COMPUTATIONAL INVESTIGATION OF THE PHOTOCYCLIZATION OF NOVEL N-ALKYLATED INDANYLIDENE PYRROLINE BIOMIMETIC SWITCHES

Mikhail N. Ryazantsev

A Dissertation

Submitted to the Graduate College of Bowling Green State University in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

August 2010

Committee:

Dr. Massimo Olivucci, Advisor

Dr. Haowen Xi, Graduate Faculty Representative

Dr. H. Peter Lu

Dr. John R. Cable
ABSTRACT

Dr. Massimo Olivucci, Advisor

The biological function of Rhodopsin (Rh), the G-protein-coupled photoreceptor responsible for twilight vision in vertebrates, is based on a very efficient process of photo-isomerization in the 11-cis retinal protonated Schiff base (PSB11) that is covalently linked to an apoprotein: opsin. The role of the opsin in the efficiency of the process is dramatical and PSB11 that is isolated from the protein environment (i.e., in methanol solution) exhibits different photo-chemical properties. When considering possible applications, the photo-isomerization is the process that can be used to change the properties of the system by means of an external light stimulus (photo-switches) or to exploit the light energy into unidirectional motion at the molecular level (molecular motors). N-alkylated indanylidene pyrroline (NAIP) switches are compounds that are designed to mimic, in solution, several aspects of the photochemistry of Rhodopsin. In this work, both the photoisomerization process of NAIP-switches, and the photochemistry of PSB11 in the opsin environment are investigated. The ab initio multi-configurational QM/MM approach (CASPT2//CASSCF/AMBER) is supported by time-resolved spectroscopy studies.

Results show that NAIP switches exhibit several properties similar to that of Rhodopsin, such as stereoselectivity of the photo-isomerization, unidirectional motion, a sub-picosecond life time, and a barierless Minimum Energy Path leading from FC point to Conical Intersection. These properties make these compounds promising photo-switches or molecular motors. However, in spite of these remarkable similarities, the quantum yield of the photoisomerization of NAIPs is 2 to 3 times lower then that of Rh. These facts suggest that NAIPs not only provide a route to new materials but that they also constitute attractive systems for the investigation of fundamental problems such as the relationship between excited-state evolution and quantum yields. The investigation of the dynamics of photo-isomerization of Rh and bathoRh provided in the second part of this thesis can be considered to be the first step in understanding the factors that are responsible for the quantum yield of the process.
ACKNOWLEDGMENTS

First and foremost I want to express my gratitude to my supervisor Prof. Massimo Olivucci, for all his help and encouragement, patience, support, and much more. During all these years he drove my scientific education and research. I would also like to thank him for giving me several opportunities to present the results of my work at conferences and meetings.

I would like to acknowledge our colleagues - Prof. Nicolas Ferre, Prof. Jan Helbing, Prof. Stefan Haacke, Prof. Majed Chergui, Prof. Vinizio Zanirato, Dr. Jeremie Leonard, Dr. Andrea Cannizzo, Dr. Elena Martin for their fundamental contribution to this work.

Many thanks go to:
All the past and present members of the Computational Photochemistry Group: Dr. Adalgisa Sinicropi, Dr. Angela Strambi, Dr. Maria Laura Parisi, Federico Melaccio, Dr. Xuebo Chen, Dr. Wan Jian Ding, Pavel Klimovich, Mark M. Huntress, Elena N Laricheva, Samer Gozem, Dr. Patrick Z. El-Khoury and Dr. Igor Schapiro
LIST OF PUBLICATION

1. A. Sinicropi, E. Martin, **M. Ryazantsev**, J. Helbing, et al.; An artificial molecular
switch that mimics the visual pigment and completes its photocycle in picoseconds,
Proceedings of the National Academy of Science of the USA, 2008, 105, 17642.

2. Masato Sumita, **Mikhail N. Ryazantsev** and Kazuya Saito; Acceleration of the Z
to E photoisomerization of penta-2,4-dieniminium by hydrogen out-of-plane motion:
thoretical study on a model system of retinal protonated Schiff base, Phys. Chem.
Chem. Phys.; 2009, 11

3. Patrick Z. El-Khoury, Alexander N. Tarnovksy, Igor Schapiro, **Mikhail N. Ryazant-
sev** and Massimo Olivucci; Structure of the Photochemical Reaction Path Populated
via Promotion of CF2I2 into Its First Excited State, J. Phys. Chem. A, 2009, 113

4. Igor Schapiro, **Mikhail N. Ryazantsev**, Wan Jian Ding, Mark M. Huntress, Federico
Melaccio, Tadeusz Andruniow, and Massimo Olivucci; Computational Photobiology
and Beyond, Aust. J. Chem., 2010, 63

5. Alfonso Melloni, Riccardo Rossi Paccani, Donato Donati, Vinicio Zanirato, Adalgisa
Sinicropi, Maria Laura Parisi, Elena Martin, **Mikhail N Ryazantsev**, et al.; Mod-
eling, Preparation and Characterization of a Dipole Moment Switch Driven by Z/E
Photoisomerization, J. Am. Chem. Soc (accepted)

6. Igor Shapiro, **Mikhail N Ryazantsev**, Luis Manuel Frutos, Nicolas Ferre, Massimo
Olivucci; The Ultrafast Photoisomerizations of Rhodopsin and Bathorhodopsin are
Modulated by Bond Alternation and HOOP driven Electronic Effects (in preparation)
# Contents

1 INTRODUCTION 1

1.1 PHOTORECEPTOR PROTEINS. RHODOPSIN 2

1.2 PHOTOCHEMICAL SWITCHES 6

1.3 ARTIFICIAL MOLECULAR MACHINES. MOLECULAR MOTORS 9

1.4 AIMS AND MAIN POINTS OF THE THESIS 15

2 METHODOLOGY 27

2.1 TOPOLOGY OF POTENTIAL ENERGY SURFACES 28

2.1.1 SEPARATION OF NUCLEAR AND ELECTRONIC MOTIONS 29

2.1.2 CONICAL INTERSECTION 32

2.2 PHOTOCHEMICAL REACTIONS 36

2.3 CASSCF AND CASPT2 METHODOLOGY 40

2.3.1 POST-SCF-METHODS 40

2.3.2 THE CASSCF METHOD 44

2.3.3 SECOND ORDER PERTURBATION THEORY 46

2.4 HYBRID METHODOLOGY. QM/MM. 50

2.5 EVOLUTION OF A SYSTEM AFTER PHOTOEXITATION. DYNAMICS 59

2.5.1 WAVEPACKET PROPAGATION METHODS 59
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.5.2 SEMICLASSICAL AND CLASSICAL TRAJECTORIES</td>
<td>63</td>
</tr>
<tr>
<td>3 THE PHOTOCHEMISTRY OF THE NAIP-1 SWITCH</td>
<td>75</td>
</tr>
<tr>
<td>4 THE PHOTOCHEMISTRY OF THE NAIP-3 SWITCH</td>
<td>97</td>
</tr>
<tr>
<td>5 ISOMERIZATION OF RHODOPSIN AND BATHORHODOPSIN</td>
<td>126</td>
</tr>
<tr>
<td>6 CONCLUSIONS</td>
<td>157</td>
</tr>
</tbody>
</table>
# List of Figures

1.1 Classes of chromophores and photosensor families ............................ 3
1.2 The photoinduced isomerization of the 11-cis to all-trans retinal ........ 4
1.3 Structures of Ab and PSB11 ......................................................... 7
1.4 Biomimetic switches: the first and the second generations. ............... 8
1.5 Schematic representation of a myosin and an ATP synthase. .............. 10
1.6 The biarylidine motor cycle. ..................................................... 12
1.7 The NAIP-3 motor cycle. ....................................................... 13
1.8 Rhodopsin. The structure and the photo-cycle. .............................. 18

2.1 A scheme of a conical intersection of two potential energy surfaces. .... 34
2.2 A scheme of some common photophysical and photochemical events. .... 37
2.3 Rhodopsin photochemical reaction path scheme ............................... 38
2.4 Scheme: single and double excited determinants ............................. 42
2.5 QM/MM description of molecular systems ..................................... 51
2.6 Schematic representation of the division of a QM-MM bond ................ 54
2.7 Scheme of the chromophore and the frontier region in rhodopsin. ......... 56

3.1 Excited-state reaction paths. .................................................... 89
3.2 Analysis of the Z →E photochemical reaction path of NAIP-1. .......... 90
3.3 Ultrafast UV-vis spectroscopy of NAIP-1. ................................... 91
3.4 Time-resolved mid-IR spectra of NAIP-1. ................................... 92
3.5  (A) FTIR absorption spectra of Z-1; (B) Transient mid-IR spectra. (NAIP-1) 93

4.1  Schematic representation of an electrostatic switch based on Z/E isomerization 100

4.2  The QM/MM model of Ala-Zw-Trp. ......................................................... 103

4.3  Equilibrium structures of S₀-Z-3 and S₀-E-3 embedded in methanol. .......... 107

4.4  Energy profiles for scans and the IRC for the Z → E → Z cycle of NAIP-3 . 117

4.5  Absorption spectra of the Z and E isomers of NAIP-3 in MeOH ............... 118

4.6  A trajectory simulating the S₁ isomerization of NAIP-3 in methanol ....... 119

4.7  Conformational changes occurring during the excited state trajectory of Z-3 120

4.8  Structure of the designed Ala-Zw-Trp tripeptide ........................................ 121

4.9  Structure of the designed Ala-Zw-Trp tripeptide in its fluorescent state ... 122

5.1  PSB11 chromophore structure................................................................. 129

5.2  Comparison of Rh and bathoRh S₀ equilibrium structures ......................... 131

5.3  The S₁ MEP of bathoRh................................................................. 146

5.4  Scaled-CASSCF/Amber trajectories for bathoRh and Rh.......................... 147

5.5  Conditions for C₁₁-C₁₂ π-bond reconstruction in Rh and bathoRh ........ 148

5.6  Effect of the HOOP on the double-bond photoisomerization stereoselectivity 149

5.7  Evolution of the S₀ π-electron density along the semiclassical trajectories .. 150

5.8  Structure of the conical intersection of Rh............................................. 151

5.9  Structure of the S₁ and S₀ PES in the CI region for Rh and bathoRh ........ 152

5.10 Scaled-CASSCF/Amber trajectory for unreactRh ..................................... 153
List of Tables

2.1 The values of the re-parametrized QM/MM point charges for Lys-296. . . . . 56
4.1 Modified MM point charges for the part of the Trp residue. . . . . . . . . . . . 104
4.2 Electronic transitions of E and Z NAIP-3 (CASPT2//CASSCF/AMBER) . . 105
4.3 Composition of the photostationary state of 3 . . . . . . . . . . . . . . . . . . 111
4.4 Computed fluorescence maxima for the conformers of the Ala-Zw-Trp . . . . 114
Chapter 1

INTRODUCTION
1.1 PHOTORECEPTOR PROTEINS. RHODOPSIN

All three domains of life, Eukarya, Bacteria and Archaea, utilize light energy for various purposes. The photoreceptor proteins that have identified by now, can be classified in different ways.\(^1\) One classification is based on the chemical structure of their light-absorbing chromophores. Accordingly to this classification, the most important families are the rhodopsins,\(^2,3\) the phytochromes,\(^4\) the xanthopsins,\(^5\) and the flavins. The flavins are further divided in the cryptochromes,\(^6\) the phototropins,\(^7\) and the BLUF proteins\(^8\) (Figure 1.1).

A light-induced change in the electronic structure of the chromophore and a subsequent nuclear rearrangement are usually the primary photochemical events that activates the photoreceptors, as summarized in Figure 1.1. An E/Z isomerization is the photoreactive process in the rhodopsins, phytochromes and xanthopsins, while, as it has been recently unveiled for the flavinic photoreceptors, a variety of reactions can serve as the primary event such as transient cysteinyl-adduct formation in the case of phototropins.\(^9\) The primary reaction induces the formation of a signaling state, which communicates the process of photon absorption to a signal transduction partner.\(^10\)

The rhodopsin family is one of the most well-characterized photoreceptor.\(^11\) The main action of this kind of proteins is as ion-pumps (phototropism), a light detection (phototaxis), and vision.\(^11\)

Rhodopsin is the photoreceptor protein (a visual pigment) in the outer segment of rod visual cell responsible for twilight vision. It is a highly specialized G protein-coupled receptor (GPCR).\(^12\) Rhodopsin is composed of seven transmembrane α-helices (opsin) and its chromophore, 11-cis retinal, which forms the Schiff base linkage with a lysine residue of the 7th helix (Lys-296 in the case of bovine rhodopsin). The Protonated Schiff Base (PSB) is stabilized by a negatively charged carboxylate (Glutamate 113 in the case of bovine rhodopsin)\(^13\) (Figure 1.8A on the page 18).

The visual process in superior animals begins with photon absorption by the retinal
Figure 1.1: Well-characterized classes of chromophores and photosensor families. The curved arrows indicate the bond subject to isomerization, in retinal, both the labeled positions can undergo isomerization. (a) Only some of the more relevant functions are listed. [Adapted from ref. 1].

chromophore. The subsequent structural changes of the protein result in the activation of the transducin, the heterotrimeric G-protein coupled to rhodopsin, which in turn activates the signal transduction cascade that generates an electric response of the receptor cells.  

Several intermediates states are formed during this process (F 1.8B on the page 18), which have been identified and characterized by electronic, vibrational, resonance Raman and infrared spectroscopy (both low-temperature and time-resolved techniques). The photoisomerization of 11-cis retinal to all-trans retinal (Figure 1.2B ) is the first step of the process and it is commonly known as the primary event of vision. This is one of the fastest photochemical reactions in nature - it leads to the formation of the primary ground-state transient called photorhodopsin in only 200 femtosends. In the first thermodynamically
Figure 1.2: Scheme of the photoinduced isomerization of the 11-cis to all-trans retinal. 

stable all-trans retinal photoprodult, called bathorhodopsin, about 60% of the incident photon energy is stored by chromophore distortion. The primary event is extremely efficient, occurring with a quantum yield of 0.67. This makes the twilight vision a highly sensitive process, enabling human rod cells to respond to a single photon absorption. While progress has recently been made in the comprehension of the excited state evolution before the decay, some key features responsible for making rhodopsin an excellent light receptor are currently not fully understood, e.g., the nature of photorhodopsin and the reaction coordinate leading to bathorhodopsin. Only recently, Mathies and co-workers have reported experimentally derived structures for photorhodopsin and bathorhodopsin, while Okada and co-workers reported the first structural view of bathorhodopsin examined by X-ray crystallography under cryogenic conditions. The structural changes occurring earlier relative to the appearance of photorhodopsin, and which prompt the excited state decay of the retinal chromophore, have been recently addressed by a computational study by Olivucci and co-workers. This work shows that the isomerization process can be described by a
space-saving photo-isomerization mechanism that resembles bicycle pedal mechanism originally proposed by Warshel.\textsuperscript{21} In spite in the progress in understanding the mechanism of photo-isomerization that has been done in the last decades, there are a number of unsolved problems that need further investigation.
1.2 PHOTOCHEMICAL SWITCHES

A molecular switch is a molecular system that can exist in two or more stable states with different physical, chemical or biological properties and can be reversibly shifted between the states in response to external stimuli, i.e. change in pH, light, temperature, electrical current or potential. A well-known example of synthetic molecular switches are pH indicators, which display distinct colors as a function of pH. A simple molecular switch can be incorporated in a complex molecular system to be used as a trigger to turn off and on important physical, chemical or biological properties of the system, such as surface wettability, polymer elasticity, host-guest recognition, catalysis, enzyme activity, fluorescence, neural activity and membrane activity for intracellular drug delivery. Another important application of molecular switches is based on the fact that two (or more) stable states, that can be distinguished by different physical properties, enable the storage of information at a molecular level. It opens new horizons in molecular electronics.

Molecular switches can be regulated by an electric field, a scanning tunneling microscope (STM) tip, or by chemical/electrochemical reaction. Since light, due to the enormous progress in optics in the last decades, can be manipulated very precisely and a photoreaction can be completed on the pico- or even femto-second time scale, the switches driven by photons are of great interest. Light-driven molecular switches that are based on the process of cis-trans-photoisomerization around a double bond constitute a common class of photoswitches.

Such a molecule has two physically distinguished stable forms (E/Z) in the ground electronic state and can be reversibly switched through a photoexcitation from one isomer to another. Azobenzene (Ab) (Figure 1.3) is one of the most widely used photoswitches of this type. A number of theoretical and experimental studies have been conducted to elucidate the photochemistry of the azobenzene and azobenzene derivatives. It lead to better understanding the mechanism of the photoisomerization of the compound and to design of new derivatives with improved quantum yield. From the other side, the design
and preparation of novel switches that differ from \textbf{Ab} in size, polarity, and photoisomerization mechanism constitutes an attractive research target with the aim to bring diversity in the field. Olivucci and coworkers have recently started interdisciplinary research to design a novel class of biomimetic molecular switches that mimic the photochemistry of the retinal protonated Shiff-base chromophore of rhodopsin, an E/Z switch shaped by biological evolution (see section 1.1). In bovine rhodopsin (\textbf{Rh}), a selective photoisomerization of the 11-cis chromophore (\textbf{PSB11}) occurs via evolution of a single $\pi \rightarrow \pi^*$ excited state ($S_1$) and yields, upon decay, the all-trans ground state ($S_0$) product with a $67\%$ quantum yield.\textsuperscript{19} A $S_1$ lifetime of \textbf{Rh} is $\approx 150$ fs, a $S_0$ transient (photorhodopsin) appearance time is $\approx 180$ fs, and the primary photoproduct (bathorhodopsin) appearance time is $\approx 6$ ps.\textsuperscript{19} Although these properties make \textbf{Rh} an excellent reference for the design of E/Z switches irradiation of \textbf{PSB11} in solution features an unselective isomerization and a picosecond excited state lifetime.\textsuperscript{13}

It has been recently demonstrated\textsuperscript{53} that conformationally-locked protonated Schiff bases (\textbf{SH}) (Figure 1.4) provide a framework for novel switches. This work has led to the design of a series of second generation switches(\textbf{NAIP-1}, \textbf{NAIP-2}, \textbf{NAIP-3})(Figure 1.4). These switches have been a focus of the research in the laboratory for the last several years and they are the main object of this thesis.
The results that have been obtained by now show that, indeed, NAIPs are able to mimic various aspects of the photoisomerization of Rh. Recently, the strategy for the synthesis of these compounds has been proposed. That allows to complement computational studies by time-resolved optical experiments. As described in the chapter 3, time-resolved spectroscopy studies show that, in methanol solution, the N-alkylated indanylidine-pyrroline Shiff base NAIP-1, indeed, mimics several aspects of the photochemistry of Rhodopsin. They display excited state properties and photodynamics similar to those of Rh-embedded PSB11. The framework of the switches imposes a preferential direction (either clockwise or counterclockwise depending on a conformation) of photoisomerization around the central double bond that occurs, according to time-resolved spectroscopy studies, on a 300 fs scale. Moreover, subsequent cooling and solvent relaxation processes take less than 10 ps. All these characteristics make NAIP switches an valuable alternative of others family of photo-switches such as azobenzenes derivatives.
1.3 **ARTIFICIAL MOLECULAR MACHINES.**

**MOLECULAR MOTORS**

A molecular machine can be defined as an assembly of a discrete number of molecular components, i.e. supramolecular structure, that is designed to perform a specific mechanical movement as a consequence of appropriate external stimuli. A switch (1.2 can be considered a simple molecular machine. One source of the inspiration for this filed originates from molecular biology studies, in particular, the discovery of a family of moving proteins and their assemblies. Such biological molecular machines play an essential role in biological events by performing directed motions in response to certain biological stimuli. Cytoplasmic proteins such as myosins, kinesins and dyneins can move unidirectionally along linear tracks of actin filaments or microtubules. The main function of these molecular motors is to transport large molecules. In the cell, relatively small molecules, such as gases and glucose, diffuse to where they are needed. Some molecules, on the other side, such as vesicles are too large to diffuse to their destination. Motor proteins transport these substrates at the expense of adenosine triphosphate (ATP) as fuel.

The Figure 1.5A shows a schematic representation of the function of a myosin. Myosins are bound to actin filament - a double-helical supramolecular polymer of actin monomers with a pitch of 36 nm. One pitch includes 13 actin monomers. Myosins, driven by a power stroke generated through the hydrolysis of ATP, can travel pitch by pitch along the active filament in one direction. Myosins comprise a superfamily consisting of 18 different classes and can perform a series of different activities such as muscle contraction, vesicle transport and so on. Kinesins and dynenins are additional examples of molecular machines that perform unidirectional sliding motion with similar functions.

Another class of biological machines that are ubiquitous in all kind of mammalian cells is ATP synthases. ATP synthases use a proton gradient for the synthesis of ATP and, inversely, if ATP is supplied, ATP synthases generate an ion gradient. The structure of the
Figure 1.5: **A.** Schematic representation of a myosin moving along an actin filament: a myosin, driven by a power stroke generated through the hydrolysis of ATP, travels pitch by pitch along the active filament in one direction; **B.** Schematic representation of an ATP synthase: F$_1$ represents the shaft part, which rotates relative to the surrounding portion F$_0$.

Main part of ATP synthases is conserved throughout the evolution. ATP synthases are large protein complexes consisting of two domains named F$_0$ and F$_1$ (Figure 1.5 B). F$_1$ represents the shaft part, which rotates relative to the surrounding portion F$_0$. The motor in several steps can accomplish a 360° rotation. The efficiency of the transformation of the ATP energy into the rotary motion of the F$_1$-motor is close to 100%.

All biological machines that are described above are characterized by the ability to harness external energy to generate unidirectional mechanical motion. The artificial analogs of biological motors are not so sophisticated yet and any development of this field is of primary importance for the future of nano-scale technology. A number of classes of synthetic molecular motors has been designed and investigated. They can be classified by the origin of the driving energy (such as chemical, thermal, electrochemical or photo-energy), the
kind of the performed motion (translation, libration or rotation) and the classes of chemical compounds and processes involved. These molecular systems include molecular propellers, \(^8^5\) shuttles, \(^8^6,^8^7\) brakes, \(^8^3\) scissors, \(^8^8\) elevators, \(^8^9\) rotating modules, \(^9^0\) muscles, \(^9^1\) rotors, \(^8^0\) ratchets\(^9^2,^9^3\) and catalytic self-propelled objects.\(^9^4\)

By now a number of synthetic molecular motors that rotate unidirectionally have been reported.\(^8^0\) They can be considered to be synthetic mimics of the ATP synthases that has been described above. Light-driven molecular rotary motors that are based on the process of cis- trans- photoisomerization around a double bond constitute a class that is one of the promising candidates. As described in section 1.2, the process of E/Z photoisomerization is widely used to switch a system from one state to another in order to change physical, chemical or biological properties. Such a molecule has at least two physically distinguished stable forms (E/Z) in the ground electronic state and can be reversibly switched through a photoexcitation from one isomer to another. Two subsequent photon absorptions can, in principle, provide a complete unidirectional 360° rotary cycle that is required for a molecular motor. However, it is impossible to achieve this goal by means of switches that have only two stable forms. Even if one succeeds in designing a molecule that is driven exclusively in one direction (clockwise or counterclockwise) after absorption of the first photon, the second photoexcitation could drive the switch in the backward direction.

To address the problem Feringa and co-workers proposed a novel chiral four station E/Z molecular switch that consists of two identical halves connected by a central carbon-carbon double bond, which acts as the axis of rotation.\(^9^5\) This biarylidine motor (Figure 1.6 possesses two stereogenic centers bearing methyl substituents and, consequentially, either of two possible enantiomers has four states determined by an isomer form (E,Z) and a helicity (M,P). In the more stable isomers, these methyl groups adopt an axial conformation to minimize steric repulsion with the other half of the molecule. In order to function, the switch has to consume both photo- and thermal energy. One full cycle of the motor starts from an enantiomerically pure (3R,3R or 3S,3S) trans- (or cis-) isomer and consists of four steps.
Figure 1.6: The biarylidene motor cycle: A is photochemically converted into B, which in turn can isomerize in the same direction only after the thermal change in helicity (BC). C is photochemically converted into D, which in turn generates A via a thermal step. Due to steric factors, the direct photoexcitation of B and D would generate A and C, respectively, by isomerization of the central double bond with a reverse rotation.

The first and third steps are an energetically uphill photoisomerization, whereas the second and forth steps are a slow energetically downhill thermal helix inversion. The helicity of an enantiomer predisposes the direction of the rotation triggered by photoisomerization. The slow (rate-limiting) helix inversion steps are required to generate a conformer allowing for a forward, rather than a reverse subsequent photoreaction. Recently it has been shown that a derivative of this motor, incorporated in a liquid-crystal film, has been able to rotate a macroscopic glass rod placed on the surface of the film during irradiation with ultraviolet light. It is the first demonstration that rotary molecular motors can perform macroscopic work. The design of faster and more efficient E/Z molecular rotors requires an improvement of both photochemical (quantum yield and rate) and thermal (rate and equilibrium con-
Figure 1.7: The NAIP-3 motor cycle: (CIS, M) is photochemically converted into (TRANS, P), which in turn can isomerize in the same direction only after the thermal change in helicity (TRANS, M); (TRANS, M) is photochemically converted into (CIS, P), which in turn generates (CIS, M) via a thermal step.

stant) steps. Feringas group has conducted a series of investigations\textsuperscript{97,98,99,100} and has shown that modifications of the framework and substitution patterns can change the rate of the rotation by several orders of magnitude. From the other side, as stated in the section 1.2, Olivucci and coworkers have recently proposed a number of novel biomimetic photochemical switches that mimic the photochemistry of the retinal protonated Shiff-base chromophore of rhodopsin, an E/Z switch shaped by biological evolution. Besides their possible applications as effective molecular switches, they can provide promising candidates for effective molecular motors. Indeed, as it has been shown in methanol solution, the N-alkylated indanylidinepyrroline Shiff base NAIP-1 (Figure 1.4 on the page 8) displays excited state properties and photodynamics similar to those of Rh-embedded PSB11\textsuperscript{55,57} The framework of NAIP-1 imposes a preferential direction (either clockwise or counterclockwise depending on a conformation) of photoisomerization around the central double bond that occurs,
according to time-resolved spectroscopy studies, on a 300 fs scale. Moreover, subsequent cooling and solvent relaxation processes take less than 10 ps.\textsuperscript{57} Considering motors design, the data suggest that a half-rotary cycle of the motor can be completed in less than 10ps (a few orders of magnitude faster than the fastest known biarylidine, 6ms).\textsuperscript{99} Therefore, it is likely that after insertion of a stereogenic center enantiomeric forms of NAIP derivatives can provide effective molecular motors. The zwitterionic switch \textbf{NAIP-3} can be considered as an example of such a motor. The Figure 1.7 shows the four-stage cycle of the \textbf{NAIP-3} motor:

(1) cis- form of the compound with the helicity M, (CIS, M), is photochemically converted into trans- form with helicity P (TRANS, P);

(2) the second thermal step is used to inverse the helicity to prepare the molecule for isomerization in the same directon (TRANS, P) \rightarrow (TRANS, M);

(3) (TRANS, M) is photochemically converted into (CIS, P);

(4) the cycle is completed by generation of (CIS, M) from (TRANS, M) via a thermal step.
1.4 AIMS AND MAIN POINTS OF THE THESIS

The aim of this thesis is to contribute to the investigation of the photo-isomerization mechanism of both NAIPs (NAIP-1 and NAIP-3) and rhodopsin (bathorhodopsin). A version of QM/MM methods that was developed in the laboratory of Massimo Olivucci and successfully implemented in a number of investigations of Rh photochemistry is used in this work.\textsuperscript{101,102} The method is based on an ab initio multi-configurational treatment of the chromophore (CASSPT2//CASSF/AMBER) electronic structure and MM treatment of the rest of the system.\textsuperscript{1} The flexibility and the precision of the multi-reference CASSPT2//CASSCF methodology have been proven by a number of investigations of exited states for various kinds of molecular systems. The coupling of the methodology with AMBER force field Molecular Mechanics (MM) gives a powerful tool to deal with chromophores in the complex environments (solvents and proteins) including the systems under investigation in this thesis.

