to stabilize the ET state, alternative to flavin protonation. Geometry optimization in the ground ET state continues until another CS/ET state crossing is reached, denoted as F3'. At F3', Q53 is almost 180° rotated with respect to its initial position. The energy of the ET state at F3' is 20 kcal/mol lower than at D'. Continuing geometry optimization in the electronic ground state from the F3' geometry yields radical recombination and spontaneous Y11 protonation (a PCET reaction). Finally the optimization leads to the photoproduct minimum, denoted as T', that contains the tautomeric rotated Q53, oxidized flavin and reduced tyrosine. At minimum T', four states were computed with (6,4)4//((6,4)4 and the switch between the (2,2)2//((2,2)2-optimized T' and the (6,4)4//((6,4)4-optimized T' was done with the energy-shifting procedure described in Section 5.2.2 for the case of geometry D/D'.

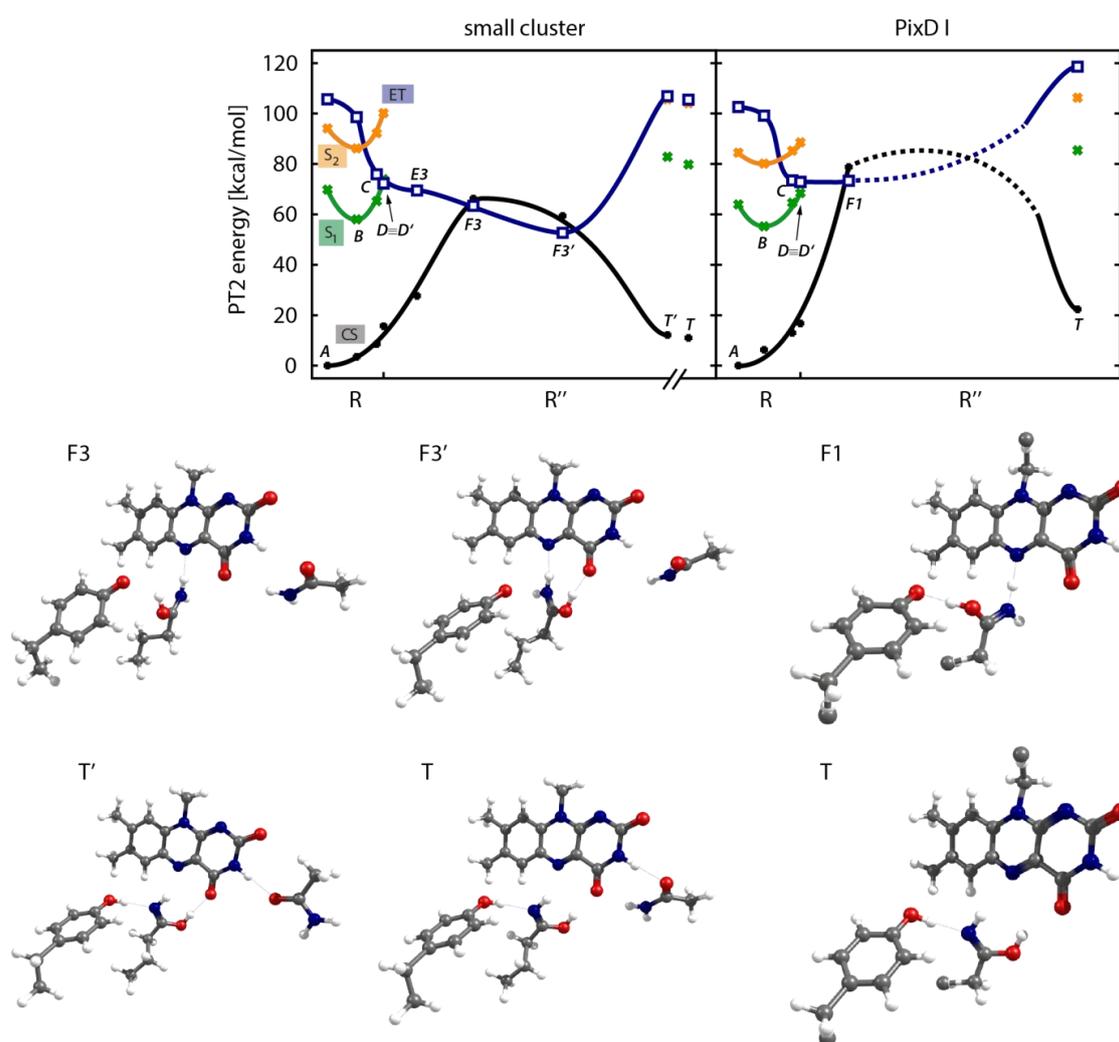


Figure 5.7. Full XMCQDPT2 and MS-CASPT2 energy profiles of the ET/PT-III pathway and structures F3/F1 and T'/T in the small cluster and PixD-I models. In the PixD-I model the energy barrier of glutamine rotation is not shown. The energies at geometries A-D were obtained with the (6,4)4//((6,4)4 calculations, at geometries D'-F3 with (2,2)2//((2,2)2, and at geometries T' and T with (6,4)4//((6,4)4, as described in Section 5.2.2.

During the course of the radical recombination the hydrogen bond between the N35-N82 and flavin-O4 atoms broke and the N35 side chain rotated away from flavin. It seems unlikely that this rotation of N35 is relevant for the BLUF photoactivation reaction because the surrounding protein (missing in the small cluster model) probably prevents such a large structural change. To obtain a photoproduct minimum in which the hydrogen bond to N35 exists consistent with the BLUF structure, the coordinates of N35 from the E3 structure were inserted in the T' structure. The geometry optimization of the prepared structure yielded the "correct" BLUF photoproduct minimum, denoted as geometry T. The ground state energy in T is 1.4 kcal/mol below that at T' and 11 kcal/mol above that at the Franck-Condon minimum A.

Starting from the ET/CS state crossing F1, the calculations of the above glutamine rotation PT-III pathway were also attempted in the PixD-I QM/MM model. Unfortunately, a technical problem with MOLCAS did not allow computations accounting for the relaxation of the environment during the energy scan of glutamine rotation, as it was not possible to freeze the internal torsional degree of freedom of the glutamine side chain. Therefore, an unrelaxed energy scan was performed that yielded a very high energy barrier, not shown in Figure 5.7. The final photoproduct minimum T, containing the rotated tautomeric Q53, was fully relaxed in the ground state and is indicated in Figure 5.7. It lies 22 kcal/mol above minimum A, considerably higher than in the cluster model. Remarkably, the protein environment does not stabilize the photoproduct in contrast to what was found in the large cluster model. Therefore, together with the results of the previous section, it is concluded that the protein environment destabilizes the whole PCET pathway in the PixD-I QM/MM model. This finding is consistent with the results of Chapter 3, where a protein-environment-induced blue shift of the ET-state energy in the PixD QM/MM model was identified and discussed.

Table 5.5. Selected interatomic distances (defined in Figure 5.3) in Å along the pathway PT-III.

small	d₇	d₈	d₁₀	d₁₁	d₁₂	d₁₃	d₁₄
geo D'	1.551	0.995	1.236	1.306	5.523	0.996	5.387
geo E3	1.251	1.112	1.253	1.296	4.922	0.998	4.895
geo F3	0.962	2.031	1.287	1.278	3.708	1.000	4.841
geo F3'	0.986	3.601	1.270	1.291	1.604	0.998	3.441
geo T'	0.958	5.229	1.322	1.264	1.879	1.922	0.962
geo T	0.957	5.185	1.326	1.263	1.991	1.927	0.961
PixD I	d₇	d₈	d₁₀	d₁₁	d₁₂	d₁₃	d₁₄
geo D'	1.443	1.030	1.235	1.306	5.864	1.003	5.364
geo F1	1.001	1.599	1.325	1.264	5.382	1.015	5.307
geo T	0.963	5.172	1.321	1.259	2.283	1.821	0.972

For the PT-III pathway a new reaction coordinate R'' is identified on the basis of the bond distance analysis. R'' features the formation of the new hydrogen bond between the tautomeric glutamine and flavin O4 at the final minimum T. Table 5.5 shows the relevant interatomic distances. The linear combination $R''=0.1 \times (-d_7+d_8-d_{12}+d_{13}-d_{14})+d_{10}-d_{11}$ was used to plot the energies in Figure 5.7.

It is noteworthy that although the characteristic conformational “softness” (in a sense of being less constrained) of the small cluster model sometimes results in structural relaxations unfeasible in BLUF (e.g. rotation of N35), it is the large structural relaxations that at the same time allow for a potentially better exploration of the potential energy landscape. The presented examples clearly show the merit of the small cluster model in the identification of new pathways: first, because of its smaller size, the small cluster model is technically easier to handle and the calculations are considerably faster; and second the PT-III pathway was found at first place with this model, providing new physical insight into the BLUF photoactivation mechanism. Yet, taking into consideration the presented results here and in Chapter 3, the large cluster model seems most suitable to study BLUF photochemistry, as it includes at least the local environment of flavin but provides enough flexibility at the same time for larger structural changes to occur. Thus, of high interest is the computation of the PT-III pathway in the large cluster model in upcoming calculations of the BLUF photoreaction. The QM/MM model, with the standard approach of freezing the coordinates of a significant part of protein, seems to impose severe limitations in addressing the geometry changes relevant to the stabilization of the excited states.

The presented PT-III pathway revealed by the small cluster model is a plausible pathway in BLUF, as the spectroscopic signatures of a neutral biradical intermediate in some BLUF proteins were not observed. Thus, it is conceivable that a neutral biradical intermediate (like H1 or H2) indeed does not form at all in, for example, AppA and BlrB, and that the photoactivation does not take the same pathway in all BLUFs: it may take the ET/PT-III route in AppA and BlrB, but the ET/PT-II route in PixD. However, the preferred route also depends on the protein dynamics which was not considered here. Importantly, independent of the formation and stability of the neutral flavin radical intermediate, the BLUF photoactivation reaction leading to the tautomerization of Q53 is a photoinduced PCET reaction. The “neutral pathway” suggested by the Meech group (Lukacs et al., 2014) thus seems rather implausible. In the presented framework their suggestion that flavin photoexcitation triggers glutamine tautomerization implies either (i) a S_1/CS state crossing along the glutamine tautomerization coordinate, which is ruled out on the basis of the presented calculations; or (ii) fluorescence decay of the red-shifted photoproduct, which experimentally is not observed; or (iii) another complex internal conversion mechanism that at present cannot be proposed on the basis of the flavin literature and which is hard to imagine based on the current models of flavin photochemistry and photophysics.

5.3.4 BLUF redox tuning with fluorotyrosine

To locate and optimize geometries A-D of the PET pathway in the fluorotyrosine models, the respective structures of the pathway computed with the native tyrosine were used as starting geometries (in the small cluster, large cluster, and the PixD-I QM/MM models), in which the respective H atom was exchanged by an F atom. The same reaction coordinate R as for the dark-state PET pathway was used to plot the energy diagrams. Table 5.6 collects the bond lengths of interest in the fluorotyrosine models and Figure 5.8 and Figure 5.9 show the obtained PET pathways in comparison to the dark state PET pathway.

Table 5.6. Selected interatomic distances (defined in Figure 5.3) in Å along the PET pathway. 2f2 abbreviates 2-fluorotyrosine-2, 3f2 abbreviates 3-fluorotyrosine-2 and so on. Models not listed here feature essentially same bond lengths as their counterparts.

small 2f2	d₁	d₂	d₃	d₄	d₆	d₈	d₁₀	d₁₁
geo A	1.279	1.462	1.277	1.490	1.004	0.957	1.215	1.334
geo B	1.275	1.436	1.336	1.453	1.009	0.959	1.222	1.326
geo C	1.304	1.424	1.350	1.447	1.020	0.988	1.233	1.311
geo D	1.307	1.423	1.350	1.447	1.023	0.998	1.237	1.305
small 3f2	d₁	d₂	d₃	d₄	d₆	d₈	d₁₀	d₁₁
geo A	1.279	1.461	1.277	1.490	1.004	0.959	1.214	1.335
geo B	1.275	1.435	1.337	1.453	1.009	0.960	1.221	1.327
geo C	1.305	1.425	1.350	1.448	1.020	0.994	1.227	1.312
geo D	1.308	1.423	1.352	1.447	1.024	1.025	1.239	1.302
large 3f2	d₁	d₂	d₃	d₄	d₆	d₈	d₁₀	d₁₁
geo A	1.295	1.454	1.278	1.491	1.003	0.956	1.213	1.335
geo B	1.292	1.425	1.340	1.456	1.007	0.959	1.222	1.327
geo C	1.319	1.415	1.354	1.448	1.013	0.984	1.229	1.316
geo D	1.325	1.412	1.357	1.445	1.015	1.008	1.235	1.309
PixD I 2f2	d₁	d₂	d₃	d₄	d₆	d₈	d₁₀	d₁₁
geo A	1.290	1.455	1.276	1.488	1.006	0.966	1.210	1.334
geo B	1.287	1.428	1.337	1.452	1.012	0.967	1.215	1.329
geo C	1.317	1.416	1.354	1.445	1.024	1.032	1.235	1.306
geo D	1.317	1.419	1.347	1.448	1.025	1.036	1.236	1.305
PixD I 3f1	d₁	d₂	d₃	d₄	d₆	d₈	d₁₀	d₁₁
geo A	1.290	1.454	1.277	1.488	1.007	0.965	1.209	1.334
geo B	1.287	1.429	1.337	1.452	1.012	0.967	1.214	1.328
geo C	1.315	1.417	1.353	1.445	1.024	1.032	1.234	1.306
geo D	1.319	1.415	1.355	1.444	1.025	1.050	1.238	1.303
PixD I 3f2	d₁	d₂	d₃	d₄	d₆	d₈	d₁₀	d₁₁
geo A	1.292	1.450	1.282	1.486	1.006	0.969	1.208	1.335
geo B	1.288	1.427	1.338	1.452	1.011	0.971	1.215	1.327
geo C	1.314	1.417	1.353	1.445	1.023	1.071	1.229	1.310
geo D	1.317	1.421	1.348	1.445	1.035	1.635	1.268	1.286

