

Alma Mater Studiorum – Università di Bologna

DOTTORATO DI RICERCA

Scienze Chimiche

Ciclo XXI

**Settore/i scientifico disciplinari di afferenza:** Settore Disciplinare: CHIM/02  
Coordinatore: Chiar.mo Prof. Giuliano Longoni

TITOLO TESI

Theoretical insight into the properties of light induced events of photochromic systems and rhodopsin proteins.

Presentata da: Tomasello Gaia

Coordinatore Dottorato

Relatore

Chiar.mo Prof. Giuliano Longoni

Chiar.mo Prof. Giorgio Orlandi

Esame finale anno 2009

# Contents:

## Chapter 1: Introduction

- 1.1. Preface
- 1.2 Inside the photochemical processes
- 1.3 Photochromic compounds: fulgides and stilbenoid systems.
- 1.4 Rhodopsin photochemistry.

References

## Chapter 2: Quantum Mechanics/Molecular Methods

- 2.1 The Schrödinger Equation and the Born-Oppenheimer Approximation
- 2.2 Molecular Orbitals and Valence Bond theories.
- 2.3 Basis Sets: The LCAO Approximation
- 2.4 Hartree-Fock Theory
- 2.5 Configuration Interaction (CI)
- 2.6 MC-SCF: Multiconfigurational Self-Consistent-Field
- 2.7 CASSCF and RASSCF
  - 2.7.1 Selecting the active space
- 2.8 Multiconfigurational Second Order Perturbation Theory Method: CASPT2
- 2.9 MultiState CASPT2
- 2.10 The evaluation of bond order in molecules
- 2.11 QMMM methods

References:

## Chapter 3: Potential Energy Surface

- 3.1 General Concepts
- 3.2 Critical Points Determination
- 3.3 Energy minimization
- 3.4 The use of derivatives in minimizations
- 3.5 Newton-Raphson method
- 3.6 The Steepest Descent and the Quasi-Newton Methods
- 3.7 Vibrational normal modes analysis
- 3.8 Transition states

- 3.9 Methods based on local information
- 3.10 Minimum Energy Path determination
- Intrinsic Reaction coordinate method
- 3.11 Methods based on interpolation between reactant and product
- 3.12 Conical Intersections: no crossing roles
- 3.13 The Physical chemistry of conical intersections
- 3.14 Conical Intersection Topology
- 3.15 Conical Intersection Optimization

#### References

### Chapter 4: Conceptual aspects of photochemistry.

- 4.1 What happens when a molecule absorbs light?
- 4.2 Adiabatic and non-adiabatic processes?
- 4.3 The “Physical Chemistry” of Non-Radiative Decay
- 4.4 Ultrafast processes
- References

### Chapter 5: Fulgide Photochromism.

- 5.1 Introduction
- 5.2 Computational method
- 5.3 Results and discussion
  - 5.3.1 Vertical excitations: the open-ring isomer E
  - 5.3.2 The open→close ring reaction
  - 5.3.3 The close-ring isomer and the open ring reaction
- 5.4 Interpretation of the results
- 5.5 Conclusions

#### References

### Chapter 6: Stilbene photo-physical properties

- 6.1 Introduction
- 6.2 Computational strategies adopted and consequent results
  - 6.2.1 Applying CASSCF//CASPT2 methods: Vertical excitations.
  - 6.2.2 Applying CASSCF//CASPT2 and CC2 methods: trans\* relaxation.
  - 6.2.3 Applying CASSCF//CASPT2 and CC2 methods: twisted region.
  - 6.2.4 Applying CASSCF//CASPT2 and CC2 methods: cis\* relaxation
- 6.3 Applying RASSCF method
- 6.4 Conclusion

#### References

- [Chapter 7: Electrostatic control of the photoisomerization efficiency in Rhodopsin](#)
- [7.1. Introduction](#)
- [7.2. Computational Details](#)
- [7.3 Result and Discussion](#)
  - [7.3.1 The Glu181 issue: neutral or protonated?](#)
  - [7.3.2 The counterion and the other polar/charged residues: counterion quenching by the protein pocket.](#)
- [7.4 Mutations and Vision Deficiencies](#)
  - [7.4.1 Rh related vision deficiencies.](#)
  - [7.4.2 Color pigments related pathologies.](#)
- [7.5. Conclusions](#)

#### References

- [Chapter 8: Steric control of photoisomerization efficiency and spectroscopic properties in Rhodopsin](#)
- [8.1 Introduction](#)
- [8.2 Computational Methods](#)
- [8.3 Results and Discussion](#)
  - [8.3.1 Steric catalysis](#)
  - [8.3.2 The trap S<sub>2</sub> state](#)
- [Transient absorption along the S<sub>1</sub> photoisomerization](#)
- [ConclusionsReferences](#)
- [Chapter 9: Conclusions](#)

#### References

- [Appendix A](#)

# -CHAPTER 1-

## Introduction

### 1.1 Preface

The light-material interaction is one of the most interesting natural phenomena, being responsible of the existence of life on the earth, and therefore originating photo-biological processes of primary importance. The absorption of a photon transfers the system from the ground state to an electronically excited state. The electronic distribution and nuclear configuration of a photochemically activated molecule generally differ substantially from those of a thermally activated molecule so that the excited state is actually an electronic isomer of the corresponding ground-state system. Further the photochemical and photophysical properties are typical of the photo-excited molecule differing from those shown on the ground state, leading the system towards new processes : photochemistry is the science that study and rationalize such processes. The aim of this introduction is to give the basic information about the photochemical concepts common interest to some of the problems. It is necessary at this point to emphasize how there are some molecular properties not really available to investigate and rationalize from an experimental point of view. Computational photochemistry arises from the necessity of associating accurate computational methodologies of simulation to advanced experimental techniques in order to clarify several aspects of photochemistry experimentally inaccessible. The modern strategies of calculations and the even more sophisticated computations techniques is nowadays possible to explore and characterize widely the topology of the potential energy surface of a chromophore excited states, and thus to properly rationalize (explain) its photoreactivity.

## 1.2 Inside the photochemical processes.

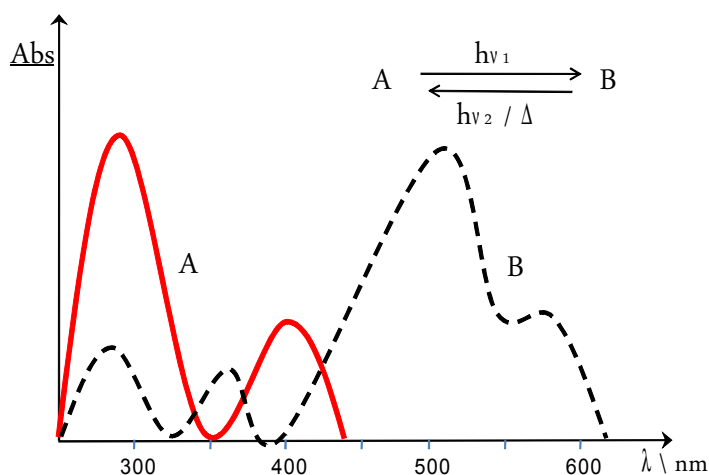
The purpose of this thesis is to investigate, through high-level computations, some of the interesting photo-biochemical processes involve conjugated  $\pi$ -systems, such as those occurring in photochromic systems<sup>2</sup> of retinal chromophores<sup>3</sup>. The goal is to control and maximize photochromic efficiency and to increase our general knowledge of the “primary events” which control the final outcome of the absorption of a photon. The computations make use of a recently implemented novel technique allowing to locate and follow all the steepest descent paths departing from the Frank-Condon point or real crossing point. Once the system is promoted to the excited state decay back to the ground state can occur by fluorescence, or by internal conversion via a radiationless process. For many years the mechanism of decay without emission of radiation was not well understood. It was thought to occur at a minimum on the excited state, in a region of the potential where the ground and excited state surfaces become close at an avoided crossing. Such processes are slow, however, and fail to explain how systems could decay on a femtosecond timescale. More recently, it has been recognized that a true crossing of potential surfaces can take place, creating a new topological feature called a conical intersection (CI)<sup>4</sup>.

The intersection acts like a funnel linking the excited and ground state surfaces, providing a fully-efficient decay path back to the ground state products. There need not be a transition state separating the Frank-Condon region and CI, but if one is present it will control the conditions under which the radiationless route is accessible. The idea that a conical intersection could play a key role in the decay mechanism for excited state reactions was proposed by Teller<sup>5</sup> more than 30 years ago. The literature now contains many theoretical studies of organic systems which find low-lying surface crossings. The photochemical events investigated in this thesis are mainly characterize from an ultrafast deactivation and observing ultra-fast radiationless decay is the main indication of a mechanism involving a CI<sup>6</sup>. Modern experimental techniques have enabled researchers to verify the existence of conical intersections. In particular laser spectroscopy can now be performed at a sub-picosecond level, giving reliable data for excited state lifetimes, decay rates, and energy thresholds for these reactions. Thus experimental evidences are interpreted and rationalized on the basis topological characteristic individuated by calculations. For example, if the radiationless decay is thermally driven, and it is ‘switched-on’ at a certain temperature, this suggests that a transition state lies on the excited state path leading to the intersection. Quantum yield measurements provide information about the product distribution, and hence the possible ground state paths accessible from the CI. Thus computational approach may model the photochemistry of several systems explaining the origin of them photochemical properties.