Due to the nature of the problem, the studies can not be limited to the computation of equilibrium properties but require a description of the entire photo-cycle such as $E \rightleftharpoons Z$ for NAIP switches and rhodopsin (Rh) $\rightleftharpoons$ bathorhodopsin (bathRh). In other words, one needs to obtain, besides the spectra and energies of stationary states, the potential energy surfaces controlling the switch $E \to Z$ and $Z \to E$ excited state evolution, its decay and ground state relaxation, and $E(P) \to E(M)$ and $Z(P) \to Z(M)$ thermal conversion, in the case of photoswitches, and the potential energy surfaces that drive the processes of Rh $\to$ bathoRh and bathoRh $\to$ Rh photo-izomerization in the case of Rh. To estimate the time of the cycle, a dynamics calculations are also needed.

The investigation of the NAIP-photoswitches is a highly interdisciplinary study that involves organic synthesis, UV and IR time-resolved spectroscopy, and computational chemistry and, therefore, a part of the research are conducted by our colleagues in several other groups\textsuperscript{ii}. Although the thesis is centered around the computational part we used the UV and

\textsuperscript{1}this methodology is described in details in Chapter 2 of the thesis
\textsuperscript{ii}we want to express, one more time, our gratitudes to our colleagues for their collaboration
IR data provided by our collaborators in the instances where it is needed to achieve clarity.

In this thesis

(i) (Chapter 3) The photocycle of NAIP-1 is investigated. The ab initio multi-configurational quantum chemistry-based QM/MM computations coupled with time-resolved spectroscopy allow to shed light on the mechanism of the light-induced isomerization of the switch in methanol. The studies show that NAIP-1 exhibits the photochemistry that is very similar to the photochemistry that has been reported for rhodopsin – the isomerization occurs on a 0.3 ps time scale and is followed by <10 ps cooling and solvation. The reactive first excited state \( S_1 \) has, like in the case of rhodopsin, charge transfer character. The main feature of the photo-isomerization mechanism (again similar to Rhodopsin) is the single-double bond inversion in the \( \pi \)-skeleton of the molecule that is followed by the rotation around the central double bond. When contemplating the design of an effective molecular rotary motor, the promising result is that a full (2-photon-powered) cycle can be completed within 20 ps. This is proved by tracing the evolution of the infrared spectrum of NAIP-1. Thus, NAIP-1 is not only an example of an effective molecular switch (two-stage system) but it also provides the basis for the design of an efficient molecular motor (that has to be a four-stage system).

(ii) (Chapter 4) The NAIP-3, that can be considered to be modified (NAIP-1), has a \(-\text{COO}^-\) group incorporated into the Shiff base framework of the switch. Due to this group NAIP-3 is an example of the four-stage switch, that can work as a molecular motor (see section 1.3). The investigation of the photochemistry of NAIP-3, which followed the NAIP-1 studies, is provided in this chapter. Moreover, we show that, due to its large dipole moment and efficient photoisomerization, such a system may constitute the prototype of a novel generation of electrostatic switches achieving a reversible light-induced dipole moment change on the order of 30 Debye. That opens up a new perspective for the light-driven conformational control of macromolecular structures determined by polar interactions.

(iii) (Chapter 5)

As described in section 1.1 the photoexcitation of Rh results in the sub-picosecond iso-
merization of PSB11 to its all-trans isomer (PSBT) that, following picosecond vibrational relaxation, leads to production of the metastable intermediate bathorhodopsin (bathoRh). Low temperature (77 K) irradiation of Rh with 580 nm light establishes, within one second, a photostationary state characterized by a Rh/bathoRh ratio of 61:1 thus showing that bathoRh can be photochemical reconverted to Rh. Recent computational studies have elucidated the excited state isomerization dynamics of Rh.\textsuperscript{20} It has been shown that the chromophore reactive $\pi$-bond completely breaks on a 100 fs timescale via a bicycle-pedal motion that leads to excited state decay. In spite of its central role in phototransduction, we are unaware of any analogous study carried out for bathoRh. A study of the photoisomerization dynamics of bathoRh appears of basic theoretical interest as the Rh $\rightarrow$ bathoRh and Rh $\rightarrow$ bathoRh isomerizations constitute the components of a photochromic cycle/equilibrium implemented at the biological level. In the present work, Rh and bathoRh quantum-mechanics/molecular-mechanics (QM/MM) models based on ab initio multiconfigurational wavefunctions are employed to investigate the light induced $\pi$-bond breaking and reconstitution occurring during the Rh $\rightarrow$ bathoRh and bathoRh $\rightarrow$ Rh isomerizations allowing description of the full cycle and providing new information concerning the mechanism of the photoisomerization of Rh and bathoRh.
Figure 1.8: A. The α-helices structure of Rh and the chromophore (PSB); B The photocycle of the visual pigment.
Bibliography


Chapter 2

METHODOLOGY
2.1 TOPOLOGY OF POTENTIAL ENERGY SURFACES

Potential Energy Surface (PES), a function that describes an electronic energy of a molecular system for any fixed nuclear configuration, $E(R^{3N})$, is one of the most important concepts in chemistry. This approach according to the Born-Oppenheimer approximation allows to picture molecular processes as a propagation of nuclei over the potential energy surface provided by the electrons. The validity of the Born-Oppenheimer approximation for the vast majority of chemical phenomena is a well-established fact. However, in many important cases the approximation breaks down and nuclear and electronic motion become coupled. This breakdown is common in the photochemistry of polyatomic molecules and, therefore, plays an important role in the description of the processes studied in the thesis. However, even in this case the language of potential energy surfaces continues to be a convenient frame needed to describe nuclear motion.
2.1.1 SEPARATION OF NUCLEAR AND ELECTRONIC MOTIONS

The central equation of the non-relativistic quantum chemistry is the Schrödinger equation

\[ i\hbar \frac{\partial \Psi(R, r, t)}{\partial t} = \hat{H}(R, r)\Psi(R, r, t), \]  

(2.1)

where \( \Psi(R, r, t) \) is the wavefunction that depends both of nuclear, \( R \), and electrons, \( r \), coordinates; the molecular Hamiltonian, \( \hat{H}(R, r) \), can be written in the form

\[ \hat{H}(R, r) = \hat{T}_n(R) + \hat{T}_e(r) + \hat{U}(R, r), \]  

(2.2)

where \( \hat{T}_n(R) \), \( \hat{T}_e(r) \) are the kinetic energy operators for nuclei and electrons, respectively, and \( \hat{U}(R, r) \) the potential energy function, a function of nuclear, \( R \), and electron, \( r \), coordinates. By setting the nuclear kinetic energy to zero the clamped-nucleus Hamiltonian, the starting point for the Born-Oppenheimer approximation, can be obtained

\[ \hat{H}(R, r) = \hat{T}_e(r) + \hat{U}(R, r), \]  

(2.3)

where the nuclear coordinates \( R \) is a parameter. Born and Huang For any value of \( R \), the eigenvalues of 2.3, \( V \), and eigenfunction, \( \Phi \), can be found from

\[ \hat{H}_{el}\Phi_i(r, R) = V_i(R)\Phi_i(r; R). \]  

(2.4)

The set of eigenfunctions are complete and orthonormal, as a set of eigenfunctions of a Hermite operator (for more details see ref.\(^1\)). This functions can be used as a basis set to expand the total wavefunction (a trial solution for the equation 2.1)

\[ \Psi(R, r) = \sum_i \chi_i(R)\Phi_i(r; R), \]  

(2.5)
where the $\chi_i$ are functions of nuclear coordinates that act as the expansion coefficients. By substituting 2.5 into the Schrödinger equation 2.1, multiplying by another electronic eigenstate and integrating over electron coordinates, the following coupled equations set for the expansion coefficients can be obtained

$$[\hat{T}_n + V_i]\chi_J - \sum_i \hat{\Lambda}_{ji}\chi_i = i\hbar \frac{\partial \chi_i}{\partial t},$$

(2.6)

where the matrix elements

$$\hat{\Lambda}_{ji} = \delta_{ji} \hat{T}_n - \langle \Phi_j | \hat{T}_n | \Phi_i \rangle$$

(2.7)

are the nonadiabatic coupling operators. They are operators in the space of coordinates and describe the dynamical interaction between the electronic and nuclear motion. This method to derive a set of coupled equations 2.6 from the Schrödinger equation 2.1 was first introduced by Born and Huang.\textsuperscript{2}

The Schrödinger equation in the form 2.6 involves the complete, infinite set of adiabatic electronic states. Taking into consideration that only states close in energy coupled significantly, the set of states included can be truncated to include only the relevant group. In this case the equation 2.6 is no longer the full Schrödinger equation but is called the group Born-Oppenheimer approximation. The dimension of the $\Lambda$ are now the number of state in the set. The Born-Oppenheimer approximation is obtained in the limit that the set contains a single state

$$[\hat{T}_n + V - \Lambda]\chi = i\hbar \frac{\partial \chi}{\partial t}$$

(2.8)

The further approximation then $\Lambda$ is completely ignored

$$[\hat{T}_n + V]\chi = i\hbar \frac{\partial \chi}{\partial t}$$

(2.9)

is called the adiabatic approximation.

The set of equations 2.6 can be transformed (see ref.$^3$) (using atomic units and and scaled
rectangular coordinates) to the matrix equation

\[
\left[ -\frac{1}{2M} (\nabla + \mathbf{F})^2 + \mathbf{V} \right] \chi = i\hbar \frac{\partial \chi}{\partial t},
\]

(2.10)

where

\[
\mathbf{F}_{ij} = \langle \Phi_i | \nabla | \Phi_j \rangle = \frac{\langle \Phi_i | \nabla \hat{H}_{el} | \Phi_j \rangle}{|V_j - V_i|} \quad \text{for} \quad i \neq j
\]

(2.11)

is the derivative coupling operator and M is a suitable mass-scale. In the equation 2.10 the kinetic energy operator for nucleus has been replaced by the dressed kinetic energy operator \((\nabla + \mathbf{F})^2\).
2.1.2 CONICAL INTERSECTION

The nonadiabatic operators in equation 2.6, \( \Lambda_{ij} \), are nonlocal, derivative operators. According equations 2.10 and 2.11 large mass leads to small \( \Lambda \) with full agreement with the idea of the adiabatic approximation. However, the nonadiabatic coupling also depends inversely on the energy gap between states. If the gap becomes small, the coupling between the nuclear motion on different surfaces occurs, i.e. a wavepacket initially localized on one surface will spread to another in a region where they are energetically close. To describe this kind of processes one has to take into consideration the derivative coupling \( F \). This nonlocal operator, that describes the coupling between electronic states by the nuclear momentum, is singular in the vicinity of an intersection between two electronic states.

An unitary transformation of the electronic basis set (for more details see ref. 4)

\[
\tilde{\Phi} = S(R)\Phi
\]

with the purpose to remove the nonadiabatic coupling from the kinetic energy operator results in the following, so-called, diabatic representation of the Shrödinger equation

\[
\left[ -\frac{1}{2M} \nabla^2 + \tilde{W} \right] \tilde{\chi} = i\hbar \frac{\partial \tilde{\chi}}{\partial t},
\]

(2.13)

where

\[
\tilde{W} = S^\dagger V S; \quad \tilde{\chi} = S\chi.
\]

(2.14)

The equation 2.13 a kinetic energy operator is diagonal and the coupling between electronic states are in the potential energy matrix. A number of different schemes have been developed to obtain the diabatic states. 4

To make a bridge between diabatic and adiabatic representations, the potential matrix elements \( W \) in equation 2.13 is expanded as Taylor expansions around a particular point,
\( R_0: \)

\[
W(R) = W^{(0)} + W^{(1)} + W^{(2)} + \ldots \tag{2.15}
\]

\( R_0 \) could be a point of degeneracy, the Franck-Condon point, or any other point of interest. At \( R_0 \) the diabatic basis can be taken to be equal to the adiabatic basis due to the fact that transformation 2.12 defines only up to a constant unitary transformation. The zero-order matrix is then diagonal, with

\[
W^{(0)} = V(R_0) \tag{2.16}
\]

It can be shown that owing to the properties of the diabatic functions, the first-order matrix elements can be expressed in the diabatic basis at \( R_0 \) as

\[
W^{(1)}_{\alpha,ss'} = \left\langle \Phi_s \left| \frac{\partial H_{el}}{\partial R_\alpha} \right| \Phi_{s'} \right\rangle R_\alpha, \tag{2.17}
\]

where \( R \) is the displacement vector relative to \( R_0 \).

The case then only two adiabatic (and, therefore two non-adiabatic) surfaces are involved in the process is of special interest. The adiabatic potential energy surface are the eigenvalues of the diabatic potential matrices

\[
V = SWS^\dagger. \tag{2.18}
\]

Thus, for a two-state system, the adiabatic surface can be analytically expressed as

\[
V_\pm = \Sigma \pm \sqrt{\Delta^2 + W_{12}^2} \tag{2.19}
\]

where

\[
\Sigma = \frac{1}{2}(W_{11} + W_{22}); \quad \Delta = \frac{1}{2}(W_{22} - W_{11}) \tag{2.20}
\]

The first-order terms can be written as

\[
W^{(1)}_{11} = k^{(1)}R; \quad W^{(1)}_{22} = k^{(2)}R \tag{2.21}
\]
Figure 2.1: A scheme of a conical intersection of two potential energy surfaces plotted along the branching space $\vec{x}_1, \vec{x}_2$.

where $k^{(s)}_a$ are the force integrals in Equation 2.17 with $s = s'$, and

$$\frac{1}{2}(W^{(1)}_{22} - W^{(1)}_{11}) = \delta \mathbf{R}; \quad W^{(1)}_{22} = \lambda \mathbf{R},$$  \hspace{1cm} (2.22)

where $\delta = \frac{1}{2}(k^{(2)} - k^{(1)})$ is a gradient difference vector and $\lambda$ the linear coupling vector, with elements defined by the integral in Equation 2.17 with $s \neq s'$

Near the degeneracy point, $\Delta \approx \delta \mathbf{R}$ and $W_{12} \approx \lambda \mathbf{R}$ and Equation 2.19 can be written

$$V_\pm = \Sigma \pm \sqrt{(\delta \mathbf{R})^2 + (\lambda \mathbf{R})^2}$$  \hspace{1cm} (2.23)

and this is the equation of an elliptic double cone (i.e., with different axes) with vertex at the origin. Hence the crossing points are called conical intersections. The $\delta$ and $\lambda$ vectors are referred to as the “branching space” and the (n-2)-dimensional subspace of the n nuclear coordinates is called the intersection space, an hyperline consisting of an infinite number of conical intersection points.
To summarize, a potential energy surface at a conical intersection has the structure of a double-cone in the plane crosssection defined by two vectors, the gradient difference and the linear coupling vector and it is degenerate in the orthogonal rest of the n-2 dimension of the configuration space (Figure 2.1). These vectors form the so-called “branching space” (on the Figure 2.1 we define $\vec{x}_1, \vec{x}_2$ as unit vectors parallel to the gradient difference and linear coupling, respectively). As we move in this plane, away from the apex of the cone, the degeneracy is lifted. The defenerate space of n-2 coordinates, the intersection space, is a hyperline consisting of an infinite number of conical intersection points.
2.2 PHOTOCHEMICAL REACTIONS

The absorption of a photon by a molecule can lead to a series of events, the most common of which are schematically represented in Figure 2.2. They can be interpreted in terms of the topology of two (or more) Potential Energy Surfaces (PES) of the electronic states that are involved in a process. The absorption of a photon and electron reorganization is the first and the fastest step ($< 10^{-15}$ sec). After an electronic transition the molecule is driven into a vibrationally excited state, so an initial vibrational relaxation may take place internally within the molecule. This is referred to as intramolecular vibrational relaxation (IRV) (Figure 2.2a). Relaxation may involve the solvent or surrounding molecules through a vibrational energy transfer (VET) process. The rate of these relaxation processes depends on the molecular size but is commonly in the range of $10^{-14} - 10^{-11}$ s.

Alternatively, a nuclear re-organization can lead to the excited state structure of another species. In the latter case, a Photoadiabatic Reaction (PAR) has taken place (Figure 2.2b), with adiabatic referring to a process involving a single PES. Once molecules reach a region of minimum on an excited state space (Figure 2.2a, b and c), they can relax through a radiative decay to the lower electronic state. Radiative emissions can involve states of the same multiplicity (fluorescence) or states of different multiplicity (phosphorescence). Emission times are ranging from $10^{-9}$ - $10^{-6}$ s for fluorescence to $10^{-3}$-$10^2$ s for phosphorescence.

On the other hand, non-radiative deactivation (radiationless decay) processes are very common, and they occur whenever molecules on an excited state find a way to access a different PES lower in energy. These processes are mediated by so-called singlet-triplet crossings (between states of different spin multiplicity-ISC in Figure 2.2c) or by conical intersections (between two PESs of the same spin multiplicity) Figure 2.2d.
Figure 2.2: Scheme of some common photophysical and photochemical events occurring in a molecular system upon light absorption.

Early thinking about nonradiative transitions was dominated by the application of Fermi’s golden rule. According it the nonradiative relaxation rate is expressed as

\[ k(S_1 \rightarrow S_2) = \frac{2\pi}{\hbar} \sum_f |V_{if}|^2 \rho(E) \]  

(2.24)

where \( i \) and \( f \) refer to the initial and final vibrational eigenstates on the \( S_1 \) and \( S_0 \) Born-Oppenheimer electronic surfaces, respectively. The matrix element \( V_{if} \) represents the electronic coupling between these Born-Oppenheimer states. Finally, the density of states of \( S_0 \) at the given energy \( E \) is denoted \( \rho(E) \). Equation 2.24 is valid as long as the coupling between electronic states is weak and low-order perturbation theory is sufficient. This is true in the
Figure 2.3: Rhodopsin photochemical reaction path scheme: model intersecting potential energy surfaces showing an excited state barrierless evolution (mapped by means of IRC computation) pointing to a conical intersection (CI). The ground state initial directions of relaxation starting at CI are indicated by arrows and they are located by means of an IRD computations.

region where two potential energy surfaces get close enough to make a transition probability not negligible but not so close to brake the approximation, i.e avoiding crossing region.

In the case of Rhodopsin, a photo-isomerization proceeds through a conical intersection (Figure 2.3) and the description of the radiationless decay requires more complicated theories that are partially described in the section 2.5.

The investigation of the topology of the potential energy surfaces that are relevant to a photochemical reaction starts from location of characteristic points of these surfaces such as minima (FC points) and CI points (see Figure 2.3) and is followed by Intrinsic Reaction Coordinate (IRC) caculations. IRC can be defined as the steepest-descent path in mass-weighted Cartesian coordinates from one characteristic point (i.e. FC) to another (i.e.
CI). Due to the fact that the defined in this way IRC is the path of the minimum distance (geodesic) between these two points, it can be considered as an invariant topological characteristic.\textsuperscript{5} The methodology that allows to find stationary points and IRC is important and well-developed part of quantum chemistry.\textsuperscript{6} To calculate IRC between a CI point and minimum points on $S_0$ one needs a direction to start the path. The small step vector defines the initial relaxation direction (IRD) towards the product or to the reactant (see Figure 2.3) can be calculated based on the vectors that defines the branching plane ($\vec{x}_1, \vec{x}_2$). The procedure is described in ref.\textsuperscript{7}


2.3 CASSCF AND CASPT2 METHODOLOGY

2.3.1 POST-SCF-METHODS

The two main drawbacks of the Hartree-Fock method, that have stimulated the development of more accurate methods, are the lack of description of a significant portion of electron correlation and the deficiency of a (restricted) Hartree-Fock wave function in describing homolytic dissociation (usually referred to as the RHF dissociation problem). The main drawback concerns the correlated motion of electrons and the reduced probability of finding an electron in the immediate vicinity of another electron. The Hartree-Fock wave function is able to take into account the Fermi correlation due to the motion of electrons of the same spin, but fails to describe the so-called Coulomb correlation due to the correlated motion of electrons of opposite spin.

The Hartree-Fock approximation for the ground state of a molecular system ($|\Psi_0>$) is represented by the single Slater determinant formed from the $N$ occupied spin orbitals with the lowest energies

\[
\Psi(x_1, x_2, \ldots, x_N) = (N!)^{-1/2} \begin{vmatrix}
\chi_i(x_1) & \chi_j(x_1) & \cdots & \chi_k(x_1) \\
\chi_i(x_2) & \chi_j(x_2) & \cdots & \chi_k(x_2) \\
\cdots & \cdots & \cdots & \cdots \\
\chi_i(x_N) & \chi_j(x_N) & \cdots & \chi_k(x_N)
\end{vmatrix}
\]  

Or in a more compact notation

\[
|\Psi_0> = |\chi_1 \chi_2 \cdots \chi_a \chi_b \cdots \chi_N >
\]  

where the normalization factor is implicit and only diagonal elements are shown,

It can be convenient to label these $N$ occupied spin orbitals by the indices $a$, $b$, $c$,... and the remaining unoccupied (from now on called virtual) $2K-N$ members of the set...
{χ_k} by the indices r, s, t,... This situation can be described by the following scheme

While such a Slater determinant is the best (in a variational sense) approximation to the ground state, it is clearly not the only one which can be formed from the total 2K spin orbitals. In fact, the number of combinations of 2K objects taken N at a time is the binomial coefficient

\[
\binom{2K}{N} = \frac{(2K)!}{N!(2K-N)!}
\]  

(2.27)

Replacing one of the orbitals in the equation 2.26 one of the unoccupied orbital gives a new determinant. These other determinants can be interpreted as approximations of excited states of the system or they can be used in linear combinations with Ψ_0 for a more accurate description of the ground state or of any state of the system. The determinants are commonly referred to as configurations.

A singly excited determinant (|Ψ_a^r >) can, for instance, be described by the scheme in Figure 2.4-left where an electron, which occupied χ_a in the Hartree-Fock ground state has been promoted to a virtual spin orbital χ_r

\[
|Ψ_a^r > = |χ_1χ_2⋯χ_rχ_b⋯χ_N >
\]  

(2.28)
Similarly, a doubly excited determinant (\( |\Psi_{ab}^{rs} > \)) is one in which electrons have been excited from \( \chi_a \) and \( \chi_b \) to \( \chi_r \) and \( \chi_s \) as shown in Figure 2.4-right.

\[
|\Psi_{ab}^{rs} > = |\chi_1 \chi_2 \cdots \chi_r \chi_s \cdots \chi_N >
\]  

(2.29)

All the other possible determinants can be similarly classified as either singly, doubly, ..., \( N \)-tuply excited determinants.

It can been shown that an exact wave function for any state of the system can be written in the form

\[
|\Phi > = c_0 |\Psi_0 > + \sum_{ra} c_a^r |\Psi_a^r > + \sum_{a<b, r<s} c_{ab}^{rs} |\Psi_{ab}^{rs} > + \sum_{a<b<c, r<s<l} c_{abc}^{rst} |\Psi_{abc}^{rst} > + \cdots
\]  

(2.30)

where the summations run over all unique doubly, triply and higher excited determinants.
The infinite set \( \{|\Psi >\} = \{|\Psi^a_r >, |\Psi^{rs}_{ab} >, |\Psi^{rld}_{abc} >, \cdots \} \) is a complete set for the expansion of any N-electron wave function. This procedure is called Full Configuration Interaction (FCI).

Within the Born-Oppenheimer approximation and in the limit of an infinite basis set, the lowest eigenvalue \( E \) obtained with this method is the exact non-relativistic ground state energy of the system. This exact energy \( E_0 \) differs from the Hartree-Fock-limit energy \( E_0 \) for a certain quantity that is called the correlation energy \( (E_{corr}) \)

\[
E_{corr} = E - E_0.
\] (2.31)

Methods that are goes beyond single electron configuration are called post-SCF-methods. The procedure known as Full CI represents the realistic case in which only one-electron basis set is truncated to a finite number, while all the possible determinants (configurations) are taken into account for the description of the electronic states. Thus, it provides exact eigenvalues and eigenvectors within the space spanned by the finite basis set. On the other hand, the cost of Full CI becomes prohibitive for all but small systems, and it is necessary to move towards more approximated methods by truncating both the one-electron and the many-electron basis sets. The type of truncation carried out, together with the class of techniques (variational principle, perturbation theory and others), characterize most of the methods currently available. A large number of systems in their ground state are qualitatively described pretty well by a single configuration, thus many approximations have been developed for which the starting point is the Hartree-Fock wave function. Examples are Muller-Plesset perturbation theory based (MP2, MP3, MP4), singles and doubles configuration interaction (CISD) as well as coupled cluster (CC) methods.

Another approach to the problem is multi-configurational Self-Consistent-Field (MCSCF) method. The MCSCF wave function is a truncated CI expansion

\[
|\Phi_{MCSCF} > = \sum_I c_I |\Psi_I >
\] (2.32)
in which both the expansion coefficients ($c_I$) and the orthonormal orbitals contained in $I$ are variationally optimized. The complete active space SCF (CASSCF) is a variant of MCSCF that has been extensively used in the present thesis and it will be discussed in the next section.

### 2.3.2 THE CASSCF METHOD

At the CASSCF level, long-range effects related to the so-called non-dynamic (static) correlation effects are taken into account, leading to the proper description of several nearly degenerate configurations. The remaining correlation effects (dynamic correlation), related to the instantaneous short-range inter-electronic interactions, are usually accounted for in a second step by using other methods like multi-reference CI (MRCI) or perturbation theory (see next section 2.3.3)

The CASSCF method is based on a selection of active electrons distributed among the active orbitals in all possible ways consistent with the spatial and spin symmetry of the electronic state. The choices of active orbitals and electrons depends on the problem under investigation.