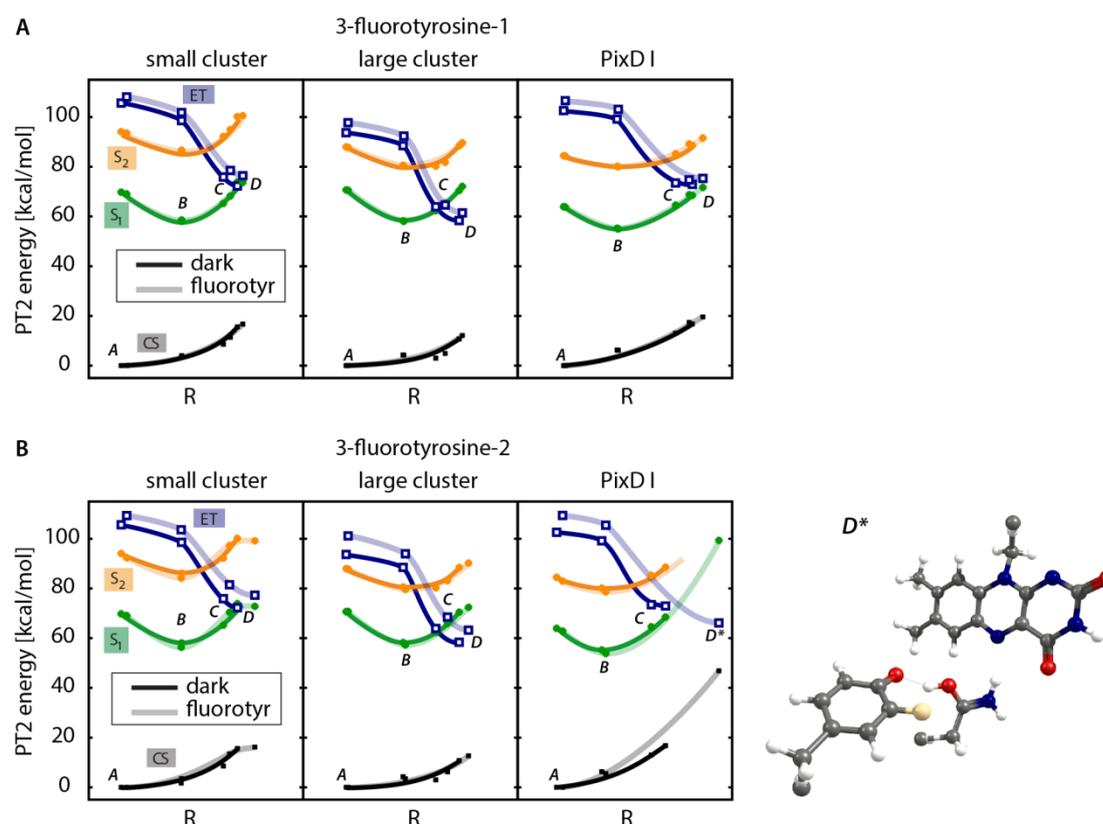


Figure 5.8. XMCQDPT2 and MS-CASPT2 energy profiles of the PET pathway in the 3-fluorotyrosine models and structure D^* in the PixD-I model. The energies at geometries A-D were obtained with (6,4)4//((6,4)4 calculations, with the following exceptions: small-cluster-3f1 geometry A (6,4)4//((4,3)3, small-cluster-3f2 geometry A (8,5)5//((4,3)3, PixD-I-3f2 geometry A (6,4)4//((4,3)3, PixD-I-3f2 geometry D^* (6,4)4//((2,2)2. The energy-shifting procedure outlined in Section 5.2.2, was used for the small-cluster-3f2 model as follows. The (8,5)5 energies at geometry A were compared to the (8,5)5 energies computed at geometry B. The energies at geometry A were then shifted such that the S_1 excitation energy at geometry B obtained with (8,5)5 matches that computed with (6,4)4. The energies from Figure 5.4 of the dark-state pathway with the native tyrosine are also plotted for comparison.

Overall, the PET pathway in the fluorotyrosine models is very similar to that of the dark-state models with the native tyrosine (in the following referred to as “wildtype” models): The S_1 and S_2 excitation energies are hardly affected by tyrosine fluorination (as expected), whereas the ET-state energies lie above those in the wildtype models at all optimized geometries A-D. This suggests a higher energy barrier for the electron-transfer reaction and thus slower S_1 fluorescence quenching. The results from the cluster and QM/MM models are consistent. It is interesting to note that a larger S_1 -ET gap at the Franck-Condon minimum correlates with a higher energy barrier of the electron transfer process upon the “chemical modification” of the model. This correlation was earlier discussed in Section 5.3.1 for the case of different dark-state models.

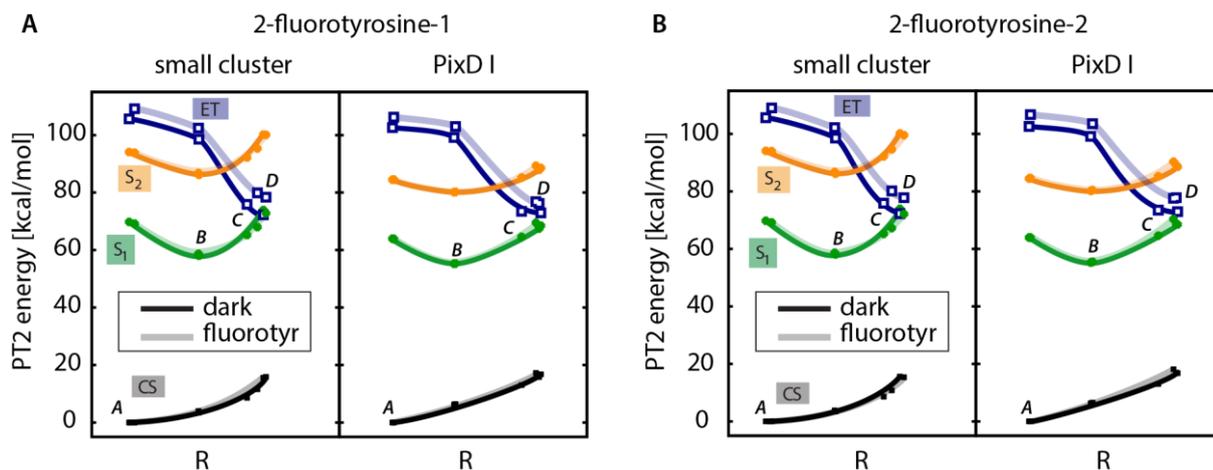


Figure 5.9. XMCQDPT2 and MS-CASPT2 energy profiles of the PET pathway in the 2-fluorotyrosine models. The energies at geometries A-D were obtained with the (6,4)4//((6,4)4 calculations, with the exception of the small-cluster-2f1 and -2f2 geometry A computed with (6,4)4//((4,3)3 and the PixD-I-2f1 and -2f2 geometry D computed with (6,4)4//((2,2)2. The energies from Figure 5.4 of the dark-state pathway with the native tyrosine are also plotted for comparison.

The increase of the ET state energy leads to a “delay” of geometries C and D along the reaction coordinate R as compared to the wildtype pathway. The fluorinated tyrosine models feature not only an increased ET state energy but also an increased acidity of the fluorinated tyrosine compared to the native one as suggested by the elongated tyrosine OH bond length (d_{OH}) in all models. In the 3-fluorotyrosine-2 QM/MM model, optimization of the ET state after the ET/S1 state crossing results in a spontaneous deprotonation of the Y11-3f2 radical instead of the formation of a zwitterionic biradical minimum. As a consequence, a minimum with the neutral Y11-3f2 radical, protonated Q53, and anionic flavin radical is formed, denoted as D* in Figure 5.8. Despite the deprotonation of the Y11-3f2 radical, the ET and CS states do not cross at D* as was the case in the wildtype dark state models at D. However, protonation of flavin N5 will lead to the ET/CS state crossing.

The obtained results, in particular the increase of the ET excited-state energy, are consistent with the experimentally observed longer flavin fluorescence life time of fluorotyrosine-containing BLUFs (Mathes, Stokkum, et al., 2012; Lukacs et al., 2014). It is noteworthy that the calculations presented here explain the longer fluorescence life time through the direct estimation of the change in the ET energy barrier upon tyrosine fluorination. This is an advantage over the experimental indirect estimation of the barrier: To estimate the barrier, the change in the redox potential of the fluorinated tyrosine in comparison to the native tyrosine has to be determined. The redox potential of tyrosine in aqueous solution is pH dependent, however, fluorination changes the acidity of tyrosine (Seyedsayamdost et al., 2006). Thus, the experimental determination of the change in the ET dynamics upon tyrosine fluorination is rather challenging.

5.3.5 PCET in the light state

Flavin in the BLUF light state features a 15-nm red-shifted absorption, shorter fluorescence life time than in the dark state and undergoes an internal photoconversion that explains its photostability. These photophysical properties and processes are investigated here using the approach that was developed to study the photoactivation of the BLUF dark state. Table 5.7 shows the spectral properties of the light state predicted by the small cluster, and the PixD-I and PixD-II QM/MM models. As expected, the S_1 excitation energy is red-shifted compared to the dark state; the QM/MM models predict a slightly more pronounced red shift compared to the cluster model. Similar to the dark state, there is only little difference in the predicted S_1 and S_2 energies and the Stokes shifts between the PixD-I and PixD-II QM/MM models. Noteworthy, the PixD II models predicts a slightly more pronounced red shift, indicating that describing the N35 residue on the QM level, as opposed to the MM level, enhances the hydrogen-bonding interactions with flavin. Concomitant to the red-shifted S_1 state, the ET excitation energy is substantially reduced in the light state compared to the dark state. The resulting smaller S_1 -ET gap at the Franck-Condon minimum suggests a faster fluorescence quenching process by electron transfer, consistent with experimental findings in (Mathes, Zhu, et al., 2012).

Table 5.7. The vertical S_1 and ET excitation energies at the FC minimum A, along with the S_1 -ET energy gap are listed together with the vertical S_1 excitation energy (emission) at the optimized S_1 minimum B and the resulting Stokes shift. XMCQDPT2 and MS-CASPT2 values are given for the cluster and QM/MM models, respectively. The red shift is the difference of the S_1 excitation energy at the dark- and light-state geometry A.

	abs [eV]	ET [eV]	S_1 -ET gap [eV]	emission [eV]	Stokes shift [nm]	red shift S_1 [nm]
small	3.00	4.17	1.18	2.36	112	4.2
PixD I	2.74	4.17	1.44	2.17	118	6.0
PixD II	2.71	4.35	1.64	2.18	111	7.8

Figure 5.10 summarizes the results of the light-state photoreaction calculations performed with the small cluster and PixD I models. The S_1 minimum B and the ET/ S_1 crossing C are similar to those in the dark state both with regard to energy and geometry. However, geometry optimization in the ET state starting from C yields a spontaneous deprotonation of the Y11 radical and leads directly to the ET/CS state crossing, geometry F. Structure F contains the neutral Y11 radical, anionic flavin radical and protonated Q53. After F, further optimization of the ET state, in the first electronic state, leads back to the Franck-Condon minimum. To reach the neutral biradical minimum H, similar to the dark state, the flavin anionic radical needs to be protonated. However, and in drastic contrast to the dark state, the ET state energy is hardly lowered at minimum H with respect to F (only by 4 kcal/mol in the small cluster). The ET state

at the crossing point F is substantially stabilized with respect to C, by 40 kcal/mol (small cluster) and by 33 kcal/mol (PixD I model), which represents an even larger stabilization than that at the final biradical minimum H in the dark-state pathway, as shown in Section 5.3.2.

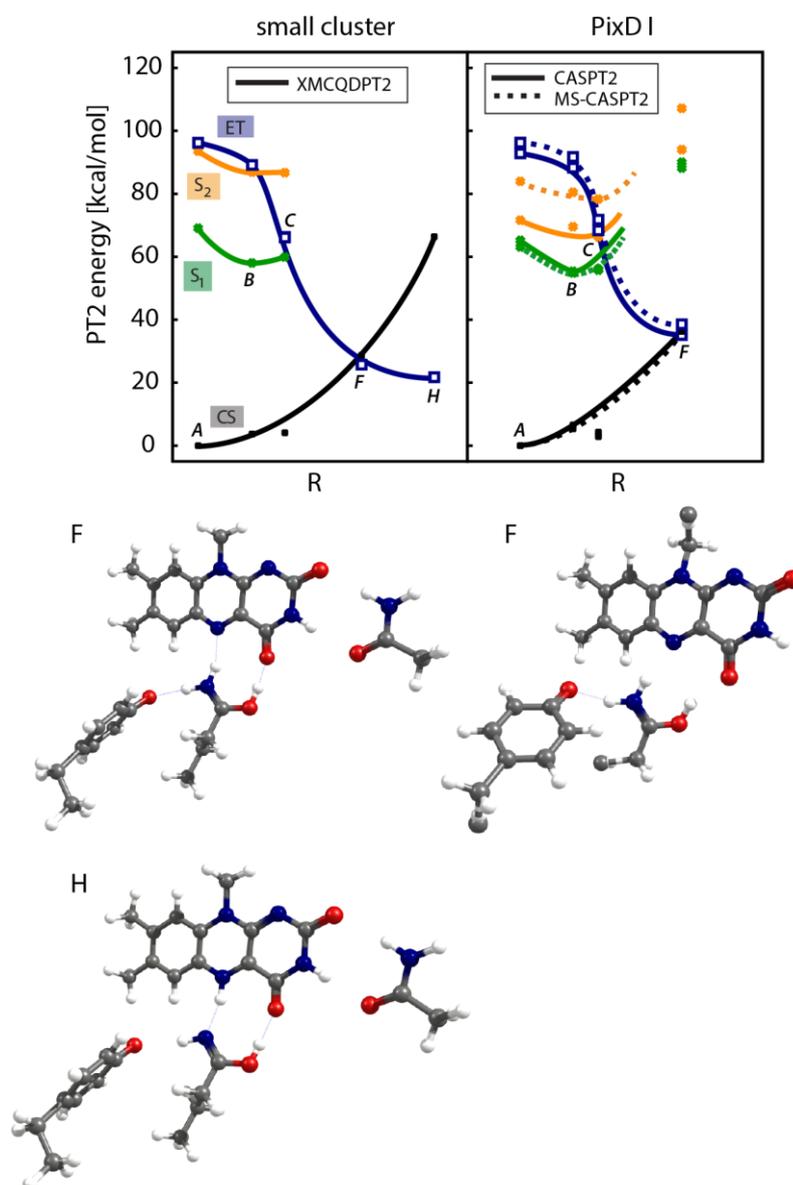


Figure 5.10. Full XMCQDPT2 and MS-CASPT2 energy profiles of the PCET pathway of the light state. Structures F and H in the small cluster and PixD-I models are also shown. The energies at geometries A-C were obtained with the (6,4)4//((6,4)4 calculations, in the small cluster model at geometries F and H with (2,2)2//((2,2)2; in the PixD-I QM/MM model at geometries A-F with (6,4)4//((6,4)4.

Similar to the dark-state PT-III reaction in the small cluster model, a hydrogen-bond rupture and rotation of the N35 residue occurs at the ET/CS state crossing F. This rotation is not expected to influence the observed energy trends to a significant extent, similarly to the case of

the dark-state photoreaction, discussed in Section 5.3.3. In the PixD-I QM/MM model the N35 conformation does not change (N35 is part of the MM subsystem). For the light state PCET pathway a slightly modified reaction coordinate was chosen to plot the energy diagram in Figure 5.10 because of the different arrangement of the hydrogen bonds compared to the dark state: $R=d_1-d_2+d_3-d_4+0.1 \times (-d_5+d_6-d_{13}+d_{14})$. Table 5.8 lists the relevant interatomic distances. The geometry changes of the respective flavin bonds are very similar to those in the dark state pathway (except for the slightly less pronounced changes in d_3 and d_4).