## 1.3 Photochromic compounds: fulgides and stilbenoid systems.

Photochromism<sup>7</sup> is a reversible transformation of a chemical species induced in one or both directions by absorption of electromagnetic radiation between two forms, A and B, having different absorption spectra (fig.1.1). Although this phenomenon is largely investigated only in the past 30 years, in reality it origins in 1800. More precisely, on 1870 Fritzsche<sup>8</sup> the bleaching of an orange-colored solution of tetracene in the daylight and the regeneration of the color in the dark. Interest in photochromism was continuous

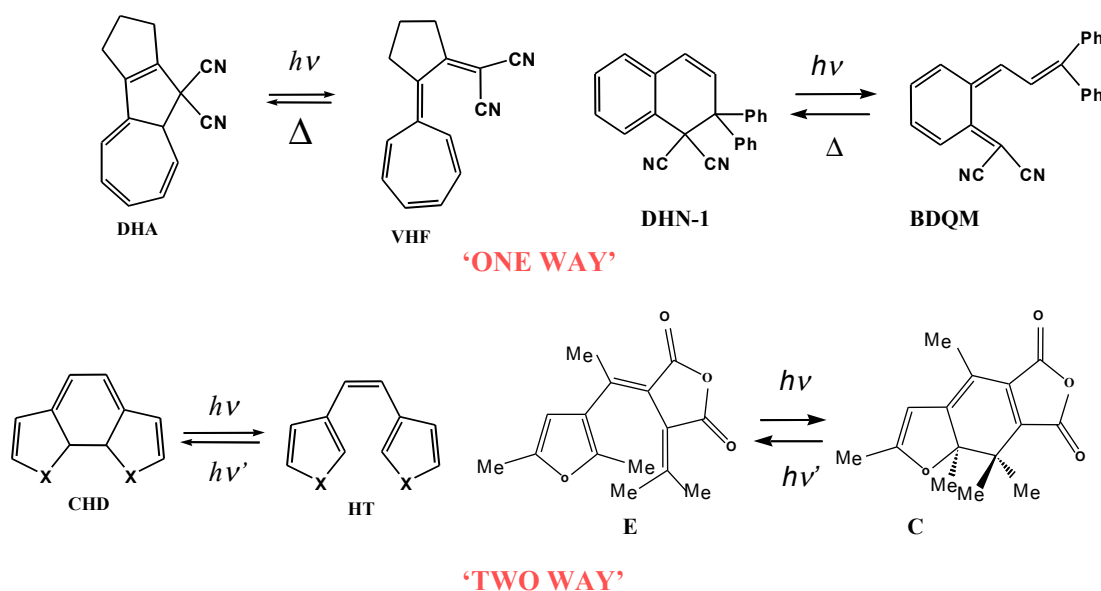
but limited until the 1940–1960 period, which saw an increase of mechanistic and synthetic studies, particularly in the research groups of Hirshberg and Fischer<sup>9</sup> in Israel. In 1950, Hirshberg suggested the term “photochromism”, from the Greek words *phos* (light) and *chroma* (color), to describe the phenomenon. This is the name used today. However, it is not limited to colored compounds; it applies to systems absorbing from the far UV to the IR, and to very rapid or very slow reactions.



**Figure 1.1** Schematic example of a reversible transformation of a single chemical species between two states, the absorption spectra of which are clearly different, the transition in at least one direction being induced by electromagnetic radiation

Photochromism expanded during the 1960s in parallel with the development of physical methods (IR, NMR, X-ray, UV, time-resolved and flash spectroscopy) and organic synthesis. Photochromic glasses became available at that period and further stimulated research. Applications, such as the photochromic micro image (PCMI) process<sup>12</sup>, which showed the possibility of reducing the a huge number of pages to about 6 cm<sup>2</sup>, attracted considerable interest. However, it appeared that the photodegradation of the known families of organic photochromic systems limited their potential for applications<sup>13</sup>. A revival of activity started in the 1980s, essentially because of the development of fatigue-resistant *spirooxazine* and *chromene* derivatives. They triggered the fabrication and commercial application of photochromic ophthalmic lenses. Since then, other commercial systems have been developed, and new photochromic systems have been discovered and explored. Examples of potential applications utilizing the physical or chemical changes that accompany the observed shift of the absorption maxima include optoelectronic system, photoswitching transport through membrane, optical information storage, photochemical switchable enzymatic systems, nonlinear optical devices. The main classes of photochromic compounds, organic compounds and biological receptors are in this introduction briefly described, since the subjects treated in this thesis belong these two classes. Among the organic compounds we have investigated the excellent photochromism of fulgides<sup>14</sup>, and the complex photochromism of stilbene<sup>15</sup>.

At the microscopic level, photochromism originates from a chemical transformation triggered by a photon absorption, such as ring-closure or ring-opening, and the widest and the most important group of photochromic system is that based on pericyclic reactions. Of the many pericyclic reactions, electrocyclizations have proved to be especially suitable as a basis for photochromism. There are mainly two photochromic species, called ‘one way’ or ‘two way’<sup>16</sup>. In this class of compounds, the photochromic transformation arises from different photochemical process (hydrogen shift, electrocyclicization, cis-trans or E/Z isomerization). Therefore if compounds like fulgides or diarylethenes are ‘two way’ photochromic systems, being photochemically converted both to the open and the closed form when irradiated at suitable different wavelengths, compounds like the dihydroazulene couple or dihydronaphthalene<sup>17</sup> couple (fig1.2) are examples of ‘one way’ photochromism since the back reaction can occur only thermally. Of the many pericyclic reactions, electrocyclizations have proved to be especially suitable as a basis for photochromism. The  $4n+2$  systems whose photochromism results from the electrocyclic 1,3,5-hexatriene/1,3-cyclohexadiene interconversion<sup>18</sup> comprise a very large number of species. The fulgides belong to the second mentioned class of photochromic compounds for which the transformation is largely driven photochemically and consequently the thermal reversal or decoloration of the colored species is forbidden (for this reason they are classified as “two way” photochromic systems). For their features and favorable photochemical properties, fulgides are photochromic systems particularly investigated for the potential industrial applications as optical devices, molecular switches, data storage. The discovery of fulgide systems has to be collocated around one century ago and is due to Hans Stobbe<sup>19</sup>, who observed for the first time triphenyl-fulgide changing color from yellow-orange to brown, under light exposition, while turning back to the original color under the dark light.

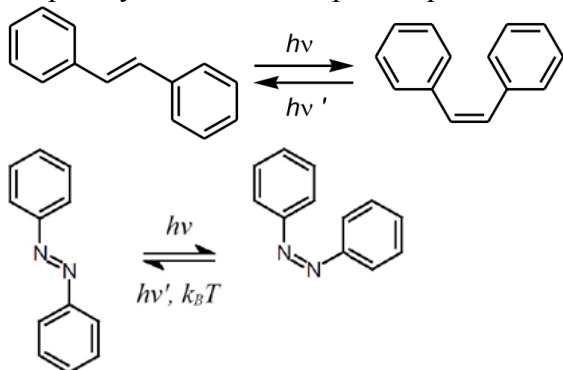


**Figure 1.3** There are reported examples of organic compounds belonging to two main species of photochromic systems. On one hand, dihydroazulene and dihydronaphthalene derivatives belong to “one way” group, where the back reaction can occur only thermally, and thus whose photochromism is



defined also of type T. On the other hand diarylethenes or fulgides belong to the “two way” group, where the back reactions occur photochemically, and whose photochromism is defined also of type P.

Initially Stobbe supposed that the origin of these special properties was correlated with some kind of crystal effect; nevertheless he changed idea while he noticed how such similar properties characterized also similar compounds. The denomination “fulgide”, arising from latin language and meaning *twinkle* has to be attributed to the same Stobbe<sup>20</sup>, the first one observing the particular of some crystal forms. Among the several species of fulgides, featured by the different chemical structure (from which photochromic properties arise), the most interesting species are those called “aromatic fulgides” or “heterocyclic-fulgides. The aromatic fulgides are distinguished depending on the kind of system condensed on the cyclohexadiene-succinic-anidrid moiety, whose photochromism results from the electrocyclic 1,3,5-hexatriene/1,3-cyclohexadiene interconversion: the experimental data are mainly concentrated on the furyl-classes. In particular compounds as indolylfulgide, furylfulgide, because of its high efficiency as potential photochromic systems, can be used as switch controls or aberchrome molecules. Remarkable, photochromic actinometers can be used repeatedly, thus obviating the need for a fresh sample for each measurement. Some thermally stable photochromic compounds, such as the following, fulfill this requirement. The fulgide we have investigated in this thesis, is suitable for being used as chemical actinometer, since the absorption wavelength are within the range of coloration (310–370 nm) or decoloration (435–545 nm) domain. Organic compounds as azobenzene and stilbene may be used in the field of chemical actinometry or photoresponsive materials. Reversible photoinduced changes of physical and chemical properties can be transferred to the micro-environment by a photochromic molecule incorporated in the system. Some examples are given from chiroptical molecular switches, where photoisomerization of photochromic units (azobenzene, spyropirans) in poly ( $\alpha$ -amino acids) is able to trigger a random coil to  $\alpha$ -helix transition. This primary photochemical event occurring in the side chains is amplified and transduced by the structural variations of the macromolecular main chains. The latter are accompanied by large and reversible variations of optical activity<sup>21</sup>. Because of their reversibility, these systems can act as “chiroptical molecular switches”<sup>22</sup> In this thesis stilbene photochemistry, has been extensively studied using accurate computational strategies with the aim of clarify the mechanism behind the radiationless deactivation to the ground state. Stilbenes photochromism is due to cis-trans isomerization, and it represents one of the most interesting challenge in spectroscopy. Stilbene and stilbenoid compounds are indeed frequently textbook examples of photochemical cis-trans isomerization.