Thus, the orbitals are classified in three categories: inactive, active and secondary orbitals (also called virtual). The inactive orbitals are the ones that remain always doubly occupied in all CASSCF configurations, while the virtual orbitals are always unoccupied. The number of electrons occupying inactive orbitals is twice the number of inactive orbitals, the rest of the electrons, the active ones, occupy the active orbitals. The CASSCF wave function is a linear combination of all possible configurations that arises from the distribution of the active electrons among the active orbitals and are consistent with a given spatial and spin symmetry. This means that, in the configuration space spanned by the active orbitals, the CASSCF function is a complete (or full) CI function. During the variational process inactive orbitals are also optimized but they are treated as in the (restricted) Hartree-Fock method. The CASSCF energy is invariant to rotations among the active orbitals. In particular,
since an active orbital may be occupied by either zero, one or two electrons in any given determinant, these CASSCF orbitals do not have unique eigenvalues associated with them. Accordingly, it is not possible to refer to the eigenvalues as orbital energies. Instead, it is possible to describe the occupation number of each such orbital $i$ as

$$\text{(occ. number)}_{i,\text{CASSCF}} = \sum_n (\text{occ. number})_{in} c_n^2,$$  \hspace{1cm} (2.33)

where the sum runs over all $C_s$ configurations and the occupation number of the $i$-th orbital in each configuration is multiplied by the percentage contribution of that configuration to the total wave function. For a normalized CASSCF wave function the sum of the squares of all configurations coefficients is unity and the contribution of any configuration to the wave function is simply the square of its expansion coefficient.

The CASSCF method is widely employed for the computation of ground state systems but has also found extensive use for electronically excited states. The State-Average CASSCF (SA-CASSCF) method is a modification of the CASSCF that is used in case of a phenomenon called root flipping occurs – a given state ordering can change during the computation of excited states, leading to an energy of a higher root which becomes lower than the energy of a lower root. Several states that belong to the same symmetry can be computed by SA-CASSCF, where the energy functional is defined as average of a number of states (I=1,M)

$$E_{\text{average}} = \sum_i \omega_i E_i,$$  \hspace{1cm} (2.34)

where $\omega_i$ is the factor of the relative weight for each state considered. The result is a set of average orbitals and a number of orthogonal wave functions equal to the number of roots used in the average process.

As already mentioned, the CASSCF method, lacking an account of dynamic correlation, is often not accurate enough to reproduce the energetics of a chemical or spectroscopic process. It becomes necessary to go beyond this approach with methods such as multi-reference-CI
or multi-configurational perturbation theory. While coupled cluster (CC) theory is the most accurate correlation treatment available for systems devoid of near-degeneracy effects, a corresponding development based on a multi-configurational reference function has so far not been possible. Therefore, in the multi-configurational case an available treatment of electron correlation is the multi-reference-CI method (MRCI). Usually, it consists of an expansion of the wave function in all singly and doubly excited configurations (MRCISD) with respect to a set of chosen reference configurations.

\[ |\Phi_{\text{MRCI}} \rangle = \sum_I (c_I |\Psi_I \rangle + \sum_{r,a} c^r_{Ia} |\Psi^r_{Ia} \rangle + \sum_{r,a,s,b} c^{rs}_{Iab} |\Psi^{rs}_{Iab} \rangle) \]  

where \( I \) indicates a set of reference configurations and the other terms contain its singly and doubly excited configurations. This can be a very accurate method for small molecules but it rapidly becomes very demanding for larger system. One strategy that has turned out to be accurate in a wide variety of applications, in particular in electronic spectroscopy, is based on perturbation theory. It starts from the CASSCF wave function and includes dynamic correlation employing second order perturbation theory by means of the so-called CASPT2 method (second-order complete active space perturbation theory). Following this protocol, the results can be accurate to within 0.1-0.2 eV (< 5 kcal mol\(^{-1}\)), which is an affordable error bar for most spectroscopic applications.

### 2.3.3 SECOND ORDER PERTURBATION THEORY

Dynamic correlation effects can be treated by using perturbation theory, which is a computationally more efficient approach than the MRCI. When the reference function is a single determinant one (Hartree-Fock function), Mller-Plesset second order perturbation theory (MP2) has been extensively used to treat electron correlation for ground states. The extension to reference functions of the multi-configurational CAS type led, in the late 80s to the CASPT2 method.\(^\text{10,11}\) Since then a large number of successful applications to organic
molecules, transition metal complexes and recently also in heavy element chemistry\textsuperscript{12,9} have been reported.

A CASPT2 calculation can be viewed as a conventional non-degenerate perturbative method, where a single reference function is considered. The particularity is that such a reference function (zeroth-order wave function) is a CASSCF wave function designed to include the electronic states of interest. The stationary Schrödinger equation can be written

\[
\hat{H}(\lambda)\psi(\lambda) = E(\lambda)\psi(\lambda)
\]

and expanded in power series of $\lambda$ as follows

\[
\hat{H}(\lambda) = \hat{H}^{(0)} + \lambda \hat{H}'
\]

\[
\Psi(\lambda) = \Psi^{(0)} + \lambda \Psi^{(1)} + \lambda^2 \Psi^{(2)} + \ldots
\]

\[
E(\lambda) = E^{(0)} + E \Psi^{(1)} + E^2 \Psi^{(2)} + \ldots
\]

where $\lambda$ is a parameter. Corrections of order $k$ to the wave function $\Psi(k)$, and to the energy $E(k)$, are obtained by consecutively solving the corresponding perturbation equations; here reported for the first few terms

\[
\hat{H}_0|\Psi_0> = E_0|\Psi_0>, \quad (\hat{H}_0 - E_0)|\Psi_1> = (E_1 - \hat{H}')|\Psi_0> ,
\]

\[
(\hat{H}_0 - E_0)|\Psi_2> = (E_1 - \hat{H}')|\Psi_0> + E_2|\Psi_0> .
\]

In the CASPT2 approach the wave function is corrected at the first-order while the correction to the energy is at the second-order. Assuming that the perturbed wave functions are orthogonal to the zeroth-order function, it is possible to write the following expressions
for the energies up to the second order:

\[ E^{(0)} = \langle \Psi^{(0)} | \hat{H} | \Psi^{(0)} \rangle \] (2.41)

\[ E^{(1)} = \langle \Psi^{(0)} | \hat{H}' | \Psi^{(0)} \rangle \] (2.42)

\[ E^{(1)} = \langle \Psi^{(0)} | \hat{H}' | \Psi^{(1)} \rangle \] (2.43)

The set of functions required to compute the first-order correction to the wave function is formed by those that interact with the zeroth-order wave function through the Hamiltonian in Rayleigh-Schrödinger perturbation theory, and belong to what it is known as the first-order interacting space. This includes functions generated by singly and doubly excited configurations with respect to the zeroth-order wave function. For a detailed description of all mathematical aspects and formalism, see references. The first-order correction of the wave function is expanded in the basis of the functions \( |j\rangle \) belonging to the first-order interacting space, hereafter called \( V_{SD} \)

\[ \Psi^{(1)} = \sum_{j=1}^{M} c_j |j\rangle \quad M \geq \text{dim}V_{SD} \] (2.44)

The procedure to determine the coefficients \( \{c_j, \, j = 1, 2, \ldots, M\} \) involves the solution of the equation for the zeroth-order wave function, that must be solved iteratively and, necessarily, the definition of the zeroth-order Hamiltonian. In the CASPT2 approach, the choice of the zeroth-order Hamiltonian is not unique, which has been the subject of much discussion and research leading to different variants. From a general point of view, it can be said that it is defined in such a way that for a closed-shell Hartree-Fock reference wave function is equivalent to the Mller-Plesset Hamiltonian. The normalized wave function corrected up to
the first order is given by

$$|\Psi > = C_0|\Psi^{(0)} > + C_1|\Psi^{(1)} >$$ (2.45)

with $C_0^2 + C_1^2 = 1$. The weight of the reference function ($C_0^2$) can be used as a simple and rapid criterion of quality for the perturbation treatment carried out. Ideally it should be close to unity, nevertheless its value depends on the number of correlated electrons. Thus, the larger the molecular system the smaller is the reference weight. Values that vary between 0.93 for ten electrons and 0.48 for 100 electrons have been found.\(^{13}\) If a specific calculation shows a reference weight much smaller than this there is a problem with intruder states in the perturbative approach. Intruder states are states in the $V_{SD}$ space that have a zeroth-order energy which is close to, or even below, the reference energy $E^{(0)}$. The corresponding first-order coefficient may then become large and a treatment based on perturbation theory is no longer valid. There are two options if such a situation occurs, one is to increase the active space to include the orbital which causes the problem, because it means that the interaction with the reference function is large and the contribution to the second-order energy cannot be neglected. If this cannot be done the CASPT2 method is not suitable for that system. The second and more common case is when the intruder state is a weak one and a level shift technique can be used to remove the intruder state. In fact, that means that the interaction with the reference state is small and the contribution to the correlation energy can be expected to be small.
2.4 HYBRID METHODOLOGY. QM/MM.

Quantum mechanical methods can be considered as the first-principal description of a system that allows to describe any molecular properties and processes. However, a higher level of the theory requires a higher computational resources. The simplest ab-initio calculations typically scale $O(N^3)$ or worse, where $N$ is the number of atoms in the system.

Molecular Mechanics methods demonstrate great performance, in case of the careful calibration of a number of parameters, and the cost of simulations scales $O(N^2)$ (or even less if one uses methodologies that allow to reduce efforts for the evaluation of electrostatic interactions term). However, in some cases the quality of the calibration is not suffice to obtain the results at the required level and, even worse, molecular mechanics can not describe intrinsically quantum mechanical properties and processes such as electron and proton reorganization.

The development of the hybrid approaches is guided by the general idea that various regions of the system often play very different roles in the process under investigation and the systems can be partitioned into an important region which requires a high level quantum chemical treatment and a remainder which can be described at a lower level of the theory. The first part can be the reactive region of the system, like the active site of an enzyme, or like a chromophore in a solution. The second part will be the protein environment in the former example, or the bulk of the solution in the latter (see Figure 2.5).
A variety of hybrid methods have been developed by now. They are conceptually quite similar but differ in a number of details. The most common methods combine a Quantum Mechanical (QM) method with a Molecular Mechanics (MM) (so-called QM/MM methods). Some hybrid methods can also combine two or more QM methodologies (QM/QM methods).

The QM/MM approach is established as a valuable tool for the modelling of biomolecular systems,\textsuperscript{14,15,16,17,18,19,20} the investigation of inorganic/organometallic\textsuperscript{21,22,23} and solid-state\textsuperscript{24,25,26} systems and for studying processes in explicit solvent.\textsuperscript{27,28,20}

Two schemes for describing the energy of a compound QM/MM system as a function of energies of component systems have been proposed. The substructive scheme follows the strategy:

1. MM calculation on the entire system (En);
2. QM calculation on the inner subsystem (In);
3. MM calculation on the inner subsystem.

Figure 2.5: QM/MM description of molecular systems. A. Partitioning of a protein system treated with a QM/MM approach. The active center is treated at the QM level and the surroundings is treated at the MM level; B. Partitioning of a solution treated with a QM/MM approach. The solute is treated at the QM level and the solvent is treated at the MM level.
The QM/MM energy of the entire system is then obtained according the equation

\[ E_{\text{sub}}^{\text{QM/MM}} = E_{\text{MM}}^{\text{En}} + E_{\text{QM}}^{\text{In}} - E_{\text{MM}}^{\text{In}}. \]  

(2.46)

ONIOM methodology\textsuperscript{29,30,31} provides an example of a substractive QM/MM scheme. The energy expression for the second QM/MM scheme (a so-called additive scheme) is given by the equation

\[ E_{\text{add}}^{\text{QM/MM}} = E_{\text{MM}}^{\text{O}} + E_{\text{QM}}^{\text{In}} + E_{\text{QM-MM}}^{\text{O-In}}, \]  

(2.47)

where the outer (O) system is the part of the entire system (En) that is not included in the inner subsystem (In) and \( E_{\text{QM-MM}}^{\text{O-In}} \) is the QM-MM coupling term. The exact form of the QM-MM coupling term \( E_{\text{QM-MM}} \) varies for different QM/MM methods. It has to include bonded, van der Waals, and electrostatic interactions between QM and MM atoms

\[ E_{\text{QM-MM}} = E_{\text{QM-MM}}^{\text{vdW}} + E_{\text{QM-MM}}^{\text{el}} + E_{\text{QM-MM}}^{\text{b}}. \]  

(2.48)

The van der Waals interaction is typically described by a Lennard-Jones potential

\[ E_{\text{QM-MM}}^{\text{vdW}} = \sum_{\text{nonbonded pairs } AB} \varepsilon_{AB} \left[ \left( \frac{\sigma_{AB}}{r_{AB}} \right)^{12} - \left( \frac{\sigma_{AB}}{r_{AB}} \right)^{6} \right], \]  

(2.49)

where \( \varepsilon_{AB} \) and \( \sigma_{AB} \) are the Lennard-Jones parameters. The Lennard-Jones parameters for QM atoms can be taken from a classical Force Field. However, some researchers prefer to re-optimize them against data obtained from QM calculations.\textsuperscript{32,33,34}

The electrostatic coupling term between the QM charge density and the charges in the MM region \( E_{\text{QM-MM}}^{\text{el}} \) can be handled at different level of the theory. According that the methodologies can be classified\textsuperscript{19} as mechanical embedding, electrostatic embedding and polarized embedding. In the case of mechanical embedding the electrostatic of interaction of QM region with MM region is treated in the same way like electrostatic interaction within MM part (see section ?? on the page ??). The main limitations of this approach are that
it does not allow to describe a polarization of the electron density of a QM part by a MM environment and that a change in distribution of the electron density during a reaction or after an excitation should be taken into consideration.

The major shortcomings of mechanical embedding can be eliminated by using an electrostatic embedding scheme. In this version of the theory the effect of MM point charges on a QM part is included as additional one-electron terms in the QM Hamiltonian

\[
\hat{H}_{\text{QM-MM}}^{\text{el}} = -\sum_i^N \sum_J^L \frac{q_i}{|\mathbf{r}_i - \mathbf{R}_J|} + \sum_\alpha^M \sum_J^L \frac{q_J Q_\alpha}{|\mathbf{R}_\alpha - \mathbf{R}_J|}.
\]

The symbols \(q_J\) are the MM point charges located at \(\mathbf{R}_J\); \(Q_\alpha\) are the nuclear charges of the QM atoms at \(\mathbf{R}_\alpha\); and \(\mathbf{r}_i\) designate electron positions. The indices \(i, J, \) and \(\alpha\) run over the \(N\) electrons, \(L\) point charges, and \(M\) QM nuclei, respectively. This strategy allows to take into account a polarization of a QM part caused by a MM part and is used in this work.

The next step in generalization, polarized embedding, emerges from the fact that the QM and MM part will polarize each other until their charge distributions are self-consistent. The self-consistency can be achieved by an iterative algorithm with electrostatic embedding and a polarizable force field for an MM part. The development of polarizable force field is in progress\(^{35,36,37,38,39,40}\) and can become a routine in the nearest future.

The last term in the equation 2.48, the bonding term \(E_{\text{QM-MM}}^b\), has to describe the covalent bonds at the QM/MM boundary. For some systems, like a solution, they can be avoided by choosing the boundary as a surface between solute and solvent molecules that does not cut any bonds. In protein systems, for example, it can not been achieved and the proper strategy has to been devised to cut covalent bonds in the QM/MM boundary and to describe the bonding energy. Schemes that deal with this issue can largely be grouped in two classes. The first, and the simpler solution, is the so-called link-atom approach,\(^{41,42}\) where the free valencies of the frontier atom are saturated with hydrogen atoms and these are included in the QM fragment. The second class of methods consists of schemes that use localized orbitals at
Figure 2.6: Schematic representation of the division of a QM-MM bond and partition of the hybrid orbitals on the MM boundary atom C. Adapted from from ref.\textsuperscript{46}

the boundary between the QM and the MM parts. In this context Rivail and co-workers\textsuperscript{43,44,45} developed a local self-consistent field (LSCF) algorithm where the bonds between the QM and MM segments are represented by a set of strictly localized bond orbitals that are kept constant throughout the SCF process. They are obtained by separate quantum chemical calculations of small model compounds and assumed to be transferable. One advantage of the LSCF treatment is that it does not require the addition of link atoms although the parameters for the localized bond orbitals have to be determined from model studies for each new system under investigation. In the spirit of the LSCF a generalized hybrid orbital (GHO)\textsuperscript{46,47} method for the treatment of QM/MM covalent bonds was also developed. In this approach, hybrid atomic orbitals are used as basis functions on the boundary atoms of the MM fragment. In particular, for applications to biopolymers, the MM boundary atom will likely be chosen as an $sp^3$ carbon, which is connected to three other MM groups. The hybridization scheme depends on the local geometry of the three MM atoms to which the boundary atom is bonded, and the parametrization is assumed to be transferable. From these
four orbitals, the one directed toward the frontier QM atom is called the active orbital and
the other hybrid orbitals are called auxiliary orbitals (Figure 2.6). All the hybrid orbitals are
included in the QM calculations but only the active orbital is optimized in the SCF process.
Consequently, the chemical bond connecting the QM and MM parts is treated in an explicit
way without introducing any artificial link-atom.

Each kind of treatment has advantages and drawbacks; the link-atom scheme is straightforward and widely used to treat the boundary region. The LSCF or GHO methods are
theoretically more rigorous but they are more complicated and not free from limitations as well. The link-atom and local-orbital approaches have been extensively tested, leading to
the conclusion that reasonably good accuracy can be obtained by both if they are carefully
used.48,49,50,51

In this work we used the QM/MM methodology that is developed and implemented in
Olivucci group.52,53 This QM/MM scheme can be classified as an Electrostatic Embedded
(EE) method with a link-atom approach for the boundary treatment in case of rhodopsin
studies and with no covalent bonding in case of switches in methanol solution.

AMBER96 force field54 were used for MM part of the rhodopsin. Water is described by
the transferable-intermolecular-potentials-3-point-charge (TIP3P) model.55 The QM/MM
frontier is set at the Cε-Cσ bond of the Lys-296 side-chain (Figure 2.7) The Hydrogen Link
Atom (HLA) scheme is used to cap the pending valence on the QM Cε atom. The link
hydrogen atom is fixed at 1 Å from Cε and kept along the Cε-Cσ axis. As it is shown in the
work,53 in order to provide a correct description of the frontier, the HLA may interact with
all the MM point charges but it cannot be involved in other MM potentials. The values
of the point charges residing on the Lys-296 atoms have been reparametrized. The lysine
residue in AMBER96 has a net charge of +1 (i.e. protonated lysine). In frame of the used
model the MM part of Lys-296 residue must have a null charge and MM point charges of
this residue has been changed according the table 2.1. The charge of the frontier MM carbon
atom Cσ is set to 0 to ensure that the QM wave function is not over-polarized by the close
Table 2.1: The values of the re-parametrized QM/MM point charges for Lys-296.

<table>
<thead>
<tr>
<th>Atom</th>
<th>N</th>
<th>C_α</th>
<th>C_carbonyl</th>
<th>H_N</th>
<th>O_carbonyl</th>
<th>H_α</th>
</tr>
</thead>
<tbody>
<tr>
<td>Charge</td>
<td>-0.3981</td>
<td>-0.2400</td>
<td>0.6840</td>
<td>0.2246</td>
<td>-0.6396</td>
<td>0.1426</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Atom</th>
<th>C_β</th>
<th>H_β</th>
<th>C_γ</th>
<th>H_γ</th>
<th>C_σ</th>
<th>H_σ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Charge</td>
<td>-0.0094</td>
<td>0.0362</td>
<td>0.0187</td>
<td>0.0103</td>
<td>0.0000</td>
<td>0.06</td>
</tr>
</tbody>
</table>

Figure 2.7: Scheme of the chromophore and the frontier region in rhodopsin.

HLA. This procedure is allowed by the small value of the original AMBER96 point charge and makes possible to use the standard MM bonded potentials (stretching, bending, etc.) for the description of the geometry of the frontier. Second, the van der Waals atomic parameters for retinal (i.e. for a conjugated hydrocarbon chain containing \( sp^2 \) carbon atoms) are not defined in the AMBER96 force field. Similarly, the \( C_{15}-N-C_ε-C_σ \) torsion potential does not exist in AMBER96. These missing parameters have been determined in such a way to best reproduce the ground state (\( S_0 \)) and first excited state (\( S_1 \)) CASSCF torsional energy profiles relative to the \( N-C_ε-C_σ-C_γ \) and \( C_{15}-N-C_ε-C_σ \) dihedral angles of the model system (Figure 2.7). The resulting van der Waals parameters are \( (R^* = 1.87\text{Å}, \varepsilon = 0.0860 \text{ kcal} \cdot \text{mol}^{-1}) \) for an
sp² carbon atom of the retinal π-system, \((R^* = 1.87\text{Å}, \varepsilon = 0.1094 \text{ kcal} \cdot \text{mol}^{-1})\) for a sp³ carbon atom of retinal (i.e. the methyl substituents in position 9 and 13 in Figure 2.7) and \((R^* = 0.92\text{Å}, \varepsilon = 0.0157 \text{ kcal} \cdot \text{mol}^{-1})\) for the hydrogen atom of retinal. The C₁₅-N-Cₑ-Cₛ torsion potential is given by:

\[
E_{\text{tor}}(\text{C}_{15}-\text{N}-\text{C}_e-\text{C}_\sigma) = 0.750[1 + \cos(\phi - 0)].
\]

The model of the switches in solution was constructed by placing the chromophore in a rectangular box of methanol molecules positioned within 10 Å from any given atom of the chromophore using xleap module of Amber 7.0 package. Methanol is described by the non-polarizable model for this organic solvent provided by the Amber 7.0. The average ground state configuration of the methanol molecules (i.e. solvent) has been determined according to the following procedure: the solvent system was minimized by 1000 conjugate-gradient minimization steps using periodic boundary conditions while keeping the chromophore (i.e. the solute) fixed in its gas-phase configuration. In this step the partial charges of the chromophore atoms were determined with GAUSSIAN03, using a Restrained ElectroStatic Potential (RESP) procedure at the HF/6-31G* level of theory. The minimized system was further relaxed keeping the solute molecule fixed by molecular dynamics simulation within isothermal-isobaric NPT ensemble (1 atm, 298 K) using the program NAMD. Since the van der Waals atomic parameters for the switch molecule are not defined in the Amber96 force-field, we have used the parameters already developed in ref. following the procedure proposed in ref. The van der Waals parameters are \((R^* = 1.8778\text{Å}, \varepsilon = 0.0860 \text{ kcal} \cdot \text{mol}^{-1})\) for a sp² carbon atom of the switch model system, \((R^* = 1.8891\text{Å}, \varepsilon = 0.1094 \text{ kcal} \cdot \text{mol}^{-1})\) for a sp³ carbon atom and \((R^* = 0.9327\text{Å}, \varepsilon = 0.0157 \text{ kcal} \cdot \text{mol}^{-1})\) for the hydrogen atoms attached to sp² or sp³ carbons. The van der Waals parameters for the oxygen atom of the -OMe group and for the carbon and oxygen atoms of the CO₂ group have been taken from the general Amber force field (GAFF). They are \(R^* = 0.6837\text{Å}, \varepsilon = 0.1700 \text{ kcal} \cdot \text{mol}^{-1}\)(for the
oxigen atom -OMe group; $R^* = 1.9080\,\text{Å}$, $\varepsilon = 0.0860 \text{ kcal} \cdot \text{mol}^{-1}$ (for the carbon atom of the CO$_2$ group); $R^* = 1.6612\,\text{Å}$, $\varepsilon = 0.2100 \text{ kcal} \cdot \text{mol}^{-1}$ (for the oxigen atom of the CO$_2$ group).

For all QM/MM calculations modified GAUSSIAN03,$^{57}$ MOLCAS 7.4,$^{63}$ and Tinker4.2$^{64}$ codes were used.

The performance of the described QM/MM methodology has been proved by the series of recent studies both for the bovine rhodopsin$^{65,66,67,68,69}$ and the biomimetic switches in methanol$^{60,69,70,71}$ and it is used in this thesis without major changes. Where needed, further details are provided in each following chapter.
2.5 EVOLUTION OF A SYSTEM AFTER PHOTOCHEMICAL
ITATION. DYNAMICS

2.5.1 WAVEPACKET PROPAGATION METHODS

After photoexcitation, the system is promoted to a non-stationary state and the questions to be answered are where does the system evolve and what is the time scale the process occurs on. The methods based on quantum chemical calculations to characterize the potential energy surfaces in terms of critical points and path (maxima, minima, transition states, conical intersections, etc.), that are described in the previous section, allow to obtain valuable information about the reaction mechanism without a direct description of the nuclear dynamics. However, non-adiabatic phenomena, that are very common in photochemistry, are inherently dynamic. For example, the evolution of the system near a conical intersection region, that often serves as a bifurcation point of the reaction path that controls directions and the quantum yield of photoreactions, is a none-adiabatic phenomenon and, therefore, the dynamics of nuclei should be taken into account. Besides that, the quantitative evaluation of the time scale of photo-processes, along with time-resolved spectroscopy studies, provides valuable information for understanding and control of photochemical reactions. The theory that can be used to address these problems, at least in cases then relativistic effects are not significant or can be taken into account by means of an effective Hamiltonian, has to be based on the time-dependent Schrödinger equation

\[ i\hbar \frac{\partial \Psi(R, r, t)}{\partial t} = \hat{H} \Psi(R, r, t), \]  

(2.51)

where \( \hat{H} \) is the Hamiltonian of the system, \( R \) and \( r \) are coordinates of nuclei and electrons, respectively.

As it was described above (see section 2.1.1 on page 29 the equation 2.51 can be modified (the group Born-Oppenheimer approximation) to the equation that describes the nuclear
motion
\[ i\hbar \frac{\partial \Phi(R, t)}{\partial t} = \left( \hat{T} + V + \Lambda \right) \Phi(R, t), \]  
(2.52)

where the wavefunction vector, \( \Phi(R) \), has a component for each electronic state, the potential matrix \( V \) is diagonal, with \( V_{ii} \) containing the potential energy function for state \( i \), \( \Lambda \) is non-adiabatic operator matrix; and the electronic Shrödinger equation

\[ \hat{H}_e(r, R)\psi(r, R) = V(R)\psi(r, R), \]  
(2.53)

where \( H_e \) is electronic Hamiltonian including the electronic kinetic energy, the Coulomb interaction between the electrons and nuclei, as well as nuclear-nuclear interaction at a specified geometry \( R \).

If electronic states are weakly coupled (i.e. in a region located far from a conical intersection), \( \Lambda \) can be neglected and equation 2.52 gives the nuclear Shrödinger equation in the Born-Oppenheimer approximation

\[ i\hbar \frac{\partial \Phi(R, t)}{\partial t} = \left( \hat{T} + V(R) \right) \Phi(R, t), \]  
(2.54)

where \( \hat{T} \) is nuclear kinetic energy operator and \( V(R) \) is the potential energy function.