Table 5.8. Selected interatomic distances (defined in Figure 5.3) in Å along the PCET pathway.

small	d1	d2	d3	d4	d5	d6	d13	d14
geo A	1.276	1.466	1.274	1.495	2.366	1.009	1.939	0.961
geo B	1.273	1.440	1.334	1.453	2.211	1.012	1.897	0.965
geo C	1.293	1.431	1.347	1.445	2.169	1.015	1.598	1.013
geo F	1.310	1.426	1.342	1.436	1.897	1.025	1.008	1.963
geo H	1.301	1.425	1.342	1.447	1.021	1.904	1.007	2.362
QM/MM	d1	d2	d3	d4	d5	d6	d13	d14
geo A	1.289	1.454	1.278	1.487	2.552	1.009	1.821	0.972
geo B	1.285	1.431	1.334	1.449	2.444	1.011	1.807	0.974
geo C	1.301	1.428	1.338	1.447	2.387	1.013	1.568	1.023
geo F	1.315	1.423	1.349	1.438	2.133	1.025	1.014	1.788

An important finding of the light-state pathway is a much stronger ET-state energy stabilization upon reaching the ET/CS state crossing geometry F than in the respective reaction of the dark state. This feature, the key difference between the dark- and light-state photoreactions, is related to the lowered excitation energy of the ET state at the light-state FC point and also to the increased basicity of the imidic Q53 =NH group interacting with the electron donor Y11. The latter characteristic is observed experimentally as a downshift of the Y11-OH stretching frequency in the light state (Iwata et al., 2011) and is reproduced by computational models (Khrenova et al., 2013). The neutral biradical minimum H is quite similar in both the dark and light state photoreactions, in accord with the experimental observations and argumentation in (Mathes, Zhu, et al., 2012). The difference between the two neutral biradical intermediates in the dark and light states is the orientation of the tautomeric glutamine: it forms a strong hydrogen bond to flavin O4 in the light state but not in the dark state. In the dark state, the tautomeric glutamine has to rotate in order to establish the hydrogen bond with the flavin O4 atom. This rotation decreases the energy of the neutral biradical by 11 kcal/mol (according to the energy difference of the ET state at H in the dark and light states). In the light state, the energy barrier to reach the Franck-Condon minimum through the state crossing F is small and thus it is possible that the system takes that route back to the light-state Franck-Condon minimum. Because of the

different topology of the ET/CS state crossing in the dark and light states, it is also conceivable that in the light state minimum H is not reached at all and that the system relaxes back to the FC directly after reaching the conical intersection. In summary, the relatively small energy difference between F and H in the light state and the relatively large energy difference between F and H in the dark state is in line with the experimentally observed differences in the two photoreactions.

In (Sadeghian et al., 2010) the photoreaction of the BLUF light state containing the tautomeric, non-rotated Q53 was computed with CC2//TD-DFT and the authors concluded that a PCET reaction underlies the photostability of the light state because of a ET/CS state crossing found “early” along the PT reaction pathway. “Early” means that the ET/CS crossing is already reached before the protonation of Q53 is completed. The results here give a more detailed view because in the present work the ET/CS state crossing was located and the stabilization of the ET state along the photoreaction coordinate was estimated, whereas in (Sadeghian et al., 2010) this was not possible due to the limitation of the single-reference methodology. The present study reveals that the photostability does not stem from an “early crossing”, because in fact a complete protonation of Q53 is required to reach the ET/CS crossing F, which mediates the decay of the excited state. The crossing geometry F in both the dark and the light states contains similar chemical species (in particular the protonated Q53). Rather the large energy stabilization of the ET state in the light state results in the crossing F being much lower in energy compared to that in the dark state. As a consequence, the formation of the neutral biradical in the light state does not lead to further stabilization, rendering the relaxation back to the light-state Franck-Condon geometry likely.

On the basis of the time-resolved ultrafast spectroscopy studies, Kennis and coworkers emphasize that in the dark state photoreaction the ET and PT reactions proceed in a stepwise manner whereas in the light state photoreaction they proceed in a concerted manner (Mathes, Zhu, et al., 2012). Their argument is mainly based on the observation of the transiently formed anionic and neutral flavin radicals in PixD. The presented results and the fact that in BLUFs other than PixD no anionic or neutral biradical intermediates could be resolved are shedding new light on the concerted versus stepwise nature of the ET and PT reactions in BLUF. The light-state photoreaction is completed in 1 ps, much faster than the dark state photoreaction (1 ns), indicating barrierless ET and PT processes and a concerted PCET reaction. In the dark-state photoreaction the ET and PT reactions may have some barrier so that intermediate species may be formed and observed, however, the presented calculations showed that proton transfer must follow the electron transfer reaction because the zwitterionic biradical D is too high in energy and must be stabilized. Thus, in that sense the dark state photoreaction can also be regarded as a concerted PCET reaction with the ET energy barrier (without the PT energy barrier) being rate-determining.

PCET is a central biological photochemical reaction with functional relevance for example in light harvesting or in the photostability of DNA, reviewed, for example, in (Migliore et al., 2014;

Sobolewski & Domcke, 2006). However, multi-configurational PT2//CASSCF calculations on photoinduced PCET reactions have only been carried out for rather simple processes that do not lead to photochemical changes. One prominent example is DNA photostability discussed in (Schultz et al., 2004; Sobolewski et al., 2005). The authors demonstrated that Watson-Crick hydrogen bonds provide the efficient deactivation channel of the excited states in DNA bases in the course of the photoinduced PCET reaction: The photoreaction is controlled by a direct proton transfer from the electron donor to the electron acceptor. Similar to the BLUF case, in a Watson-Crick complex of two nucleobases the ET state lies higher in energy than the locally-excited state at the Franck-Condon minimum, but during the reaction the ET and locally-excited states cross. The subsequent proton-transfer reaction leads to the ET/CS state crossing, from where the system relaxes back to the Franck-Condon minimum. The photostable reaction of the BLUF light state is similar yet more complex because the direct hydrogen-bonding contact between electron donor and acceptor is broken by the conserved glutamine Q53. Moreover, in the BLUF dark state the same photoinduced PCET reaction is tuned for photoactivation, where the special hydrogen-bonding network allows for opening the photoactive relaxation channel along which the photoproduct is formed. The tautomeric glutamine in the BLUF light state “closes” the photoactive channel and ensures photostability. This work gives first hints how the interactions with the Q53 side chain switch the PCET reaction in BLUF between photoactive and photostable ones. The specific hydrogen-bonding network is a critical element of this mechanism. To date, no similar example in biology was studied and explained, to the best of my knowledge. Given that there is a significant variation in dynamics inside the chromophore-binding pocket among BLUF proteins, it is interesting to elaborate how the dynamics “selects” the right pathway for the photoreaction.

5.3.6 Electron transfer rate estimates using excited-state energies and Marcus theory

To deduce electron transfer rate estimates from the presented pathways, the seminal Marcus rate theory (Marcus & Sutin, 1985) is used. The ET rate in the non-adiabatic, high-temperature limit is given by

$$k_{ET} = \frac{2\pi}{\hbar} |H_{DA}|^2 \frac{1}{\sqrt{4\pi\lambda k_B T}} e^{-\frac{(\Delta G^0 + \lambda)^2}{4\lambda k_B T}} \quad (5.1)$$

where H_{DA} is the electronic coupling matrix element between the electron donor and acceptor, λ is the reorganization energy, ΔG^0 is the standard free energy of the reaction. The parameter λ is the energy needed for the reorganization of the environment without electron transfer; its typical value ranges between 0.7 and 1.7 eV in biological systems (Moser et al., 1992; Venturoli et al.,

1998; Liu et al., 2012; Woiczikowski et al., 2011). The magnitude of the coupling H_{DA} is reported to change in a large range, between 0.3 meV and 0.1 eV (Venturoli et al., 1998; Blancafort & Voityuk, 2007; Liu et al., 2012; Woiczikowski et al., 2011). The non-trivial dependence of the three parameters H_{DA} , λ , and ΔG^0 on the properties of the system, and possibly also existing correlations between them, discussed for example in (Venturoli et al., 1998), makes their experimental and theoretical evaluation challenging. Therefore, from Equation (5.1) only an order of magnitude estimate for the ET rate can be obtained. Thus, the logarithmic version of the formula is used here to estimate the trends. With $k_B T = 0.0257 \text{ eV}$ (room temperature) and assuming $-0.5 \log(4\pi\lambda k_B T) \approx [0.1 \dots 0.4]$, the logarithmic rate expression reads:

$$\log(k_{ET}) \approx 16.2 + 2 \log(H_{DA}) - 4.22 \frac{(\Delta G^0 + \lambda)^2}{\lambda} \quad (5.2)$$

In the ‘‘optimal’’ case with $\Delta G^0 = -\lambda$ turning the last term into zero, the physical limit of the maximum ET rate is reached. This case is used here to estimate the coupling H_{DA} . The fastest ET rate at a van der Waals contact constitutes about 10^{-13} s^{-1} (Moser et al., 1992), leading to the requirement $2 \log(H_{AB}) = -3.2$, from which the coupling $H_{DA} = 0.025 \text{ eV}$ is deduced as an upper limit.

Equation (5.2) with the estimated H_{DA} is used to determine the electron transfer rate, controlling the flavin fluorescence life time in BLUF. The energy estimates from the results of the large cluster model are used. The reorganization energy λ is estimated as the XMCQDPT2 energy difference between the S_1 state at geometry D and B (see Figure 5.4) and ΔG^0 is estimated as the XMCQDPT2 energy difference between the ET state at geometry D and the S_1 state at geometry B, ignoring the zero-point vibrational energy and entropic contributions. The latter assumption implies that the change in the total energy (computed here) is the dominant term in the change of the free energy G and for an order-of-magnitude estimate of the ET rate this approximation should be valid. Thus, with $\lambda = 0.53 \text{ eV}$ and $\Delta G^0 = 0.0 \text{ eV}$, an ET rate of the order of 10-100 ps is obtained (for H_{DA} ranging between 0.009 eV and 0.025 eV), which, as a matter of fact, is in good agreement with experimental observations. In the case of the 3-fluorotyrosine-1 model, both parameters determining the electron-transfer activation energy slightly increase, $\lambda = 0.61 \text{ eV}$ and $\Delta G^0 = 0.15 \text{ eV}$, yielding a one order of magnitude slower ET rate, which is also in agreement with experimental observations. Yet in the presented calculations the effect of tyrosine fluorination on the ET rate is somewhat overestimated: the 3-fluorotyrosine-containing PixD and AppA proteins feature five-times and three-times slower fluorescence quenching rates than the wildtype proteins, respectively (Mathes, Stokkum, et al., 2012; Lukacs et al., 2014). It is noteworthy that also the reorganization energy λ changes in the fluorotyrosine-containing model, in contrast to the assumption often made (for example in (Lukacs et al., 2014)) that λ stays constant and only ΔG^0 is modified by mutations or other alterations in a given protein.

5.4 Summary and conclusions

For the first time, photoreaction pathways in BLUF along which the excited flavin chromophore relaxes were characterized and analyzed in detail in this chapter. This was feasible because of the developed PT2//CASSCF protocol that enabled mapping the excited-state potential energy surfaces of the relevant flavin local-excited S_1 and the tyrosine-flavin ET states, in both the dark state and the light state of the photoreceptor and also in fluorotyrosine-modified models in a totally consistent manner. The calculations reveal that the photoreaction in BLUF is rather complex (the most complex photoreaction hitherto characterized using excited-state calculations) and involves a sequence of PCET reactions. The photoreaction starts with the population of the S_1 state through blue-light absorption. At the FC geometry, the ET state lies significantly higher in energy than S_1 . However, geometry relaxation in the S_1 and ET states proceeds, to a certain extent, along the same reaction coordinate and a pathway along which the S_1 and ET states cross was identified with the ET energy being significantly lowered at the crossing point with respect to that at the FC point. This pathway features a photoinduced electron transfer reaction, yielding a zwitterionic (charge-separated) radical-pair intermediate. The energy barrier separating the excited flavin S_1 and the radical-pair minima is estimated at the ET/ S_1 crossing geometry obtained upon geometry optimization in the ET state starting from the FC minimum. The height of the energy barrier (the energy of the ET/ S_1 crossing) depends on the ET-state energy in a particular BLUF model. The estimates of the electron-transfer rates obtained here suggest that this barrier determines the life time of flavin fluorescence in BLUF. In particular, in the photoreaction of the fluorotyrosine-containing BLUFs, the higher ET-state energy results in a higher energy barrier for the PET reaction thus explaining the measured longer fluorescence life time. Analogously, the lower ET-state energy results in a lower PET energy barrier and thus a shorter flavin fluorescence life time. The photoinduced electron transfer is followed by proton transfer from the electron donor tyrosine to the excited electron acceptor flavin, mediated by hydrogen bonds involving the conserved glutamine residue. The special position of glutamine with respect to tyrosine and flavin opens several possible PT channels, leading either to the formation of a photoproduct (in the case of the photoactive BLUF dark state) or to the recovery of the initial FC structure (in the case of the photostable BLUF light state).

In the dark-state photoactivation reaction, three PT pathways were computationally characterized in this chapter. All these PT reactions stabilize the ET state and concomitantly increase the energy of the closed-shell ground state so that eventually the system reaches the ET/CS conical intersection. At the conical intersection, two channels (the ET and the CS channels) compete: The CS-state channel corresponds to the dark-state recovery, which reduces the efficiency of photoactivation, whereas the ET channel eventually leads to the formation of the photoproduct, which is the 15-nm red-shifted light-state of BLUF. The mechanism

controlling the branching between these two channels is yet to be unraveled. Two PT reactions in the photoactive channel after the ET/CS state crossing produce a neutral radical-pair intermediate containing the neutral flavin and tyrosyl radicals and a tautomeric glutamine on the ground-state potential energy surface. In the third PT reaction, the protonated glutamine spontaneously rotates approximately 180° which prevents the protonation of the flavin anionic radical and thus the formation of the neutral flavin radical. Instead, mediated by the rotated protonated glutamine, a second PCET reaction leads to the recombination of the radicals and the formation of the photoproduct (i.e. the BLUF light state) in the CS ground state. The photoproduct contains the reduced tyrosine, oxidized flavin and the tautomeric glutamine. The S_1 absorption of this species is 15 nm red-shifted compared to the dark state minimum.

In the photostable reaction of the light-state, only one PT pathway was identified. The proton transfer (from tyrosine to the tautomeric glutamine) stabilizes the ET state to a substantially larger extent than it was found during the dark-state photoreaction. Similar to the dark-state photoreaction, a neutral biradical intermediate may form after the system passes through the ET/CS conical intersection. However, in contrast to the dark-state reaction, this intermediate lies close in energy to the crossing geometry and therefore is unlikely to be populated to the same extent as during the dark-state reaction. The different topology of the excited-state surface in the vicinity of the respective ET/CS crossings in the dark- and light-state reactions gives initial hints to how photoactivation and photostability are governed by the PCET photorelaxation processes in BLUF.

The thermal dark-state recovery reaction, involving the reverse tautomerization of the glutamine in the electronic ground state, has not been a subject of computational studies so far. However, on the basis of the presented calculations, a sequence of PCET reactions in the ground state can be proposed to accomplish the reverse tautomerization of the glutamine side chain. Therefore, the redox potential of the tyrosine-flavin redox pair should control also the dark-state recovery.