**Figure1.3** On the top Stilbene and down azobenzene

Their reactivity has an important role as model compound of biological phototropic systems. They, also serve as building blocks for organic materials, whose properties could be used in optical and electro-optical applications such as optical data storage, laser dyes, nonlinear optics, or photochemically, cross-linked polymers. Stilbene photochemistry has always been particularly difficult to rationalize: when rotate to twisted geometry stilbene arrives at a minimum of potential energy, called the “phantom state”, from which it decays to the ground state through nonadiabatic transitions. Understanding the nature of this “phantom state” may permit to amply the knowledge of photochemical properties of more extended system, as stiff-stilbene, or the so called Feringa molecular motors. Concerning to this, it is necessary, at this stage to briefly described one of the most interesting application in this field. One of the main current scientific challenges is the bottom-up construction of systems that represent nanosize analogues of switches, devices and motors. The control of chirality in a molecular switching system allows inter-conversion between molecules of opposite helicity using different wavelengths of light. Such bistable chiral switches are of potential use in optical data storage and processing at the molecular level. The control of molecular chirality is even more subtle in the case of molecular motor systems. The exquisite control of chirality using light as an energy source has resulted in a controlled, repetitive 360° unidirectional rotation in two generations of molecular motor systems. It is hard to imagine daily the life without motors. Controlled rotary and linear motor is at the heart of the variety of machinery our society depend on. For most individual it goes unnoticed that at any instant numerous molecular motors are also engaged in pour body to control such delicate processes catalyst, transport and muscle movement. The fascinating molecular motors discovered in biological systems such as ATP- ase rotary and muscle linear motors offer a great source of inspiration to design synthetic motor systems in which controlled translational or rotary motion can be accomplished. In designing molecular motors one has to take in consideration that, compared to ordinary motors, where energy input induces motion, in molecular motors input has also to restrain motion. Rotary motion around a single carbon-carbon bond in alkane for instance is extremely fast with an energy barrier of only 20 kJ/mol and there is not control over the directionality of the rotary motion. The directionality may be established with a double bond carbon-carbon, which is broken upon light irradiation leading the system isomerization on excited state following a precise direction. The problem of mutual random motion of molecules has recently emphasized in discussion or ratchet-type molecular motors and nano-machines. In synthesis approaches towards artificial machinery a variety of molecular or supra-molecular systems have been designed in which an external chemical, electrochemical or photochemical stimulus induces a switching process or movement within the molecule or triggers a change in shape or assembly of molecules.

---

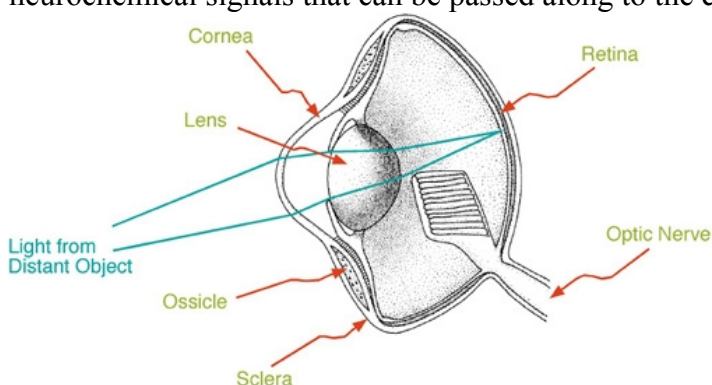
#### [1.4 Rhodopsin photochemistry.](#)

Photochemical processes are not only interesting for technological and industrial applications, but many biological systems are photochromic, with the photo-induced reactions of fundamental interest for understanding biochemical mechanism. Usually the overall reactivity of the protein, is due to the non-proteic moiety of the entire molecule, which is the chromophore, responsible of light absorbing and of the

photochemical process featuring the protein function.

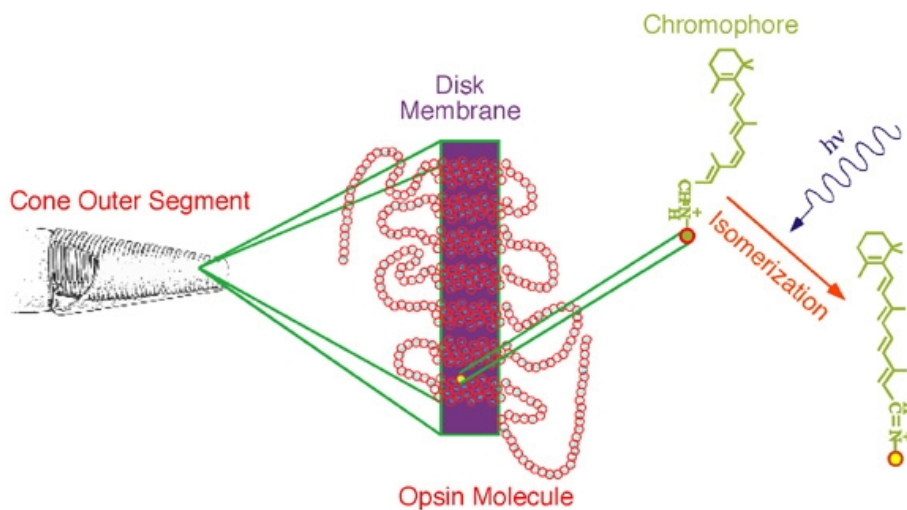
Many intriguing photobiochemical processes involve conjugated  $\pi$ -systems. Examples are the photoisomerization of the retinal chromophores, which trigger the conformational changes underlying the activity of rhodopsin proteins<sup>21</sup>, and the set of photochemical reactions characterizing the bio-genesis of vitamin D<sup>22</sup>. The theoretical chemist who wants to perform an accurate computational study of these reactions, would face the unovercoming problem due to the size of his target. Then, to apply accurate computational methods, it is necessary to reduce the dimension of the molecule, by selecting a model-system with the same chemical reactivity but a smaller size. It is usually a good approximation to limit our study only on the polyenic moiety of the system, that in the case of rhodopsin proteins is retinal chromophore, covalently linked to the protein via a lysine through a protonated Schiff base. Among the several proteins, retinal chromophore belong mainly to two different biological photo-receptors: bacteriorhodopsin, a light-activated proton pump that supports phototrophic growth in the oxygen-deficient conditions and rhodopsin, of crucial interest in biological photochemistry, since the reactivity is at basis of mechanism of vision process.

The production of a neural image of the light converging upon an animal begins with the focusing of that light onto a sheet of receptors. Figure depicts a generalized bird eye. In terrestrial animals most of the focusing power of the eye is the air/cornea interface. Light impinging upon the eye from particular directions is redirected to particular locations on the retina, the tissue lining the inside of the eye. Within the retina are photoreceptors, cells specialized for transducing relative light intensities into neurochemical signals that can be passed along to the central nervous system.



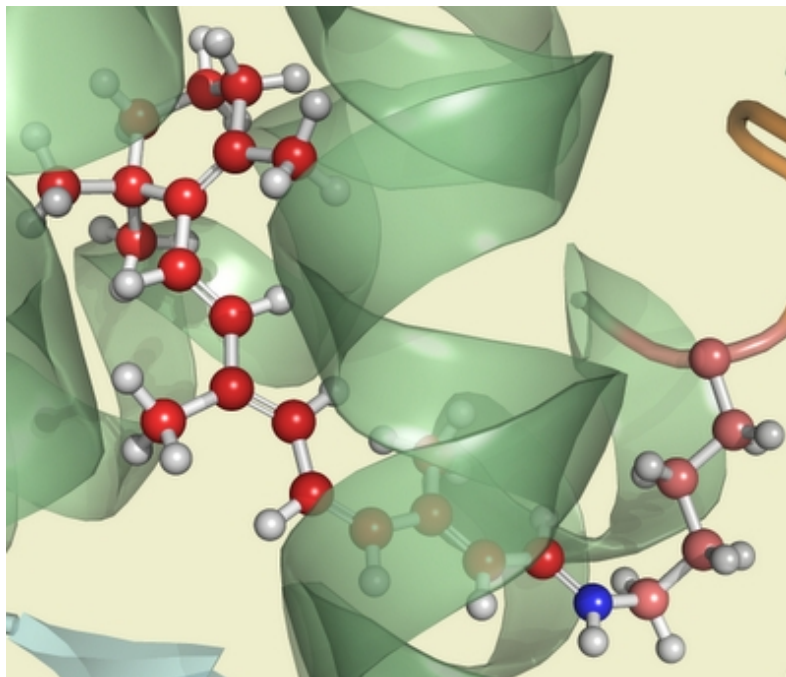
**Figure 1.4** A vertebrate eye. Although there is considerable diversity among the eyes of various vertebrates, they all share some common features. In particular they have optical elements such as the cornea and lens which focus light in order to form an image on the retina.

At the basis of phototransduction mechanism there are the photoreceptors outer segment, where the light is absorbed and transduced in vertebrates. Embedded in the disk membranes are large numbers of photopigment molecules. Each photopigment consists of a chromophore covalently bound to a protein called an opsin. Phototransduction processes described here are the same in all animals with eyes. Opsin does not absorb visible light, but when it is bonded with 11-cis-retinal to form rhodopsin, which has a very broad absorption band in the visible region of the spectrum. The peak of the absorption is around 500 nm, which matches the output of the sun closely. Upon absorption of a photon of light in the visible range, cis-retinal can isomerize to all-trans-retinal. The shape of the molecule changes as a result of this isomerization.



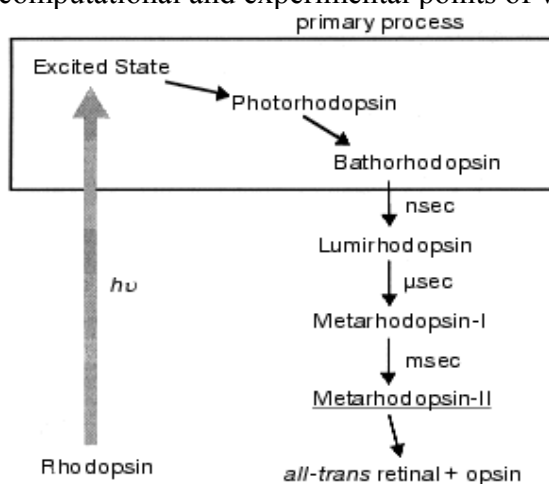
**Figure 1.5** There is represented a scheme summarizing the phototransduction process. The diagram in the center schematically depicts how a photopigment is embedded within a disk membrane of a photoreceptor outer segment. The filled yellow circle is a lysine to which is attached the chromophore retinal. The absorption of a photon may cause the chromophore to convert from the 11-cis form to the all trans form as shown by the chemical reaction. This conformational change in the chromophore causes a change in the shape of the opsin molecule. The opsin's change in shape converts it from an inactive to an activated enzyme, and thus is light absorption converted to a biochemical signal in a photoreceptor.