The simple form of the Born-Oppenheimer approximation 2.54 is often sufficient to describe the nuclear dynamics in studies of ground-state properties and reactions, there the process takes place on a single potential energy surface, i.e. adiabatic processes and can be used for excited states in the regions there coupling of potential energy surfaces are relatively weak. Besides that, the theories based on the equation 2.54 provide the frame for more sophisticated non-adiabatic studies.

The potential that drives nuclei is a function of nuclear coordinates \( R \) and can be obtained as a solution of a set of the stationary electronic Shrödinger equation 2.53 for all possible geometries (coordinates \( R \)). In actual practice, one can calculate discrete number of point
and to fit this grid by a proper function to obtain an approximation to the exact potential energy surface. This approach is practical for relatively small systems and quickly becomes prohibitive with increasing of the number of nuclei. Based on the fact that an evolution of a system is defined only by the structure of the part of the potential energy surface in the region where the system actually goes, i.e. locally but not globally, the direct dynamics methods were proposed. In direct dynamics methods, the electronic Schrödinger equation 2.53 is solved only for the region of the potential energy surface over which the nuclei move (“on-the-fly”). This idea allows increase the size of the system available for calculation.

The fundamental method to solve the equation 2.54 for a given initial function $\Phi(R, t = 0)$ is wavepacket dynamics. The nuclear wavefunction can be written in the form

$$\Phi(R, t) = \sum_{j_1} \ldots \sum_{j_f} A_{j_1 \ldots j_f}(t) \varphi_{j_1}^{(1)}(R_1) \ldots \varphi_{j_f}^{(f)}(R_f), \quad (2.55)$$

where each of $f$ the modes of nuclear motion is expanded in a set of $j$ basis functions

$$\Phi_f(R_f, t) = \sum_{n=1}^j c_{jf} \varphi_{jf}(R_f).$$

If the basis functions are time-independent, solving the time dependent Schrödinger equation is equivalent to solving the linear equation of motion for the expansion coefficients

$$i \hbar \dot{A} = HA, \quad (2.56)$$

where $H$ is the Hamiltonian matrix basis set

$$H_{ij} = \left< \varphi_{i_1}^{(1)} \ldots \varphi_{i_f}^{(f)} \right| \hat{H} \left| \varphi_{j_1}^{(1)} \ldots \varphi_{j_f}^{(f)} \right> \quad (2.57)$$

The computation can be done in three parts - to describe the initial wave function by a set
of numbers (matrix of coefficients \( \mathbf{A} \)), to evaluate the matrix elements 2.57, to propagate the initial wave function in time (to solve the equation 2.56). A series of suitable orthogonal basis functions are chosen according to the type of motion. For bound-state problems where the system moves close to the bottom of the potential energy surface, vibrational degrees of freedom are often efficiently described by basis set expansion using, for example, harmonic or Morse oscillator eigenfunctions. Although in practice one has to truncate the basis set expansion, in principle, the set of functions can be chosen as close to complete set of functions as needed. Efficient techniques for the direct solution of equation 2.56 and evaluation of the matrix elements 2.57 have been developed.\(^{72,73,74}\) However, the number of coefficients in the product 2.55 grows exponentially with the number of modes in the system and the size of the system that can be treated by the method is up to three or four atoms. Anyway, the method can be used as a reference and as a starting point for approximate methods.

The multi-configuration time-dependent Hartree method (MCTDH\(^{75,76,77}\)) that was introduced in 1990 by Meyer, Manthe and Gederbaum\(^ {78}\) uses not fixed but flexible time-dependent functions which follows variational equation of motion. A time-dependent basis is more efficient because it follows the evolving wavepacket, and does not waste effort on describing of regions of empty space. This method has been used to treat dynamics of systems with up to 24-modes. Problems that were investigated by the method include, among others, simple reactions for 5 atoms molecules,\(^ {79,80,81}\) tunneling splitting and vibrational states of intra- and intermolecular proton transfer systems as malonaldehyde\(^ {82}\) and \( \text{H}_5\text{O}_2^+ \) cluster,\(^ {83,84,85,86}\) the vibronic dynamics of photoexcited pyrazine\(^ {87}\) and of photoionized butatriene.\(^ {88}\)

Recently, the new version of MCTDH, the multilayer MCTDH, has been proposed.\(^ {89}\) The multilayer MCTDH approach has proved its efficiency in rigorous quantum dynamics calculations including 100-1000 correlated coordinates and has been used to study system-bath models of electron\(^ {90,91,92,93}\) and proton\(^ {91,94}\) transfer processes. A big advantage of all quantum dynamics methods is that they are able to treat non-adiabatic phenomena in the same way as adiabatic processes as all quantum effects are included in the nuclear wavepacket.
2.5.2 SEMICLASSICAL AND CLASSICAL TRAJECTORIES

Another possible approach to nuclear dynamics is based on the fact that for the wide range of phenomena the quantum nature of nuclei is negligible and one can use classical molecular dynamics. In this case the system can be propagated according to the laws of classical mechanics. To provide the necessary background to the theory one has to use a formulation of the quantum mechanics that allows transparent transition to the classical limit. One well-known alternative formulation of the quantum mechanics, so-called the hydrodynamic formulation, proposed by D. Bohm.\textsuperscript{95,96}

A nuclear wavefunction can be always written in the form

\[ \Phi(\mathbf{R}, t) = A(\mathbf{R}, t)e^{i\frac{\hbar}{\mu}S(\mathbf{R}, t)}. \]  

(2.58)

Inserting this wave function in the nuclear Shrödinger equation 2.54 leads to a couple of equations

\[ \dot{S} + \frac{\nabla S^2}{2m} + V + Q = 0 \]  

(2.59)

\[ \dot{A}^2 + \frac{1}{m} \nabla \cdot (A^2 \nabla S) = 0 \]  

(2.60)

where

\[ Q = \frac{\hbar^2}{2m} \frac{\nabla^2 A}{A} \]  

(2.61)

is a non-local quantum potential and the classical limit can be obtained by setting this term equal to zero (\( \hbar \rightarrow 0 \)). \( P = A^2 \) is the electronic density and equation 2.60 can be thought as a hydrodynamic equation that describes the evolution of the quantum fluid that is consisted of pseudo-particles. The velocity field of this fluid is

\[ v = \frac{\nabla S}{m} \]  

(2.62)
The substitution 2.62 into the classical limit of the equation 2.59 gives an equation of the classical mechanics that can be written in the Newton formulation as

\[ m\ddot{v} = F \]  

(2.63)

Therefore, the problem is reduced to solving Newton’s equations of motion for an ensemble of pseudo-particles with initial coordinates and velocities that describes an initial wavepacket. The force experienced by a pseudo-particle is

\[ F_\alpha(R) = -\nabla_\alpha V(R) = -\nabla_\alpha \langle \psi(r, R)|H|\psi(r, R) \rangle. \]  

(2.64)

In this, so-called semiclassical, molecular dynamics, the delocalized quantum mechanical wavepacket is represented by a set of the classical pseudoparticles and their classical trajectory approximated the evolution of the wavepacket. The next level of approximation is pure classical molecular dynamics. One can take only one point in phase space that describes pure classical nuclei and propagate it according equation 2.64.

The semiclassical or classical dynamics is a case when a direct dynamics approach can be involved. The electronic wave function and gradients are calculated at time \( t \) and the system is propagated forward a step, and the process is repeated. In majority cases the thermal effect of the environment is important and the theory has to be elaborated by using the density matrix approach in case of quantum theory description and the phase-space distribution function in case of classical description.

If one uses quantum description of the initial nuclear wavepacket and semiclassical dynamics, the question arises how to map the initial wavepacket onto a function that has the form of classical phase-space distribution function \( \rho(p, q) \). The uncertainty principle makes the concept of phase space in quantum mechanics problematic. Because a particle cannot simultaneously have a well defined position and momentum, one cannot define a probability that a particle has a position \( q \) and a momentum \( p \), i.e. one cannot define a classical analog
of true phase space probability distribution for a quantum mechanical particle. A number of methods to produce the quasi-probability distribution functions has been proposed. However, the translation is not trivial (see ref for details). If one assumes that the system behaves classically (i.e. this is the case of pure classical dynamics) the appropriate ensemble of classical thermodynamics can be invoked. The normal-mode sampling algorithm, for example, uses a microcanonical ensemble to distribute energy into the internal modes of the system by assigning a desired energy to each normal mode as a harmonic amplitude

\[ A_\alpha = \frac{\sqrt{2E_\alpha}}{\omega_\alpha}, \quad (2.65) \]

where \( \omega_\alpha \) is the harmonic frequency. The initial momentum and position are then generated by adding a random phase, \( \xi_\alpha \)

\[ Q_\alpha = A_\alpha \cos(2\pi\xi_\alpha) \quad (2.66) \]

\[ \dot{Q}_\alpha = -A_\alpha \sin(2\pi\xi_\alpha). \quad (2.67) \]

An important advantage of this methodology that it can be used to include the effects of zero-point energy into a classical calculation.

A measure of the quantum character of a given particle is its de Broglie wavelength, which is inversely proportional to the particle momentum (see 2.61). It means that semiclassical dynamics is not applicable to the light atoms especially at low temperature and important effects such as tunneling are left out of scope of the methodology. To deal with this kind of problems and to keep the computational simplicity of classical dynamics, at the same time, the range of mixed quantum-classical dynamics methods where proposed. Any mixed quantum-classical dynamics methods is an approximation to the true full quantum dynamics - the critical degree of freedoms of the system are treated at the quantum level and the rest by classical mechanics.

Another problem that can not be solved by this model is the description of non-adiabatic
effects that are very common in photochemistry. The simplest way to deal with the problem
is to use surface hopping method. In this method non-adiabatic effects are added to
the swarm of trajectories by allowing a trajectory starting on one electronic state change to
another. One of the widely used schemes is the fewest switches algorithm. According this
algorithm, the probability of changing from state 2 to state 1 an time $T$ is given by

$$P_{2\rightarrow 1} = -\frac{d}{dt} \log |c_2|^2,$$  \hfill (2.68)

where $c_2$ is the coefficient for the electronic state $\psi_2$ at that time in the total electronic
wavefunction

$$\psi(t) = c_1(t)\psi_1(t) + c_2\psi_2(t).$$ \hfill (2.69)

This is obtained by integrating the equation

$$i\hbar \dot{c}_i = c_i V_i - i\hbar \sum_j \dot{R} F_{ij} c_j,$$ \hfill (2.70)

where $\dot{R}$ is nuclear velocity of the pseudo-particle and $F_{ij} = \langle \psi_i | \nabla \psi_j \rangle$ is the derivative
coupling vector. The method has been implemented for a number of systems (for examples
see references in). The comparisons have been made between surface hopping and full
quantum dynamics for the butatriene cation and ozone. The results show that the
methods give a similar description of the initial part of the dynamics, with a similar time
scale for the interstate crossing on the lower potential energy surface. However, the dynamics
after the crossing are different - the trajectories spread out far more than in wavefunction
obtained from the full quantum dynamics calculation.
Bibliography


Chapter 3

THE PHOTOCHEMISTRY OF THE NAIP-1 SWITCH
Single molecules that act as light-energy transducers (e.g., converting the energy of a photon into atomic-level mechanical motion) are examples of minimal molecular devices. Here, we focus on a molecular switch designed by merging a conformationally locked diarylidene skeleton with a retinal-like Schiff base and capable of mimicking, in solution, different aspects of the transduction of the visual pigment Rhodopsin. Complementary ab initio multiconfigurational quantum chemistry-based computations and time-resolved spectroscopy are used to follow the light-induced isomerization of the switch in methanol. The results show that, similar to rhodopsin, the isomerization occurs on a 0.3 ps time scale and is followed by <10 ps cooling and solvation. The entire (2-photon-powered) switch cycle was traced by following the evolution of its infrared spectrum. These measurements indicate that a full cycle can be completed within 20 ps.

\[1\] The chapter based on the article – An artificial molecular switch that mimics the visual pigment and completes its photocycle in picoseconds Adalgisa Sinicropi, Elena Martin, Mikhail Ryazantsev, Jan Helbing, Julien Briand, Divya Sharma, Jrmie Lonard, Stefan Haackee, Andrea Cannizzo, Majed Chergui, Vinicio Zanirato, Stefania Fusi, Fabrizio Santoro, Riccardo Basosi, Nicolas Ferr, and Massimo Olivucci “An artificial molecular switch that mimics the visual pigment and completes its photocycle in picoseconds”, PNAS 105, 2008
INTRODUCTION

Molecular switches based on photochemical E/Z isomerizations have been used in different contexts to convert light energy into mechanical motion at the molecular level. For instance, switches based on azobenzene have been used to control ion complexation, electronic properties, catalysis, and the folding of peptides whereas diarylidenes have provided the framework for the construction of rotary motors and transmissions. The computer modeling of switches that differ in size, polarity, and isomerization mechanism represents an attractive research target yielding building blocks to be used in diverse molecular environments. However, this cannot be limited to the computation of equilibrium properties but requires the description of the entire photocycle. In other words, one needs to compute the potential energy surfaces controlling the switch E → Z and Z → E excited-state evolution, its decay and ground state relaxation, and the competing thermal E/Z isomerization in the proper environment (e.g., in solution or in a biomolecule backbone). The complexity of these calculations impedes the study of candidates that are intractable with accurate quantum chemical methods (allowing comparison with spectroscopic data) or that feature, as for azobenzene and diarylidenes, more than one low-lying excited state, leading to a plethora of reaction paths to be computed.

The retinal protonated Schiff-base chromophore of rhodopsins constitutes an example of an E/Z switch shaped by biological evolution that can be modeled with quantitative computations. In bovine rhodopsin (Rh), a selective photoisomerization of the 11-cis chromophore (PSB11) occurs via evolution of a single $\pi \rightarrow \pi^*$ excited state ($S_1$) that survives for only 150 fs and yields, upon decay, the all-trans ground state ($S_0$) product with a 67 % quantum yield. Although these properties make Rh an excellent reference for the design of E/Z switches, irradiation of PSB11 in solution features an unselective isomerization and a picosecond excited-state lifetime prompting a search for artificial Rh-mimetic molecules. To address this problem, we have shown that, in methanol solution, the N-alkylated indanylidene-pyrroline Schiff base NAIP-1 displays excited-state properties
similar to those of Rh-embedded PSB11. We have also shown that NAIPs can be prepared in good yields (> 80%) through a cyclopropyl ring-opening/nitrilium ring-closing tandem reaction.22

In the present report, complementary computational and experimental tools are used to characterize the photoisomerization of NAIP-1 (from now on E-1 and Z-1). We show that quantum-mechanics/ molecular-mechanics (QM/MM) calculations based on ab initio (i.e., first-principle) multiconfigurational second-order perturbation theory (CASPT2), and complete-active-space self-consistent-field (CASSCF) quantum chemical methods point to E-1 → Z-1 and Z-1 → E-1 S\textsubscript{1} isomerization paths that intercept distinct conical intersections on a 200 fs time scale. Such features are probed via time-resolved UV-vis absorption, emission, and mid-IR spectroscopy and show that the Z → E → Z photocycle (Inset in Figure 3.1 on the page 89) is completed in picoseconds.

MATERIALS AND METHODS

Computations. The ab initio CASSCF method\textsuperscript{23} is a multiconfigurational quantum chemical method providing the 0th-order wave function for subsequent CASPT2\textsuperscript{24} computations leading to a quantitative evaluation of electronic energies. The ab initio CASPT2//CASSCF protocol (where geometries and electronic energies are determined at the CASSCF and CASPT2 levels, respectively, with an active space comprising the full 12-electron in 11-orbital π-system of 1) has been recently implemented in a QM/MM scheme\textsuperscript{20,25} allowing for the evaluation of the excitation energies and barriers of chromophores (treated quantum mechanically) embedded in explicit solution or protein environments (described by the AMBER force field). The quality of the used CASPT2//CASSCF/AMBER methanol solution model is probed by computing the Z-1 and E-1 S\textsubscript{0} → S\textsubscript{1}, S\textsubscript{0} → S\textsubscript{2}, and S\textsubscript{0} → S\textsubscript{3} vertical excitation energies. These values reproduce the observed quantities with a < 3 kcal mol\textsuperscript{-1} error, and the corresponding oscillator strength values reproduce the observed transition
intensity patterns. These data are confirmed via ASEP/MD computations\textsuperscript{26} that, via a self-consistent protocol, yield a congruous description of the solvent shells of the associated average electrostatic fields and of the Z-1 and E-1 equilibrium structures. The ASEP/MD Z-1 and E-1 $S_0$ equilibrium structures differ from the one computed with the approximate solvent model of 0.030 and 0.019Å, respectively (RMS deviation). Similarly, the two sets of excitation energies differ by 0.1 and 2.2 kcal mol$^{-1}$ for the Z-1 and E-1 isomers, respectively.

The $S_1$ reaction coordinates on Figure 3.1 on the page 89 are computed via unconstrained geometry optimizations starting at $S_0$-Z-1 and $S_0$-E-1 and yielding $S_1$ relaxed structures. From these structures, a relaxed scan along the $C_9'$-$C_1'$-$C_4$-$C_5$ angle is computed with steps of 10° (for Z-1) or +10° (for E-1). This protocol based on a scan [rather then an Intrinsic Reaction Coordinate (IRC) calculation] avoids problems associated with the flatness of the energy surface. Its validity has been assessed for Z-1 by computing the initial part of the IRC. The IRC path intercepts the optimized $S_1$ point at 17° twisting and then continues along the points of the scan indicating that the presented paths provides a mechanistically valid approximations of minimum energy paths. The $S_1$ paths computed with a frozen or a relaxed solvent shell (with respect with the initial $S_0$ Z-1 shell) show only small differences.

**Ultrafast UV-Vis and UV-Mid-IR Spectroscopy.** Z-1, prepared as described in ref.,\textsuperscript{22} was dissolved in methanol dried over molecular sieves for the IR experiments and in spectroscopy grade MeOH for the UV-vis pump-probe experiments. Solutions of $\approx$10 mmol/L were used for time-resolved IR spectroscopy in a 100-µm-thick flow cell (OD 0.8-1.2 at the $S_0 \rightarrow S_1$ absorption maximum at 390 nm). More-diluted samples were used for the UV-Vis pump-probe and fluorescence experiments performed in 0.5-mm path-length flow cells (OD $\approx$0.5). Thermal relaxation from the E to the Z forms takes place with a 17.5-h time constant at room temperature. Therefore, to minimize E-1 accumulation during the measurements (typically 1-2 h for a complete dataset), a sufficiently large amount of sample was circulated in a flow cell at a rate high enough to exchange the sample after each laser shot. Sub-100-fs pulses at 388-420 nm were obtained by frequency-doubling the output of
commercial Ti:Sapphire laser systems and used to excite <5% of the molecules in the laser focus. Home-built setups were used for femtosecond fluorescence up-conversion,\textsuperscript{27} UV-vis, and UV-mid–IR pump–probe measurements. The temporal resolution was 80-120 fs in the former two experiments and 300 fs in the latter.

RESULTS AND DISCUSSIONS

Reaction Path Computation and Analysis. In Figure 3.1, we report the CASPT2//CASSCF/AMBER photochemical reaction paths (a combination of a $\pi \rightarrow \pi^*$ $S_1$ reaction path followed by a $S_0$ relaxation path) for the chloride-Z-1 (left) and chloride-E-1 (right) ion pairs (from now on Z-1 and E-1, respectively) in methanol solution. The relative stability of Z-1 and E-1 has been determined in terms of the difference in Gibbs free energy between the solvated ion pairs by using averaged solvent electrostatic potential/molecular dynamics (ASEP/MD) computations.\textsuperscript{26} The results indicate that, at 298 K, E-1 is 0.7 kcal mol$^{-1}$ higher in free energy than Z-1. This value compares well with the 1.4 kcal mol$^{-1}$ free-energy difference (Z/E ratio is 92:8) derived from HPLC analysis of the thermally equilibrated mixture.

From inspection of Figure 3.1 on the page 89, it is apparent that the $S_0 \rightarrow S_1$ photoexcitation of Z-11 and E-1 releases these systems along barrierless paths that evolve toward two distinct Z-CI-1 and E-CI-1 conical intersections. As found for PSB11 in Rh, these paths span a single excited-state energy surface with a charge transfer character (i.e., positive charge translocates toward the phenyl group) that increases moving from the Franck-Condon region to the intersections. The existence of the additional intersection points Z-CI-$81^\circ$ and E-CI-$92^\circ$ indicates that Z-CI-1 and E-CI-1 are part of a smooth intersection space segment spanning the bottom of the $S_1$ energy surface and providing prompt access to the $S_0$ energy surface.

Figure 3.2A on the page 90 reports the values of the computed oscillator strength (f)
along the $Z \rightarrow E$ path for the transition between $S_0$ and the first three excited singlet states. Such data indicate a prompt and intense $S_1 \rightarrow S_0$ emission in the $-17^\circ$ to $-20^\circ$ region of the path. This emission ($\alpha_{max} \approx 530$ nm, see the $54.5$ kcal mol$^{-1}$ $S_1$-$S_0$ energy differences in Figure 3.1) must arise immediately after photoexcitation and shift to the red as a function of time.

In Figure 3.2B, we report the energy profile along the $S_0$ relaxation path departing from Z-CI-1 and pointing toward the E-1 photoproduct. The path (we use a fixed solvent shell to model the initial relaxation where the solute structure changes faster than solvent reorientation) leads to a flat region of the $S_0$ energy surface centered on an approximately $-150^\circ$ twisted structure ($I_{Z\rightarrow E}$). However, upon relaxation of the solvent, $I_{Z\rightarrow E}$ further relaxes to an approximately $180^\circ$ twisted minimum corresponding to $S_0$-E-1. In fact, a limited pyrroline out-of-plane distortion allows for a path where the distance between the methyl groups at C$_2'$ and C$_5$ is large enough to avoid steric clash. To estimate the excited state lifetime of Z-1, we take advantage of the $Z \rightarrow E$ reaction path data. Analysis of the $S_1$ reaction coordinate indicates that there are 2 modes driving the reaction. The stretching mode corresponds to the expansion of the double bonds and contraction of the single bonds. The torsional mode is the reactive mode and describes the twisting of the central double bond. We assume that a 2-dimensional potential energy surface suffices to describe the $S_1$ population dynamics. Accordingly, we have fitted a suitable function of the $C_{1'y}$-$C_4$ bond length and $C_{9'y}$-$C_1'y$-$C_4$-$C_5$ twist against the $Z \rightarrow E$ data of Figure 3.1. As shown in Figure 3.2 C, the resulting (effective) potential energy surface incorporates the $S_1$ CASPT2/CASSCF/AMBER $Z \rightarrow E$ energy profile. The $S_1$ population motion is simulated generating a “classical” wave packet of 1,000 trajectories defined by constructing a Wigner distribution. We find that the population oscillates across the reaction-path valley and broadens but does not split. The population maximum (i.e., the average trajectory) reaches the energy minimum associated with the Z-CI-1 decay point on a 200 fs time scale pointing to a subpicosecond lifetime for $S_1$.

The model above predicts that the stretching mode is immediately excited and, provided
little kinetic energy redistribution occurs, it maintains $\approx 5$ kcal mol\(^{-1}\) of kinetic energy for
the rest of the $S_1$ evolution. In contrast, the isomerization mode accelerates slowly until
$\approx 100$ fs but then quickly incorporates kinetic energy reaching a $>25$ kcal mol\(^{-1}\) value upon
$S_1 \rightarrow S_0$ decay at $\approx 200$-fs delay. This type of behavior has also been found for Rh\(^{28}\)
where the kinetic energy evolution of its PSB11 chromophore was determined by computing
scaled-CASSCF/AMBER trajectories of a molecular model of the entire pigment.

**Time-Resolved UV-Vis Spectroscopy.** Figure 3.3 A shows the time-resolved emission data of Z-1. The fluorescence excited via the $S_0 \rightarrow S_1$ transition at 400 nm rises instantaneously and covers most of the visible spectral region. It shows a blue-centered band at time zero that red-shifts and disappears in $<1$ ps. The fitting of the kinetics traces with exponential functions convoluted with the instrument response function (solid lines) reveals a biphasic behavior at wavelengths $<600$ nm [time constants $\tau_1 < 40$ fs (i.e., an instrumental response-limited decay) and $\tau_2 = 300 \pm 30$ fs] and a single-exponential decay with $\tau_2$ at wavelengths $>600$ nm]. This fluorescence is assigned to the $S_1 \rightarrow S_0$ transition because: (i) it lies at the red of the singlet absorption band, (ii) appears promptly, and (iii) at early times is centered at 530 nm (53.8 kcal mol\(^{-1}\)), consistent with the $S_1$-$S_0$ energy gap and f values for approximately 20° torsions. The observed dynamics is consistent with the $S_1$ Z $\rightarrow$ E reaction path of Figure 3.1 and modeled $\approx 200$ fs decay. The initial relaxation from the Franck-Condon ($S_0$-Z-1) window along the stretching mode is probably reflected by the $\tau_1$ component, whereas the slower torsional motion, which depopulates the small-angle fluorescent window with high oscillator strength gives rise to the slower $\tau_2$ component.

We show transient absorption spectra (3.3 B) and kinetic traces (3.3C) after Z-1 excitation at 400 nm. The first 200 fs are dominated by the $S_1$ dynamics and later times by the $S_0$ relaxation, leading ultimately to relaxed E-1 and Z-1 conformations. This equilibrium situation is reached for $t < 30$ ps, and is best borne out by the Z/E difference spectrum at 100 ps. In the spectrum recorded after 0.1 ps, a broad stimulated emission (SE) band is observed at $\approx 500$-$520$ nm (55 kcal mol\(^{-1}\)), as expected from the above fluorescence. In
addition, we detect excited state absorption (ESA) at ≈ 430 nm (66.3 kcal mol\(^{-1}\)). Because the computed \(S_1 \rightarrow S_3\) transition has a negligible \(f\) value along the path (Figure 3.2A), such signal is assigned to a \(S_1 \rightarrow S_n\) transition. The negative signal in the 360 to 390 nm range is due to removal of molecules from \(S_0\) (ground-state bleach, GSB). Fitting of the ESA and SE kinetic traces at 420 and 560 nm, respectively, yields a sub-100 fs decay time that parallels the fast \(\tau_1\) relaxation time retrieved from the fluorescence decay.