6 Conclusions and Outlook

Over the recent years, computations proved an indispensable complement to experimental studies on photoreceptor proteins aiding the interpretation of spectroscopy data, providing detailed models of protein dynamics and facilitating our understanding of the photochemical mechanisms. Quantum-chemical methods allow following the relaxation of the excited chromophore and treating its interactions with the protein environment. Sophisticated methods and models are indeed needed to explain the observed complex reactions used by photoreceptor proteins to sense light. With the discovery of the flavin-binding photoreceptors LOV, cryptochrome, and BLUF a new type of photoreaction was identified, which invokes PCET rather than the cis-trans photoisomerization of the chromophore like in the rhodopsin photoreceptors. PCET in general is a key reaction in many biologically relevant processes from respiration to cell signaling and photoinduced PCET plays a major role not only in photoactive reactions but also in mechanisms underlying the photostability of biomolecules. The BLUF protein is an exceptionally interesting system as it features both photoactive and photostable states, among which it can switch, as this thesis identified, by means of the tautomerization of the conserved glutamine residue. Both the photoactive dark state and the photostable light state contain the oxidized flavin chromophore. In the light state, the flavin absorption is 15 nm red-shifted, indicative of the altered hydrogen-bonding network that, as proposed here, contains the tautomeric glutamine. In this work I presented the first detailed and systematic computational study of photoinduced PCET underlying both photoactivation and photostability in BLUF, covering all most important experimental observations.

In order to perform such a complex study, first the computational method had to be established. The major computational requirements to study PCET in BLUF are (i) the possibility to treat the electronic ground and excited states of different chemical character (for example the flavin S_1 excited state and the tyrosine-flavin ET state) in a balanced way; (ii) the possibility for geometry optimization in any low-lying excited state of interest to map the pathway along which the energy is dissipated after photon absorption; and (iii) a multi-reference treatment in order to localize state-crossing geometries and to describe the formation and recombination of biradicals.

These requirements are only fulfilled by the CASSCF method. The CASSCF wave function includes all possible electronic configurations that can be constructed from distributing the electrons among the MOs within the active space. This procedure yields a qualitatively correct wave function that accounts for static electron correlation. In order to also describe dynamic electron correlation, multi-reference second-order perturbation theory is used with the CASSCF wave function as the zero-order wave function. The PT2//CASSCF approach is a computationally highly expensive method and its conventional way of selecting the active space (the full $\pi\pi^*$ valence shell) limits the calculations to small molecular systems. In Chapter 2, I demonstrated that choosing only the principal orbitals describing the single-electronic excitations of interest (that is flavin HOMO, LUMO and tyrosine HOMO in the case of the BLUF photoreaction), makes the PT2//PO-CASSCF method also practical in rather large molecular complexes as the BLUF photoactive triad. Both energies and geometries computed with POCAS were shown to be consistent and in agreement with those computed with other methods.

A suitable model for the treatment of the intermolecular interactions in the BLUF photoreceptor had to also be established. To this end, an extensive comparison between various cluster and QM/MM models of BLUF was carried out in Chapter 3. The evaluation of the intermolecular interaction effects in the computational models aids the BLUF mechanistic studies with accurate energy computations. Comparing different models, the energy of the flavin S_1 state was found to be insensitive to the flavin environment, whereas the ET-state energy is highly sensitive on the geometry of the active site, as well as to the local and also the more distant protein environment. The local environment in the cluster models red- or blue-shifts the ET state energy, depending on the specific intermolecular interactions. In the largest size cluster model the ET state is substantially red-shifted by the electrostatic interactions and also because of charge-transfer interactions between the photoactive triad and the environment. In the QM/MM models, the energy of the S_1 state does not change significantly upon QM/MM embedding, however, the ET-state energy is strongly blue-shifted, because of the high polarity of this transition. The energy of polar transitions is highly dependent on the employed electrostatic atomic-charge model, i.e. on the MM point charges. The employed charge model is difficult to validate against experimental data as the ET excitation energies are not observed directly. In that view, a superiority of the QM/MM model over the cluster model, which does not account for the major part of the protein, cannot be established. At the same time, the full QM treatment of the local interactions of the chromophore in the cluster model certainly results in an improvement of the model. Therefore, for comparative studies that are more focused on the effect of the local environment of flavin (the hydrogen-bonding interactions), large cluster models including all relevant side chains may be superior to the QM/MM models.

Next, using cluster models including a rather extended part of the protein environment around the photoactive triad, I examined the structural inhomogeneity in BLUF and evaluated, from a theoretical point of view by formulating suitable criteria, the conflicting assignments of

the dark-state structure leading to different mechanistic proposals in Chapter 4. The two main hypotheses discussed in the literature consider light-induced rotation as opposed to tautomerization of the conserved glutamine in the photoactive triad. The changes of the glutamine structure are considered concomitant with a putative conformational switch between the W_{in} and W_{out} conformations of the $\beta 5$ strand. To resolve the controversy, three criteria for the BLUF functional states were formulated: (i) the agreement of the proposed stationary dark structure with the X-ray electron density; (ii) the red shift of the light state absorption as compared to the dark state, which has to be reproduced; and (iii) the existence of a high energy barrier between the dark- and light-state structures in the ground state to ensure photosensitivity (low thermal noise) of BLUF activation. A comparison of the optimized models with the experimental electron density and complete dihedral-angle energy scans along the glutamine rotation coordinate rendered the glutamine rotation mechanism clearly inconsistent with the three criteria, whereas the glutamine tautomerization mechanism is in agreement with them. The dihedral-angle energy scans revealed that glutamine rotation may also occur in the dark state (i.e. without photoexcitation) because of the low energy barriers separating the rotamers. The comparison of energy scans in the W_{in} and W_{out} models hints at the W_{out} conformation to correspond to the dark state and the W_{in} conformation to correspond to the light state of BLUF. Further extended models including the whole $\beta 5$ strand and both tryptophan and methionine side chains in both W_{in} and W_{out} models will allow the computation of the relative energies and, on that basis, a more definite conclusion concerning the distinct conformations of the functional states.

Glutamine tautomerization is a highly peculiar reaction because so far there are no examples of the imidic forms of glutamine to play a role in biochemistry. A low energy barrier separates the imidic forms from the commonly known lower-energy amide form in a protic environment like water. However, in the special protein environment of BLUF, the energy barrier increases and the imidic glutamine can be stable for seconds to hours. In fact, the tautomerization of the glutamine side chain ensures that a chemically distinct species is formed in the BLUF photoproduct as compared to the dark state. Therefore, glutamine tautomerization defines the photoreaction coordinate in BLUF.

After establishing the methods, models, and general theoretical considerations regarding a photoreaction involving the tautomerization of the glutamine side chain, I investigated the complete excited-state photoreaction pathway. Figure 6.1 summarizes the photoreaction pathway calculations presented in Chapter 5.

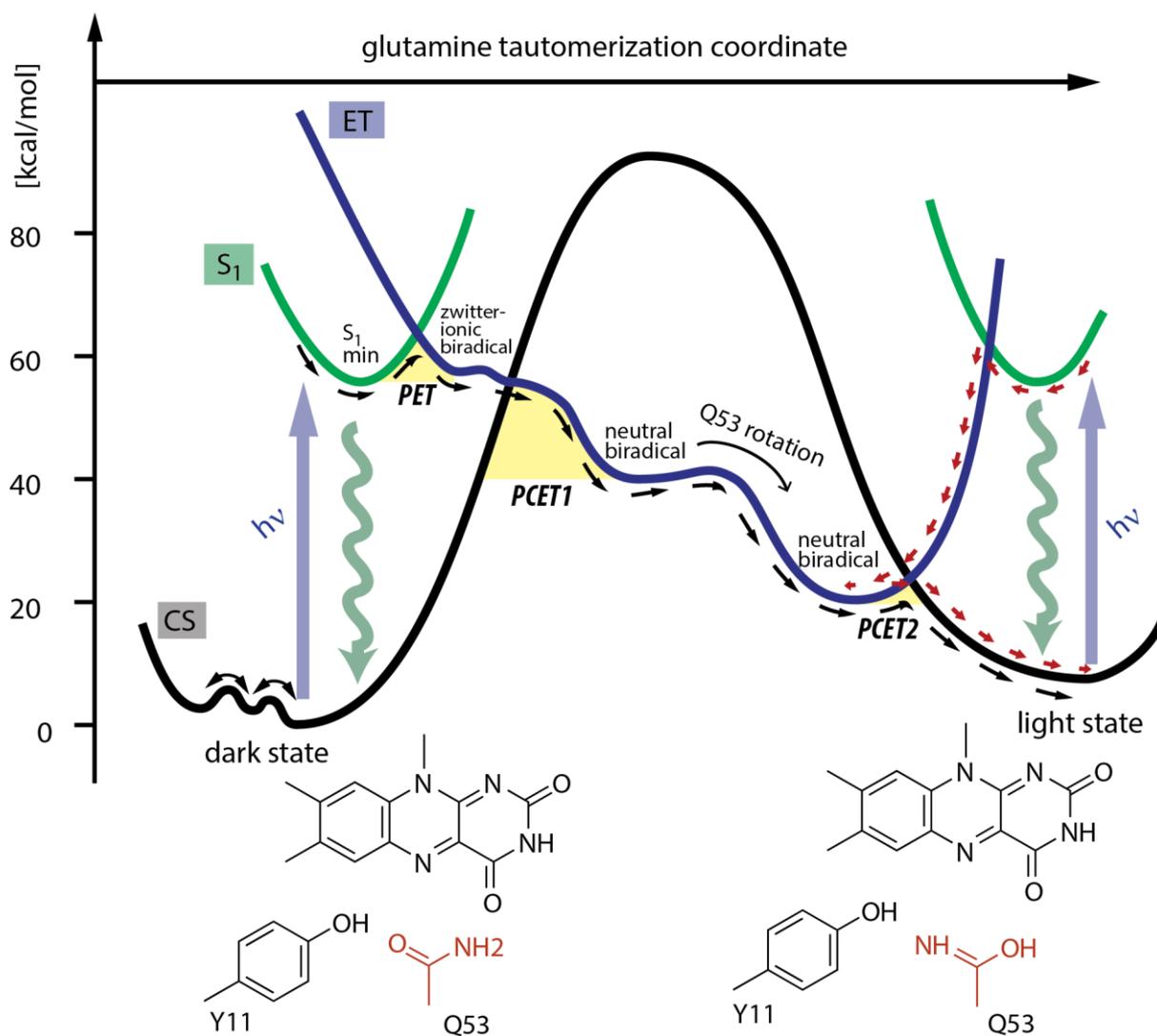


Figure 6.1. Photoreaction of the BLUF dark and light states: black arrows indicate the photoactive dark-state pathway, red arrows the photostable light-state pathway.

The photoactive reaction of the dark state starts with flavin light absorption followed by electron transfer from the conserved tyrosine residue to the excited flavin, quenching the S_1 fluorescence and leading to a zwitterionic (charge-separated) biradical intermediate. The ET/ S_1 state crossing mediates this process. Subsequent PT reactions involving the glutamine side chain lead to the ET/CS state crossing, which is rather close in energy to the zwitterionic biradical minimum. Two PT reactions – deprotonation of the tyrosine electron donor and protonation of the flavin electron acceptor mediated by the glutamine – yield the neutral biradical intermediate in the ET state, which contains the tautomeric glutamine. This intermediate is further stabilized by the rotation of the tautomeric glutamine, to finally form the hydrogen bond via its OH group with the flavin O4 atom. Recombination of the neutral radical pair involves again two PT steps, leading to the formation of the photoproduct light state and eventually the signaling state of the

protein. The light state is stabilized by the interactions between the flavin and the tautomeric glutamine and has a slightly red-shifted flavin absorption maximum and also a red-shifted ET state. Following blue light absorption, the same ET process as in the dark state quenches the S_1 fluorescence in the light-state, yet with the faster kinetics that is determined by the lowered ET-state energy. The concomitant PT reactions yield the ET/CS state crossing much lower in energy than that in the dark state, whereas the energy of the neutral radical pair intermediate does not change (because, chemically, it is the same intermediate formed in both cases). The lower energy ET/CS crossing reached by the photoexcitation of the light state apparently has a higher probability to decay back to the light-state Franck-Condon minimum. The modulation of the redox potential of the flavin/tyrosine pair comes about as an up- or downshift of the ET-state energy curve in Figure 6.1, whereas the CS and S_1 energy curves stay practically unaffected. Using the Marcus theory, Section 5.3.6 demonstrated that the upshift of the ET energy upon tyrosine fluorination, is consistent with the experimentally observed longer flavin fluorescence life time.

Consistent mechanistic conclusions were obtained with several cluster and QM/MM models, which provide further validation of the chosen computational approach and ensures a certain robustness of the mechanistic predictions derived in this extensive study. Disappointingly, for questions addressing the specific protein environment, the standard QM/MM models built based on the crystal structures (without accounting for protein dynamical effects) proved inappropriate or difficult to handle, indicating that the existing QM/MM protocols require further improvements. One of the examples is the failure to reproduce the flavin fluorescence life times in PixD and BlrB. The fluorescence quenching rate by PET was predicted to be shorter in BlrB than in PixD, in total contrast to experiment. Another important aspect is that the structural relaxation of the QM subsystem in the QM/MM geometry optimizations was rather limited compared to the cluster models, which complicated the computations of the photochemical pathways. Overall, the pathways computed in the QM/MM models were similar to the ones in the cluster models with the same composition of the QM subsystem, so that no protein-environment induced effect that may influence the photoreaction yield could be identified.

The presented photoreaction pathways lay the foundation for further studies elucidating the details of the BLUF photoactivation mechanism: conical intersection analysis, photodynamics, redox tuning etc. All of these upcoming studies must rely on the knowledge of the potential energy surfaces, which is established in great detail here. The developed computational framework allows for the exploration of other pathways involving electron-transfer states from alternative electron donors to flavin, like tryptophan or methionine. In BLUF, photoinduced electron transfer involving other side chains and competing energy-dissipation channels are relevant for the explanation of the observed complex fluorescence decay kinetics. It is interesting to elaborate on how the protein dynamics contributes to selecting the pathway ensuring the highest quantum yield of photoproduct formation or preventing photochromism.

In conclusion, the studies presented in this thesis have shown that an energetically accessible optical-dark ET state mediates the decay of the flavin spectroscopically-observed S_1 state during the BLUF dark and light state photoreactions. The ET state is strongly stabilized by two proton transfer reactions which eventually lead to glutamine tautomerization. Along the proton-transfer pathway, the ET state connects the S_1 state with the ground state via two state crossings. The intermolecular ET reaction coupled to specific PT reactions plays the same role in BLUF as the cis-trans photoisomerization of the chromophores in rhodopsin, phytochrome, and xanthopsin photoreceptors. The same PT reaction underlies the photostability of the BLUF light-state. BLUF is therefore a highly interesting example where PCET is realized as the fundamental mechanism for both photoactivation and photostability. Further in-depth investigations of the properties of this unique switch operating through the peculiar reaction of glutamine tautomerization will help to obtain a more complete understanding of the fundamental role of PCET reactions in biology.