Rhodopsin is the only photoreceptor protein (a visual pigment) in the outer segment of rod visual cell responsible for twilight vision. It has been extensively studied more than any other visual pigments such as cone pigments responsible for color vision because of relative ease of preparation and abundance. Rhodopsin has 11-cis retinal as its chromophore, which is embedded inside a single peptide transmembrane protein called opsin. The role of rhodopsin in the signal transduction cascade of vision is to activate transducin, a heterotrimeric G-protein, upon absorption of light<sup>23</sup>. Therefore, a central question in rhodopsin is how light energy is used to change protein structure through photochemical reaction of the retinal chromophore. Rhodopsin (opsin), a member of G-protein coupled receptor family, is composed of 7-transmembrane helices. The 11-cis retinal forms the Schiff base linkage (see fig.1.6) with a lysine residue of the 7th helix (Lys296 in the case of bovine rhodopsin), and the Schiff base is protonated, which is stabilized by a negatively charged carboxylate (Glu113 in the case of bovine rhodopsin)<sup>24</sup>. The beta-ionone ring of the retinal is coupled with hydrophobic region of opsin through hydrophobic interactions<sup>26</sup>. Thus, the retinal chromophore is fixed by three kinds of chemical bonds in the retinal binding pocket of rhodopsin. Absorption of a photon by the chromophore induces a primary photoreaction, followed by conformational changes of protein, and eventually activates transducin. This is called the “bleaching process” because rhodopsin loses its color. Several intermediate states are formed during the bleaching process, which have been identified by visible spectroscopy (fig. 1.7).



**Figure.1.6** The chromophore retinal covalently bound to lysine 296 through a protonated Schiff base.

Among the bleaching processes of rhodopsin, we define the “primary process” as the event of formation of bathorhodopsin (fig. 1.7). Bathorhodopsin is formed and stable in the picosecond time domain. It is also stable at low temperature ( $<-140^{\circ}\text{C}$ ). Unlike bathorhodopsin, photorhodopsin cannot be stabilized at low temperature, as well as the excited state of rhodopsin. There has been a fundamental question on the nature of the primary photoprocess in rhodopsin: What is the primary reaction in vision? To address this question, the primary process of rhodopsin has been extensively studied from both computational and experimental points of view.



**Figure 1.7** Photobleaching process of bovine rhodopsin. The primary process is shown in the frame, which involves excited state of rhodopsin, photorhodopsin, and bathorhodopsin. Metarhodopsin-II that activates transducin is underlined.

Recently, bovine rhodopsin was crystallized by Okada et al.<sup>27</sup>, and its three-dimensional structure was determined with 2.8 Å resolution. These accomplishments promise better understanding of the primary reaction mechanism in rhodopsin, and existing data by various techniques have to be reexamined on the basis of the structural background. Because of the limited space, we would like to make brief comments in the following. Extensive studies by means of ultrafast spectroscopy of rhodopsin have provided an answer to the question relative to the primary event mechanism. Femtosecond spectroscopy of rhodopsin eventually captured the excited state of rhodopsin, and as the consequence, we know that unique photochemistry takes place in our eyes. Namely, the cis-trans isomerization in vision is a coherent reaction through the barrierless potential surface, in which formation of photorhodopsin occurs in 200 fs. Such amazingly fast reaction is facilitated in the protein environment, and vibrational analysis of primary intermediates, such as resonance Raman and infrared spectroscopies, have provided insight into the isomerization mechanism from structural background. The atomic structure of bovine rhodopsin has further encouraged detailed understanding of the mechanism.

Remarkable progress has taken place during the past 10 years, and this progress is shown in this article. What will be the progress in the next 10 years? Regarding the dynamics, we still do not know the origin of the nonreactive pathways. On the other hand, we have started to study the role of the protein environment during the retinal isomerization from the structural background. Further efforts by means of spectroscopy, diffraction, and theory will provide us with new insights into the structure and dynamics of photoisomerization in rhodopsin.

### References:

- 1) Nicholas J. Turro 'Modern Molecular Photochemistry', Yoshizawa, T.; Kuwata, O. In *CRC Handbook of Organic Photochemistry and Photobiology*;
- 2) Irie, M. *Chem. Rev.* 2000, *100*, 1685. Bouas-Laurent, H. Bouas-Laurent, H. ; Durr, H. *Pure Appl. Chem.* 2001, *73*, 639
- 3) a) Kandori, H.; Shichida, Y.; Yoshizawa, T. *Biochemistry-Moscow* 2001, *66*, 1197. Needleman, R. b) Bacteriorhodopsin and Rhodopsin. In *CRC Handbook of Organic Photochemistry and Photobiology*; ed.; Horspool, W. M., Song, P.-S., Eds.; CRC Press: Boca Raton, FL, 1995; Vol. pp 1508-1515.
- 4) Herzberg, G. & Longuet-Higgins, H.C. *Trans. Faraday Soc.*, 1936, *35*, 77. Bernardi, F.; Olivucci, M. & Robb, M.A. *Chem. Soc. Rev.*, 1996, *25*, 321-328.
- 5) Teller, E. *J. Phys. Chem.*, 1937, *41*, 109.
- 6) Zewail, A. H. *Femtochemistry: Ultrafast Dynamics of the Chemical Bond*; World Scientific: Singapore 1994; Vols. I and II.
- 7) Irie, M. *Chem. Rev.* 2000, *100*, 1685. Bouas-Laurent, H. Bouas-Laurent, H. ; Durr, H. *Pure Appl. Chem.* 2001, *73*, 639
- 8) Fritzsche, J. *Compt. Rend. Acad. Sci.*, Paris 1867, *69*, 1035.
- 9) Fisher, E. Hirshberg, Y. *J. Chem. Soc.* 1952, 4522.
- 10) Hirshberg, Y. *Compt. Rend. Acad. Sci. Paris* 1950, *231*, 903.

- 11) 1. Yoshizawa, T.; Kuwata, O. In *CRC Handbook of Organic Photochemistry and Photobiology*; Horspool, W. M., Song, P.-S., Eds.; CRC Press: Boca Raton, FL, 1995; pp 1493-1499.
- 12) 6. Photochromism, G. H. Brown (Ed.), Wiley-Intersciences, New York 1971.
- 13) 7. R. C. Bertelson. *Mol. Cryst. Liq. Cryst.* 246, 1 (1994)
- 14) Heller, H. G.; Szewczyk, M. *J. J. Chem. Soc. Perkin Trans.* 1 1974, 12, 1487.
- 15) (1).a) Syage, J.A.; Felker, P.M.; Zewail, A.H. *Chem Phys* 1981, 11, 4685; b) Syage, J.A.; Lambert, W.R.; Felker, P.M.; Zewail, A.H.; Hochstrasser, R.M. *Chem. Phys. Lett.* 1982, 88, 266. c) Felker, P.M.; Zewail, J. *Phys. Chem.*, 1985, 89, 5402. d) Orlandi, G.; Siebrand, W. *Chem. Phys. Lett.* 1975 30, 352
- 16) 1. Irie, M. *Chem. Rev.* 2000, 100, 1685. Bouas-Laurent, H.; Bouas-Laurent, H.; Durr, H. *Pure Appl. Chem.* 2001, 73, 639
- 17) Gaia Tomasello, Francois Ogliaro, Michael J. Bearpark, Michael A. Robb, and Marco Garavelli *J. Phys. Chem. A*; 2008; 112(41) pp 10096 - 10107;
- 18) Celani, P.; Bernardi, F.; Robb, M. A.; Olivucci, M. *J. Phys. Chem.* 1996, 100, 19364-19366.
- 19) Stobbe, H. *Die Fulgide*, *Annalen*, 380, 1, 1911.
- 20) Stobbe, H. *Phototropiererscheinungen bei Fulgiden und Anderen Stoffen*, *Annalen*, 359, 1, 1908.
- 21) A. Fissi, O. Pieroni, G. Ruggeri, F. Ciardelli. *Macromol.* 28, 302 (1995) and references therein. B. L. Feringa, W. F. Jager, B. de Lange. *J. Chem. Soc., Chem. Commun.* 288 (1993).
- 22) Yoshizawa, T.; Kuwata, O. In *CRC Handbook of Organic Photochemistry and Photobiology*; Horspool, W. M., Song, P.-S., Eds.; CRC Press: Boca Raton, FL, 1995; pp 1493-1499
- 23) (a) Jacobs, H. J. C.; Havinga, E. *Photochemistry of Vitamin D and Its Isomers and of Simple Trienes*. In *Advances in Photochemistry*; Pitts, J. N., Jr., Hammond, G. S.; Gollnick, K. Eds.; John Wiley & Sons: New York, 1979; Vol. 11, pp 305-373; (b) Dauben, W. G.; McInnis, E. L.; Mincho, D. M. *Photochemical rearrangements in trienes*. In *Rearrangements in ground and excited states*; De Mayo, P. Ed.; Academic Press: London, 1980; Vol. 3, pp 91-129.
- 24) a) Khorana, H. G. (1992) *J. Biol. Chem.*, 267, 14. b) Hofmann, K. P., and Helmreich, E. J. M. (1996) *Biochim. Biophys. Acta*, 1286, 285-322.
- 25) a) Sakmar, T. P. (1998) *Prog. Nucleic Acid Res. Mol. Biol.*, 59, 133. b) Shichida, Y., and Imai, H. (1998) *CMLS, Cell. Mol. Life Sci.*, 54, 1299-1315.
- 26) Matsumoto, H., and Yoshizawa, T. (1975) *Nature*, 258, 523-526.
- 27) Okada, T., Trong, I. L., Fox, B. A., Behnke, C. A., Stenkamp, R. E., and Palczewski, K. (2000) *J. Struct. Biol.*, 130, 73-80.