The SE does not show the \(\tau_2\) component of the fluorescence, because, on this time scale, the data are dominated by a broad \(S_0\) absorption signal (GSA). The GSA is clearly visible in the 0.3 ps spectrum and gives rise to a sharp feature in the kinetic traces (>420 nm). It has a delayed onset (≈ 220 fs with respect to the excitation pulse) and decays with \(\tau_3 = 200 \pm 40\) fs. [Fitting the GSA without a delayed onset yields unsatisfactory results. At 540 nm, we found that this delay time is 180 ± 20 fs. It shifts to later times as one goes from 580 to 440 nm (Figure 3.3C)]. Because the model on Figure 3.2 C and the fluorescence data indicate \(S_1\) decay occurring on the same or even faster time scales, we suggest that the GSA is associated with a repopulation of \(S_0\) after decay at the conical intersection Z-CI-1. The small-amplitude long-wavelength wing (600-720 nm) of this band (0.3 ps spectrum in Figure 3.3 B) would represent molecules with larger twist angles. The wavelength-dependent onset time is consistent with a wave-packet-like evolution triggered by decay at Z-CI-1, with the 180 to 250 fs onset delay giving an approximate value for the internal conversion time. Although the GSA is most probably dominated by vibrationally hot Z-1, torsionally unrelaxed E-1 products (tentatively associated with the computed 445 nm absorbing \(I_{Z\rightarrow E}\) structure) are simultaneously formed at the Z-CI-1 bifurcation. The isomerization time is thus ≈ 200 fs, i.e., close to the 180 fs time observed for the production of photorhodopsin: the primary and nonisolable \(S_0\) intermediate of Rh.\(^{17}\)

At times <1 ps, the GSA signal is blue-shifted and is limited to wavelengths <500 nm, being attributed to vibrationally excited E-1 or Z-1. The average decay time of the spectrally integrated signal of this hot \(S_0\) absorption (GSA) is ≈ 5 ps, as expected for vibrational energy
relaxation in methanol and in agreement with the mid-IR results (see below). Fits of the 
GSB recovery \( \approx 380 \) nm yield a fast \( \approx 200 \) fs time and a slower \( \approx 5 \) ps component. The 
latter is the counterpart of the GSA decay due to vibrational relaxation, whereas the former 
reflects the \( S_0 \) recovery in both \( Z \) and \( E \) forms (isomerization). For delay times \(< 25 \) ps, 
a quasistationary spectrum is observed (100 ps spectrum, Figure 3.3 B) that displays the 
bleach of \( Z-1 \) (428 nm) and absorption due to the \( E \) isomers (380 nm). The signal level is 
consistent with the reported\(^{21} \) 0.20 \( Z \) to \( E \) isomerization quantum yield.

In Figure 3.3D, we summarize the relationship between the signals discussed above and 
the computed reaction path. The early sub-100 fs parts of fluorescence, SE (and ESA) 
are assigned to molecules with a twisting angle of less than \( 30^\circ \), located in the region of 
highest \( S_1-S_0 \) \( f \) value. This is corroborated by the agreement of the calculated and observed 
emission energies (Figure 3.1). Progression toward the conical intersection is predicted to 
occur on a \( \approx 90 \) fs time scale in accord with the \( \tau_1 \) decay time, but molecules with smaller 
kinetic energies populate larger twist angles (with lower \( S_1-S_0 \) \( f \) values) and show the weaker 
red-shifted fluorescence decaying with a \( \tau_2 = 300 \) fs. The early \( \tau_1 \) decay of \( S_1 \) triggers the 
fast-appearing GSA that dominates the long-wavelength transients at 0.3 ps. This early GSA 
is associated with highly twisted \( Z \)- and \( E \)-like molecules. This assignment is rationalized by 
the calculations predicting a transition energy of 40 kcal mol\(^{-1} \) (715 nm) for approximately 
\(-40^\circ \) and approximately \(-140^\circ \) twisting angles. \( S_0 \) cooling then leads to nearly equilibrated \( Z \) 
and \( E \) structures on a \( \approx 5 \) ps time scale.

**Time-Resolved Mid-IR Spectroscopy.** UV-pump-mid-IR-probe spectra were recorded 
in the 1,450- to 1,650- cm\(^{1} \) region, where the \( C_1-C_4 \) stretch band of \( Z-1 \) at 1,575 cm\(^{1} \) is spec- 
trally isolated from the \( E-1 \) band (Figure 3.4A). A smaller band at 1,610 cm\(^{1} \) can be assigned 
to the \( N-C_5 \) stretch mode based upon vibrational frequency calculations.\(^{21} \) Immediately after 
\( Z-1 \) excitation, a bleach spectrum is observed that is a replica of the absorption spectrum 
(Figure 3.4B), in line with the expectation that the \( C_1-C_4 \) stretch should disappear from \( S_1 \) 
because of the loss of double bond character. Within 1 ps, a broad band grows in at the
lower-frequency side of the original spectrum. This band continuously narrows and shifts to higher frequencies, overlapping with the original negative bleach (Figure 3.4C). After 30 ps, the spectrum is stable and closely resembles the steady-state E-Z difference spectrum, recorded after continuous 435 nm irradiation in a FTIR spectrometer (Figure 3.4D).

Growth and shift of the photoproduct band become more evident after subtracting from all subsequent signals the transient absorption spectrum at zero time delay. This reveals that the integrated intensity of the C$_1$′-C$_4$ stretch is almost completely recovered on a time scale of a few hundred femtoseconds in agreement with the fluorescence decay time and consistent with the fact that the double-bond character is restored on S$_0$. Thus, the induced absorption signal is assigned to S$_0$ species (Figure 3.3D). Because of the S$_1$ population branching occurring upon decay at Z-CI-1, both molecules relaxing to E-1 and molecules returning to Z-1 contribute to the signal. Indeed, 1 ps after photoexcitation, the excess energy from the S$_1$-S$_0$ decay results in a nonequilibrium excitation of low-frequency vibrational modes, which anharmonically couple to the C$_1$′-C$_4$ stretch, leading to a red-shift and spectral broadening. The steady-state difference spectrum is observed only after dissipation of this excess energy to the solvent. A moment analysis of the pump–probe data shows that, next to a small subpicosecond component, the main shift of the photoproduct band takes place on a 6 to 9 ps time scale, when only very small changes in the integrated band intensity are observed. This cooling/solvation time scale is in good agreement with the spectral dynamics observed for the UV-vis GSA signal. The quantum yield of the Z $\rightarrow$ E isomerization was calculated by comparing the magnitudes of initial bleach and final difference spectrum to the calibrated FTIR absorption and difference absorption spectra$^{29}$ yielding 0.21 $\pm$ 0.03. This is consistent with steady-state HPLC measurements$^{21}$.

The reverse E $\rightarrow$ Z photoreaction [occurring with a 0.35 quantum yield$^{21}$] was followed after transferring 87% of all molecules to the E form by continuous irradiation at 454.5 nm. By setting the UV-pump-laser wavelength at 388 nm, where the E-1 absorption is 30% stronger than the Z-1 absorption, we made sure that at least 90% of the photoexcited
molecules were initially in the E form. In Figure 3.5 we compare the transient spectra for E-1 to those for Z-1 under identical conditions. The spectral changes in both isomerization directions occur in parallel. In particular, the growth of the photoproduct spectrum and its shift on a < 10 ps time scale are observed in both cases consistently with the barrierless paths of Figure 3.1. However, for E-1, the spectral shift due to the isomerization is of opposite direction with respect to the initial (heat-induced) red-shift, and the isomerization process can be better distinguished. Indeed, a small shoulder at the spectral position of the Z-1 band can be seen in the transient spectra already 0.5-1 ps after E-1 photoexcitation (see arrows in Figure 3.5, confirming the ultrafast formation of Z-1. The sharpening and increase of this feature beyond 1 ps is again related to the vibrational relaxation of the S₀ species.

**CONCLUSIONS**

The development of single–molecule E/Z switches and rotary motors (chiral switches that undergo consecutive unidirectional Z → E or E → Z photoisomerizations) would require building blocks featuring faster response times and higher quantum yields. Studying the properties of a series of biarylidene motors, Feringa and coworkers¹⁴ have unveiled the complexity of the response time concept whose value may span 1 or more orders of magnitude as a function of framework and substitution pattern. In these systems, the isomerization time scale is set by a conformational helix inversion (e.g., from a P to a M helix) of either the Z or E form of the molecule:

\[
Z(M) \xrightarrow{h\nu} E(P) \overset{\triangle}{=} E(M) \xrightarrow{h\nu} Z(P) \overset{\triangle}{=} Z(M).
\]

The first and third steps are photoisomerizations, whereas the second and fourth steps are slow (rate-limiting) thermal helix inversions. The slow inversion steps are required to generate a conformer allowing for a forward, rather than reverse, subsequent photoreaction.

NAIPs are “chimerical” switches that incorporate into the Feringa’s biarylidenes skeleton
a protonated or alkylated Schiff base function. Although the application of NAIPs to the construction of rotary motors will require further research work, our computational and spectroscopic data indicate that their photoinduced dynamics replicates the dynamics of the PSB11 isomerization in Rh. This protein features a $S_1$ lifetime of $\approx 150$ fs, a $S_0$ transient (photorodopsin) appearance time of 180 fs, and a primary photoproduct (bathorhodopsin) appearance time of $\approx 6$ ps. These time scales match the observed and computed 0.1 to 0.3 ps $S_1$ lifetime of Z-1, the 200 to 250 fs $S_0$ replenishment, and the 6 to 9 ps time scale for the E-1 photoproduct cooling/solvation. Such properties suggest that NAIP-based motors may complete a half-rotary cycle in $< 10$ ps, i.e., a few orders-of-magnitude faster than the fastest ($\approx 6$ s for half-cycle) known biarylidene.

The replacement of the slow biarylidene helix inversion step with an impulsive reaction shuttling the primary photoproduct population (e.g., E(P) and Z(P) in the cycle) toward the stable conformers (e.g., E(M) and Z(M), respectively) may lead to much faster light-powered rotary motors. However, this can be implemented only in switches where an ultrafast photochemical reaction produces a product with a nonstatistical distribution of the vibrational energy polarized on the reactive ($Z \rightarrow E$ or $E \rightarrow Z$) mode. It is likely that, after insertion of a stereogenic center in position 2, 3, or 2’ (see Figure 1.4) [the synthetic routes do not exclude such possibility$^{22}$], NAIP switches could provide this type of system. In fact, as shown above, upon decay from $S_1$, $< 25$ kcal mol$^{-1}$ of kinetic energy are predicted to be located on the isomerization mode. Such momentum, together with the additional torque provided by the following $S_0$ reconstitution of the central double bond, may push part of the population well beyond the initial conformation, vibrational relaxation completing the process within 10 ps.

Despite the fact that the excited-state populations of 1 and Rh both evolve along a single path intercepting a conical intersection, the corresponding reaction coordinates indicate different isomerization mechanisms. Indeed, the $S_1$ evolution of PSB11 in Rh occurs via an asynchronous crankshaft mode involving twisting of 2 adjacent double bonds.$^{28}$ Because of
its locked skeleton, such motion cannot be accomplished in 1 that isomerizes by twisting a single double bond. This difference is accompanied by a lower value of the kinetic energy deposited, upon decay to $S_0$, in the reactive modes that is $< 10 \text{ kcal mol}^{-1}$ in Rh, in contrast to the above $25 \text{ kcal mol}^{-1}$ for 1. It is also important to point out that, relative to 1, the photoisomerization quantum yield of Rh is 2 to 3 times larger. These facts suggest that NAIPs not only provide a route to new materials but that they also constitute attractive systems for the investigation of fundamental problems such as the relationship between excited-state evolution and quantum yields.
Figure 3.1: Excited-state reaction paths. (Inset) Overview of the studied parts of the photocycle. CASPT2/CASSCF/6-31G*/AMBER S\(_0\) (diamonds), S\(_1\) (squares), S\(_2\) (triangles), and S\(_3\) (crosses) energy profiles along the photoisomerization coordinate (the dihedral angle C\(_5\)-C\(_4\)-C\(_1\)′-C\(_9\)′) of Z-1 and E-1. Notice that the protonated (RGraphicH) rather than methylated (RGraphicMe) derivative is used for all computations. S\(_0\)-Z-1 and S\(_0\)-E-1 correspond to the S\(_0\) equilibrium structure. The Z-CI-1, Z-CI\(_{81°}\), E-CI-1, and E-CI\(_{92°}\) correspond to conical intersection structures. (The Z and E conical intersections have different solute and solvent shell structures). I\(_{Z\rightarrow E}\) corresponds to an intermediate generated via a relaxation starting at Z-CI-1 with a fixed solvent shell (dotted lines, see details in Figure 3.2B). The corresponding S\(_0\) path starting at E-CI-1 has also been computed (data not shown) but leads to a structure close to S\(_0\)-Z-1. The stream of arrows departing from S\(_0\)-Z-1 and ending at IZ \(\rightarrow\) E indicates the computed Z \(\rightarrow\) E photochemical reaction path. The dashed arrow on the bottom right indicates that relaxation of the I\(_{Z\rightarrow E}\) solvent shell produces E-1. The vertical arrows indicate the electronic transition (absorption and fluorescence) probed experimentally. The double arrow indicates the S\(_0\) Gibbs free energy difference of S\(_0\)-Z-1 and S\(_0\)-E-1. The computed and observed (italics) energies are given in parentheses. The S\(_0\)-Z-1 and S\(_0\)-E-1 observed excitation energies are from ref. \(^{21}\)
Figure 3.2: Analysis of the Z → E photochemical reaction path. (A) Change in CASSCF/6-31G*/AMBER S₀ → S₁ (diamonds), S₀ → S₂ (triangles), and S₀ → S₃ (squares) oscillator strengths. (B) CASPT2//CASSCF/6-31G*/AMBER S₀ (diamonds), S₁ (squares) energy profiles along the Z → E S₀ relaxation path computed with a fixed solvent shell and starting at Z-CI-1. (C) Simulation of the population dynamics along a 2D-model of the S₁ energy surface. The arrow indicates the initial direction of the gradient. The values in parentheses refer to the kinetic energy component along the torsion (i.e., isomerization) mode.
Figure 3.3: Ultrafast UV-vis spectroscopy. (A) Time0 and wavelength-resolved fluorescence of Z–1 with color-scale-coded intensity. The maximum at 520-530 nm and a biphasic decay are apparent. (Inset) Kinetic traces at selected wavelengths (data points are vertically shifted for the sake of clarity). Solid lines are best-fits to the data. (B) Time-resolved absorption spectra (change in absorbance, $\Delta A$), showing $S_1 \rightarrow S_0$ absorption (ESA, 425 and 330 nm), $S_1 \rightarrow S_0$ stimulated emission (SE, 460-700 nm) and the $S_0$ absorption (GSA, 440-700 nm), blue-shifting at later times (420-480 nm). The 100 ps data are the E-Z difference spectrum. Data points with residual pump light scatter (395-405 nm) are removed. (C) Selected kinetic traces, vertically shifted by 30 mOD. The dashed line is a guide to the eye highlighting the blue-shifting early GSA feature (D) Schematic view of the observed spectral transitions on the $S_0$ (solid lines) and $S_1$ (dashed lines) potential-energy surfaces as a function of the ($Z$ /to $E$) reaction coordinate. Arrows indicate the course of the reaction as suggested by the observed sequence ESA/FS/SE $\rightarrow$ GSA $\rightarrow$ hot-GSA. GSA and hot-GSA arise, most probably, from both Z and E conformers. The phases of the photoreaction probed by time-resolved IR spectroscopy are also indicated.
Figure 3.4: Time-resolved mid-IR spectra. (A) FTIR absorption spectra of Z-1 (dashed line) and E-1 (solid line). (B) Transient spectra at different delays (0, 0.2, 0.4, 0.6, and 1 ps) after the excitation of Z-1 at 420 nm. (C) Transient spectra at delays 1, 2, 5, 10, and 20 ps, all scaled by a factor of 1.5 for better visibility. (D) Difference between the FTIR spectra shown in A reproducing the transient spectra at delays > 30 ps.
Figure 3.5: Ultrafast bidirectional isomerization: (A) FTIR absorption spectra of Z-1 (dashed line) and E-1 (solid line). (B) Transient mid-IR spectra at different delays after the excitation of Z-1 (dashed line) and E-1 (solid line) at 388 nm.


Chapter 4

THE PHOTOCHEMISTRY OF THE NAIP-3 SWITCH
ABSTRACT

We report the results of a multidisciplinary research effort where the methods of computational photochemistry and retrosynthetic analysis/synthesis, have contributed to the preparation of a novel N-alkylated indanylidene-pyrroline Schiff base featuring an exocyclic double bond and a permanent zwitterionic head. We show that, due to its large dipole moment and efficient photoisomerization, such system may constitute the prototype of a novel generation of electrostatic switches achieving a reversible light-induced dipole moment change of the order of 30 Debye. The modeling of a peptide fragment incorporating the zwitterionic head into a conformationally rigid side-chain, shows that the switch can effectively modulate the fluorescence of a tryptophan probe.

1 The chapter based on the article – Alfonso Melloni, Riccardo Rossi Paccani, Donato Donati, Vinicio Zanirato, Adalgisa Sinicropi, Maria Laura Parisi, Elena Martin, Mikhail Ryazantsev, Wan Jian Ding, Luis Manuel Frutos, Riccardo Basosi, Stefania Fusi, Loredana Latterini, Nicolas Ferr and Massimo Olivucci “Modeling, Preparation and Characterization of a Dipole Moment Switch Driven by Z/E Photoisomerization” (accepted in J. Am. Chem. Soc.)
INTRODUCTION

We have recently shown that N-alkylated indanylidene-pyrroline (NAIP) Schiff bases provide a class of biomimetic switches that replicate different aspects of the Z/E photoisomerization of rhodopsin (i.e. the visual pigment of superior animals).\textsuperscript{1,2} Below we report the results of a multidisciplinary effort where the methods of computational photochemistry and retrosynthetic analysis/synthesis, have contributed to the preparation of the NAIP-3 switch (from now on switch 3) featuring a stable zwitterionic head. We show that due to its large dipole moment, one may constitute the prototype of a generation of electrostatic switches achieving a light-induced dipole moment inversion of the order of 30 Debye.

Photochemical switches are bistable compounds that can be interconverted between two different isomers (states) via light irradiation. Literature examples of electrostatic photoswitches are compounds of the spiropyran-type that can be changed between a neutral state and a zwitterionic state via a photochemical ring-opening reaction. Such property has been employed in different experiments to reversibly modulate the activity of enzymes\textsuperscript{3,4,5} and channel proteins\textsuperscript{6} and to achieve novel sensors.\textsuperscript{7} Furthermore, it has been shown that the strong permanent dipole moment in the zwitterionic form of nitrospiropyrane units attached to specific peptide residues is responsible for light-induced $\alpha$-helix$\rightarrow$random coil conformational transition.\textsuperscript{8}

In principle, the Z/E photoisomerization of olefins can provide the basis for the development of electrostatic photoswitches where a large dipole moment change is achieved via a ca. 180° rotation of a functionalized alkylidene unit. As shown in Figure 4.1a the placement of opposite charges on unit A and along an axis orthogonal to the central double bond provides a permanent dipole moment that is exactly inverted (with respect to unit B) upon double-bond isomerization. Notice that, provided that the unit B is spatially constrained, the molecular dipole moment inverts sign rather than changes intensity (as in spiropyrans where it goes from a small to a large value without direction control) leading to a maximum electrostatic potential change on the surrounding environment. As a working hypothesis we
propose that a system of the type described above can be achieved via functionalization of the known NAIP photoswitch $1$. The photochemical and spectral characterization of $1$ (a chloride-alkylated Schiff base ion-pair) in methanol reveals a prompt wavelength control of the Z/E photostationary composition as well as a $Z \rightarrow E \rightarrow Z$ photocycle that is completed in ca. 20 picoseconds. The same studies also reveal that ab initio multiconfigurational quantum chemical methods coupled with a molecular mechanics force field allow a realistic modeling of the excited states of the system as well as of other solution-phase NAIPs.

The zwitterion $3$ is a derivative of $1$ in which a non-conjugating internal carboxylate (replacing the external chloride) is substituted at $C_2$. Nakanishi and coworkers have reported the synthesis and characterization of different zwitterionic alkylated Schiff bases. These were used to investigate the point charge model for the control of the absorption wavelength in visual receptors. However, to our knowledge, $3$ represents the first zwitterionic Schiff base where two relatively rigid frameworks (i.e. five-membered rings) are connected by a single

---

Figure 4.1: a) Schematic representation of an electrostatic switch based on Z/E isomerization. The $180^\circ$ rotation of unit A will invert the dipole moment vector with respect to unit B. b) Inversion (relative to the superimposed indanylidene unit) of the dipole moment of $S_0$-$Z$-$3$ and $S_0$-$E$-$3$. The computed dipole moment values are 15.7 and 14.8 Debye respectively.
exocyclic double bond providing full conformational control.

In the following we report the result of a four-stage work starting with the construction and analysis of a quantum-mechanics/molecular-mechanics (QM/MM) model of 3. Since the model indicates that 3 has the same favourable photochemical and spectral features of 1, during the second stage of the work we design and perform the synthesis of a zwitterionic derivative. The characterization of the prepared compound shows that the predicted features are supported by spectral and photochemical data. Through the continuation of the synthesis work, we then show that 1 can be turned into an unnatural α-amino acid featuring a quaternary α-carbon at C₂. Finally, since the results above support the possibility to prepare semi-synthetic proteins incorporating 3 as a rigidly oriented side-chain, we build a QM/MM model of a tripeptide fragment incorporating a tryptophan probe. The simulation of the tryptophan emission maximum for the Z and E form of the peptide is used to quantify the ability of the switch to control the optical properties of the fragment.

METHODS

Spectroscopy and photochemistry: Photoisomerization was carried out with a 900W irradiator, f/3.4 monochromator (Applied Photophysics) apparatus and was followed by 1H NMR spectroscopy (Bruker AC 200 and 400 spectrometers at 200.13 and 400.13 MHz respectively). The composition of the photostationary state at different irradiation wavelengths was evaluated from the area ratio of the signal of the aromatic proton in the ortho position with respect to the exocyclic double bond (see above and the Supporting Information. UV/Vis measurements were performed by using a Hewlett Packard 8423 spectrophotometer.

Computations: The model of the Z and E switches in solution was constructed by placing the chromophore in a rectangular box of methanol molecules positioned within 10 Å from any given atom of the chromophore by using the xleap module of the Amber package. The average ground-state configuration of the methanol molecules (that is, the sol-
vent) was determined according to the following procedure. The solvent was relaxed by 1000 conjugate-gradient minimization steps using periodic boundary conditions while keeping the chromophore (i.e. the solute) fixed in its gas-phase configuration. In this step the partial charges of the chromophore atoms were determined with GAUSSIAN03, using a Restrained ElectroStatic Potential (RESP) procedure at the HF/6-31G* level of theory. The minimized system was further relaxed (keeping the solute molecule fixed) using molecular dynamics simulation within isothermal-isobaric NPT ensemble (1 atm, 298 K) using the program NAMD. In the next step we performed CASSCF/6-31G*/AMBER geometry optimization to relax the coordinates of the QM chromophore and of all solvent molecules featuring at least one atom less than 4.5 Å away from any solute atom. The positions of the remaining solvent molecules, more distant from the chromophore, were kept frozen. The QM calculations were based on a CASSCF/6-31G* level including an active space of 12 electrons in 11 π orbitals (that is, the full π system of the solute). The chosen 6-31G* basis set represented a cost/accuracy compromise yielding an excitation energy error of less than 3 kcal mol$^{-1}$ for rhod and solution-phase PSB11. CASSCF/6-31G*/AMBER geometry optimization was carried out with GAUSSIAN03 and TINKER 4.2. To account for dynamic correlation energy, four root-state average CASPT2 calculations were carried out by using the MOLCAS6.7 software.

The QM/MM model of the gas-phase tripeptide Ala-$\text{Zw}$-Trp (see below) was constructed using the CASSCF(10,9)/6-31G* QM level to describe the 3-methyl indole (3-MI) fluorescent moiety of the Trp residue and the generalized amber force field (GAFF) point charges to describe the rest of the peptide at the MM level including the unnatural Zw residue (see Result and Discussion section). Accordingly the QM/MM frontier was set at the weakly polarized $\text{C}_\alpha$-$\text{C}_\beta$ bond of Trp. The hydrogen link-atom (HLA) scheme was used to saturate the QM $\text{C}_\beta$ atom. The hydrogen link atom was fixed at 1.0 Å from $\text{C}_\beta$ and along the $\text{C}_\alpha$-$\text{C}_\beta$ bond axis (see Figure 4.2). In order to avoid overpolarization of the QM wavefunction by the close point charges, the charge of the frontier $\text{C}_\alpha$ carbon atom was set to zero, and
Figure 4.2: The QM/MM model of Ala-\textit{Zw}-Trp.

the charges of the other MM atoms of Trp were modified since Trp is a natural residue (Table 4.1). This procedure is allowed by the small values of the original AMBER99 point charges, which also make it possible to use the standard MM van der Waals and bonded potentials parameters. The point charges of all other MM atoms (except those of standard Ala residue) were determined via HF/6-31G* RESP calculations.

As for the solution system described above, the QM/MM calculations were carried out with a modified version of GAUSSIAN03,\textsuperscript{11} linked with a modified version of TINKER4.2.\textsuperscript{15} The guess structures were drawn in two conformations (Conf1 and Conf2) that allowed a stabilizing stacking interaction between the \textit{Zw} and Trp side-chains and hydrogen bonding between the -COO- function of the Z form of \textit{Zw} and the N-H group of the 3-MI moiety. The corresponding equilibrium structures were computed via QM/MM geometry optimization. Based on the QM/MM optimized ground-state (GS) structure, a CASPT2 single point calculation was conducted with MOLCAS6.7\textsuperscript{16} to evaluate the excitation energies as well the associated oscillator strength f. The emission energies from the second excited state (that
corresponds to the emitting $^1$La state) were estimated by computing the GS-$^1$La vertical energy gaps via CASPT2 computations using a zeroth order three-root state-average CASSCF wavefunction at the $^1$La equilibrium geometry.

Table 4.1: Modified MM point charges for the part of the Trp residue.

<table>
<thead>
<tr>
<th>Atom</th>
<th>N</th>
<th>C$_\alpha$</th>
<th>C$_\text{carbonyl}$</th>
<th>H$_N$</th>
<th>O$_\text{carbonyl}$</th>
<th>H$_\alpha$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Charge</td>
<td>-0.4157</td>
<td>-0.0000</td>
<td>0.5973</td>
<td>0.24335</td>
<td>-0.5679</td>
<td>0.08375</td>
</tr>
</tbody>
</table>

RESULTS AND DISCUSSIONS

Recently$^{14,18}$ we have implemented the ab initio CASPT2//CASSCF protocol (where equilibrium geometries and electronic energies are determined at the CASSCF$^{19}$ and CASPT2$^{20}$ levels respectively) in a QM/MM scheme allowing for the evaluation of the excitation and emission energy of neutral or charged chromophores (treated quantum mechanically) embedded in protein or solution environments (described by the AMBER force field) with a few kcal mol$^{-1}$ errors. Using this ab initio CASPT2//CASSCF/AMBER protocol we were able to show$^{1,2}$ that the observed absorption and fluorescence maxima of 1 can be reproduced within a few kcal mol$^{-1}$. In the following the same methodology is used to achieve a description of the excited state (e.g. vertical excitation energy, nature of the spectroscopic state, geometrical relaxation) of 3 in methanol solution.