Abbreviations

BLUF	blue light using FAD sensor
BSSE	basis set superposition error
CAS	complete active space
CASPT2	complete active space perturbation theory of second order
CASSCF	complete active space self-consistent field
CS	closed shell
DFT	density functional theory
ESPF	electrostatic potential fitting
ET	electron transfer
FC	Franck-Condon
Fl	flavin
GS	ground state
HOMO	highest occupied molecular orbital
LA	link atom
LF	lumiflavin
LUMO	lowest unoccupied molecular orbital
MD	molecular dynamics
MO	molecular orbital
NMR	nuclear magnetic resonance (method)
PCET	proton-coupled electron transfer
PDB	protein data bank
PET	photoinduced electron transfer
POCAS	principal-orbital complete active space
PT	proton transfer
PT-I	proton-transfer pathway I
PT-II	proton-transfer pathway II

PT-III	proton-transfer pathway III
PT2	perturbation theory of second order
RMSD	root mean square displacement
QM/MM	hybrid quantum-mechanical molecular-mechanical model
RF	riboflavin
S ₁	first excited singlet state
S ₂	second excited singlet state
TD-DFT	time-dependent density functional theory
W _{in}	BLUF structure with tryptophan pointing into the flavin-binding pocket
W _{out}	BLUF structure with the conserved tryptophan solvent exposed
XMCQDPT2	extended multi-configuration quasi-degenerate perturbation theory (2 nd o.)

One-letter codes for the amino acids

A	alanine
R	arginine
N	asparagine
D	aspartic acid
C	cysteine
E	glutamic acid
Q	glutamine
G	glycine
H	histidine
I	isoleucine
L	leucine
K	lysine
M	methionine
F	phenylalanine
P	proline
S	serine
T	threonine
W	tryptophan
Y	tyrosine
V	valine

References

- ALTOÈ, P., CEMBRAN, A., OLIVUCCI, M. & GARAVELLI, M. (2010). Aborted double bicycle-pedal isomerization with hydrogen bond breaking is the primary event of bacteriorhodopsin proton pumping. *Proceedings of the National Academy of Sciences* **107**, 20172–20177.
- ANANIKOV, V. P., MUSAEV, D. G. & MOROKUMA, K. (2010). Real size of ligands, reactants and catalysts: Studies of structure, reactivity and selectivity by ONIOM and other hybrid computational approaches. *Journal of Molecular Catalysis A: Chemical* **324**, 104–119.
- ANDERSON, S., DRAGNEA, V., MASUDA, S., YBE, J., MOFFAT, K. & BAUER, C. (2005). Structure of a Novel Photoreceptor, the BLUF Domain of AppA from *Rhodobacter sphaeroides*. *Biochemistry* **44**, 7998–8005.
- ANDERSSON, K., MALMQVIST, P.-Å. & ROOS, B. O. (1992). Second-order perturbation theory with a complete active space self-consistent field reference function. *The Journal of Chemical Physics* **96**, 1218–1226.
- ANGELI, C., CIMIRAGLIA, R., EVANGELISTI, S., LEININGER, T. & MALRIEU, J.-P. (2001). Introduction of n-electron valence states for multireference perturbation theory. *The Journal of Chemical Physics* **114**, 10252–10264.
- AQUILANTE, F., DE VICO, L., FERRÉ, N., GHIGO, G., MALMQVIST, P.-ÅKE, NEOGRÁDY, P., PEDERSEN, T. B., PITOŇÁK, M., REIHER, M., ROOS, B. O., SERRANO-ANDRÉS, L., URBAN, M., VERYAZOV, V. & LINDH, R. (2010). MOLCAS 7: The Next Generation. *Journal of Computational Chemistry* **31**, 224–247.
- BARENDS, T. R. M., HARTMANN, E., GRIESE, J. J., BEITLICH, T., KIRIENKO, N. V., RYJENKOV, D. A., REINSTEIN, J., SHOEMAN, R. L., GOMELSKY, M. & SCHLICHTING, I. (2009). Structure and mechanism of a bacterial light-regulated cyclic nucleotide phosphodiesterase. *Nature* **459**, 1015–1018.
- BECKE, A. D. (1993). Density-functional thermochemistry. III. The role of exact exchange. *The Journal of Chemical Physics* **98**, 5648–5652.

- BERENDSEN, H. J. C., POSTMA, J. P. M., VAN GUNSTEREN, W. F. & HERMANS, J. (1981). Interaction models for water in relation to protein hydration. In *Intermolecular forces*, Pullman, B. (Ed.), pp. 331–342. Dordrecht: D. Reidel Publishing Company.
- BERNARDI, F., OLIVUCCI, M. & ROBB, M. A. (1996). Potential energy surface crossings in organic photochemistry. *Chemical Society Reviews* **25**, 321–328.
- BERNINI, C., ANDRUNIÓW, T., OLIVUCCI, M., POGNI, R., BASOSI, R. & SINICROPI, A. (2013). Effects of the Protein Environment on the Spectral Properties of Tryptophan Radicals in *Pseudomonas aeruginosa* Azurin. *Journal of the American Chemical Society* **135**, 4822–4833.
- BLANCAFORT, L. & VOITYUK, A. A. (2007). MS-CASPT2 Calculation of Excess Electron Transfer in Stacked DNA Nucleobases. *The Journal of Physical Chemistry A* **111**, 4714–4719.
- BONETTI, C., MATHES, T., STOKKUM, I. H. M. VAN, MULLEN, K. M., GROOT, M.-L., GRONDELLE, R. VAN, HEGEMANN, P. & KENNIS, J. T. M. (2008). Hydrogen Bond Switching among Flavin and Amino Acid Side Chains in the BLUF Photoreceptor Observed by Ultrafast Infrared Spectroscopy. *Biophysical Journal* **95**, 4790–4802.
- BONETTI, C., STIERL, M., MATHES, T., VAN STOKKUM, I. H. M., MULLEN, K. M., COHEN-STUART, T. A., VAN GRONDELLE, R., HEGEMANN, P. & KENNIS, J. T. M. (2009). The Role of Key Amino Acids in the Photoactivation Pathway of the *Synechocystis* Slr1694 BLUF Domain. *Biochemistry* **48**, 11458–11469.
- BORN, M. & OPPENHEIMER, R. (1927). Zur Quantentheorie der Molekeln. *Annalen der Physik* **389**, 457–484.
- BOYS, S. F. & BERNARDI, F. (1970). The calculation of small molecular interactions by the differences of separate total energies. Some procedures with reduced errors. *Molecular Physics* **19**, 553–566.
- BRUST, R., HAIGNEY, A., LUKACS, A., GIL, A., HOSSAIN, S., ADDISON, K., LAI, C.-T., TOWRIE, M., GREETHAM, G. M., CLARK, I. P., ILLARIONOV, B., BACHER, A., KIM, R.-R., FISCHER, M., SIMMERLING, C., MEECH, S. R. & TONGE, P. J. (2014). Ultrafast Structural Dynamics of BlsA, a Photoreceptor from the Pathogenic Bacterium *Acinetobacter baumannii*. *The Journal of Physical Chemistry Letters* **5**, 220–224.
- BRUST, R., LUKACS, A., HAIGNEY, A., ADDISON, K., GIL, A., TOWRIE, M., CLARK, I. P., GREETHAM, G. M., TONGE, P. J. & MEECH, S. R. (2013). Proteins in Action: Femtosecond to Millisecond Structural Dynamics of a Photoactive Flavoprotein. *Journal of the American Chemical Society* **135**, 16168–16174.
- CASIDA, M. E. & HUIX-ROTLANT, M. (2012). Progress in Time-Dependent Density-Functional Theory. *Annual Review of Physical Chemistry* **63**, 287–323.
- CHENG, A. A. & LU, T. K. (2012). Synthetic Biology: An Emerging Engineering Discipline. *Annual Review of Biomedical Engineering* **14**, 155–178.
- CHRISTIANSEN, O., KOCH, H. & JØRGENSEN, P. (1995). The second-order approximate coupled cluster singles and doubles model CC2. *Chemical Physics Letters* **243**, 409–418.

- CHRISTIE, J. M., GAWTHORNE, J., YOUNG, G., FRASER, N. J. & ROE, A. J. (2012). LOV to BLUF: Flavoprotein Contributions to the Optogenetic Toolkit. *Molecular Plant* **5**, 533–544.
- CLIMENT, T., GONZÁLEZ-LUQUE, R., MERCHÁN, M. & SERRANO-ANDRÉS, L. (2006). Theoretical Insight into the Spectroscopy and Photochemistry of Isoalloxazine, the Flavin Core Ring. *The Journal of Physical Chemistry A* **110**, 13584–13590.
- COE, J. D. & MARTÍNEZ, T. J. (2006). Ab Initio Molecular Dynamics of Excited-State Intramolecular Proton Transfer around a Three-State Conical Intersection in Malonaldehyde. *The Journal of Physical Chemistry A* **110**, 618–630.
- CORNELL, W. D., CIEPLAK, P., BAYLY, C. I., GOULD, I. R., MERZ, K. M., FERGUSON, D. M., SPELLMEYER, D. C., FOX, T., CALDWELL, J. W. & KOLLMAN, P. A. (1995). A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *Journal of the American Chemical Society* **117**, 5179–5197.
- CRAMER, C. J. (2004). *Essentials of Computational Chemistry: Theories and Models*. Auflage: 2. Auflage. Chichester, West Sussex, England ; Hoboken, NJ: John Wiley & Sons.
- DEISSEROTH, K. (2011). Optogenetics. *Nature Methods* **8**, 26–29.
- DOMCKE, W. & SOBOLEWSKI, A. L. (2013). Peptide deactivation: Spectroscopy meets theory. *Nature Chemistry* **5**, 257–258.
- DOMCKE, W. & YARKONY, D. R. (2012). Role of Conical Intersections in Molecular Spectroscopy and Photoinduced Chemical Dynamics. *Annual Review of Physical Chemistry* **63**, 325–352.
- DOMCKE, W., YARKONY, D. R. & KÖPPEL, H. (2011). *Conical Intersections: Theory, Computation and Experiment*. WORLD SCIENTIFIC.
- DOMRATCHEVA, T. (2011). Neutral Histidine and Photoinduced Electron Transfer in DNA Photolyases. *Journal of the American Chemical Society* **133**, 18172–18182.
- DOMRATCHEVA, T., FEDOROV, R. & SCHLICHTING, I. (2006). Analysis of the Primary Photocycle Reactions Occurring in the Light, Oxygen, and Voltage Blue-Light Receptor by Multiconfigurational Quantum-Chemical Methods. *Journal of Chemical Theory and Computation* **2**, 1565–1574.
- DOMRATCHEVA, T., GRIGORENKO, B. L., SCHLICHTING, I. & NEMUKHIN, A. V. (2008). Molecular Models Predict Light-Induced Glutamine Tautomerization in BLUF Photoreceptors. *Biophysical Journal* **94**, 3872–3879.
- DOMRATCHEVA, T., UDVARHELYI, A. & SHAHI, A. R. M. (2014). Computational Spectroscopy, Dynamics, and Photochemistry of Photosensory Flavoproteins. In *Flavins and Flavoproteins, Methods in Molecular Biology*, Weber, S. & Schleicher, E. (Eds.), pp. 191–228. Springer New York.
- DRAGNEA, V., ARUNKUMAR, A. I., LEE, C. W., GIEDROC, D. P. & BAUER, C. E. (2010). A Q63E Rhodobacter sphaeroides AppA BLUF Domain Mutant Is Locked in a Pseudo-Light-Excited Signaling State. *Biochemistry* **49**, 10682–10690.

- DRAGNEA, V., ARUNKUMAR, A. I., YUAN, H., GIEDROC, D. P. & BAUER, C. E. (2009). Spectroscopic Studies of the AppA BLUF Domain from *Rhodobacter sphaeroides*: Addressing Movement of Tryptophan 104 in the Signaling State. *Biochemistry* **48**, 9969–9979.
- DRAGNEA, V., WAEGELE, M., BALASCUTA, S., BAUER, C. & DRAGNEA, B. (2005). Time-resolved spectroscopic studies of the AppA blue-light receptor BLUF domain from *Rb. sphaeroides*. *Biochemistry* **44**, 15978–15985.
- DREUW, A., WEISMAN, J. L. & HEAD-GORDON, M. (2003). Long-range charge-transfer excited states in time-dependent density functional theory require non-local exchange. *The Journal of Chemical Physics* **119**, 2943–2946.
- FALKLÖF, O. & DURBEEJ, B. (2013). Modeling of phytochrome absorption spectra. *Journal of Computational Chemistry* **34**, 1363–1374.
- FENNO, L., YIZHAR, O. & DEISSEROTH, K. (2011). The Development and Application of Optogenetics. *Annual Review of Neuroscience* **34**, 389–412.
- FERRE, N. & ANGYAN, J. G. (2002). Approximate electrostatic interaction operator for QM/MM calculations. *Chemical Physics Letters* **356**, 331–339.
- FINLEY, J., MALMQVIST, P.-Å., ROOS, B. O. & SERRANO-ANDRÉS, L. (1998). The multi-state CASPT2 method. *Chemical Physics Letters* **288**, 299–306.
- FORESMAN, J. B., HEAD-GORDON, M., POPLE, J. A. & FRISCH, M. J. (1992). Toward a systematic molecular orbital theory for excited states. *The Journal of Physical Chemistry* **96**, 135–149.
- FORSBERG, N. & MALMQVIST, P.-Å. (1997). Multiconfiguration perturbation theory with imaginary level shift. *Chemical Physics Letters* **274**, 196–204.
- FRÄHMCKE, J. S., WANKO, M., PHATAK, P., MROGINSKI, M. A. & ELSTNER, M. (2010). The Protonation State of Glu181 in Rhodopsin Revisited: Interpretation of Experimental Data on the Basis of QM/MM Calculations. *The Journal of Physical Chemistry B* **114**, 11338–11352.
- FRUTOS, L. M., MARKMANN, A., SOBOLEWSKI, A. L. & DOMCKE, W. (2007). Photoinduced Electron and Proton Transfer in the Hydrogen-Bonded Pyridine–Pyrrole System. *The Journal of Physical Chemistry B* **111**, 6110–6112.
- FUJIYOSHI, S., HIRANO, M., MATSUSHITA, M., ISEKI, M. & WATANABE, M. (2011). Structural Change of a Cofactor Binding Site of Flavoprotein Detected by Single-Protein Fluorescence Spectroscopy at 1.5 K. *Physical Review Letters* **106**, 078101.
- FUKUSHIMA, Y., OKAJIMA, K., SHIBATA, Y., IKEUCHI, M. & ITOH, S. (2005). Primary Intermediate in the Photocycle of a Blue-Light Sensory BLUF FAD-Protein, Tll0078, of *Thermosynechococcus elongatus* BP-1. *Biochemistry* **44**, 5149–5158.
- GAUDEN, M., GRINSTEAD, J. S., LAAN, W., VAN STOKKUM, I. H. M., AVILA-PEREZ, M., TOH, K. C., BOELEN, R., KAPTEIN, R., VAN GRONDELLE, R., HELLINGWERF, K. J. & KENNIS, J. T.