# (THEORETICAL BACKGROUND)

## -CHAPTER 2-

### Quantum Mechanics/Molecular Methods

When studying photochemical systems one encounters several fundamental problems: the detailed processes as the bond-breaking / bond-making process, the cis-trans photo-isomerization, which occur in these reactions must be modeled using some form of quantum-mechanical potential surface. Furthermore for some of the systems analyzed in this thesis, environment influence of the protein must be considered together with the chromophore reaction. Recently the strategy of QMMM method is one of the most adopted. In the next chapter it will be briefly described in what this method consists. For the moment it is sufficient to emphasize how through this strategy is possible to treat larger molecule, as entire proteins or enzyme, by studying separately the small part of system (of limited number of atoms) responsible of the reaction with a quantum mechanical potential, and the surrounding protein with molecular mechanics potential.

The next section briefly describes the QM methods that have been adopted to investigate the electronic properties of system responsible of the reaction (as a chromophore in the proteic environment); through the potential energy surface of ground state and the excited state ( computed by solving the Schroedinger equation) is possible to map the reaction path of the molecule after photon absorption and the following relaxation channels on the ground state populated after the decay. The remainder of the thesis documents applications of the method to some of important photochemical reactions.

#### 2.1 The Schrödinger Equation and the Born-Oppenheimer Approximation

The development of quantum mechanics<sup>1</sup> helped to consolidate the theoretical foundations of the experimental chemical sciences. The total energy of a system is given by the non-relativistic, time-independent, ignoring the spin-orbit coupling, Schredinger equation:

$$\hat{H}\Psi(R,r) = E\Psi(R,r) \quad (2.1)$$

In this equation  $\Psi$  is the wave function,  $R$  and  $r$  denote respectively the relative positions of the nuclei and of the electrons within the molecule, and the Hamiltonian  $\hat{H}$



operator can be decomposed in 5 terms: the kinetic energy of the electrons ( $\hat{T}_e$ ) and nuclei ( $\hat{T}_N$ ), the attraction of the electrons to the nuclei ( $\hat{V}_{eN}$ ), and the interelectronic ( $\hat{V}_e$ ) and internuclear repulsion ( $\hat{V}_N$ ):

$$\hat{H} = - \sum_i \frac{\hbar^2 \nabla_i^2}{2m_e} - \sum_k \frac{\hbar^2 \nabla_k^2}{2m_k} - \sum_i \sum_k \frac{e^2 Z_k}{r_{ik}} + \sum_{i < j} \frac{e^2}{r_{ij}} + \sum_{k < l} \frac{e^2 Z_k Z_l}{r_{kl}} \quad (2.2)$$

$$\hat{H} = \hat{T}_e + \hat{T}_N + \hat{V}_{eN} + \hat{V}_e + \hat{V}_N \quad (2.3)$$

where  $i$  and  $j$  run over electrons,  $k$  and  $l$  run over nuclei. The energy and many properties of the system can be obtained by solving the Schrödinger equation for  $\Psi$ , subject to the appropriate boundary conditions. Many different wavefunctions are solutions to it, corresponding to different stationary states of the system. If the potential energy  $V$  is not a function of time, the Schrödinger equation can be simplified using the mathematical technique known as separation of variables. For the problems we are interested on, this separation is valid, and we focus entirely on the familiar time-independent Schrödinger equation. Under typical physical conditions, the nuclei of molecular systems are moving much more slowly than the electrons. Therefore the electronic ‘relaxation’ with respect to nuclear motion is considered instantaneous and it becomes possible to decouple these two motions, and compute electronic energies for *fixed* nuclear positions. This assumption is at the base of the *adiabatic Born-Oppenheimer approximation*<sup>2</sup>, and within this approximation it will be defined the concept of Potential Energy Surface (PES). Rather than attempting to solve the Schrödinger equation in the space of all the variables, the Born-Oppenheimer approximation may be formulated by writing down the (time-independent) Schrödinger equation for only the electrons in the field of fixed nuclei, using the electronic Hamiltonian instead of the full Hamiltonian:

$$\hat{H}_{elec} \Psi_{elec}(\vec{r}, \vec{R}) = E_{eff}(\vec{R}) \Psi_{elec}(\vec{r}, \vec{R}) \quad (2.4)$$

In this formulation  $\hat{H}_{elec}$  is the electronic Hamiltonian, differing from the full, only for neglecting the kinetic energy term of the nuclei. The eigenvectors of the equation (2.4) are the adiabatic electronic wavefunctions  $\Psi_{elec}(\vec{r}, \vec{R})$ , where the electronic coordinates ( $\vec{r}$ ) being independents from the velocity of the nuclei, vary parametrically with nuclear configuration ( $\vec{R}$ ). Consequently, we can write the total wavefunction  $\Psi(R, r)$  as the product of the electronic wavefunction  $\Psi_{elec}(\vec{r}, \vec{R})$  and a nuclear wavefunction  $\Psi_{nucl}(\vec{R})$ :  $\Psi(R, r) = \Psi_{elec}(\vec{r}, \vec{R}) \cdot \Psi_{nucl}(\vec{R})$ . Solving the equation (2.4) for a particular nuclear geometry gives, in general, an infinite set of eigenvalues  $E_{eff}$ . Moving in the configuration space  $R$  then maps out what is called the potential energy surface (PES) for each electronic state. The latter is one of the most important concepts in computational chemistry. The lowest energy solution gives the ground state surface.

The other roots generate the excited state surfaces, although we might only be concerned with the lowest of these.

In most situations, the B-O approximation is reliable, but for one stage of a photochemical reaction it fails. When two potential surfaces are nearly degenerate, and the nuclear velocity is non-zero, nuclear and electronic motion become coupled, and the molecule no longer feels the effect of a single electronic state. The idea of a surface hop is then used as a way of dealing with this phenomenon in a semi-classical way, whilst still retaining the classical trajectory model.

The main task of theoretical studies of electronic structures is to solve, at least approximately, the Schrödinger equation for the electronic wavefunction and hence find the effective nuclear potential function  $E^{eff}(\mathbf{R})$ . The potential surface  $E^{eff}(\mathbf{R})$  is fundamental to the quantitative description of chemical structures and reactivity, as we will see in Chapter 3.

## 2.2 Molecular Orbitals and Valence Bond theories.

We can only solve analytically the Schrödinger equation for the simplest systems, like the hydrogen atom or molecular hydrogen cation,  $\text{H}_2^+$  (1-electronic systems). The problem with the systems with more electrons is due to the interelectronic repulsion

energy  $\sum_{i < j} \frac{e^2}{r_{ij}}$ ; this term in the Hamiltonian does not permit the variables separation for

each electron, through which it is possible to solve the Schrödinger equation. Since it is not possible to solve the Schrödinger equation for systems with two and more electrons, approximations have to be introduced. The procedure to obtain an energy and a wavefunction as an approximate solution of the Schrödinger equation should be completely specified in terms of nuclear positions and the number and spins of electrons in the molecule.

The first approximation we will consider arises from the interpretation of  $|\Psi|^2$  as a probability density for the electrons within the system. Molecular Orbital (MO) theory is an approach to molecular quantum mechanics which uses one-electron functions (or orbitals) to approximate the full wavefunction  $\Psi$ . The square of the wavefunction,  $\Psi^2$  (or  $|\Psi|^2$  if  $\Psi$  is complex), is interpreted as a measure of the probability density for the particles described. Therefore we require  $\Psi$  to be normalized: if we integrate over all space, the probability should be unity (the probability of finding all the particles anywhere in space is unity). Secondly, according with Pauli's principle<sup>3</sup>,  $\Psi$  must also be antisymmetric, meaning that it must change sign when two identical particles (such as electrons) are interchanged. For an electronic wavefunction, antisymmetry is a physical requirements following from the fact that the electrons are fermions.<sup>†</sup> More specifically, this requirement means that any valid wavefunction must satisfy the following condition:

$$\Psi(\mathbf{r}_1, \dots, \mathbf{r}_i, \dots, \mathbf{r}_j, \dots, \mathbf{r}_n) = -\Psi(\mathbf{r}_1, \dots, \mathbf{r}_j, \dots, \mathbf{r}_i, \dots, \mathbf{r}_n) \quad (2.5)$$

---

<sup>†</sup> Fermions are particles that have the properties of antisymmetry and half-integral spin quantum number, among others.

The orbital description of atoms and molecules electronic structures is probably the most powerful unifying tool to provide such understanding. There are two main theories, which use the orbitals description: Molecular Orbital (MO) theory and Valence Bond theory (VB).

Molecular orbital (MO) theory is an approach to molecular quantum mechanics which uses one-electron functions (or orbitals) to approximate the full wavefunction  $\Psi$ . This approximate treatment of electron distribution and motion assigns individual electrons to one-electron functions termed spin-orbitals. These comprise a product of a spatial function, termed molecular orbital,  $\phi_1(x,y,z)$ ,  $\phi_2(x,y,z)$ ,  $\phi_3(x,y,z)$ , ..., and either  $\alpha$  or  $\beta$  spin function (describing the spin component of the electron). The spin-orbitals are allowed complete freedom to spread throughout the molecule, their exact form being determined variationally to minimize the total energy. The variational method, yield a total energy that is an upper bound to that which would result from exact solution of the full Schrödinger equation.