In Table 4.2 we report the computed absorption maxima for the equilibrium $S_0$ structure of both the Z and E isomers of 3 ($S_0$-Z-3 and $S_0$-E-3 in Figure 4.3a). Consistent with the oscillator strength values, the spectroscopic state corresponds to $S_1$. Figure 4.3b) gives the $S_0$ and $S_1$ charge distribution computed for the ground-state equilibrium structure $S_0$-Z-3 which points to a charge-transfer nature of the spectroscopic state. Accordingly, upon $S_0 \rightarrow S_3$ vertical excitation, $S_0$-Z-3 undergoes a 29% charge translocation through its reactive $C_1'=C_4$ bond from the pyrroline to the indanylidene moiety. On $S_1$ the positive charge is stabilized in a region away from the Schiff base function ($C_5$=N) by delocalization on the
phenyl ring. As found for the parent photoswitch \(1^2\), the p-OMe group further enhances the positive charge stabilization in the indanylidene moiety.

Table 4.2: CASPT2//CASSCF/AMBER absorption (\(\lambda_{\text{max}}\)), change in dipole moment (\(\Delta\mu\)), oscillator strength (\(f\)) and charge translocation (\(\Delta q\)) [a] Oscillator strengths are calculated using the CASSCF 0th-order wavefunction and CASSCF energies. [b] Charge transfer from the pyrroline to the indanylidene ring for \(3\) in MeOH.

<table>
<thead>
<tr>
<th>Exc</th>
<th>(\lambda_{\text{max}}) (nm)</th>
<th>(\Delta\mu) (Debyes)</th>
<th>(f^{[a]})</th>
<th>(\Delta q^{[b]}) (au)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Z S_0 \rightarrow S_1)</td>
<td>390</td>
<td>9.4</td>
<td>0.646</td>
<td>-0.29</td>
</tr>
<tr>
<td>(Z S_0 \rightarrow S_2)</td>
<td>304</td>
<td>0.8</td>
<td>0.026</td>
<td>-0.02</td>
</tr>
<tr>
<td>(Z S_0 \rightarrow S_3)</td>
<td>250</td>
<td>1.1</td>
<td>0.210</td>
<td>-0.03</td>
</tr>
<tr>
<td>(E S_0 \rightarrow S_1)</td>
<td>371</td>
<td>10.0</td>
<td>0.667</td>
<td>-0.27</td>
</tr>
<tr>
<td>(E S_0 \rightarrow S_2)</td>
<td>301</td>
<td>0.8</td>
<td>0.033</td>
<td>-0.02</td>
</tr>
<tr>
<td>(E S_0 \rightarrow S_3)</td>
<td>248</td>
<td>1.5</td>
<td>0.0255</td>
<td>-0.04</td>
</tr>
</tbody>
</table>

In Figure 4.4a, we report the relative CASPT2//CASSCF/AMBER energies for the \(S_0-Z-3\) and \(S_0-E-3\) and for two associated \(S_1/S_0\) conical intersections (CI-Z-3 and CI-E-3) as a function of the torsional deformation along the reactive \(C_1'-C_4\) bond. Similar to the \(1\) ion-pair\(^2\) we have located two very shallow (actually unstable) \(S_1\) energy minima (\(S_1-Z-1\) and \(S_1-E-1\)) describing the region of the excited state energy surface reached immediately after relaxation form the corresponding Franck-Condon points. In other words, CI-Z-3 is located at the bottom of an initially flat \(S_1\) potential energy surface valley reached via relaxation of the \(S_0-Z-3\) structure and developing out of the \(S_1-Z-3\) region. This is demonstrated in Figure 4.4b that incorporates the result of intrinsic reaction coordinate (IRC) calculations starting at \(S_0-Z-3\) and reaching CI-Z-3 after a barrierless path. Similarly, the second conical intersection is located at the bottom of the \(S_1\) valley starting at \(S_1-E-3\). As for \(1^2\), these structures are analogues of the excited state minimum and conical intersection reported for the visual pigment rhodopsin (Rh).\(^{14}\) While the \(S_1\) energy surface always has a dominating charge transfer character, at \(S_1-Z-3\) the \(S_1-S_2\) energy gap is \(\leq 10\) kcal mol\(^{-1}\). Therefore, it is consistent with a weak coupling between \(S_1\) and the diradical (dark) state \(S_2\).

The data in Figure 4.4a (on the page 117) suggest, for both the \(Z\rightarrow E\) and \(E\rightarrow Z\) reactions, an isomerization mechanism (see full arrows for the \(Z\rightarrow E\) process) similar to the one fully
documented for 1 and Rh.\textsuperscript{14} After $S_0 \rightarrow S_1$ photoexcitation, 3 initially relaxes along an $S_1$ path dominated by double-bond/single-bond inversion and only a limited torsional deformation. This initial event is followed by a torsional relaxation about the $C_1\equiv C_4$ bond. Indeed, upon excitation of the $S_0$-$Z$-3 structure, which has a -13$^\circ$ pre-twisted exocyclic $C_1\equiv C_4$ bond, it relaxes to a $S_1$-$Z$-3 structure featuring an inverted $\pi$-bond order and a -20$^\circ$ twisted $C_1\equiv C_4$ bond. From the $S_1$-$Z$-3 structure, the system evolves toward the conical intersection Z-3-CI which is approximately -80$^\circ$ twisted (in the modelled R enantiomer) and it is accessed through a substantially barrierless path. After $S_1 \rightarrow S_0$ decay at the conical-intersection, the $S_0$-E-3 photoproduct (featuring a +177$^\circ$ pre-twisted $C_1\equiv C_4$ bond) can be reached via further torsional relaxation in the same direction. At this point a photocycle can be achieved via subsequent photoisomerization of $S_0$-E-3 (dashed arrows in Figure 3). Indeed, the data presented in Figure 4.4a support the existence of a $S_0$-E-3 excited-state energy surface similar to that of $S_0$-$Z$-3 and leading to $S_1$ evolution towards the E-3-CI intersection featuring an approximately 90$^\circ$ twisted $C_1\equiv C_4$ bond.

Similar to the parent compound 1, the relative stability of $S_0$-$Z$-3 and $S_0$-E-3 in solution has been determined in terms of the difference in Gibbs free energy between the solvated systems using averaged solvent electrostatic potential/molecular dynamics (ASEP/MD) computations.\textsuperscript{21} The results indicate that, at 298 K, $S_0$-E-3 is 4.1 kcal mol$^{-1}$ higher in free energy than $S_0$-$Z$-3. This difference can be ascribed to steric hindrance factors: in fact, in the case of the E isomer the methyl substituent on the C$_5$ of the pyrroline ring is closer to the methyl substituents on C$_{2'}$ of the indanylidene moiety. The larger difference in 3 relative to 1 could be due to diaxial interaction of the carboxylate and methyl groups.

As already mentioned above, 3 presents a permanent dipole moment that, with respect to the indanylidene unit, can be inverted via light-induced Z/E isomerization. The computed solution $S_0$ dipole moments are 15.7 and 14.8 Debyes for the $S_0$-$Z$-3 and $S_0$-E-3 respectively. As shown in Figure 4.1b, the superposition of the indanylidene units of the two isomers displays dipole moment vectors nearly orthogonal to the $C_1\equiv C_4$ axis and forming a 157$^\circ$
Figure 4.3: a) CASSCF/AMBER computed equilibrium structures (hydrogens not shown) of $S_0$-Z-3 and $S_0$-E-3 embedded in a box (not shown) of methanol molecules. The computations are carried out for the R enantiomer and show that the $\text{COO}^-$ group is stable in an axial (Ax) rather than equatorial position in both forms (the equatorial positions appear to be unstable in methanol). The values in degrees refer to the $C_9'-C_1'C_4'-C_5$ torsional angle. Gas-phase and preliminary solution computations indicate that these values correspond to lowest energy conformations. (See also below. Solution conformations may exist with a $C_9'-C_1'C_4'-C_5$ torsional angle of opposite sign. These conformers will be investigated in a future work). Bond lengths are given in Å. b) Change in the charge distribution for the $S_0$-Z-1 structure upon $S_0 \rightarrow S_1$ excitation.

angle thus yielding a difference module of 29.8 Debye. The dipole moment values, obtained within an explicit solvent representation (i.e. with a suitably constructed box of methanols) are confirmed by an implicit treatment of the solvent. Indeed, when using an independent solution model based on B3LYP/6-31+G* density functional theory with the polarizable continuum model (PCM) implemented in GAUSSIAN03,$^{11}$ the computed values of the $S_0$-Z-3 and $S_0$-E-3 dipole moments are 15.8 and 13.5 Debyes respectively. Therefore, both explicit and implicit solvent representations appear to yield large and consistent dipole moment values.
Since the dipole moments of 3 could only be predicted and not measured, we benchmarked the employed computational methods indirectly. Using the B3LYP/6-31+G* and PCM methods we have evaluated the dipole moments of 3 in dioxane finding slightly decreased values with respect to methanol (13.7 and 11.6 Debyes for the Z and E form respectively). With exactly the same method, we also computed the dipole moments of an experimentally investigated merocyanine-spirobenzopyrane system in dioxane and compared the results with the observed values available in the literature\textsuperscript{22}. The observed dipole moments 17.7 and 4.3 Debye for the open merocyanine and closed spiropyran form respectively) compare reasonably well with the computed data (15.9 and 6.9 Debyes respectively) pointing to moderately underestimated predictions.

As shown in Figure 4.5 (on the page 118), the absorption spectra of Z-3 and E-3 in methanol display two bands absorbing above 250 nm. The observed $\lambda_{\text{max}}$ values and intensities were compared with the predicted properties in Table 4.2 to validate our computational models and, in turn, to assign the bands to given electronic transitions. In fact, the computed $S_0 \rightarrow S_1$ absorption maximum for the Z isomer ($\lambda_{\text{max}}$) falls within 10 nm of the intense band at 397 nm, which yields a computational error of less than 2 kcal mol$^{-1}$ in excitation energy.

Similarly, the observed weaker band at $\lambda_{\text{max}} = 263$ nm falls close to the computed $S_0 \rightarrow S_3$ transition. Furthermore, the $S_0 \rightarrow S_1$ and $S_0 \rightarrow S_3$ values of the oscillator strength $f$ (0.646 and 0.210, respectively) are qualitatively consistent with the observed absorbance pattern. The $S_0 \rightarrow S_2$ transition is predicted to correspond to a weak band ($f=0.026$) most probably hidden below the shoulder near 300 nm, as indicated in Figure 4.5.

The computational results displayed in Figure 4.4 indicate that $S_0$-Z-3 and $S_0$-E-3 can be photochemically interconverted via barrierless, and therefore ultrafast (i.e. with a timescale below 1 picosecond) reactions. This conclusion is in agreement with the results of a recently reported combined femtosecond fluorescence up-conversion, UV-Vis and IR transient absorption spectroscopic study of 3. Indeed, the fluorescence lifetime for the Z$\rightarrow$E
process is found to be 140 fs with an excited state absorption persisting over 230 fs in the form of a vibrational wavepacket according to twisting of the isomerising double bond. It is also shown that the hot photoproduct appears on the ground state on a 400 fs time-scale and that the reaction is substantially completed after 600 fs from photon absorption.

To assign the observed time-scale we have used a recently proposed computational strategy to simulate the excited state motion of 3 in methanol. Accordingly we computed a scaled-CASSCF/AMBER trajectory\textsuperscript{23} of our Z-3 solvated model (i.e. the same used to reproduce the reaction path of Figure 4.4b) starting at a point close to the Franck-Condon structure and describing the relaxation of the first solvent shell together with the solute. In this calculation the excited state CASSCF gradient is scaled in such a way to simulate the more accurate (but computationally unpractical) CASPT2 gradient along the entire reaction path of Figure 4.4b). In other words such a trajectory simulates, at the cost of CASSCF gradients, the excited state evolution on a CASPT2-like potential energy surface.\textsuperscript{23} This protocol has reproduced successfully the magnitude of the observed excited state lifetime of the retinal chromophore in Rh and provided a mechanism in line with its ultrafast isomerization.\textsuperscript{23} Notice that it has not been possible to start the trajectory from the Franck-Condon structure due to the incorrect order (with respect to the corresponding CASPT2 values) of the S\textsubscript{2} and S\textsubscript{3} states computed at the CASSCF level of theory. The order becomes correct only at ca. 30° double bond twisting deformation (i.e. ca. 20° more twisted than the Franck-Condon structure). Thus, the computed trajectory provides an approximated lower limit of the reaction time-scale that, according to Figure 4.6a and assuming a ca. 50 fs initial relaxation, is predicted to be in the 300-400 fs range. In spite of this uncertainty the ultrafast nature of the isomerization and the magnitude of the predicted S\textsubscript{1} lifetime are consistent with the observed excited state dynamics\textsuperscript{24} that points to an excited state lifetime between 250 and 400 fs. Notice that transient fluorescence is predicted to occur, assuming a ca. 50 fs initial relaxation time, during the first 200 fs as indicated by the values of the S\textsubscript{0}-S\textsubscript{1} oscillator strength (see Figure 4.6a).
The scaled-CASSCF/AMBER trajectory also reveals that the solvent cavity cannot stop
the excited state photoisomerization motion that, according to our analysis, is facilitated
by the ring inversion of both the indanylidene and pyrrolinium envelope-like five-member
rings. Indeed, we could observe (see movie in the Supporting Information) an inversion of
the axial/equatorial position of the two methyl and hydrogen groups at C2′ and C3 respec-
tively. Because of the inversions the double bond can undergo a twisting deformation of
90° and decay at the conical intersection without moving the bulkiest parts of the molec-
ular framework. Indeed, as shown in Figure 4.7a, both the phenyl ring and the highly
solvated COO− and N(Me)+ groups remain substantially fixed during the entire S1 mo-
tion. The counter-clockwise rotation about the reactive C1′-C4 double bond is accompa-
nied by clockwise twisting (associated with the ring inversion) of the C2′-C3′ and C2-C3 bonds.
These directions are imposed by the pre-twisted of the C1′-C4 double bond and initial ring
conformation (i.e. imposing a framework with M,M helicity) of S0-Z-3 and, according to
our computations, are a consequence of the system R configuration. This is reminiscent of
the type of cooperative bond-twisting space-saving mechanism documented for the retinal
chromophore embedded in the chiral cavity of Rh.23

Since, according to Figure 4.4 and the experimental data presented in ref.,23 both the
Z → E and E → Z are found to be ultrafast and the absorption λmax values of Z-3 and E-3
are close but not identical (see Table 4.2), it is predicted that, upon continuous irradiation, a
photostationary state will be rapidly generated whose composition depends on the irradiation
wavelength, leading to the possibility to control the dipole moment value. Comparison of
the Z-3 and E-3 spectra suggests, consistently with the computed λmax data, a decrease of
the Z/E ratio upon an increase in the wavelength (see Table 4.3) since the S0 → S1 band
of E-3 appears to be slightly blue-shifted and more intense than the corresponding Z band.
In particular, at 440 nm, the E form becomes dominant. On the other hand, due to a
relatively low E → Z thermal isomerization barrier, which has not yet been measured or
computed, the photogenerated ratio slowly returns to the original Z/E mixture (a Z/E ratio
larger than 99:1) when the irradiation is interrupted at room temperature. This is consistent with the more than 4 kcal mol\(^{-1}\) computed energy difference between \(S_0\)-Z-\(\mathbf{3}\) and \(S_0\)-E-\(\mathbf{3}\). Notice that, while the photostationary state is rapidly reached (e.g., 12 min irradiation at the isosbestic point), this room-temperature thermal relaxation/interconversion is not a fast process. A relaxation time of about 8 hours could be estimated, allowing the isomers to be fully characterized. The systems could undergo irradiation-dark relaxation cycles several times without showing any degradation. This indicates that \(\mathbf{3}\) can indeed be employed in applications where the photoswitch triggers molecular events that are completed on the timescale shorter than a few hours or when one wants to recover the original Z/E composition (and be ready to repeat a photoswitching cycle) after several hours or more.

Table 4.3: Composition of the photostationary state of \(\mathbf{3}\) in methanol as a function of the irradiation wavelength determined by NMR analysis.

<table>
<thead>
<tr>
<th>(\lambda_{\text{max}}) (nm)</th>
<th>Z (±0.1)</th>
<th>E (±0.1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>340</td>
<td>1</td>
<td>0.5</td>
</tr>
<tr>
<td>390</td>
<td>1</td>
<td>0.6</td>
</tr>
<tr>
<td>440</td>
<td>1</td>
<td>1.2</td>
</tr>
</tbody>
</table>

The similarity between the reaction path in Figure 4.4 and that of the parent \(\mathbf{1}\) photoswitch and Rhodopsin, suggests that Z-\(\mathbf{3}\) and E-\(\mathbf{3}\) could potentially display good photoisomerization quantum yields. These quantities were measured via HPLC and spectroscopic analysis and found to be 0.20 and 0.23 for the Z-\(\mathbf{3}\) and E-\(\mathbf{3}\) isomerization respectively. The E \(\rightarrow\) Z quantum yield was determined by considering the effect of the thermal E \(\rightarrow\) Z isomerization on the isomer concentrations determined by HPLC. The fact that an ultrafast reaction driven by a barrierless path leads to a moderate quantum yield points to a decay mechanism that deviates from the simple one-dimensional Landau-Zener picture.\(^{25}\) This could be connected with the multimode nature of the excited state reaction path involving both ring puckering and bond twisting and/or to the restraints imposed, after the excited state decay, by the solvent shell on the isomerization motion. Furthermore, the limited magnitude of the quantum yields may be related to the fact that immediately after the decay
the system needs to reshape the solvent shell to accomplish full double bond isomerization.

The successful photoisomerization documented above for textbf3 featuring a photoswitchable side-chain, strongly support the possibility to achieve semi-synthetetic peptides and proteins incorporating a dipole switch in a conformationally locked orientation. While the preparation of such a system goes beyond the scope of the present research, below we report on the computational design of a tripeptide (Ala-\textit{Zw}-Trp) that incorporate \textit{Zw} as the central residue and tryptophan as the N-terminal residue. The tryptophan residue constitutes an internal fluorescence probe sensing the changes in the electrostatic field produced in its environment. In virtue of the fact \textit{Zw}, features a quaternary \(\alpha\)-carbon, the side-chain of this residue is conformationally rigid and has a well-defined orientation with respect to the peptide backbone. In particular, as displayed in Figure 4.8b (right), this side-chain would be orthogonal to the local backbone axis and is positioned off to the left or to the right of the same axis as a function of the stereochemistry (R or S) of the \(\alpha\)-carbon. Here our target is to assess if the Z/E photoisomerization reverting the dipole moment of the central residue would significantly change the tryptophan emission wavelength. In other words we want to find out if the dipole moment inversion would expose the 3-MI fluorophore of tryptophane to a different electrostatic field, thus modulating its GS-\textsuperscript{1}La energy gap. In fact, as shown in Figure 4.8b, the fluorescent \textsuperscript{1}La state of tryptophan has a charge-transfer character (ca. 30 \% of the pyrrole ring \(\pi\)-electron density is shifted towards the phenyl ring upon photoexcitation).\textsuperscript{26} As a consequence an electrostatic potential that stabilizes a positive charge on the 3-MI pyrrole ring or a negative charge on the phenyl ring will stabilize the \textsuperscript{1}La state with respect to the ground state leading to a red shift of the fluorescence \(\lambda_{\text{max}}\).

In the following the tripeptide is assumed to represents a fragment of the interior of a protein and that is not in contact with the solvent. Therefore the calculations are carried out using isolated models (i.e. no solvent molecules are considered). As displayed in Figure 4.8, we have investigated one stereoisomer of the switch (the configuration of the \(C'\_2\) quaternary \(\alpha\)-carbon of \textit{Zw} is R and the configuration of the \(C\_2\) carbon is S) allowing for a shorter
distance between the Zw and Trp residues. We consider two specifically designed conformers (Conf1-Ala-Zw-Trp and Conf2-Ala-Zw-Trp) of the stereoisomer both featuring stacked and hydrogen bonded Trp and Zw side-chains (the Z forms of Figure 4.9). This situation is reminiscent of the documented Arg-Trp stacking documented for cytokine-binding proteins and where, loosely, the positively charge arginine unit is replaced by the pyrrolinium moiety.

Conf1 features an extended backbone while Conf2 has a β-turn-like backbone. As mentioned above, Conf1-Ala-Zw-Trp and Conf2-Ala-Zw-Trp are assumed to represent different realistic arrangements of the chosen tripeptide fragment in a protein. The effect of the photoisomerization on the tryptophan fluorescence is determined in the following way. The structure of the corresponding fluorescent conformers (i.e. featuring the 3-MI moiety in its 1La state) is optimized at the CASSCF/6-31G*/AMBER QM/MM level. Accordingly the QM part corresponds to the 3-MI side-chain of tryptophan while the rest of the molecule is described at the MM level (see Method section for details). The central residue corresponding to the unnatural amino acid Zw and the Ala residue are described at the MM level with a GAFF potential and features parametrized RESP charges (computed at the HF/6-31G* level) for both the Z and E forms.

The energy gaps between the fluorescent state and the ground state are computed via CASPT2//CASSCF/6-31G*/AMBER single point computations and correspond to the S0-S2 transition (the 1La state features the largest oscillator strength with respect to the ground state). In Table 4.4 we show that, according to our QM/MM models, the tryptophan fluorescence can be modulated via Z → E photoisomerization. In particular the Conf1-Ala-Zw-Trp and Conf2-Ala-Zw-Trp emissions are significantly blue-shifted (28 nm and 25 mm respectively) upon photoisomerization. Notice that these effects correspond to an increase in the S0-1La energy gap of more than 8 kcal mol⁻¹. It is also remarkable to find out that the emission maximum of the Z and E forms are predicted to be respectively strongly red-shifted and blue-shifted with respect to the isolated 3-MI.

In order to demonstrate that the above changes are dominated by the different electro-
Table 4.4: Computed fluorescence maxima ($\lambda_{\text{max}}^f$) for the Conf1 and Conf2 conformers of the Ala-Zw-Trp tripeptide and for the gas-phase fluorophore 3-MI (from ref. 26). The active space used in the calculation comprises the full $\pi$-system of the 3-MI fluorophore. The oscillator strength ($f$) and the change in the dipole moment ($\Delta \mu$) upon the vertical electronic transition are also given.

<table>
<thead>
<tr>
<th>$\Delta E$ (kcal mol$^{-1}$)</th>
<th>$\lambda_{\text{max}}^f$ (nm)</th>
<th>$f$</th>
<th>$\Delta \mu$ (Debyes)</th>
</tr>
</thead>
<tbody>
<tr>
<td>97.1</td>
<td>295</td>
<td>0.11</td>
<td>5.9</td>
</tr>
<tr>
<td>90.0 (96.4)</td>
<td>318</td>
<td>0.10</td>
<td>5.2</td>
</tr>
<tr>
<td>98.5 (97.3)</td>
<td>290</td>
<td>0.11</td>
<td>4.9</td>
</tr>
<tr>
<td>91.9 (96.5)</td>
<td>311</td>
<td>0.11</td>
<td>5.2</td>
</tr>
<tr>
<td>100.0 (98.0)</td>
<td>286</td>
<td>0.12</td>
<td>4.8</td>
</tr>
</tbody>
</table>

static field created by the change in orientation of the Zw-head moiety, we have repeated the energy gap calculation using the same 1La equilibrium peptide structures in the absence of the MM point charges of the models, which are responsible for the electrostatic field acting on the fluorophore. The results of these calculations (see values in parenthesis in Table 4.4) fully confirm that the energy gap change is due to a change in electrostatics. Indeed the resulting energy gaps are all within 1 kcal mol$^{-1}$ from the corresponding isolated 3-MI value.

The results above can be readily explained considering the typical GS and 1La charge distributions seen in Figure 4.8b as well as the geometry of the peptide conformers seen in Figure 4.9. In fact, both Z-isomers feature a hydrogen bond between the ionized -COO$^-$ group of the switch, placed directly above the 3-MI unit and the indole N-H bond. This situation would strongly stabilize the excited state with respect to the ground state due to the short distance between the -COO$^-$ negative charge to the pyrrole moiety positive charge of 1La 3-MI. The removal of this interaction in both E-isomers must therefore lead to a blue-shift of the absorption due to a less stabilized 1La state. The analysis of the E and Z conformers also provide information on the possible structural effects induced upon photoisomerization and can be due to both mechanical (i.e. bond breaking) and electrostatic changes. For instance, in Conf1-Ala-Zw-Trp one the photoisomerization breaks the -COO$^-$···H-N hydrogen bond which also results in the formation of an intra-backbone hydrogen bond stabilizing the conformer. Interestingly the breaking of the -COO$^-$···H-N hydrogen bond in
Conf2-Ala-$\text{Zw}$-Trp leads, in our model, to a backbone rearrangement involving the breaking of the hydrogen bond A and the formation of the new hydrogen bond B of Figure 4.9.

**CONCLUSIONS**

The photoinduced dipole moment change of spiropyran merocyanine electrostatic photoswitches has been measured for the case of indolino-spyrobenzopyranes. The results indicate a maximum dipole moment change of ca. 17 Debye due to the low dipole moment (ca. 3 Debye) of the spyropyrane closed-form with respect to the merocyanine open-form (ca. 15 Debye). In a situation where the indanylidene ring is rigidly oriented, the presented NAIP Schiff base 3 provides an electrostatic photoswitch displaying a nearly two-fold dipole change. This is due to the spatial inversion (re-orientation) of a large (ca. 15 Debye) permanent dipole moment (see Figure 4.1).