- M. (2007). On the Role of Aromatic Side Chains in the Photoactivation of BLUF Domains. *Biochemistry* **46**, 7405–7415.
- GAUDEN, M., VAN STOKKUM, I. H. M., KEY, J. M., LÜHRS, D. C., VAN GRONDELLE, R., HEGEMANN, P. & KENNIS, J. T. M. (2006). Hydrogen-bond switching through a radical pair mechanism in a flavin-binding photoreceptor. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 10895–10900.
- GAUDEN, M., YEREMENKO, S., LAAN, W., VAN STOKKUM, I. H. M., IHALAINEN, J. A., VAN GRONDELLE, R., HELLINGWERF, K. J. & KENNIS, J. T. M. (2005). Photocycle of the Flavin-Binding Photoreceptor AppA, a Bacterial Transcriptional Antirepressor of Photosynthesis Genes. *Biochemistry* **44**, 3653–3662.
- GHIGO, G., ROOS, B. O. & MALMQVIST, P.-Å. (2004). A modified definition of the zeroth-order Hamiltonian in multiconfigurational perturbation theory (CASPT2). *Chemical Physics Letters* **396**, 142–149.
- GOMELSKY, M. & KLUG, G. (2002). BLUF: a novel FAD-binding domain involved in sensory transduction in microorganisms. *Trends in Biochemical Sciences* **27**, 497–500.
- GONZÁLEZ, L., ESCUDERO, D. & SERRANO-ANDRÉS, L. (2012). Progress and Challenges in the Calculation of Electronic Excited States. *ChemPhysChem* **13**, 28–51.
- GORDON, M. S., SMITH, Q. A., XU, P. & SLIPCHENKO, L. V. (2013). Accurate First Principles Model Potentials for Intermolecular Interactions. *Annual Review of Physical Chemistry* **64**, 553–578.
- GOZEM, S., SCHAPIRO, I., FERRÉ, N. & OLIVUCCI, M. (2012). The Molecular Mechanism of Thermal Noise in Rod Photoreceptors. *Science* **337**, 1225–1228.
- GRANOVSKY, A. A. (2011). Extended multi-configuration quasi-degenerate perturbation theory: The new approach to multi-state multi-reference perturbation theory. *The Journal of Chemical Physics* **134**, 214113–214113–14.
- GRANOVSKY, A. A. (2013). *Firefly version 8.0*. <http://classic.chem.msu.su/gran/firefly/index.html>.
- GRIMME, S. (2011). Density functional theory with London dispersion corrections. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **1**, 211–228.
- GRINSTEAD, J. S., AVILA-PEREZ, M., HELLINGWERF, K. J., BOELENS, R. & KAPTEIN, R. (2006). Light-Induced Flipping of a Conserved Glutamine Sidechain and Its Orientation in the AppA BLUF Domain. *Journal of the American Chemical Society* **128**, 15066–15067.
- GRINSTEAD, J. S., HSU, S.-T. D., LAAN, W., BONVIN, A. M. J. J., HELLINGWERF, K. J., BOELENS, R. & KAPTEIN, R. (2006). The Solution Structure of the AppA BLUF Domain: Insight into the Mechanism of Light-Induced Signaling. *ChemBioChem* **7**, 187–193.
- GROENHOF, G., SCHÄFER, L. V., BOGGIO-PASQUA, M., GRUBMÜLLER, H. & ROBB, M. A. (2008). Arginine52 Controls the Photoisomerization Process in Photoactive Yellow Protein. *Journal of the American Chemical Society* **130**, 3250–3251.

- GROMOV, E. V., BURGHARDT, I., KÖPPEL, H. & CEDERBAUM, L. S. (2007). Electronic Structure of the PYP Chromophore in Its Native Protein Environment. *Journal of the American Chemical Society* **129**, 6798–6806.
- HAN, Y., BRAATSCH, S., OSTERLOH, L. & KLUG, G. (2004). A eukaryotic BLUF domain mediates light-dependent gene expression in the purple bacterium *Rhodobacter sphaeroides* 2.4.1. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 12306–12311.
- HASEGAWA, J., BUREEKAEW, S. & NAKATSUJI, H. (2007). SAC-CI theoretical study on the excited states of lumiflavin: Structure, excitation spectrum, and solvation effect. *Journal of Photochemistry and Photobiology A: Chemistry* **189**, 205–210.
- HASEGAWA, K., MASUDA, S. & ONO, T. (2004). Structural Intermediate in the Photocycle of a BLUF (Sensor of Blue Light Using FAD) Protein Slr1694 in a Cyanobacterium *Synechocystis* sp. PCC6803. *Biochemistry* **43**, 14979–14986.
- HASEGAWA, K., MASUDA, S. & ONO, T. (2006). Light Induced Structural Changes of a Full-length Protein and Its BLUF Domain in YcgF(Blrp), a Blue-Light Sensing Protein That Uses FAD (BLUF). *Biochemistry* **45**, 3785–3793.
- HELGAKEK, T., JORGENSEN, P. & OLSEN, J. (2000). *Molecular Electronic-Structure Theory*. John Wiley & Sons Ltd, West Sussex, England.
- HENZLER-WILDMAN, K. & KERN, D. (2007). Dynamic personalities of proteins. *Nature* **450**, 964–972.
- HOHENBERG, P. & KOHN, W. (1964). Inhomogeneous Electron Gas. *Physical Review* **136**, B864–B871.
- VAN DER HORST, M. A. & HELLINGWERF, K. J. (2004). Photoreceptor Proteins, ‘Star Actors of Modern Times’: A Review of the Functional Dynamics in the Structure of Representative Members of Six Different Photoreceptor Families. *Accounts of Chemical Research* **37**, 13–20.
- HSIAO, Y.-W., GÖTZE, J. P. & THIEL, W. (2012). The Central Role of Gln63 for the Hydrogen Bonding Network and UV–Visible Spectrum of the AppA BLUF Domain. *The Journal of Physical Chemistry B* **116**, 8064–8073.
- HYPERCUBE, INC. (1985). *HyperChem*. Hypercube, Inc. <http://www.hyper.com/?tabid=360>.
- ISBORN, C. M., GÖTZ, A. W., CLARK, M. A., WALKER, R. C. & MARTÍNEZ, T. J. (2012). Electronic Absorption Spectra from MM and ab Initio QM/MM Molecular Dynamics: Environmental Effects on the Absorption Spectrum of Photoactive Yellow Protein. *Journal of Chemical Theory and Computation* **8**, 5092–5106.
- ISBORN, C. M., LUEHR, N., UFIMTSEV, I. S. & MARTÍNEZ, T. J. (2011). Excited-State Electronic Structure with Configuration Interaction Singles and Tamm–Dancoff Time-Dependent Density Functional Theory on Graphical Processing Units. *Journal of Chemical Theory and Computation* **7**, 1814–1823.

- ISEKI, M., MATSUNAGA, S., MURAKAMI, A., OHNO, K., SHIGA, K., YOSHIDA, K., SUGAI, M., TAKAHASHI, T., HORI, T. & WATANABE, M. (2002). A blue-light-activated adenylyl cyclase mediates photoavoidance in *Euglena gracilis*. *Nature* **415**, 1047–1051.
- IWATA, T., WATANABE, A., ISEKI, M., WATANABE, M. & KANDORI, H. (2011). Strong Donation of the Hydrogen Bond of Tyrosine during Photoactivation of the BLUF Domain. *The Journal of Physical Chemistry Letters* **2**, 1015–1019.
- JENSEN, F. (2011). *Introduction to Computational Chemistry: Second Edition*. Auflage: 2. Auflage. Chichester, England ; Hoboken, NJ: JW.
- JORGENSEN, W. L., CHANDRASEKHAR, J., MADURA, J. D., IMPEY, R. W. & KLEIN, M. L. (1983). Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics* **79**, 926–935.
- JUNG, A., DOMRATCHEVA, T., TARUTINA, M., WU, Q., KO, W., SHOEMAN, R. L., GOMELSKY, M., GARDNER, K. H. & SCHLICHTING, I. (2005). Structure of a bacterial BLUF photoreceptor: Insights into blue light-mediated signal transduction. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 12350–12355.
- JUNG, A., REINSTEIN, J., DOMRATCHEVA, T., SHOEMAN, R. L. & SCHLICHTING, I. (2006). Crystal Structures of the AppA BLUF Domain Photoreceptor Provide Insights into Blue Light-mediated Signal Transduction. *Journal of Molecular Biology* **362**, 717–732.
- KAMERLIN, S. C. L., HARANCZYK, M. & WARSHEL, A. (2009). Progress in Ab Initio QM/MM Free-Energy Simulations of Electrostatic Energies in Proteins: Accelerated QM/MM Studies of pKa, Redox Reactions and Solvation Free Energies. *The Journal of Physical Chemistry B* **113**, 1253–1272.
- KENNIS, J. T. M. & MATHES, T. (2013). Molecular eyes: proteins that transform light into biological information. *Interface Focus* **3**, 20130005.
- KHRENOVA, M. G., DOMRATCHEVA, T., SCHLICHTING, I., GRIGORENKO, B. L. & NEMUKHIN, A. V. (2011). Computational Characterization of Reaction Intermediates in the Photocycle of the Sensory Domain of the AppA Blue Light Photoreceptor. *Photochemistry and Photobiology* **87**, 564–573.
- KHRENOVA, M. G., NEMUKHIN, A. V. & DOMRATCHEVA, T. (2013). Photoinduced Electron Transfer Facilitates Tautomerization of the Conserved Signaling Glutamine Side Chain in BLUF Protein Light Sensors. *The Journal of Physical Chemistry B* **117**, 2369–2377.
- KHRENOVA, M. G., NEMUKHIN, A. V., GRIGORENKO, B. L., KRYLOV, A. I. & DOMRATCHEVA, T. M. (2010). Quantum Chemistry Calculations Provide Support to the Mechanism of the Light-Induced Structural Changes in the Flavin-Binding Photoreceptor Proteins. *Journal of Chemical Theory and Computation* **6**, 2293–2302.
- KITA, A., OKAJIMA, K., MORIMOTO, Y., IKEUCHI, M. & MIKI, K. (2005). Structure of a Cyanobacterial BLUF Protein, Tll0078, Containing a Novel FAD-binding Blue Light Sensor Domain. *Journal of Molecular Biology* **349**, 1–9.

- KLAUMÜNZER, B., KRÖNER, D. & SAALFRANK, P. (2010). (TD-)DFT Calculation of Vibrational and Vibronic Spectra of Riboflavin in Solution. *The Journal of Physical Chemistry B* **114**, 10826–10834.
- KLESSINGER, M. (1995). Conical Intersections and the Mechanism of Singlet Photoreactions. *Angewandte Chemie International Edition in English* **34**, 549–551.
- KOHN, W. & SHAM, L. J. (1965). Self-Consistent Equations Including Exchange and Correlation Effects. *Physical Review* **140**, A1133–A1138.
- KRAFT, B. J., MASUDA, S., KIKUCHI, J., DRAGNEA, V., TOLLIN, G., ZALESKI, J. M. & BAUER, C. E. (2003). Spectroscopic and Mutational Analysis of the Blue-Light Photoreceptor AppA: A Novel Photocycle Involving Flavin Stacking with an Aromatic Amino Acid. *Biochemistry* **42**, 6726–6734.
- LEE, C., YANG, W. & PARR, R. G. (1988). Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Physical Review B* **37**, 785–789.
- LEVERENTZ, H. R., GAO, J. & TRUHLAR, D. G. (2011). Using multipole point charge distributions to provide the electrostatic potential in the variational explicit polarization (X-Pol) potential. *Theoretical Chemistry Accounts* **129**, 3–13.
- LEVINE, B. G. & MARTÍNEZ, T. J. (2009). Ab Initio Multiple Spawning Dynamics of Excited Butadiene: Role of Charge Transfer. *The Journal of Physical Chemistry A* **113**, 12815–12824.
- LIN, H. & TRUHLAR, D. G. (2007). QM/MM: what have we learned, where are we, and where do we go from here? *Theoretical Chemistry Accounts* **117**, 185–199.
- LIU, Z., GUO, X., TAN, C., LI, J., KAO, Y.-T., WANG, L., SANCAR, A. & ZHONG, D. (2012). Electron Tunneling Pathways and Role of Adenine in Repair of Cyclobutane Pyrimidine Dimer by DNA Photolyase. *Journal of the American Chemical Society* **134**, 8104–8114.
- LOSI, A. & GÄRTNER, W. (2011). Old Chromophores, New Photoactivation Paradigms, Trendy Applications: Flavins in Blue Light-Sensing Photoreceptors. *Photochemistry and Photobiology* **87**, 491–510.
- LOSI, A. & GÄRTNER, W. (2012). The Evolution of Flavin-Binding Photoreceptors: An Ancient Chromophore Serving Trendy Blue-Light Sensors. *Annual Review of Plant Biology* **63**, 49–72.
- LUKACS, A., BRUST, R., HAIGNEY, A., LAPTENOK, S. P., ADDISON, K., GIL, A., TOWRIE, M., GREETHAM, G. M., TONGE, P. J. & MEECH, S. R. (2014). BLUF Domain Function Does Not Require a Metastable Radical Intermediate State. *Journal of the American Chemical Society* **136**, 4605–4615.
- LUKACS, A., HAIGNEY, A., BRUST, R., ZHAO, R.-K., STELLING, A. L., CLARK, I. P., TOWRIE, M., GREETHAM, G. M., MEECH, S. R. & TONGE, P. J. (2011). Photoexcitation of the Blue Light Using FAD Photoreceptor AppA Results in Ultrafast Changes to the Protein Matrix. *Journal of the American Chemical Society* **133**, 16893–16900.
- MARCUS, R. A. & SUTIN, N. (1985). Electron transfers in chemistry and biology. *Biochimica et Biophysica Acta (BBA) - Reviews on Bioenergetics* **811**, 265–322.