For simplification reasons molecular orbitals are chosen to be orthogonal to each other and normalized (normalization correspond to the physical requirement that the probability of finding the electron anywhere in space is unity):

$$\int \phi_i^* \phi_i d\tau = 1 \quad (2.6)$$

$$\int \phi_i^* \phi_j d\tau = 0, \quad i \neq j$$

In the simplest version of the theory, as in the *restricted Hartree-Fock* (RHF)<sup>4</sup> method, a single assignment of electrons to orbitals is made: they are ordered by energy and the electrons are used to populate starting from the lowest one, featuring the so called *electron configuration*. The orbitals are then combined together to form a suitable *many-electron wavefunction*  $\Psi$  which is the simplest MO approximation to the Schrödinger equation. The simplest possible way of making  $\Psi$  for the description of an n-electron molecular system would be in the form of a simple product of spin orbitals (the so called *Hartree product*). Since the electronic wave function has to satisfy the Pauli's exclusion principle, it must be antisymmetric. To ensure this condition the wave function is written as a Slater determinant:

$$\Psi(x_1, x_2, \dots, x_N) = \begin{vmatrix} \Phi_1(x_1) & \Phi_2(x_1) & \cdots & \Phi_N(x_1) \\ \Phi_1(x_2) & \Phi_2(x_2) & \cdots & \Phi_N(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ \Phi_1(x_N) & \Phi_2(x_N) & \cdots & \Phi_N(x_N) \end{vmatrix} \quad (2.7)$$

Trough this formulation the property of antisymmetry is ensured ; in fact swapping two rows of the determinant corresponds to swapping two electrons, having the effect of changing the determinant sign. Furthermore according with the Pauli<sup>3</sup> exclusion principle, it is not possible for a molecular orbital  $\phi_i$  to be occupied by two electrons of the same spin. This is also guaranteed by determinantal formalism because the determinantal wavefunction vanishes if two columns are identical. Hence orbitals may be classified as doubly occupied, singly occupied or empty. Most molecules have an

even number of electrons and their ground states may be represented by closed-shell wavefunction with orbitals either doubly occupied or empty. The open-shell wavefunction represents an electronic configuration with the last orbitals partially occupied, as mono-occupied.

The MO approach is particularly useful for describing excited states and is quite popular in spectroscopy, because an electronic transition can be visualized like an electronic excitation from the molecular orbital  $\Phi_j$  to the  $\Phi_{j+1}$ . In VB theory one starts with the occupied atomic orbitals of the atoms and constructs a many-electrons wave function to describe bonding directly in terms of these atomic orbitals. VB theory is most useful for describing reactions, bond dissociation and spin recoupling, because the many-electrons states are built into VB. Important chemical concepts such as resonance are based on VB concepts. Indeed while the molecular orbitals theory uses delocalized orbitals, the valence bond theory uses localized orbitals to build the state wave function. For instance the benzene wave function is described by the six resonance structures (each combination of Slater determinants); two of them are:



However, VB is considered less accurate and less reliable than MO and it is much less obvious how to describe excited states.

### [2.3 Basis Sets: The LCAO Approximation](#)

In MO theory one considers combinations of the atomic orbitals to generate molecular orbitals: thus the electronic state wave function is constructed using the molecular orbitals. In reality a further approximation involves expressing the molecular orbitals  $\phi_i$  as linear combination of a pre-defined finite set of  $N$  one-electron functions known as basis functions.

$$\phi_i = \sum_{\mu=1}^N c_{\mu i} \chi_{\mu} \quad (2.8)$$

where the coefficients  $c_{\mu i}$  are known as the molecular orbital expansion coefficients.

The basis functions (also chosen to be normalized) constitute the basis set. These basis functions (which are defined in the specification of the model) are usually centered on the atomic nuclei and so bear some resemblance to atomic orbitals. If the basis functions are the atomic orbitals for the atoms making up the molecule, the previous equation is often described as the linear combination of atomic orbitals (LCAO) approximation, and is frequently used in qualitative descriptions of electronic structure. However, the actual mathematical treatment is more general than this, and any set of appropriately defined functions may be used.

To provide a basis set that is well defined for any nuclear configuration and therefore useful for theoretical model, it is convenient to define a particular set of basis functions associated with each nucleus, depending only on the charge on that nucleus. Such functions may have the symmetry properties of atomic orbitals, and may be classified as  $s$ ,  $p$ ,  $d$ ,  $f$ ,... types according to their angular properties. Two types of basis functions have

received widespread use: Slater-type functions<sup>5</sup> and gaussian-type functions<sup>6</sup>. The former provide a reasonable representations to atomic orbitals but are not well suited to numerical work, and their use in practical molecular orbital computations has been limited. The latter are less satisfactory as representation of atomic orbitals (particularly because they do not have a cusp at the origin), nevertheless they have the important advantage that all integrals in the computations can be evaluated explicitly (analytically) without recourse to numerical integration.

Actually, linear combination of gaussian functions are used to define the basis functions (for example an  $s$ -type basis function  $\chi_\mu$  can be expanded in terms of  $s$ -type gaussians  $g_i$ ):

$$\chi_\mu = \sum_i d_{\mu i} g_i \quad (2.9)$$

where the  $d$  coefficients (called contraction coefficients) are fixed constants within a given basis set. Basis functions of this type are called contracted gaussians, the individual  $g$  being termed primitive gaussians.

All of the construction result in the following expansion for molecular orbitals:

$$\phi_i = \sum_{\mu=1}^N c_{\mu i} \sum_n d_{\mu n} g_n \quad (2.10)$$

where  $m$  represents the generic angular type ( $s, p, d, f, \dots$ ).

Given the basis set, the unknown coefficients  $c_{\mu i}$  are determined so that the total electronic energy calculated from the many-electron wavefunction is minimized and, according to the variational theorem, is as close as possible (an upper bound) to the energy corresponding to the exact solution of the Schrödinger equation. The Self Consistent Field (SCF) procedure is the iterative methods through which expansion coefficients  $c_{\mu i}$  are optimized to gain the minimum value of the energy (a threshold in the energy gap between consecutive iterations is chosen, under which the procedure is considered “convergent”). This energy and the corresponding wavefunction represent the best that can be done within the chosen approximation, that is, the best given the constraints imposed by: (a) the use of a limited (not infinite) basis set, and (b) the constraint in the wavefunction extension (for the RHF methods being the use of a single-determinant wavefunction).

## 2.4 Hartree-Fock Theory

Molecular orbitals for a system can be obtained by solving Hartree-Fock Self Consistent Field (SCF) equations:

$$\hat{F}\Psi = E\Psi \quad (2.11)$$

where the Hartree-Fock operator  $\hat{F}$  is defined as:

$$\hat{F} = -\frac{1}{2} \nabla_i^2 - \sum_{k=1}^M \frac{Z_k}{r_{ij}} + V_i \{j\} \quad (2.12)$$

By comparison with the (2.12) it is evident that the Fock operator does not include the simultaneous inter-electronic repulsion. The final term represents a potential interaction with all of the other electrons occupying orbitals  $\{j\}$  (each electron feels the average field of all other electrons in the molecule) and may be computed as ;

$$V_i \{j\} = \sum_{j \neq i} \int \frac{\rho_j}{r_{ij}} dr \quad (2.13)$$

$$\rho_j = |\Psi_j|^2 \quad (2.14)$$

where  $\rho_j$  is the charge (probability) density associated with electron  $j$ .

Thus it is apparent how the Hartree-Fock operator is a function of its own solutions, and so why it is not possible to determine directly the eigenfunctions  $\Psi$ . To finesse this problem, Hartree proposed an iterative “self consistent field “ method, in which one itguesses the wave function for all occupied MO and uses to construct the necessary Hartree-Fock operator. Solution of each differential equation provides a new set of eigenfunctions, different from the initial guesses. Finally when the difference between a newly determined set and the preceding set falls below some threshold criterion, we refer to the final set of  $\Psi$  as the converged SCF orbitals. Computationally to guess the initial wave function means to introduce an initial set of orbitals expansion coefficients that are optimized through a variational procedure, leading to the lowest possible energy.

To solve the HF equation with a wave function constructed using the Slater determinant and LCAO scheme, in a self consistent procedure, the Roothaan-Hall equations<sup>7</sup> must be introduced, corresponding to solve the following eigenvalues problem:

$$FC = \epsilon SC \quad (2.15)$$

where  $F$  is the Fock matrix  $F_{ij} = \langle \phi_i | \hat{F} | \phi_j \rangle$  corresponding to all the interacting terms between the orbitals,  $C$  is the expansion coefficient matrix,  $S$  is the overlap matrix and  $\epsilon$  is the eigenvalues vector. The Roothaan-Hall equations need to be solved iteratively comparing the energies of the iteration  $n$ th to that of iteration  $n$ th-1: if they differ less than the selected threshold, then the procedure stops and it is reached converged.

The Hartree-Fock wave function is able to describe only ground states (closed shell systems as well as the open shell systems using RHF, and UHF). In fact this wave function is approximate in the form of an only one Slater determinant, and it is not appropriate to study any excited state. Further, there are not described the dynamic correlation as well as the non-dynamic electronic correlation. In fact the chief error of HF approximation consists in ignoring the correlated motion of each electron with each other. This kind of electron correlation is called dynamical correlation which is particular important for the core orbitals. Moreover the non-dynamical correlation arises

from the possibility of presence of near-degeneracy configurations. Since the HF wave function is only single determinant, the difference between the HF energy and the exact energy is called correlation energy:

$$E^{corr} = E^{exc} - E^{HF} \quad (2.16)$$

The two main limitations of this method (it is not able to describe any electronic excited state and it is not able to take in account the correlation energy) were partially solved with the Post-Hartree Fock methods.

## 2.5 Configuration Interaction (CI)

After the Hartree Fock method there were mainly developed two theories to solve its problems, based on two different approaches: the variational theory and the perturbation theory. One of the most used methods able to supply some of HF limitations and better describing the potential energy surface, is the Configuration Interaction method<sup>8</sup>(CI). The basic idea of this theory is to use a multiconfigurational wavefunction generated from a linear combination of Slater determinants, able to study, simultaneously, both ground and excited states. The wave function is composed as linear combination of Configuration State Functions (CFS) that are generate from single, double and all other excitations from all the occupied orbitals to the unoccupied orbitals:

$$\Psi = \alpha_0 \Phi_{HF} + \sum_i \sum_r^{occ\ vir} \alpha_i^r \Phi_i^r + \sum_{i < j} \sum_{r < s}^{occ\ vir} \alpha_{ij}^{rs} \Phi_{ij}^{rs} + \dots \quad (2.17)$$

where  $\alpha_i^r$  are the CI expansion coefficients, referred to excited states determinant and thus representing the correction to the HF model. In fact while  $\Phi_{HF}$  represents the HF wavefunction (corresponding to the ground state electronic configuration where all the electrons lies to the lowest occupied orbitals) and the following  $\Phi_i^r, \Phi_{ij}^{rs}$  represent the wavefunction generated from single and double electronic excitations. The CI method being a variational method, use a procedure where the CI  $\alpha_i^r$  expansion coefficients are variationally optimized until a suitably description of the considered state is reached. The CI  $\alpha_i^r$  expansion coefficients reflect the weight of different configurations and they ensure the normalization procedure also.