The ca. 30 Debye dipole moment change of 3 opens up a new perspective for the light-driven conformational control of macromolecular structures determined by polar interactions. In a situation where the indanylidene ring is held in a fixed orientation (e.g. when part of a protein backbone or supramolecular scaffold), light can be used to invert the dipole, yielding a dramatic change in the local electrostatic field and, ultimately, to a conformational change thus leading to a destabilization of the original equilibrium conformation. The construction of QM/MM models of the Z and E forms of the Ala-$\text{Zw}$-Trp tripeptide strongly supports this hypothesis. In fact, we have provided computational evidence that the Ala-$\text{Zw}$-Trp photoisomerization leads to easily detectable (< 20 nm) changes in fluorescence maximum thus providing an excellent basis for future experimental work. Furthermore, given the presence of the charged carboxylate and imminium centers in the rotating head (Zw-head in Figure 6a), the same photoisomerization could be used to break and reconstitute hydrogen bonds and salt-bridges in suitably designed peptides. However, as mentioned above, to implement these concepts one needs further functionalization of 3 at the level of the indanylidene ring.
(introducing, for instance, -NH$_2$ and -COOH groups) to graft the molecule to the macro-molecule backbone. Synthesis work in this direction is currently being carried out in our laboratories. A different perspective may be opened by the preparation of enantiomerically pure 1 (e.g. via chiral resolution of the available racemate). Similar to the well-known chiral diarylenes, the R or S enantiomers of 3 display molecular helicity. According to our computation, the lowest energy solution conformers of Z-3 and E-3 have axial carboxylates and pre-twisted double bonds in the same direction (clockwise). This establishes that the R(S) configuration at C$_2$ induces an M,M (P,P) helicity (see Figures 4.4 for the case of the R enantiomer) in the lower energy diastereomeric conformer of both the Z and E forms. This leads to the possibility of a preferential direction of isomerization and, ultimately, to a two-photon induced unidirectional rotary motion of the zwitterionic head with respect to the indanylidene unit. Indeed the IRC of Figure 4.4b displays a clockwise rotation of the S$_0$-Z-3 pyrrolinium head that would continue upon relaxation to the S$_0$-Z-3 intermediate. Future computational work will explore this perspective.
Figure 4.4: a). $S_1$ (diamonds) and $S_0$ (squares) relative CASPT2/CASSCF/AMBER energies for the $S_0$, $S_1$ and CI points of 3 (full symbols) characterizing its $Z \rightarrow E \rightarrow Z$ photocycle in methanol solution. The energies of the corresponding NAIP-1 points (open symbols)$^1$ are also given for comparison. The structures (hydrogens not shown) provide a perspective showing the helicity of the molecular framework. The values in degrees refer to the $C_9'-C_1'-C_4-C_5$ torsional angle. Bond lengths are given in Å. Notice that the difference between $S_0$-E-1 and $S_0$-Z-1 (and the corresponding NAIP-1 isomers) is calculated using averaged solvent electrostatic potential/molecular dynamics computations. Notice also that, due to the C$_2$ stereogenic center, the ca. 90° twisted CI-Z-3 and the CI-E-3 structures are distinct distereoisomers that relax along stereochemically distinct $S_0$ paths connecting CI-Z-3 to $S_0$-E-3 and CI-E-3 to $S_0$-Z-3. The ca. 40 kcal mol$^{-1}$ gap between the $S_0$-E-3 and $S_0$-Z-3 equilibrium structures and the corresponding CI-Z-3 and the CI-E-3 is not related to the thermal isomerization barrier that is smaller and requires the knowledge of the corresponding solvated transition structure to be evaluated. b). $S_1$ energy profile (diamonds) along the computed $S_1$ IRC path for the $Z \rightarrow E$ photoisomerization of $S_0$-Z-3. It is shown that the $S_0$-Z-3, $S_1$-Z-3 and the CI-Z-3 are connected by a barrierless reaction coordinate. The IRC value is given in a.u. (amu$^{1/2}$/bohr). The $S_0$ energy profile (squares) along the computed $S_1$ coordinate and the M,M helicity of the computed (R)-$S_0$-Z-3 conformer is also shown. In the formula at the top R indicates the absolute configuration of the C$_2$ center and $A_x$ the axial position of the -COO$^-$ substituent.
Figure 4.5: Room temperature experimental absorption spectra of the Z isomer (blue line) and E isomer (red line) of 3 in MeOH.
Figure 4.6: a) Scaled-CASSCF/AMBER trajectory simulating the S\textsubscript{1} isomerization of 3 in methanol solution. The system is predicted to reach the conical intersection region on a 300 fs timescale. The open triangles indicate single-point CASPT2//CASSCF/AMBER energies demonstrating that the scaled-CASSCF energy profile is only slightly off the CASPT2 energy. The CASSCF/AMBER value of the oscillator strength f along the trajectory is also given (full circles). b) Evolution of the two main geometrical parameters (C\textsubscript{1′}-C\textsubscript{4} length and C\textsubscript{9′}-C\textsubscript{1′}-C\textsubscript{4}-C\textsubscript{5} dihedral) characterizing the S\textsubscript{1} structure of the switch.
Figure 4.7: Conformational changes occurring during the excited state trajectory of Z-3 in methanol and allowing a space-saving isomerization motion.
Figure 4.8: a) Structure of the designed Ala-Zw-Trp tripeptide. The incorporated unnatural amino acid Zw is framed. The position of the 3-MI fluorophore treated at the QM level with respect to the peptide backbone and to the rigidly oriented zwitterionic side-chain of the Zw residue (tube representation) of Conf1-Ala-Zw-Trp. b) The ground state (GS) and fluorescent state (\(^{1}L_a\)) charged distribution of 3-MI from ref.\(^{26}\)
Figure 4.9: Structure of the designed Ala-Zw-Trp tripeptide in its fluorescent state (i.e. at the $^1$La equilibrium geometry). The incorporated unnatural amino acid Zw is in its R configuration (see straight arrow) while the COO$^-$ group resides on an asymmetric carbon with S configuration. The curly arrows indicate the position of the isomerising double bond. The hydrogen bonds present in the structures are indicated with dashed lines. The incipient or broken hydrogen bonds are indicated with dotted lines. The distance values are given in Å.
Bibliography


Chapter 5

ISOMERIZATION OF RHODOPSIN
AND BATHORHODOPSIN
ABSTRACT

Rhodopsin (Rh) and bathorhodopsin (bathoRh) quantum-mechanics/molecular-mechanics models based on an ab initio multiconfigurational wavefunctions are employed to look at the light induced $\pi$-bond breaking and reconstitution occurring during the Rh→bathoRh and bathoRh→Rh isomerizations. More specifically, semiclassical trajectory computations are used to compare the excited ($S_1$) and ground ($S_0$) state dynamics characterizing the opposite steps of the Rh/bathoRh photochromic cycle during the first 200 fs following photoexcitation. We show that the information contained in these data provide an unprecedented insight into the sub-picosecond $\pi$-bond reconstitution process at the basis of the reactivity of the protein embedded 11-cis and all-trans retinal chromophores. More specifically, the data point to the phase/amplitude of the skeletal bond length alternation stretching mode as the key factor switching the chromophore to a bonding state. It is also found/confirmed that the phase of the hydrogen-out-of-plane mode controls the stereochemical outcome of the forward and reverse photoisomerizations.

\footnote{The chapter based on the article – Igor Shapiro, Mikhail N Ryazantsev, Luis Manuel Frutos, Nicolas Ferre, Massimo Olivucci ”The Ultrafast Photoisomerizations of Rhodopsin and Bathorhodopsin are Modulated by Bond Alternation and HOOP driven Electronic Effects” (to be submitted)}
INTRODUCTION

Rhodopsin (Rh) is the G-protein-coupled photoreceptor responsible for twilight vision in vertebrates.\textsuperscript{1,2,3} It comprises an opsin apoprotein and the 11-cis retinal protonated Schiff base (PSB11) chromophore covalently linked to the opsin core.\textsuperscript{1} The photoexcitation of Rh results in the sub-picosecond isomerization of PSB11 to its all-trans isomer (PSBT) that, following picosecond vibrational relaxation, leads to production of the metastable intermediate bathorhodopsin (bathoRh).\textsuperscript{1,2,3} Low temperature (77 K) irradiation of Rh with 580 nm light establishes, within one second, a photostationary state characterized by a Rh/bathoRh ratio of 61:1 thus showing that bathoRh can be photochemical reconverted to Rh. The efficiency of the Rh/bathoRh interconversion is consistent with the fact that these species have close molar extinction coefficients (4.06 \cdot 10^4 M^{-1} cm^{-1} at 498 nm for Rh and 3.52 \cdot 10^4 M^{-1} cm^{-1} at 535 nm for bathoRh)\textsuperscript{4,5} and that the quantum yields for the Rh $\rightarrow$ bathoRh and bathoRh $\rightarrow$ Rh photoreactions are 0.67 and 0.49 respectively.\textsuperscript{5} Notice that, while bathoRh is stable at cryogenic temperatures, at room temperature it relaxes (via a series of intermediates), to metarhodopsin II that, in turn, triggers phototransduction by activating the G-protein transducin.\textsuperscript{2,3} Therefore the photoisomerization of bathoRh competes with phototransduction. Intensive illumination of Rh also leads to the formation of isorhodopsin,\textsuperscript{4,5} a Rh isomer containing the 9-cis retinal chromophore. However, this reaction is significantly less efficient than the Rh/bathoRh interconversion.\textsuperscript{6}

Recent computational studies have elucidated the excited state isomerization dynamics of Rh.\textsuperscript{7,8} It has been shown that the chromophore reactive $\pi$-bond completely breaks on a 100 fs timescale via a bicycle-pedal motion that leads to excited state decay. In spite of its central role in phototransduction, we are unaware of any analogue study carried out for bathoRh. A study of the photoisomerization dynamics of bathoRh appears of basic theoretical interest as the Rh $\rightarrow$ bathoRh and Rh $\rightarrow$ bathoRh isomerizations constitute the components of a photochromic cycle/equilibrium implemented at the biological level. In Figure 5.1 we outline the simplest mechanism for a photochromic cycle based on two ultrafast
Figure 5.1: PSB11 chromophore structure and the elementary (simplest) mechanism for ultrafast Rh/bathoRh photochromic interconversion.

(barrierless) $\pi$-bond photoisomerizations. The central mechanistic element of the cycle is a conical intersection featuring a fully broken $\pi$-bond and mediating excited state decay in both direction. At the conical intersection, the E/Z stereochemical signature is lost. From the figure it is apparent that the simulation of the dynamics of the direct and reverse reaction is expected to provide information on the time-dependent structural and electronic features that make an isomerization successful (trajectory 1 and 2) or aborted (trajectory 3). Most important, these simulations may unveil the factors that control the stereochemistry of the $\pi$-bond reconstitution as trajectories 1 and 2 accomplish this process in opposite directions. Below we focus on the efficient photochemical steps of the Rh/bathoRh photochromic cycle to address these issues.

The above target, appears of basic importance in the perspective of manipulating the Rh/bathoRh isomerization efficiency with tailored excitation light-pulses. By focusing on the
related retinal protein bacteriorhodopsin, Prokhorenko, Miller and coworkers\(^9\) have shown that by manipulating the phases and amplitudes of a laser pulse it is possible to enhance or suppress by 20\% the quantum efficiency of the retinal chromophore isomerization. The pulses are apparently able to steer the isomerization through constructive and destructive interference effects involving vibrational modes displaced along the reaction coordinate and within the time scale of coherent chemistry (i.e. prior to intermolecular vibrational energy redistribution). As mentioned in ref.\(^9\) the shaped pulse must alter the mode amplitude and phase relationship imposed by the potential energy surfaces accessed via photoexcitation. The analysis of these relationships appears of paramount importance if an understanding of the achieved quantum efficiency is needed.

In spite of the fact that few computational studies have focused on the photoisomerization dynamics of Rh, we are unaware of any simulation based on models comprising the entire chromophore and opsin components. In the present work, Rh and bathoRh quantum-mechanics/molecular-mechanics (QM/MM) models based on an ab initio multiconfigurational wavefunctions are employed to look at the light induced \(\pi\)-bond breaking and reconstitution occurring during the Rh \(\rightarrow\) bathoRh and bathoRh \(\rightarrow\) Rh isomerizations. More specifically, QM/MM semiclassical trajectory computations are used to compare the excited (\(S_1\)) and ground (\(S_0\)) state dynamics characterizing the opposite steps of the Rh/bathoRh photochromic cycle during the first 200 fs following photoexcitation. We show that the information contained in these data provide an unprecedented insight into the sub-picosecond \(\pi\)-bond reconstitution process at the basis of the reactivity of the protein embedded PSB11 and PSBT chromophores. More specifically, the data point to the phase/amplitude of the skeletal bond alternation stretching mode as the key factor switching the chromophore to a bonding state. It is also found/confirmed that the phase of the hydrogen-out-of-plane mode controls the stereochemical outcome of the direct and reverse photoisomerizations.
MODELS AND METHODS

QM/MM models based on ab initio multiconfigurational second-order perturbation theory have been shown to provide structures and excitation energies consistent with the observed data. For instance, using a CASPT2//CASSCF/6-31G*/AMBER protocol we have shown that a suitably constructed Rh model: (i) yields a PSB11 conformation consistent with experiment; (ii) reproduces the \( \lambda_{\text{max}} \) change for a small set of modified rhodopsins; (iii) features \( S_0 \rightarrow S_1 \) and \( S_0 \rightarrow S_2 \) \( \lambda_{\text{max}} \) values (478 and 327 nm) 3 kcal mol\(^{-1}\) off the experimental values (498 and 340 nm) a computed 14.6 D change in dipole moment that falls within the observed 13-15 D range and an \( S_0 \rightarrow S_1 \) oscillator strength value (0.8) that compares well with the experimental quantity (1.0). The same protocol has also been successfully used to evaluate the \( \lambda_{\text{max}} \) of PSB11 in methanol yielding a kcal mol\(^{-1}\) shift from the experimental value.
Our CASPT2/CASSCF/6-31G*/AMBER protocol is fully described in ref.\textsuperscript{10,11} The method is based on a hydrogen link-atom\textsuperscript{12} and electrostatic embedding\textsuperscript{10} schemes with the frontier placed at the $C_\sigma C_\varepsilon$ bond of the Lys296 side chain.

The active space comprises the full $\pi$-system of PSB11 (12 electrons in 12 $\pi$-type orbitals). The AMBER charges account for $S_0$ polarization effects in a mean-field way.\textsuperscript{13} The same charges are used for the $S_1$ computations with no ad hoc dielectric constant added. The models used in the present work are constructed on the basis of the 2.2 Å resolution 1U19 crystal structure\textsuperscript{14} of Rh. In order to interpret the bathoRh trajectory data and estimate the reaction/relaxation time scales we also compute the bathoRh CASSCF/AMBER $S_1$ minimum energy path (MEP) starting from its Franck-Condon (FC) structure. The MEP was obtained in mass-weighted Cartesian coordinates using the method\textsuperscript{15} implemented in GAUSSIAN03.

The semi-classical trajectories are computed with an extended version of the scaled-CASSCF//AMBER protocol recently employed to investigate the $S_1$ dynamics of Rh.\textsuperscript{7} The scaling factors used for correcting the gradient/time are 0.795 and 0.663 for Rh and bathoRh respectively. Two-root state average $S_1$ and $S_0$ QM/MM force fields are computed using the MOLCAS developer version 7.5 coupled with the Tinker 4.2 molecular mechanics program. The velocity Verlet algorithm and forces obtained from the same QM/MM setup were used to propagate the Newtons equation of motions. Once the $S_1$ and $S_0$ state potentials approach, a surface hopping algorithm was employed to detect non-adiabatic transitions between the states. We use an algorithm based on the change of the state-averaged wavefunctions that has been tested for different systems.\textsuperscript{16,17,18} Briefly: at each time step the $S_1$ configuration state function coefficient vector is compared with the corresponding $S_0$ vector of the previous time step. A sudden increase in their scalar product indicates that the orthogonality is no longer conserved pointing to a region featuring a large non-adiabatic (i.e. derivative) coupling and hop probability. In practice, we simulate the hop by changing the root selected for gradient calculation when a 0.5 product value is detected. The energy differences resulting from
hopping were recovered in terms of kinetic energy by scaling the velocity vector.

In Figure 5.2, we show the superposition of the Rh and bathoRh chromophore framework. The changes are localized in the C₉-C₁₀-C₁₁-C₁₂-C₁₃ segment (see Figure 5.1) of the backbone. In contrast, the position of the β-ionone ring and that of the protonated Schiff base segment remain substantially unchanged (including the Glu113-chromophore framework). This finding can be rationalized by the fact that the ring is hosted in a tight protein pocket (mainly formed by Thr265, Phe212, and Glu122) and that the acidic hydrogen of the Schiff base is hydrogen-bounded to the counterion. This conclusion is supported both by recent time-resolved Raman measurements demonstrating that the C=N bonds of Rh and bathoRh have substantially the same frequency and by the x-ray data. Resonance Raman data indicate that bathoRh has an all-trans structure. Smith et al. concluded from the solid state 13C-NMR data that the C₁₀-C₁₁ bond is twisted in agreement with this structure. Recent investigations refined the 13C-NMR data but did not question the conformation. The circular dichroism spectrum of bathoRh shows a negative peak at 540 nm which is opposite to that of Rh. This was assigned to the change at the C₉-C₁₃ portion of the chromophore with agrees with an all-trans configuration. The available x-ray crystal structure of bathoRh (PDB ID: 2G87) has a 2.8 Å resolution. This limited resolution makes necessary a computational refinement that, for the chromophore is quantum mechanical. Recently, the SCC-DFTB, a self-consistent charge density-functional tight-binding method, has been used to refine the chromophore bathoRh crystal structure. In Figure 5.2 we compare dihedral angles that were reported in ref. and ref. to our bathoRh model. The CASPT2/CASSCF/6-31G*/AMBER energy calculation for bathoRh gives energy of S₀ → S₁ 538 nm consistently with the observed value (535 nm).
RESULTS AND DISCUSSION

In Figure 5.3A (page 146) we report the CASPT2//CASSCF/6-31G*/AMBER energy profile along the $S_1$ MEP of bathoRh. This is a barrierless path connecting the FC point to a $S_1/S_0$ conical intersection (CI$_\text{bathoRh}$). As shown in Figure 5.3B the charge transfer character of the $S_1$ wavefunction (revealed by a 0.4-0.5 a.u. increase in the charge residing on the $\beta$-ionone half of the chromophore) is maintained along the path. The oscillations occurring near CI$_\text{bathoRh}$ correctly describe the rapidly changing wavefunction in this region (see also below). The analysis of the reaction coordinate in Figure 5.3C shows that the initial relaxation (from 0 to 2.6 a.u.) is dominated by a mode describing a backbone bond order inversion (double-bond expansion and single bond contraction). However, notice that this is immediately accompanied by HOOP motion at the reactive C$_{11}$-C$_{12}$ bond (see legend of Figure 5.3 for a definition of HOOP). After the initial relaxation (beyond 2.6 a.u.) the C$_{10}$-C$_{11}$-C$_{12}$-C$_{13}$ dihedral describing the backbone isomerization becomes the driving mode (coupled with changes in the C$_8$-C$_9$-C$_{10}$-C$_{11}$ and C$_{12}$-C$_{13}$-C$_{14}$-C$_{15}$ dihedrals). A similar MEP coordinate was reported for Rh.\textsuperscript{8} However, after the initial stretching relaxation, the Rh $S_1$ energy profile becomes flat and shows a tiny energy barrier. The intercepted conical intersection (CI$_\text{Rh}$) is less twisted. In a previous report we have provided evidence that intersection structures like (CI$_\text{Rh}$) and CI$_\text{bathoRh}$ are part of the same $S_1/S_0$ intersection space.\textsuperscript{8}

Two QM/MM semi-classical trajectories for bathoRh and Rh were calculated starting at the corresponding FC points on the $S_1$ energy surface and propagated until unambiguous product identification. The initial velocities are set to zero. It is assumed that the trajectories represent the evolution of the center of the excited state population generated by a laser pulse (i.e. the vibrational wavepacket) and provides a representation of the average dynamics of the systems. Given the short (less 200 fs) time scale simulated in this work, we kept the opsin backbone and side-chains fixed at their crystal structure that is assumed to provide a representation of the average opsin environment. The chromophore-Lys296 side-chain and
the internal waters are left free to evolve during the simulation for a total of 198 vibrational degrees of freedom.

We have found that both trajectories are reactive and generate the photoproduct within the simulation time. The $S_1$ and $S_0$ energies computed along the bathoRh and Rh trajectories are reported in Figure 5.4A (on the page 147) and 5.4B respectively. For bathoRh the $S_1$-$S_0$ gap is found to decrease of ca. kcal mol$^{-1}$ during the first 15 fs. As previously reported$^{27}$ and consistently with the MEP in Figure 5.3C, the initial relaxation is dominated by bond order inversion along the chromophore backbone. During this time the $C_{10}$-$C_{11}$-$C_{12}$-$C_{13}$ dihedral angle describing the isomerization remains substantially constant (see Figure 5.4C). In contrast, the H-$C_{11}$-$C_{12}$-H dihedral undergoes a significant +60° change indicating a large (-60°) hydrogen-out-of-plane (HOOP) motion that leads to large pyramidalization at $C_{11}$ and $C_{12}$. This is consistent with a prompt acceleration of the hydrogen atoms at $C_{11}$ and $C_{12}$ along a direction that decreases the orbital overlap (and conjugation) across the $C_{11}$-$C_{12}$ bond. This process is coupled with the elongation of the $C_{11}$-$C_{12}$ bond from 1.36 to 1.54 Å that becomes, effectively, a single bond. During the entire trajectory the HOOP and H-$C_{11}$-$C_{12}$-H dihedral motion oscillate out-of-phase (see Figure 5.4C). The amplitude of these oscillations is much larger than the oscillations in the $C_{10}$-$C_{11}$-$C_{12}$-$C_{13}$ dihedral that are in-phase with the HOOP mode.

Before describing the $S_1$ decay and the reconstitution of the broken $C_{11}$-$C_{12}$ π-bond, we discuss the stereo-electronic factors involved in such basic processes. As illustrated in Figure 5.5 on the page 148, the formation of a π-bond along a conjugated chain occurs when two conditions are met: (i) the wavefunction allows for electron pairing (a bonding state) between the two centers of the conjugated moieties forming the π-bond and (ii) the overlap of the fragment π-orbitals spanning the two conjugated moieties forming the π-bond is not zero.

Condition (i) is met when the electronic structure of the chromophore is dominated by a resonance formula displaying the positive charge on the -NH=CH- fragment as found for
the $S_0$ state of bathoRh (see Figure 5.3B) where the ca. 88% of the positive charge is placed on the $-N-C_{15}-C_{14}-C_{13}-C_{12}-$ moiety. The condition is not met when the electronic structure is dominated by a charge transfer configurations. This is the case of the $S_1$ state of the bathoRh chromophore where 66% of the positive charge is placed on the $-C_7-C_8-C_9-C_{10}-C_{11}-$ moiety. As a consequence, in a situation where the $C_{11}-C_{12}$ bond is fully twisted (i.e. as in the reference structure $R$), a change in the bond alternation (BAS) of the $-N-C_{15}-C_{14}-C_{13}-C_{12}-$ moiety can control the electronic structure of the fragment and thus its bonding or antibonding status (see Figure 5.5A). Such status determines if, upon $S_1$ decay, the system will immediately form a double bond or if it will relax to $S_0$ in the form of a transient charge transfer state that will eventually change its electronic structure before a photoproduct is generated.

Condition (ii) controls the stereochemistry of the decay to $S_0$. Indeed, as illustrated in Figure 5.5B, this condition can be met when one of two distinct stereo-electronic requirements is satisfied. In the figure a reference structure (R) has near zero overlap between the fragment $\pi$-orbitals (represented by the p-orbitals at $C_{11}$ and $C_{12}$). Twisting about the $C_{11}-C_{12}$ bond in a clockwise or counterclockwise direction induces $\pi$-bond formation in stereochemically distinct configurations (E or Z) and independently from the stereochemistry of the reactant $\pi$-bond. Such process is at the core of the reaction efficiency in the sense that bond reconstitution leading to the original chromophore stereoisomer (e.g. Z for Rh and E for bathoRh) will result in a decreased quantum yield (and increased internal conversion).

As shown in Figure 5.4A, bathoRh enters the $S_1/S_0$ intersection space ca. 60 fs after photoexcitation when a surface-hop occurs delivering the system to $S_0$. As reflected by the sudden change in the charge distribution along the chromophore backbone (see Figure 5.4A), the hop is accompanied by the switching of the electronic structure from an antibonding to a bonding state. According to the diagram of Figure 5.5A this change can be induced by specific displacements of the BAS mode. Indeed, as shown in Figure 5.4E, at the decay point one has a small and decreasing values for the $N-C_{15}$ and $C_{14}-C_{13}$ bonds and larger and
expanding values for the C_{15}-C_{14} bond. As anticipated above this kind of progression biases the electronic structure towards a region of the S_0 energy surface characterized by radical states for the two moieties of the twisted chromophore and, in turn, leading to facile C_{11}-C_{12} π-bond reconstitution (via recoupling of the radical centers).

An important stereoelectronic event seems to occur during the first 20 fs following S_1 decay. While during this time the C_{10}-C_{11}-C_{12}-C_{13} torsional deformation remains substantially constant at a ca. -85°, the H-C_{11}-C_{12}-H dihedral changes rapidly going from a -90° to -30° value. As a consequence the C_{11} and C_{12} centers becomes highly pyramidal yielding a -90° value for the HOOP mode. In Figure 5.6 (page 149) we report a schematic representation of this process and show how the hybridization of the p-orbitals at C_{11} and C_{12} can control the stereochemistry (Z or E) of the reconstituting -bond. Starting from the reference structure R (already seen in Figure 5.5 and representing the decay point) the decrease of the HOOP value leads to opposite sp^3 hybridizations at C_{11}, C_{12} and, in turn, to an increase in overlap between the p-orbitals at these centers. The negative HOOP value results into a displacement prompting the formation of a π-bond with a 11-cis stereochemistry (in Figure 5.6 lobe a interact with lobe b while lobe a interacts with lobe b). This conclusion is strongly supported by the π-electron density analysis reported in Figure 5.7A (page 150) which indicates that the value of the HOOP mode at the decay point is critical for efficient photoproduct formation.

According to our trajectory, the photolysis of bathoRh leads to an extremely distorted form of the Rh 11-cis π-bond already 20 fs after the decay to S_0 and 80 fs after photoexcitation to S_1. During the following 20 fs the trajectory describes a planarization of the C_{11}-C_{12} bond that is completed on a 100 fs time scale. Notice that, at this point, the PSB11 chromophore does not have the structure of an equilibrated Rh yet. The computed excitation energies during the last 10 fs of the trajectory (between 90 and 100 fs) correspond to an absorption between 540 and 610 nm and thererfore considerably red-shifted with respect to our relaxed Rh model (478 nm).
The 100 fs bathoRh isomerization described above takes place in the confined space of the protein binding pocket and must thus adopt a space-saving reaction coordinate. Consistently with the computed S\(_1\) MEP, the bathoRh trajectory shows that the clockwise twisting around the C\(_{11}\)−C\(_{12}\) reactive bond is accompanied by partial 30°-40° counterclockwise twists of the adjacent C\(_9\)−C\(_{10}\) and C\(_{13}\)−C\(_{14}\) bonds (see Figure 5.4C). These twists are required to accomplish the isomerization without significant displacement of the β-ionone ring and of the Schiff-base moiety bounded to Lys296. We also have found that such motion is substantially reversed with respect to the previously reported\(^7\) bicyclo-pedal motion driving the photoisomerization of Rh (see below).

The energy profiles along the semi-classical trajectories of Rh (Figure 5.4B) has features in common with that of bathoRh (Figure 5.4A). For instance the S\(_1\)−S\(_0\) energy gap falls by more than 50 kcal mol\(^{-1}\) in the first 15 fs. The cause of this evolution is the inversion of the bond length alternation which takes place on the same time frame. In addition, similar to bathoRh, the changes in the H-C\(_{11}\)−C\(_{12}\)−H dihedral (see Figure 5.4D) are much earlier than the changes in the C\(_{10}\)−C\(_{11}\)−C\(_{12}\)−C\(_{13}\) dihedral. However, in Rh the S\(_1\) decay event is achieved on a slower, 100 fs, time-scale. Accordingly, notice that before decay the Rh chromophore performs three complete HOOP oscillations with amplitudes significantly smaller than in bathoRh.