- MARQUES, M. A. L. & GROSS, E. K. U. (2004). Time-Dependent Density Functional Theory. *Annual Review of Physical Chemistry* **55**, 427–455.
- MASUDA, S. & BAUER, C. E. (2002). AppA Is a Blue Light Photoreceptor that Antirepresses Photosynthesis Gene Expression in *Rhodobacter sphaeroides*. *Cell* **110**, 613–623.
- MASUDA, S., HASEGAWA, K., ISHII, A. & ONO, T. (2004). Light-Induced Structural Changes in a Putative Blue-Light Receptor with a Novel FAD Binding Fold Sensor of Blue-Light Using FAD (BLUF); Slr1694 of *Synechocystis* sp. PCC6803. *Biochemistry* **43**, 5304–5313.
- MASUDA, S., HASEGAWA, K. & ONO, T. (2005). Tryptophan at Position 104 is Involved in Transforming Light Signal into Changes of β -sheet Structure for the Signaling State in the BLUF Domain of AppA. *Plant and Cell Physiology* **46**, 1894–1901.
- MASUDA, S., TOMIDA, Y., OHTA, H. & TAKAMIYA, K. (2007). The Critical Role of a Hydrogen Bond between Gln63 and Trp104 in the Blue-Light Sensing BLUF Domain That Controls AppA Activity. *Journal of Molecular Biology* **368**, 1223–1230.
- MATHES, T., VAN STOKKUM, I. H. M., BONETTI, C., HEGEMANN, P. & KENNIS, J. T. M. (2011). The Hydrogen-Bond Switch Reaction of the Blrb Bluf Domain of *Rhodobacter sphaeroides*. *The Journal of Physical Chemistry B* **115**, 7963–7971.
- MATHES, T., STOKKUM, I. H. M. VAN, STIERL, M. & KENNIS, J. T. M. (2012). Redox Modulation of Flavin and Tyrosine Determines Photoinduced Proton-coupled Electron Transfer and Photoactivation of BLUF Photoreceptors. *Journal of Biological Chemistry* **287**, 31725–31738.
- MATHES, T., ZHU, J., VAN STOKKUM, I. H. M., GROOT, M. L., HEGEMANN, P. & KENNIS, J. T. M. (2012). Hydrogen Bond Switching among Flavin and Amino Acids Determines the Nature of Proton-Coupled Electron Transfer in BLUF Photoreceptors. *The Journal of Physical Chemistry Letters* **3**, 203–208.
- MEIER, K., THIEL, W. & VAN GUNSTEREN, W. F. (2012). On the effect of a variation of the force field, spatial boundary condition and size of the QM region in QM/MM MD simulations. *Journal of Computational Chemistry* **33**, 363–378.
- MELACCIO, F., FERRÉ, N. & OLIVUCCI, M. (2012). Quantum chemical modeling of rhodopsin mutants displaying switchable colors. *Physical Chemistry Chemical Physics* **14**, 12485–12495.
- MELACCIO, F., OLIVUCCI, M., LINDH, R. & FERRÉ, N. (2011). Unique QM/MM potential energy surface exploration using microiterations. *International Journal of Quantum Chemistry* **111**, 3339–3346.
- MIGLIORE, A., POLIZZI, N. F., THERIEN, M. J. & BERATAN, D. N. (2014). Biochemistry and Theory of Proton-Coupled Electron Transfer. *Chemical Reviews* **114**, 3381–3465.
- MIURA, R. (2001). Versatility and specificity in flavoenzymes: Control mechanisms of flavin reactivity. *The Chemical Record* **1**, 183–194.
- MO, Y., BAO, P. & GAO, J. (2011). Energy decomposition analysis based on a block-localized wavefunction and multistate density functional theory. *Physical Chemistry Chemical Physics* **13**, 6760–6775.

- MONARI, A., RIVAIL, J.-L. & ASSFELD, X. (2013). Theoretical Modeling of Large Molecular Systems. Advances in the Local Self Consistent Field Method for Mixed Quantum Mechanics/Molecular Mechanics Calculations. *Accounts of Chemical Research* **46**, 596–603.
- MOSER, C. C., KESKE, J. M., WARNCKE, K., FARID, R. S. & DUTTON, P. L. (1992). Nature of biological electron transfer. *Nature* **355**, 796–802.
- MOUGHAL SHAHI, A. R. & DOMRATCHEVA, T. (2013). Challenges in Computing Electron-Transfer Energies of DNA Repair Using Hybrid QM/MM Models. *Journal of Chemical Theory and Computation* **9**, 4644–4652.
- MUSSI, M. A., GADDY, J. A., CABRUJA, M., ARIVETT, B. A., VIALE, A. M., RASIA, R. & ACTIS, L. A. (2010). The Opportunistic Human Pathogen *Acinetobacter baumannii* Senses and Responds to Light. *Journal of Bacteriology* **192**, 6336–6345.
- NEISS, C., SAALFRANK, P., PARAC, M. & GRIMME, S. (2003). Quantum Chemical Calculation of Excited States of Flavin-Related Molecules. *The Journal of Physical Chemistry A* **107**, 140–147.
- NEMUKHIN, A. V., GRIGORENKO, B. L., TOPOL, I. A. & BURT, S. K. (2003). Flexible effective fragment QM/MM method: Validation through the challenging tests. *Journal of Computational Chemistry* **24**, 1410–1420.
- VON NEUMANN, J. & WIGNER, E. (1929). Über merkwürdige diskrete Eigenwerte. *Physikalische Zeitschrift* **30**, 467.
- OBANAYAMA, K., KOBAYASHI, H., FUKUSHIMA, K. & SAKURAI, M. (2008). Structures of the Chromophore Binding Sites in BLUF Domains as Studied by Molecular Dynamics and Quantum Chemical Calculations. *Photochemistry and Photobiology* **84**, 1003–1010.
- OKAJIMA, K., YOSHIHARA, S., FUKUSHIMA, Y., GENG, X., KATAYAMA, M., HIGASHI, S., WATANABE, M., SATO, S., TABATA, S., SHIBATA, Y., ITOH, S. & IKEUCHI, M. (2005). Biochemical and Functional Characterization of BLUF-Type Flavin-Binding Proteins of Two Species of Cyanobacteria. *Journal of Biochemistry* **137**, 741–750.
- OLIVUCCI, M. (2005). *Computational Photochemistry*. Amsterdam: Elsevier Ltd.
- OLSEN, S., LAMOTHE, K. & MARTÍNEZ, T. J. (2010). Protonic Gating of Excited-State Twisting and Charge Localization in GFP Chromophores: A Mechanistic Hypothesis for Reversible Photoswitching. *Journal of the American Chemical Society* **132**, 1192–1193.
- OOSTENBRINK, C., VILLA, A., MARK, A. E. & VAN GUNSTEREN, W. F. (2004). A biomolecular force field based on the free enthalpy of hydration and solvation: The GROMOS force-field parameter sets 53A5 and 53A6. *Journal of Computational Chemistry* **25**, 1656–1676.
- PAL, R., SEKHARAN, S. & BATISTA, V. S. (2013). Spectral Tuning in Halorhodopsin: The Chloride Pump Photoreceptor. *Journal of the American Chemical Society* **135**, 9624–9627.
- PENZKOFER, A., STIERL, M., HEGEMANN, P. & KATERIYA, S. (2011). Photo-dynamics of the BLUF domain containing soluble adenylate cyclase (nPAC) from the amoebflagellate *Naegleria gruberi* NEG-M strain. *Chemical Physics* **387**, 25–38.

- PERRAKIS, A., MORRIS, R. & LAMZIN, V. S. (1999). Automated protein model building combined with iterative structure refinement. *Nature Structural & Molecular Biology* **6**, 458–463.
- PEVERATI, R. & TRUHLAR, D. G. (2014). Quest for a universal density functional: the accuracy of density functionals across a broad spectrum of databases in chemistry and physics. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **372**, 20120476.
- POLLI, D., ALTOÈ, P., WEINGART, O., SPILLANE, K. M., MANZONI, C., BRIDA, D., TOMASELLO, G., ORLANDI, G., KUKURA, P., MATHIES, R. A., GARAVELLI, M. & CERULLO, G. (2010). Conical intersection dynamics of the primary photoisomerization event in vision. *Nature* **467**, 440–443.
- PONDER, J. W. & RICHARDS, F. M. (1987). An efficient newton-like method for molecular mechanics energy minimization of large molecules. *Journal of Computational Chemistry* **8**, 1016–1024.
- POPLE, J. A., BINKLEY, J. S. & SEEGER, R. (1976). Theoretical models incorporating electron correlation. *International Journal of Quantum Chemistry* **10**, 1–19.
- PRONK, S., PÁLL, S., SCHULZ, R., LARSSON, P., BJELKMAR, P., APOSTOLOV, R., SHIRTS, M. R., SMITH, J. C., KASSON, P. M., SPOEL, D. VAN DER, HESS, B. & LINDAHL, E. (2013). GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics* **29**, 845–854.
- RAJAGOPAL, S., KEY, J. M., PURCELL, E. B., BOEREMA, D. J. & MOFFAT, K. (2004). Purification and Initial Characterization of a Putative Blue Light-regulated Phosphodiesterase from *Escherichia coli*. *Photochemistry and Photobiology* **80**, 542–547.
- RAJAMANI, R., LIN, Y.-L. & GAO, J. (2011). The opsin shift and mechanism of spectral tuning in rhodopsin. *Journal of Computational Chemistry* **32**, 854–865.
- RHEE, Y. M. & HEAD-GORDON, M. (2007). Scaled Second-Order Perturbation Corrections to Configuration Interaction Singles: Efficient and Reliable Excitation Energy Methods. *The Journal of Physical Chemistry A* **111**, 5314–5326.
- RIEFF, B., BAUER, S., MATHIAS, G. & TAVAN, P. (2011). DFT/MM Description of Flavin IR Spectra in BLUF Domains. *The Journal of Physical Chemistry B* **115**, 11239–11253.
- RINALDI, S., MELACCIO, F., GOZEM, S., FANELLI, F. & OLIVUCCI, M. (2014). Comparison of the isomerization mechanisms of human melanopsin and invertebrate and vertebrate rhodopsins. *Proceedings of the National Academy of Sciences* **111**, 1714–1719.
- RIZZINI, L., FAVORY, J.-J., CLOIX, C., FAGGIONATO, D., O'HARA, A., KAISERLI, E., BAUMEISTER, R., SCHÄFER, E., NAGY, F., JENKINS, G. I. & ULM, R. (2011). Perception of UV-B by the Arabidopsis UVR8 Protein. *Science* **332**, 103–106.
- ROBB, M. A., BERNARDI, F. & OLIVUCCI, M. (1995). Conical intersections as a mechanistic feature of organic photochemistry. *Pure and Applied Chemistry* **67**.
- ROOS, B. (1972). A new method for large-scale CI calculations. *Chemical Physics Letters* **15**, 153–159.

- ROOS, B. O. (1987). The Complete Active Space Self-Consistent Field Method and its Applications in Electronic Structure Calculations. In *Advances in Chemical Physics*, Lawley, K. P. (Ed.), pp. 399–445. John Wiley & Sons, Inc.
- ROOS, B. O., TAYLOR, P. R. & SIEGBAHN, P. E. M. (1980). A complete active space SCF method (CASSCF) using a density matrix formulated super-CI approach. *Chemical Physics* **48**, 157–173.
- RUNGE, E. & GROSS, E. K. U. (1984). Density-Functional Theory for Time-Dependent Systems. *Physical Review Letters* **52**, 997–1000.
- SADEGHIAN, K., BOCOLA, M. & SCHÜTZ, M. (2008). A Conclusive Mechanism of the Photoinduced Reaction Cascade in Blue Light Using Flavin Photoreceptors. *Journal of the American Chemical Society* **130**, 12501–12513.
- SADEGHIAN, K., BOCOLA, M. & SCHÜTZ, M. (2010). A QM/MM study on the fast photocycle of blue light using flavin photoreceptors in their light-adapted/active form. *Physical Chemistry Chemical Physics* **12**, 8840–8846.
- SADEGHIAN, K. & SCHÜTZ, M. (2007). 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. *Journal of the American Chemical Society* **129**, 4068–4074.
- SALZMANN, S., SILVA-JUNIOR, M. R., THIEL, W. & MARIAN, C. M. (2009). Influence of the LOV Domain on Low-Lying Excited States of Flavin: A Combined Quantum-Mechanics/Molecular-Mechanics Investigation. *The Journal of Physical Chemistry B* **113**, 15610–15618.
- SALZMANN, S., TATCHEN, J. & MARIAN, C. M. (2008). The photophysics of flavins: What makes the difference between gas phase and aqueous solution? *Journal of Photochemistry and Photobiology A: Chemistry* **198**, 221–231.
- SCHAPIRO, I., MELACCIO, F., LARICHEVA, E. N. & OLIVUCCI, M. (2011). Using the computer to understand the chemistry of conical intersections. *Photochemical & Photobiological Sciences* **10**, 867.
- SCHMIDT, M. W., BALDRIDGE, K. K., BOATZ, J. A., ELBERT, S. T., GORDON, M. S., JENSEN, J. H., KOSEKI, S., MATSUNAGA, N., NGUYEN, K. A., SU, S., WINDUS, T. L., DUPUIS, M. & MONTGOMERY, J. A. (1993). General atomic and molecular electronic structure system. *Journal of Computational Chemistry* **14**, 1347–1363.
- SCHMIDT, M. W. & GORDON, M. S. (1998). The Construction and Interpretation of Mcscf Wavefunctions. *Annual Review of Physical Chemistry* **49**, 233–266.
- SCHULTZ, T., SAMOYLOVA, E., RADLOFF, W., HERTEL, I. V., SOBOLEWSKI, A. L. & DOMCKE, W. (2004). Efficient Deactivation of a Model Base Pair via Excited-State Hydrogen Transfer. *Science* **306**, 1765–1768.
- SEKHARAN, S., MOONEY, V. L., RIVALTA, I., KAZMI, M. A., NEITZ, M., NEITZ, J., SAKMAR, T. P., YAN, E. C. Y. & BATISTA, V. S. (2013). Spectral Tuning of Ultraviolet Cone Pigments: An Interhelical Lock Mechanism. *Journal of the American Chemical Society* **135**, 19064–19067.

- SEKHARAN, S. & MOROKUMA, K. (2011). QM/MM Study of the Structure, Energy Storage, and Origin of the Bathochromic Shift in Vertebrate and Invertebrate Bathorhodopsins. *Journal of the American Chemical Society* **133**, 4734–4737.
- SEKHARAN, S., WEI, J. N. & BATISTA, V. S. (2012). The Active Site of Melanopsin: The Biological Clock Photoreceptor. *Journal of the American Chemical Society* **134**, 19536–19539.
- SENN, H. M. & THIEL, W. (2009). QM/MM Methods for Biomolecular Systems. *Angewandte Chemie International Edition* **48**, 1198–1229.
- SERRANO-ANDRÉS, L. & MERCHÁN, M. (2005). Quantum chemistry of the excited state: 2005 overview. *Journal of Molecular Structure: THEOCHEM* **729**, 99–108.
- SEYEDSAYAMDOST, M. R., REECE, S. Y., NOCERA, D. G. & STUBBE, J. (2006). Mono-, Di-, Tri-, and Tetra-Substituted Fluorotyrosines: New Probes for Enzymes That Use Tyrosyl Radicals in Catalysis. *Journal of the American Chemical Society* **128**, 1569–1579.
- SHIBATA, Y., MURAI, Y., SATOH, Y., FUKUSHIMA, Y., OKAJIMA, K., IKEUCHI, M. & ITOH, S. (2009). Acceleration of Electron-Transfer-Induced Fluorescence Quenching upon Conversion to the Signaling State in the Blue-Light Receptor, TePixD, from *Thermosynechococcus elongatus*. *The Journal of Physical Chemistry B* **113**, 8192–8198.
- SIEGBAHN, P. E. M. & HIMO, F. (2011). The quantum chemical cluster approach for modeling enzyme reactions. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **1**, 323–336.
- SINGH, U. C. & KOLLMAN, P. A. (1986). A combined ab initio quantum mechanical and molecular mechanical method for carrying out simulations on complex molecular systems: Applications to the CH₃Cl + Cl⁻ exchange reaction and gas phase protonation of polyethers. *Journal of Computational Chemistry* **7**, 718–730.
- SINICROPI, A., PISCHEL, U., BASOSI, R., NAU, W. M. & OLIVUCCI, M. (2000). Conical Intersections in Charge-Transfer Induced Quenching. *Angewandte Chemie International Edition* **39**, 4582–4586.
- SOBOLEWSKI, A. L. & DOMCKE, W. (2006). The chemical physics of the photostability of life. *Europhysics News* **37**, 20–23.
- SOBOLEWSKI, A. L., DOMCKE, W. & HÄTTIG, C. (2005). Tautomeric selectivity of the excited-state lifetime of guanine/cytosine base pairs: The role of electron-driven proton-transfer processes. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 17903–17906.
- SOLOV'YOV, I. A., DOMRATCHEVA, T., MOUGHAL SHAHI, A. R. & SCHULTEN, K. (2012). Decrypting Cryptochrome: Revealing the Molecular Identity of the Photoactivation Reaction. *Journal of the American Chemical Society* **134**, 18046–18052.
- SOLOV'YOV, I. A., DOMRATCHEVA, T. & SCHULTEN, K. (2014). Separation of photo-induced radical pair in cryptochrome to a functionally critical distance. *Scientific Reports* **4**.
- SONG, S.-H., DICK, B. & PENZKOFER, A. (2007). Photo-induced reduction of flavin mononucleotide in aqueous solutions. *Chemical Physics* **332**, 55–65.