In order to get the eigenvalues and eigenvectors corresponding to the considered states, the fundamental equations we need to solve are the secular equations:

$$HC = ESC \quad (2.18)$$

In this equation  $H$  indicates the Hamiltonian matrix, where the generic element is  $H_{KL} = \langle \Phi_K | \hat{H} | \Phi_L \rangle$  representing the interacting term between two configurations;  $C$  is the coefficients matrix and  $E$  is the energy matrix collecting the eigenvalues of the examined states.

Even if the mathematic formalism is quite simple we cannot perform a FCI (Full CI) calculation even for small systems, since the number of configuration increases incredibly, in fact the number of CSF is given by:

$$CSF = \frac{m!(m+1)!}{\frac{n!}{2} \binom{\frac{n}{2}+1} \binom{m-\frac{n}{2}} \binom{m-\frac{n}{2}+1}} \quad (2.19)$$

Where  $m$  and  $n$  are the number of orbitals and electrons in the molecule. So we must use truncated CI functions where we consider only those configurations (CSF) fundamental for a reliable wave function. The most common way is to included only the determinants referred to a certain type of excitations, like single excitations (CIS), double (CID) or single and double (CISD) and so on. The CIS method is quite fast, but not very accurate to describe electronic excitations because it can calculate only single-excitation states. While CISD is more accurate, but the computational performance dramatically lowers down as the number of basis functions increases.

The ground state energy obtained using a complete basis set (full CI) is the exact non relativistic energy of the system within the BO approximation and with that specific basis set, it considers all the correlation energy. The limiting and best value for this procedure is obtained when an infinite basis set and a full CI are considered (within the system dimension): this corresponds to solve the exact non relativistic time-independent Schrödinger equation of the system within the BO approximation. It is apparent that infinite basis sets cannot be handle and have no practical meaning.

Moreover the CI procedure is not “size-consistent” that means that there is not correspondence between separate systems and united systems, therefore is not possible to describe with the same accuracy all the points of PES associated to products, reactants or transition states. While it is generally not possible to satisfy this condition fully, it is often possible to construct models that are *size-consistent for infinitely separated systems*. This means that applications of the model to a system of several molecules at infinite separation will yield properties that equal the sum of these same properties for the individual molecules.

## [2.6 MC-SCF: Multiconfigurational Self-Consistent-Field](#)

A further development of the CI theory is the MC-SCF (Multiconfigurational Self-Consistent-Field) theory. The following theory represents an improvement respect to the precedent one, especially for studying the excited states. If with the CIS method the CI optimization is a process slowly convergent, the MCSCF has adopted an optimization procedure able to reach faster (quickly) the convergence criteria and thus to get a more reliable energy eigenvalues. This formalism can be considered as a natural extension of SCF method, and it is based on two main idea solving the two CI limitations. According with the CI theory the MC-SCF wavefunction is still obtained as a linear combination of several Slater determinants: nevertheless with this procedure the CI coefficients  $\alpha_i^r$  (eq 2.17) are optimized simultaneously with all the atomic orbitals coefficients  $c_{\mu i}$  (eq.2.8). Furthermore the artful truncated CI wave functions is

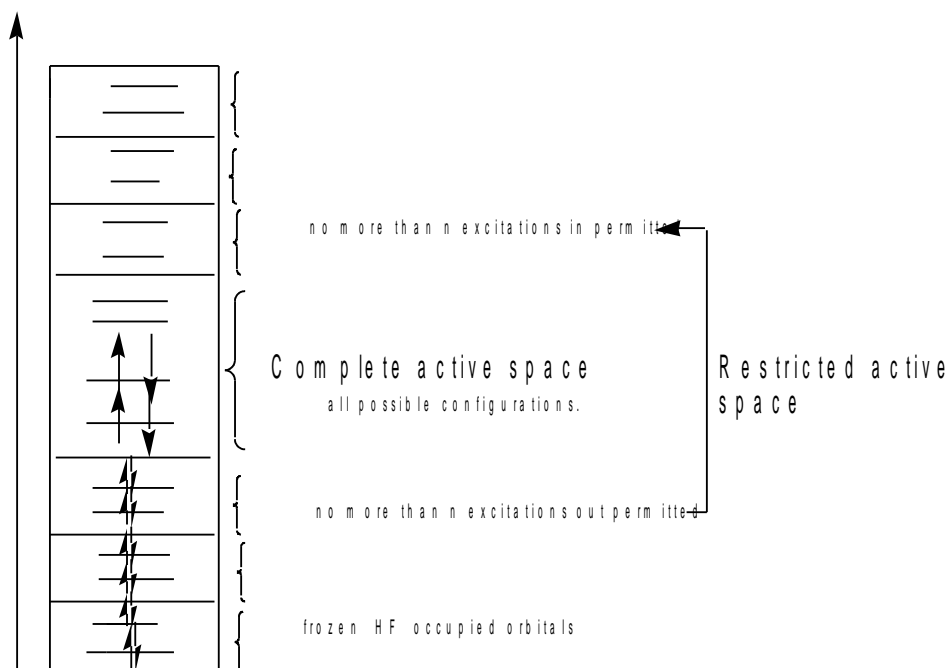


avoided by a full CI calculation limited to a subset of orbitals accurately selected. Using this approach it is possible to get accurate results with a limited number of determinants, moreover the MC-SCF wave function as composed reflect also the property of “size-consistency”. However even if the number of used configurations is still limited, it is possible supply at latter with optimization of the molecular orbitals, (changing variationally the coefficients) minimizing the overall energy.

The mathematical machinery to solve the MCSCF equations is very complicate and lots of different ways have been developed: the easiest way, is the one optimizing separately the two type of coefficients. The most important MC-SCF calculation methodologies are Complete Active Space Self Consistent Field (CASSCF) and restricted active space self consistent field (RASSCF ) methods: in both case the excitations are considered only between a defined number of orbitals.

## 2. 7 CASSCF and RASSCF

CASSCF (Complete Active Space Self Consisted Field)<sup>9</sup> calculations are a combination of an SCF computation with a full CI calculation involving a subset of the orbitals. The orbitals used to built the CI space are known as the active orbitals. In this way the MC-SCF methodology optimizes the orbitals appropriately for the exited states and does not use orbitals SCF as the full CI does. Selecting the correct active space requires a good understanding of the chemical problem under investigation. In a CASSCF calculation once choose the appropriate active space; the different Slater determinants reflecting all possible configurations are generated from electronic excitations in a limited number of active orbitals. Instead with CAS calculations the total number of CSF would be incredibly high, when the symmetry is neglected. Various schemes exist to try to reduce the number of CSF in the expansion in a rational way. Shrink the size of the CAS calculation is what a RASSCF (Restricted Active Space Self Consistent Field) calculation allows to do, where the CSF describe only a limited number of configurations generated from a narrow number of electronic excitations into a different space which is called restricted active space. Thus while all possible configurations of electrons in the CAS space are permitted , only a limited number of RAS configurations is possible. Remaining occupied and virtual orbitals are restricted to occupation number of exactly two and zero, respectively. It is important emphasize a fundamental step between CAS and RAS methodologies. Running a RAS calculation the shapes of the core orbitals are frozen to those determined at HF level.

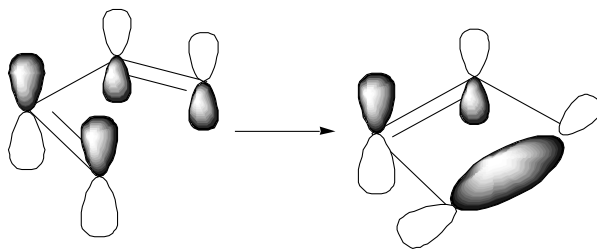


**Scheme 2.1** There are represented all possible electronic configuration generated by the allowed excitations in CASSCF and RASSCF.

In conclusion, as it is clear from the scheme 2.1 both methods are characterized from the orbitals divided in three regions: a) the core orbitals, always double occupied and treated at the HF level, b) the active orbitals, within them there are considered all the excitations explicitly and c) the virtual orbitals, always empty and as the core orbitals, treated at the HF level. The difference between CASSCF and RASSCF is the type of excited determinants used: while the CASSCF method is computationally more expensive with a wavefunction generated by all the possible excitations in the active space (corresponding to a full CI for the active orbitals region), the RASSCF being computationally less expensive, may be used if it is necessary to include more orbitals into the active space. In fact, the RASSCF active space is divided into three parts (scheme 2.1): 1) RAS1: occupied orbitals where only a certain number of electrons are allowed to be excited, like in truncated CI methods, 2) RAS2: occupied and empty orbitals, this part is treated at the CASSCF level, 3) RAS3: empty orbitals, similar to RAS1, but in this case only a defined number of electrons can occupy these orbitals.

### 2.7.1 Selecting the active space

Along a reaction path it is necessary to identify those normal modes dominating the reaction coordinates and involving a substantial changes in bonding. Thus it is critical to ensure that an appropriate choice of orbitals have to be done, in order to guarantee enough flexibility in the wave function. For instance the electrocyclozation reaction of 1,3-butadiene in cyclobutene transforms the two  $\pi$  bonds of butadiene (frontier orbitals) into one different  $\pi$  bond and a new  $\sigma$  bond. Thus to study the product (cyclobutene) a reasonable selection of active orbitals would comprise the two  $\pi$  bonds (bonding and antibonding), and the two  $\sigma$  bonds of the new single bond.