The largest differences between bathoRh and Rh are seen after the decay and during S\(_0\) relaxation. Immediately after the decay Rh undergoes a 20 fs evolution in a region where the S\(_0\) and S\(_1\) states are close and the S\(_0\) energy is dominated by an antibonding wavefunction (see the positive charge evolution in Figure 5.4B) and thus does not display the electronic structure required for π-bond reconstitution. This region of the Rh potential energy surface has been discussed in ref.\(^{28}\) that reports on the shape of the S\(_0\) and S\(_1\) potential energy surfaces in the vicinity of a -90° twisted conical intersection CI (i.e. close to CI\(_{\text{bathoRh}}\) and CI\(_{\text{Rh}}\)). The results are summarized in 5.8A by displaying the CASPT2//CASSCF energy profile along a loop centered on the intersection point.
Before reporting on the evolution of Rh after the decay to $S_0$, we need to revise the main results of that study. It is shown that along a small loop centered on the CI point and lying on the plane defined by the branching plane vectors (that correspond, roughly, to the BAS mode and HOOP mode), one locates two entry ($M^{++}$ and $M^{-+}$) and two exit ($M^+$ and $M^-$) channels. These are defined by the minima of the $S_1$ and $S_0$ energy profiles along the loop (see Figures 5.8A left). The two exit channels are oriented at ca. $180^\circ$ with respect to each other (i.e. in opposite direction with respect to the CI) and describe displacements leading towards Rh and bathoRh respectively. The two entry channels are located almost above the exit channels and support a shape of the $S_1$ and $S_0$ energy surfaces consistent with that reported in Figures 5.8B left. The exit channel ($M^-$) leading to Rh is entered after progression along a $S_1$ entry channel ($M^{++}$) driving bathoRh relaxation and featuring a small $S_1$-$S_0$ energy gap. This is consistent with the terminal part of the MEP energy profiles of Figure 5.3A and of the trajectory energy profiles of Figure 5.4 ($\leq 10$ kcal mol$^{-1}$ $S_1$-$S_0$ gap). It has also been reported that along the loop the wavefunction change twice by passing from a $S_0$-like diradical character to an $S_1$-like charge transfer character (see Figure 5.3A right). The change in the nature of the wavefunction is such that the exit channel driving the system towards Rh is dominated by a bonding (FC-$S_1$-like) character. In contrast, the exit channel pointing to bathoRh is almost in correspondence of a mixed FC-$S_1$-like/FC-$S_0$-like character. The change in wavefunction along the loop defines two $S_0$ potential energy regions surrounding the conical intersection that are characterized by different electronic structures. These regions are schematically visualized on the coordinate plane of Figure 5.8B where the region dominated by the FC-$S_1$-like (antibonding) character marked by a dashed area. The results in Figure 5.8A (i.e in ref.$^8$) appear to be consistent with the energy profiles and changing charge distribution reported in Figure 5.4A and B for bathoRh and Rh respectively. In fact, according to Figure 5.4A, the $S_1$ and $S_0$ energy profiles of bathoRh rapidly separate immediately after the decay to $S_0$ and the system displays a bonding character leading to rapid $C_{11}$-$C_{12}$ $\pi$-bond reconstitution. In contrast, the Rh data (see Figure 5.4B) shows that
after the decay the $S_1$ and $S_0$ energy profiles remain very close and the electronic structure conserves an FC-$S_1$-like character.

In Figure 5.4B we see that during the first 20 fs following the decay of $S_1$ Rh the $C_{10}$-$C_{11}$-$C_{12}$-$C_{13}$ dihedral changes from $-80^\circ$ to $-90^\circ$. However, the analysis of the electron density during this time, shows that the reconstitution of the $C_{11}$-$C_{12}$ double bond has not started yet. As anticipated above, this is explained by the fact that along this part of $S_0$ trajectory/energy surface the FC-$S_1$-like, charge-transfer character of the wavefunction does not allow for $C_{11}$-$C_{12}$ $\pi$-bond reconstitution. Indeed, no $\pi$-electron density has developed along the $C_{11}$-$C_{12}$ bond that remains unbounded (e.g. see the 104 fs structure of Figure 5.7B).

Finally, 130 fs after photoexcitation the system exits the near degeneracy region and within 10 fs displays the typical $S_0$ charge distribution that is then maintained for the rest of the trajectory. This process corresponds to the evolution from the point of decay (open circle at ca. 100 fs) within the dashed region of the coordinate plane of Figure 5.8B to the point in which the dashed region is abandoned (full circle at ca. 130 fs). A vibrationally excited form of bathoRh is achieved in 200 fs and features a chromophore in a distorted all-trans form. It features an $S_1$-$S_0$ energy gap corresponding to a 670 and 710 nm absorption and a $S_2$-$S_0$ between 450 and 530 nm absorption. The $S_1$-$S_0$ absorption is red-shifted with respect to the 538 nm found for our bathoRh model. Therefore, there is the possibility that further cooling occurring after our 200 fs trajectory would generate a ca. 600 nm absorption closer to that of the photoRh intermediate reported in the literature.\textsuperscript{2}

As shown in Figure 5.7B the E or Z stereochemical decision is achieved in the 140 fs time scale when we clearly see a build up of the electron density between the $C_{11}$ and $C_{12}$ centers. At this point in time and in contrast to bathoRh, the phase of the HOOP mode is such that the HOOP is increasing rapidly and reaches positive values (i.e. the absolute value of the H-$C_{11}$-$C_{12}$-H dihedral is larger than the $C_{11}$-$C_{12}$-$C_{13}$-$C_{14}$ dihedral). According to the diagram in Figure 5.6 such a situation prompts a large overlap between lobes b and b. This overlap drives the $\pi$-bond reconstitution that, as a consequence of the positive HOOP value,
displays an all-trans stereochemistry. Notice, that the 20 fs evolution occurring immediately after the decay corresponds to a half HOOP period during which the chromophore maintains an antibonding wavefunction (see Figure 5.5A). This lack of reactivity is maintained long enough to set the HOOP to positive values thus allowing the system to achieve the correct stereochemistry when a full FC-$S_0$-like character is reached.

**CONCLUSIONS**

In a recent work, Weingart has analyzed the results of different sets of gas-phase semiclassical CASSCF trajectories for the minimal retinal chromophore model 3-cis penta-3,5-dieniminum cation. It has been shown that a general behavior holds for all trajectories: when the $S_1$ decay occurs during an increase of the $H-C_3-C_4-H$ dihedral, a 3-trans product is formed while when the decay occurs during a decrease of the $H-C_3-C_4-H$ dihedral the chromophore returns to the original configuration. The author concludes that the HOOP mode has a directing effect on the isomerization efficiency. The data reported above using suitable QM/MM models of the full Rh and bathoRh systems confirm such conclusion and clarify it in terms of the electronic effect controlling the stereochemical decision (i.e. the E/Z branching) taken after the $S_1$ decay. However the same data also demonstrate that HOOP phase control can only be effective after the change of the chromophore bonding status (i.e. of its electronic structure) which is effectively controlled by the BAS mode. As shown for Rh, the change in bonding may not occur upon decay as almost assumed in previous studies. But occurs whenever the BSA mode oscillates towards a ground state bonding pattern.

The bathopRh and Rh semiclassical trajectories discussed above point to a photochromic cycle mechanism completely consistent with the simple mechanistic diagram of Figure 5.2. The potential energy surface schemes of Figure 5.9A and Figure 5.9B provide a basis for the interpretation of the trajectories data. After photoexcitation bathoRh travels along a $S_1$ channel placed ca. 10 kcal mol$^{-1}$ above the $S_0$ energy surface and it is funneled, immediately
after decay at the intersection, along a $S_0$ exit channel lying below the $S_1$ energy surface and featuring a diradical (FC-$S_0$-like) wavefunction. In contrast, the entry channel that drives the $S_1$ relaxation of Rh is well separated from $S_0$. However, the corresponding exit channel has a small $S_1$-$S_0$ gap and is initially dominated by a charge transfer (FC-$S_1$-like) wavefunction. The opposite progressions of bathoRh and Rh suggest that their reaction coordinates describe opposite space-saving deformations. In fact, Figure 5.5C and Figure 5.5D show fully inverted retinal chromophore motion ultimately involving torsional deformations along the C$_{13}$-C$_{14}$, C$_{11}$-C$_{12}$ and C$_9$-C$_{10}$ bonds. As shown by the 59 fs and 104 fs structures in Figure 5.9A and Figure 5.9B this motion ultimately corresponds to a bicycle or crankshaft counterclockwise-clockwise-counterclockwise motion in bathoRh and to a clockwise-counterclockwise-clockwise motion in Rh.

In bathoRh and Rh the decay to $S_0$ and the chemical event corresponding to the change from an antibonding to a bonding character of the wavefunction occur with different time scales and, consequently, in different parts of the trajectory. In bathoRh these two events are simultaneous. As shown in Figure 5.4E, at the hop the length of the N-C$_{15}$ and C$_{14}$-C$_{13}$ bonds are quickly contracting while the C$_{15}$-C$_{14}$ bond length is expanding. This BSA phase prompts a bonding (diradical) character so that 20 fs after the hop the C$_{11}$-C$_{12}$ $\pi$-bond reconstitution has started (see Figure 5.7A). Indeed, the stereochemical event deciding between the evolution towards a Z or E stereoisomer is taken in the same time frame and, as described in Figure 5.6 and proposed by Weingart, is controlled by the phase and sign of the HOOP mode.

The data above indicate that the interplay between the initial value and number of oscillations of the BAS and HOOP mode s may highly impact the quantum yield of the reaction. In Figure 5.4B we show that, for bathoRh, the HOOP mode starts at a maximum (positive phase) and perform one full oscillation before decaying at value near 0°. As a consequence, when bond reconstitution begins (triggered by the change in the wavefunction at the decay point) the HOOP value is rapidly decreasing to -80° prompting formation
of a Z stereoisomer. This mechanism suggests that a HOOP phase inversion (e.g. as a result of a delay or acceleration of half a period) would change the reaction stereochemical outcome. Notice that the stereochemical outcome also makes the difference between ultrafast photoproduct formation and ultrafast internal conversion.

In Rh the evolution of the HOOP mode is such that the motion starts at a minimum and, therefore, it is half a period delayed with respect to bathoRh. The HOOP mode oscillates two times and a half before decay and therefore reaches the $S_0$ state immediately after an oscillation maximum. In this situation one would expect a decrease in the HOOP value that becomes negative leading to internal conversion via an aborted isomerization back to the E stereoisomer. However, the fact that after the decay Rh conserves a FC-$S_1$-like wavefunction for ca. 20 fs allows for half HOOP oscillation to be carried out before the electronic structure starts to become suitable for double bond reconstitution (i.e. before exiting the dashed area in the schemes of Figure 5.8B). This situation is controlled by the phase of the BAS mode. In fact, as apparent from inspection of Figure 5.4F, at the hop the length of the N-C$_{15}$ and C$_{14}$-C$_{13}$ bonds are quickly expanding while the C$_{15}$-C$_{14}$ bond length is contracting. His BSA mode phase prompts an antibonding (charge transfer) character that prevent the C$_{11}$-C$_{12}$ $\pi$-bond reconstitution. However, 20-25 fs after the decay the BSA mode is such to allow for a bonding situation. At the same time the HOOP mode have performed almost three oscillations and when the wavefunction changes and is rapidly growing towards positive values. Ultimately this leads to the formation of the E bond.

In conclusion, it appears that the phase and value of the HOOP mode, taken at the moment of the change of the $S_0$ wavefunction character (controlled by the phase of the BAS mode) from charge transfer to diradical, represents a critical event for the control of the stereoselectivity of the photoisomerization of protonated Schiff bases. The protein environment may exploit this mechanisms to enhance the isomerization efficiency in different ways. For instance it could bias the initial $S_1$ relaxation out of the FC point to either a positive or negative HOOP phase. Alternatively, it could delay the time of the decay to $S_0$
or, as for Rh it may use a more complex strategy by conserving the reacting chromophore in the $S_1$ electronic structure (unfavorable to bond reconstitution) for the time necessary to achieve a favorable phase. Notice, that in principle, also the amplitude of the BAS and HOOP oscillation can affect the ratio between successful and unsuccessful events. A lower or higher amplitude shall change the dynamics at the point of decay and during the initial evolution on the $S_0$ energy surface owing to the fact that in these regions the motion is highly anharmonic.

In order to provide further computational evidence supporting the electronic effects described above, we have constructed a different Rh model (unreactRh) that does not complete the isomerization. The analysis of the trajectory of the reactive Rh model investigated above shows that the $H_{\text{link}}$-C-NH-C_{15}-C_{14}-C_{13}$ chromophore fragment (incorporating the hydrogen link atom) is displaced towards the counterion upon twisting of the reactive C_{11}-C_{12} bond. The unreactRh model features a spatially fixed $H_{\text{link}}$ atom (and a mobile counterion side chain) that restrains the fragment displacement ultimately inducing quantitative changes in the BLA mode phase of the -N-C_{15}-C_{14}-C_{13}-C_{12} fragment. In Figure 5.10 we display the computed 160 fs semiclassical trajectory of unreactRh. Comparison of the Rh and unreactRh trajectory date in Figures 5.4B, 5.4D and 5.10A, 5.10B shows that the energy profiles and twisting of the chromophore backbone on $S_1$ is substantially the same for both systems. In particular, as displayed in Figure 5.4D and 5.10B this is true for the oscillatory motion of the HOOP mode and for the progression of the C_{10}-C_{11}-C_{12}-C_{13} dihedral. However, in contrast to Rh but similar to bathoRh (see Figures 5.4B and 5.4A respectively), unreactRh changes electronic structure immediately after decay to $S_0$ (see Figure 5.10A) and consequently initiates the C_{11}-C_{12} $\pi$-bond reconstitution immediately after the hop and not after a 20 fs delay (see Figure 5.4B). The differences between the unreactive and reactive trajectory of Rh and unreactRh are schematically illustrated in Figure 5.4B (right).

It is apparent that, at the hop, the length of the N-C_{15} and C_{14}-C_{13} bonds of unreactRh have reached a maximum or are contracting while the C_{15}-C_{14} bond length has reached a minimum and it is ready to re-expand (see Figure 5.10C). This situation rapidly (within 5
leads to a bonding (diradical) character of the wavefunction consistently with the change in charge distribution displayed in Figure 5.10A. In Rh the situation is exactly opposite. As displayed in 5.4F the C15-C14 is contracting and the N-C15 and C14-C13 are expanding biasing the electronic structure towards a charge transfer state. The HOOP mode in unreactRh performs slightly more than two and a half oscillations before decay to S0. At this point the C10-C11-C12-C13 dihedral has a ca. -90° value, while the HOOP mode is decreasing and points to negative values. According to the general rule established above such a situation would lead to formation of the Z stereoisomer. Indeed, the trajectory displays a fast change in the C10-C11-C12-C13 dihedral that reverts its direction from counterclockwise to clockwise 120 fs after photoexcitation. In conclusion, the unreactRh trajectory shows that a change in the phase of the HOOP mode during the electronic structure change from antibonding to bonding results into an aborted isomerization. We believe that this observation open interesting perspectives for the control of the isomerization stereochemistry and, ultimately, for the understanding of the factors controlling the reaction quantum yield.
Figure 5.3: The $S_1$ MEP of bathoRh. A. Energy profiles along the MEP. Geometrical structures of the Rh and bathoRh conical intersections ($\text{CI}_{\text{Rh}}$ and $\text{CI}_{\text{bathoRh}}$) along the backbone segment of Fig. 5.2A. The values are given in degrees. B. Change in the value of the total positive charge on the $\beta$-ionone containing fragment (framed) along the MEP. C. Analysis of the reaction coordinate in terms of the $\text{C}_{11}-\text{C}_{12}$ bond length and HOOP values. The HOOP deformation is described as the deviation of the H-$\text{C}_{11}$-$\text{C}_{12}$-H dihedral angle from the $\text{C}_{10}$-$\text{C}_{11}$-$\text{C}_{12}$-$\text{C}_{13}$ value (i.e. subtracting the value of the first dihedral from the value of the second).
Figure 5.4: Scaled-CASSCF/Amber trajectories for bathoRh and Rh. A. The $S_1$ and $S_0$ energy profiles along the trajectory of bathoRh. The change in the positive charge of the $\beta$-ionone containing fragment displays a sudden change in the electronic structure of the chromophore. B. The same data for Rh. C. Evolution of the main twisting angles and HOOP mode for the bathoRh chromophore. D. The same data for Rh. E. The change in the bond length along the -N-C$_{15}$-C$_{14}$-C$_{13}$-C$_{12}$-C$_{11}$ fragment of bathoRh. F. The same data for Rh. In all cases the vertical dashed line indicate the decay (hop) time.
Figure 5.5: Conditions for \( C_{11}-C_{12} \) \( \pi \)-bond reconstruction in Rh and bathoRh. A. Relationship between the bond alternation stretching mode of the N-C\(_{15}\)-C\(_{14}\)-C\(_{13}\)-C\(_{12}\) moiety and the electronic structure of the fully twisted reference structure R. The compression of the N-C\(_{15}\) and C\(_{14}\)-C\(_{13}\) bonds stabilize a diradical structure that correlates with the S\(_0\) state at the chromophore FC point. The expansion of the same two bonds (and compression of the C\(_{15}\)-C\(_{14}\) and C\(_{13}\)-C\(_{12}\) bonds) leads to a closed-shell configuration that correlates with the S\(_1\) state at the FC point. B. Evolution of the overlap between the \( \pi \)-orbitals at C\(_{11}\) and C\(_{12}\) from R. The Newman projection of R shows a situation in which the HOOP is zero (i.e. the C\(_{10}\)-C\(_{11}\)-C\(_{12}\)-C\(_{13}\) and H-C\(_{11}\)-C\(_{12}\)-H dihedral are both ca. -90°) and the C\(_{11}\) and C\(_{12}\) centers are planar \( sp^2 \) hybrids. Clockwise and counterclockwise twisting induces equivalent overlaps prompting the formation of the E (overlap between lobes a and a/b and b) or Z (overlap between lobes b and a/a and b) isomer respectively.
Figure 5.6: Effect of the HOOP (and H-C\textsubscript{11}-C\textsubscript{12}-H dihedral) on the double-bond photoisomerization stereoselectivity.  R refers to a situation in which the HOOP is zero (i.e. the C\textsubscript{10}-C\textsubscript{11}-C\textsubscript{12}-C\textsubscript{13} and H-C\textsubscript{11}-C\textsubscript{12}-H dihedral are both ca. -90°).  At R the overlap of the p-orbitals at C\textsubscript{11} and C\textsubscript{12} centers is close to zero.  Oscillation along the H-C\textsubscript{11}-C\textsubscript{12}-H dihedral (top and bottom arrows) induces overlaps and prompts the formation of the Z or E isomer respectively.  Vibrational deformation towards positive HOOP and negative HOOP values (opening and closing of the H-C\textsubscript{11}-C\textsubscript{12}-H dihedral.  Negative HOOP prompt, exclusively, formation of the E isomer (large b-b and a-a overlap) while positive HOOP prompt formation of the Z isomer (large b-a and b-a overlap).
Figure 5.7: Evolution of the $S_0$ $\pi$-electron density along the semiclassical trajectories of Fig. 5.6 A. bathoRh to Rh. The 59 fs is the detected $S_1$ to $S_0$ decay (hop) point. B. Rh to bathoRh. The 104 fs is the detected $S_1$ to $S_0$ decay point. In both cases the comparison between the top and bottom density plot shows that the bond formation process has began with a stereochemistry leading to PSB11 (Z) or PSBT (E) for bathoRh and Rh respectively.
Figure 5.8: Structure of the conical intersection of Rh. A. $S_0$ (full squares) and $S_1$ (open triangles) energy profiles (left) and charge distributions on the $\beta$-ionone containing half of the chromophore (right) along a small loop centered on a conical intersection and lying on the branching plane. B. Schematic representation of the structure of the $S_0$ and $S_1$ potential energy surfaces surrounding the conical intersection (CI). The circle corresponds to the branching plane loop of part A. The vectors corresponding to the branching plane coordinates are molecular modes dominated by a single-bond/double bond inversion stretching (BAS) mode and to a complex mode containing components of the HOOP and isomerization coordinates. This last mode is assumed to represent the local reaction coordinate. The BAS mode changes the $S_0$ electronic structure effectively. The top and bottom structures on the right represents, in terms of resonance formula, the FC-$S_1$-like (top) and FC-$S_0$-like (bottom) electronic structures dominating the different regions of the area of the $S_0$ potential energy surface defined by the circle.
Figure 5.9: Schematic structure of the $S_1$ and $S_0$ potential energy surfaces in the -90° conical intersection region for the systems under investigation. A. Schematic trajectory for bathoRh. B. Schematic trajectories for Rh (full line) and unreactRh (dashed line, see below). The molecular structures at the bottom correspond to the decay (hop) point in the vicinity of the CI. The corresponding geometrical parameters are given in degrees.
Figure 5.10: Scaled-CASSCF/Amber trajectory for unreactRh and corresponding evolution of the $S_0$ $\pi$-electron density. **A.** The $S_1$ and $S_0$ energy profiles along the trajectory. The positive charge of the $\beta$-ionone containing fragment displays a sudden change in the electronic structure of the chromophore. **B.** Evolution of the main twisting angles and HOOP mode. The arrows point to the HOOP slope at decay and at the point of the isomerization inversion. **C.** The change in the bond length along the $-N-C_{15}-C_{14}-C_{13}-C_{12}-C_{11}$ fragment. The orbital diagrams describe the change in electronic structure completed in 10 fs. The arrows point to the slope of the critical dihedral at decay. In all parts (parts A, B and C) the vertical dashed line indicate the decay (hop) time. **D.** The 108 fs is the detected $S_1$ to $S_0$ decay (hop) point. The comparison between the top and bottom density plot shows that the bond formation process has began already 20 fs after the decay and leading to a PSB11 (Z) stereochemistry.
Bibliography


Chapter 6

CONCLUSIONS
The work that has been done in this thesis aims to contribute to three main research directions:

(i) the investigation of the mechanism of the photo-isomerization of Shiff bases in complex environment such solutions or proteins;

(ii) the design of new effective photo-switches;

(iii) the design of new four-stage molecular rotary motors that are driven by external alternate photo- and thermal stimuli.

(i) In spite of the significant progress that has been achieved in the last decades, the mechanism of the photoizomerization of Rhodopsin chromophore is still a matter of debate. The main question that has to be solved is what is the reason for the high quantum yield of the chromophore in Rhodopsin (~66%) which is not obtained for the chromophore in the solution (< 20%). It is important to stress that the low quantum yield for solvated chromophore is accompanied by non-selective isomerization and a picosecond time scale in the contrary to sub-picosecond process of photoisomerization for the same chromophore in protein. The investigation of the photochemistry of N-alkylated indalydene pyrroline Shiff bases (NAIP-1 and NAIP-3) that is given in chapters 3 and 4 reveals that this molecules exhibit photo-isomerization on sub-picosecond time scale, like in the case of Rhodopsin. However, the photoproduct quantum yield of these compounds remains lower (<35%) than that of rhodopsin. This demonstrates that a high speed of the process is not sufficient for the high quantum yield. Besides that, as it was shown in chapter 3 and 4 the mechanism of the relaxation from FC point to the region of the conical intersection both for NAIP-1 and NAIP-3 is similar to that documented for rhodopsin – the single- double-bonds inversion following by the torsional motion. This result suggests that the main difference in the mechanism of the photoisomerization, that responsible for the difference in the efficiency of the process, might reveals in the CI region.
The investigation of rhodopsin and bathorhodopsin photo-isomerization dynamics, which is given in chapter 5, shed light on this problem. It appears that the phase and value of the hydrogen out of plane (HOOP) mode, taken at the moment of the change of the $S_0$ wave-function character at the conical intersection from charge transfer to diradical, represents a critical event for the control of the stereoselectivity of the photoisomerization of protonated Schiff bases. The protein environment may exploits this mechanism to enhance the isomerization efficiency in different ways. It is important to investigate an analog of the HOOP motion in the case of the photoisomerization of N-alkylated indaldehyde pyrroline Shiff bases that drives the processes in forward or reverse directions. This allows not only to make progress toward understanding the reason for the high quantum yield in the rhodopsin but gives a hint for a design more efficient switches.

(ii) The ideal molecular photo-switch would be characterized by a large photoisomerization yield, a large photochromic contrast (absorption difference of E and Z form), and the ability for optical resettling ($E \rightarrow Z \rightarrow E$). As it is shown in chapter 3 and 4, NAIP-1 and NAIP-3 exhibit selective isomerization at sub-picosecond time scale and the mechanism of the photoisomerization resembles that of rhodopsin. The moderate quantum yield ($<35\%$) of this compounds is an issue to be understood and improved.

NAIP-3, besides that, has an additional remarkable property: here the isomerization leads to a difference in dipole moment as large as $\sim30$ Debye. This opens new potential applications for this switch, transducing light into both mechanical and electrostatic energy, and will be addressed in the future work.

As mentioned in the chapter 1, an important step towards the design of bioactive materials is represented by the development of the ability to change the specific conformation of peptides or proteins. This suggests the future work that focuses on the design, synthesis and computational investigations functionalized derivatives of the biomimetic photoswitchable unit that can be incorporated into bulk materials.

(iii) As shown in chapters 3 and 4, NAIP-based motors may complete a half-rotary cycle
in < 10 ps, i.e., a few orders-of-magnitude faster than the fastest (≈ 6 s for half-cycle) known biarylidene. The replacement of the slow biarylidene helix inversion step with an impulsive reaction shuttling the primary photoproduct population (e.g., E(P) and Z(P) in the cycle) toward the stable conformers (e.g., E(M) and Z(M), respectively) may lead to much faster light-powered rotary motors. However, this can be implemented only in switches where an ultrafast photochemical reaction produces a product with a nonstatistical distribution of the vibrational energy polarized on the reactive (Z → E or E → Z) mode. It is likely that NAIP switches could provide this type of system. In fact, as shown in chapter 3, upon decay from the S1 state of NAIP-1, < 25 kcal mol\(^{-1}\) of kinetic energy are predicted to be located on the isomerization mode. Such momentum, together with the additional torque provided by the following S0 reconstitution of the central double bond, may push part of the population well beyond the initial conformation, vibrational relaxation completing the process within 10 ps. The first step in this direction is the design and synthesis of zwitterionic NAIP-3 that exhibits similar photochemistry to NAIP-1 (as shown in chapter 4) but allows the preparation of enantiomerically pure R or S enantiomer (e.g., via chiral resolution of the available racemate). Similar to the well-known chiral diarylidenes, the R or S enantiomers of NAIP-3 display molecular helicity and can be driven through four possible stages in the rotary cycle. Future work in this direction is required and may lead to a design of an effective molecular motor.