- SOUSA, S. F., FERNANDES, P. A. & RAMOS, M. J. (2007). General Performance of Density Functionals. *The Journal of Physical Chemistry A* **111**, 10439–10452.
- STELLING, A. L., RONAYNE, K. L., NAPPA, J., TONGE, P. J. & MEECH, S. R. (2007). Ultrafast Structural Dynamics in BLUF Domains: Transient Infrared Spectroscopy of AppA and Its Mutants. *Journal of the American Chemical Society* **129**, 15556–15564.
- STRAMBI, A., DURBEEJ, B., FERRÉ, N. & OLIVUCCI, M. (2010). Anabaena sensory rhodopsin is a light-driven unidirectional rotor. *Proceedings of the National Academy of Sciences* **107**, 21322–21326.
- TAKAHASHI, R., OKAJIMA, K., SUZUKI, H., NAKAMURA, H., IKEUCHI, M. & NOGUCHI, T. (2007). FTIR Study on the Hydrogen Bond Structure of a Key Tyrosine Residue in the Flavin-Binding Blue Light Sensor TePixD from *Thermosynechococcus elongatus*. *Biochemistry* **46**, 6459–6467.
- TELLER, E. (1937). The Crossing of Potential Surfaces. *The Journal of Physical Chemistry* **41**, 109–116.
- TOH, K. C., VAN STOKKUM, I. H. M., HENDRIKS, J., ALEXANDRE, M. T. A., ARENTS, J. C., PEREZ, M. A., VAN GRONDELLE, R., HELLINGWERF, K. J. & KENNIS, J. T. M. (2008). On the Signaling Mechanism and the Absence of Photoreversibility in the AppA BLUF Domain. *Biophysical Journal* **95**, 312–321.
- TOZER, D. J., AMOS, R. D., HANDY, N. C., ROOS, B. O. & SERRANO-ANDRES, L. (1999). Does density functional theory contribute to the understanding of excited states of unsaturated organic compounds? *Molecular Physics* **97**, 859–868.
- TROFIMOV, A. B. & SCHIRMER, J. (1995). An efficient polarization propagator approach to valence electron excitation spectra. *Journal of Physics B: Atomic, Molecular and Optical Physics* **28**, 2299.
- TSCHOWRI, N., BUSSE, S. & HENGGE, R. (2009). The BLUF-EAL protein YcgF acts as a direct anti-repressor in a blue-light response of *Escherichia coli*. *Genes & Development* **23**, 522–534.
- TURBOMOLE GMBH (1989). *TURBOMOLE V6.3 2011*. TURBOMOLE GmbH <http://www.turbomole.com>.
- TYAGI, A., PENZKOFER, A., GRIESE, J., SCHLICHTING, I., KIRIENKO, N. V. & GOMELSKY, M. (2008). Photodynamics of blue-light-regulated phosphodiesterase BlrP1 protein from *Klebsiella pneumoniae* and its photoreceptor BLUF domain. *Chemical Physics* **354**, 130–141.
- UDVARHELYI, A. (2009). ‘Small- and wide-angle X-ray scattering experiments on blue light photoreceptors’. Heidelberg, Germany: Heidelberg University.
- UDVARHELYI, A. & DOMRATCHEVA, T. (2011). Photoreaction in BLUF Receptors: Proton-coupled Electron Transfer in the Flavin-Gln-Tyr System. *Photochemistry and Photobiology* **87**, 554–563.

- UDVARHELYI, A. & DOMRATCHEVA, T. (2013). Glutamine Rotamers in BLUF Photoreceptors: A Mechanistic Reappraisal. *The Journal of Physical Chemistry B* **117**, 2888–2897.
- UNNO, M., MASUDA, S., ONO, T. & YAMAUCHI, S. (2006). Orientation of a Key Glutamine Residue in the BLUF Domain from AppA Revealed by Mutagenesis, Spectroscopy, and Quantum Chemical Calculations. *Journal of the American Chemical Society* **128**, 5638–5639.
- UNNO, M., SANO, R., MASUDA, S., ONO, T. & YAMAUCHI, S. (2005). Light-Induced Structural Changes in the Active Site of the BLUF Domain in AppA by Raman Spectroscopy. *The Journal of Physical Chemistry B* **109**, 12620–12626.
- VALSSON, O., CAMPOMANES, P., TAVERNELLI, I., ROTHLSBERGER, U. & FILIPPI, C. (2013). Rhodopsin Absorption from First Principles: Bypassing Common Pitfalls. *Journal of Chemical Theory and Computation* **9**, 2441–2454.
- VENTUROLI, G., DREPPER, F., WILLIAMS, J. C., ALLEN, J. P., LIN, X. & MATHIS, P. (1998). Effects of Temperature and ΔG° on Electron Transfer from Cytochrome c2 to the Photosynthetic Reaction Center of the Purple Bacterium *Rhodobacter sphaeroides*. *Biophysical Journal* **74**, 3226–3240.
- VERYAZOV, V., MALMQVIST, P. Å. & ROOS, B. O. (2011). How to select active space for multiconfigurational quantum chemistry? *International Journal of Quantum Chemistry* **111**, 3329–3338.
- WANG, J., WANG, W., KOLLMAN, P. A. & CASE, D. A. (2006). Automatic atom type and bond type perception in molecular mechanical calculations. *Journal of Molecular Graphics and Modelling* **25**, 247–260.
- WANG, M. & FRITCHIE JNR, C. J. (1973). Geometry of the unperturbed flavin nucleus. The crystal structure of 10-methylisalloxazine. *Acta Crystallographica Section B Structural Crystallography and Crystal Chemistry* **29**, 2040–2045.
- WANKO, M., HOFFMANN, M., FRÄHMCKE, J., FRAUENHEIM, T. & ELSTNER, M. (2008). Effect of Polarization on the Opsin Shift in Rhodopsins. 2. Empirical Polarization Models for Proteins. *The Journal of Physical Chemistry B* **112**, 11468–11478.
- WANKO, M., HOFFMANN, M., FRAUENHEIM, T. & ELSTNER, M. (2008). Effect of Polarization on the Opsin Shift in Rhodopsins. 1. A Combined QM/QM/MM Model for Bacteriorhodopsin and Pharaonis Sensory Rhodopsin II. *The Journal of Physical Chemistry B* **112**, 11462–11467.
- WARSHEL, A. & LEVITT, M. (1976). Theoretical studies of enzymic reactions: Dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *Journal of Molecular Biology* **103**, 227–249.
- WEIGEL, A., DOBRYAKOV, A., KLAUMÜNZER, B., SAJADI, M., SAALFRANK, P. & ERNSTING, N. P. (2011). Femtosecond Stimulated Raman Spectroscopy of Flavin after Optical Excitation. *The Journal of Physical Chemistry B* **115**, 3656–3680.
- WERNER, H.-J. & MEYER, W. (1981). A quadratically convergent MCSCF method for the simultaneous optimization of several states. *The Journal of Chemical Physics* **74**, 5794–5801.

- WINKLER, A., HEINTZ, U., LINDNER, R., REINSTEIN, J., SHOEMAN, R. L. & SCHLICHTING, I. (2013). A ternary AppA–PpsR–DNA complex mediates light regulation of photosynthesis-related gene expression. *Nature Structural & Molecular Biology* **20**, 859–867.
- WINKLER, A., UDVARHELYI, A., HARTMANN, E., REINSTEIN, J., MENZEL, A., SHOEMAN, R. L. & SCHLICHTING, I. (2014). Characterization of Elements Involved in Allosteric Light Regulation of Phosphodiesterase Activity by Comparison of Different Functional BlrP1 States. *Journal of Molecular Biology* **426**, 853–868.
- WITEK, H. A., CHOE, Y.-K., FINLEY, J. P. & HIRAO, K. (2002). Intruder state avoidance multireference Møller–Plesset perturbation theory. *Journal of Computational Chemistry* **23**, 957–965.
- WOICZIKOWSKI, P. B., STEINBRECHER, T., KUBARČ, T. & ELSTNER, M. (2011). Nonadiabatic QM/MM Simulations of Fast Charge Transfer in Escherichia coli DNA Photolyase. *The Journal of Physical Chemistry B* **115**, 9846–9863.
- WOLF, M. M. N., SCHUMANN, C., GROSS, R., DOMRATCHEVA, T. & DILLER, R. (2008). Ultrafast Infrared Spectroscopy of Riboflavin: Dynamics, Electronic Structure, and Vibrational Mode Analysis. *The Journal of Physical Chemistry B* **112**, 13424–13432.
- WU, Q. & GARDNER, K. H. (2009). Structure and Insight into Blue Light-Induced Changes in the BlrP1 BLUF Domain. *Biochemistry* **48**, 2620–2629.
- WU, Q., KO, W.-H. & GARDNER, K. H. (2008). Structural Requirements for Key Residues and Auxiliary Portions of a BLUF Domain. *Biochemistry* **47**, 10271–10280.
- WU, Y. I., FREY, D., LUNGU, O. I., JAEHRIG, A., SCHLICHTING, I., KUHLMAN, B. & HAHN, K. M. (2009). A genetically encoded photoactivatable Rac controls the motility of living cells. *Nature* **461**, 104–108.
- YARKONY, D. R. (1996). Diabolical conical intersections. *Reviews of Modern Physics* **68**, 985–1013.
- YUAN, H., ANDERSON, S., MASUDA, S., DRAGNEA, V., MOFFAT, K. & BAUER, C. (2006). Crystal Structures of the Synechocystis Photoreceptor Slr1694 Reveal Distinct Structural States Related to Signaling. *Biochemistry* **45**, 12687–12694.
- YUAN, H., DRAGNEA, V., WU, Q., GARDNER, K. H. & BAUER, C. E. (2011). Mutational and Structural Studies of the PixD BLUF Output Signal That Affects Light-Regulated Interactions with PixE. *Biochemistry* **50**, 6365–6375.
- ZIRAK, P., PENZKOFER, A., LEHMPFUHL, C., MATHES, T. & HEGEMANN, P. (2007). Absorption and emission spectroscopic characterization of blue-light receptor Slr1694 from Synechocystis sp. PCC6803. *Journal of Photochemistry and Photobiology B: Biology* **86**, 22–34.
- ZIRAK, P., PENZKOFER, A., SCHIEREIS, T., HEGEMANN, P., JUNG, A. & SCHLICHTING, I. (2005). Absorption and fluorescence spectroscopic characterization of BLUF domain of AppA from Rhodospirillum rubrum. *Chemical Physics* **315**, 142–154.

- ZIRAK, P., PENZKOFER, A., SCHIEREIS, T., HEGEMANN, P., JUNG, A. & SCHLICHTING, I. (2006). Photodynamics of the small BLUF protein BlrB from *Rhodobacter sphaeroides*. *Journal of Photochemistry and Photobiology B: Biology* **83**, 180–194.
- ZOLTOWSKI, B. D. & GARDNER, K. H. (2011). Tripping the Light Fantastic: Blue-Light Photoreceptors as Examples of Environmentally Modulated Protein–Protein Interactions. *Biochemistry* **50**, 4–16.

Acknowledgements

The work presented in this thesis would not have been possible without the support of many others. I am grateful for their various contributions.

Tatiana Domratcheva – Thank you for your excellent supervision of my projects, for introducing me to quantum chemistry, for constantly offering valuable advice and support, and for your tireless corrections of my manuscripts. Thank you also for giving me the freedom to follow my research interests, for encouraging and supporting my participation in various conferences and summer schools, and for letting me use generous amounts of computing time. Спасибо!

Ilme Schlichting – for introducing me to the BLUF challenge, for the continuous support throughout my time as PhD candidate, and for acting as first referee of my dissertation.

Horst Köppel – for stimulating discussions on excited-state problems and kindly agreeing to act as second referee; and to **Ulrich Schwarz** and **Christian Enss** for acting as members of my disputation committee.

Boehringer Ingelheim Fonds (BIF) – for financial and non-material support over three years, and for the amazing and inspiring Hirschegg and Lautrach seminars. I am also grateful to the **Max Planck Society** for funding and to the Heidelberg Graduate School of Mathematical and Computational Methods for the Sciences (**HGS**) for financial support for the Sostrup summer school.

Massimo Olivucci – for kindly hosting me in Bowling Green and giving me the opportunity to deepen my knowledge on QM/MM modeling and photochemistry. I thank all Olivucci lab members for my great time in the United States. I am grateful in particular to **Samer, Federico,** and **Igor** for answering many questions and offering computational advice and help with MOLCAS/Tinker-related issues, as well as to **Mark** for help with the calculations of the RESP charges.

Chris Roome – for excellent computer support at any time of the day, which saved me so much time and allowed me to focus on my research.

Thomas Barends and **Roman Fedorov** – for computing the X-ray electron-density maps used in Chapter 4.

Andreas Schäfer – for serving on my Thesis Advisory Committee (TAC).

Andreas Winkler – for insightful discussions on BLUF, the fruitful collaboration on the BlrP project, and helpful comments on the glutamine-rotamers manuscript.

Martina Blochmann, Vanessa Würges, and Fabienne Höfer-Elfner – for their kind help with organizational and administrative matters.

Deniz Tuna – for lively discussions on photochemistry, and a fruitful collaboration on the melanin project.

the whole BMM department – for sharing and discussing ideas, a stimulating multi-disciplinary atmosphere, and the opportunity to “being close” to the experimental research. I thank Alhosna and Rehaman for encouraging discussions. I enjoyed the Thursday cakes (especially the chocolate ones), departmental parties and barbeques, and the intensive table soccer games with the BMM champions Aytaç, Stephan, Lukas, Kristina and Madeleine!

Kristina, Cathleen, Stephan, Frauke, Jule, Jörgen, and Fabian – for being wonderful friends! Thank you for many memorable moments, the big baby-shower surprise, delicious fruit smoothies, and a lot of fun during our skiing vacation.

Maria – for your honest friendship that means a lot to me. Ich danke Dir, dass Du immer ein offenes Ohr für mich hast, vor allem, wenn es besonders nötig ist.

My family – for the great and unconditional support throughout my life. Köszönöm a folyamatos biztatást és, hogy hittetek bennem. A very special thanks goes to Anya and Apa, Inge, and Kata for cooking, babysitting, and dealing with my ‘mood swings’.

Benjamin and Ildi – Thank you, my wonderful little family, for your love, support, comprehension, patience, the Dulsão hot chocolates, and everything else one can wish for. I wouldn't have made it without you! ☺