**Scheme.2.2**

Frontier orbitals of cis-butadiene      active orbitals of cyclobutene  
 $\pi_1, \pi_2, \pi_3, \pi_4$                                    $\pi, \pi^*, \sigma, \sigma^*$

Typically a certain amount of error is usual in the selection of an active space, and a chemical experience is often required. A criterion to select the “correct” orbitals to include in CAS space consists in examine the occupation number that should not be higher than 1.98 and less than 0.02 values.

## [2.8 Multiconfigurational Second Order Perturbation Theory Method: CASPT2](#)

The MCSCF is one of the method most used being a quite reliable method for topological studies on potential energy surfaces both in the ground and in the excited states, providing accurate descriptions of chemical reactivity problems. Nevertheless for a correct evaluation of energy in chemical and photochemical problems, contributions due to dynamic electronic correlation have to be considered. Now CASSCF potential does account for the correlation energy of the active orbitals only; these depends on its inadequacy to describe the instantaneous inter-electronic repulsion due to the  $1/r_{ij}$  terms in the Hamiltonian. If on one hand CASSCF wavefunction, being multi-configurational, reproduces only non-dynamical correlation, associated with the “valence” electrons, on the other hand the dynamical correlation, associated with the core electrons (belonging to the inner shell) is not taken in account. The correction due to this terms is necessary to have correct (comparable to experimental data) energy gaps and activation barriers. Second order perturbation theory methods<sup>10</sup> compute these contributions perturbatively and in particular, the CASPT2 method is one of the most powerful ones available<sup>11</sup>. Within this theory, CASSCF wavefunction is perturbed by adding more terms to the multi-configurational expansion of  $\Psi$ , including also Slater determinants of core-virtual excitation type, valence-virtual excitation type, as described from the following expression

$$\widehat{H} = \widehat{H}_0 + \lambda \widehat{H}_1 \quad (2.20)$$

where  $\widehat{H}_0$  and  $\widehat{H}_1$  are the zero<sup>th</sup> order and the perturbation Hamiltonians, and  $\lambda$  the weight of the perturbation. In the same way are defined the wave functions and the energies:

$$E_0 = \langle \psi_0 | \widehat{H}_0 | \psi_0 \rangle$$

$$\begin{aligned}
E_1 &= \langle \psi_0 | \hat{H}_1 | \psi_0 \rangle \\
E_2 &= \langle \psi_0 | \hat{H}_1 | \psi_1 \rangle
\end{aligned}
\tag{2.21}$$

Where the condition of orthogonality between the zeroth order wavefunction and perturbation is implied. The  $\hat{H}_0$  is defined as :

$$\hat{H}_0 = \hat{F} = \sum_i f_i
\tag{2.22}$$

Usually the Fock operator is assumed diagonal in a HF space orbitals, but this is not the case for CASSCF reference wave function. The CASSCF wavefunction is invariant against rotations among the core orbitals, the active space, and the virtual orbitals. This freedom can be used to simplify the operator, dividing it in diagonal subset:

$$\hat{F} = f_{cc} + f_{aa} + f_{vv} + f_{ac} + f_{vc} + f_{va}
\tag{2.24}$$

where the labels  $c, a, v$  are referred respectively to *core*, *active*, and *virtual* orbitals spaces. Defined the operator  $\hat{F}$ , from this point the solution of the equations follow the normal perturbation theory, where the Hamiltonian of the perturbation is defined as the difference between the full Hamiltonian and  $\hat{H}_0$ , and the correction to the wave function:

$$\psi_1 = \sum_{\mu} C_{\mu} \phi_{\mu}
\tag{2.25}$$

Where  $\phi_{\mu}$  are all the configurations in the CI space, except those belonging to the active space. The coefficients are defined as:

$$C_{\mu} = - \frac{\langle \phi_{\mu} | \hat{H}_1 | \psi_0 \rangle}{E_{\mu} - E_0}
\tag{2.26}$$

The CASPT2 gives very accurate results, but it has an enormous computational cost. The gradient calculation are available only in Molpro package but yet in Molcas<sup>12</sup> for example. In fact it is usually applied to correct (through single point computational) the energy of a previous CASSCF geometry optimizations of the geometry.

## 2.9 MultiState CASPT2

An extension of the multiconfigurational second-order perturbation approach CASPT2 is suggested, where several electronic states are coupled at second order via an effective-Hamiltonian approach. However, the CASPT2 method can be inadequate at avoided crossing. In such instances, the CASSCF wave function is not a good reference state for the perturbation calculation and strong mixing occurs between the reference state and one or more secondary-space CASCI states. Unfortunately, CASPT2 cannot account for this mixing. This is simple to understand if we remember how the CASPT2 formalism is defined: CASPT2 method perturbatively computes the dynamic correlation using a formalism that employs a single CASSCF, and a non-diagonal zeroth-order Hamiltonian  $H$ . The CASSCF states that are orthogonal to the reference state are part of the secondary space. However, these states do not contribute in a second-order treatment, since, by definition, they do not interact with the reference state via the total Hamiltonian  $H$ . The CASPT2 method can also run into problems when valence-Rydberg mixing occurs at the CASSCF level. As in avoided crossings, in such instances the CASSCF wave function is not a good zeroth-order description for either the valence or Rydberg states. A natural remedy to this problem, and also applicable to avoided crossing regions, is to use MultiState CASPT2 (MS-CASPT2)<sup>13</sup>. MS-CASPT2 uses a multi-dimensional reference space that is spanned by two or more state-average SA-CASSCF states. An effective Hamiltonian  $H_{\text{eff}}$  is computed perturbatively and diagonalized within the reference space, permitting the SA-CASSCF states to interact via  $H_{\text{eff}}$ .

Basically if one wants to mix together  $nstates$  of a CASPT2 wavefunctions, a  $nstates$  single-state, single-reference CASPT2 calculations must be calculated. At the end of each calculation, the CASPT2 wavefunction is stored, and at the end of the last CASPT2 calculation the Bloch Hamiltonian and the corresponding overlap matrix are automatically assembled and printed. The Hamiltonian is diagonalized after symmetrization, as well as with simple half-sum (averaging) and Multi-state CASPT2, provides the new eigenvectors and relative eigenvalues.

There are two different modes in which MS-CASPT2 calculations can be performed:

- 1) Each of the states to be mixed is computed independently, and finally all states are mixed. In the following, such calculations will be denoted SS-SR-CASPT2 (single-state, single reference CASPT2). There is one contracted reference state for each CASPT2 calculation that is specific for the state under consideration. This is the cheapest method, but there are no gradients available in this case.

- 2) All  $nstates$  states are treated together, with  $nstates$  contracted reference states. This is more expensive, but should give a more balanced description since the different reference states can mix in the CASPT2. It is required that  $nstates$  equals the number of states specified on the `state` directive. For this case, denoted "MS-MR-CASPT2" (multi-state multi reference CASPT2), analytical energy gradients are available.

The CASPT2 method has been extended to be able to treat cases, where the CASSCF wave function is not a good reference function for the perturbation treatment, due to coupling between different electronic states through the second-order energy. It has been demonstrated that such a coupling can easily be introduced in the CASPT2 formalism in the spirit of quasidegenerate perturbation theory. The resulting computational method is not much more time consuming than a series of single-state calculations using the old program.

## 2.10 The evaluation of bond order in molecules

One of the classic problems of quantum chemistry is how to interpret the calculated wave function and take off its physical meaning. The MC-SCF wavefunction are very difficult to interpret qualitatively: the familiar PMO /Frontier Orbitals methods are inapplicable because one use many configurations. At this point is important recall that as a technical point, a configuration state function (CSF) refers to the molecular spin and the occupation number of the orbitals. So if for closed-shell singlet, CSF can always be represented as single determinants, for many open-shell systems CSF can only be represented by a combination of two or more determinants. At the present, the molecular (MO) theory is undoubtedly the most used computational method in quantum chemistry, while the VB theory has more limited applicability. Nevertheless, the VB theory offers a great deal of insight into a variety of chemical problems, and it would be desirable to develop VB methodologies with increased capabilities. For instance the CASSCF implementation of a valence bond based method uses the spin-exchange density matrix  $\mathbf{P}$  with a localized orbitals, where the determinants of the CASSCF wave function become VB-like determinants with different spin coupling patterns. The delocalized orbitals of the active space are localized on the atoms. For a given pair of electrons, the  $P_{12}$  is given by:

$$P_{12} = \left\langle 2 [S(1)S(2)] + \frac{1}{2} 1 \right\rangle \quad (2.27)$$

Where  $[S(1)S(2)]$  is the spin-exchange operator.

The most significant result from derivation of an expression for the spin-exchange density matrix  $\mathbf{P}$  is that the final value of the  $\mathbf{P}_{ij}$  elements is the difference between  $\alpha\alpha$  and  $\alpha\beta$  terms (non-bonding and bonding contributions):

$$P_{ij} = P_{ij}^{\alpha\alpha} - P_{ij}^{\alpha\beta} \quad (2.28)$$

The values of the matrix elements  $P_{ij}$  allow us to infer something about the physical content of the VB wave function. Thus we could imagine how become simpler evaluate resonance formulas, directly from index  $P_{ij}$  that describe the contributions of the determinants to the CASSCF wave function. The MC-SCF wave function can be projected into a Valence Bond Space consisting of simple Heitler-London configurations generated from an optimized minimum basis of non-orthogonal atomic orbitals. In this basis, the transition from reactants to products is associated with a simple change in spin-coupling (from the appropriate to the reactants to that of the products). The stumbling block for efficient developments in an *ab initio* VB theory is the use of non-orthogonal orbitals that results in enormous computational effort. Nevertheless this difficulty can be reduced by use of core-valence separation technique, which leaves only a few valence electrons to be treated at the VB level; these are

generally the “active” electrons that are involved in chemical bond, and that lead to a compact VB wavefunction, with a few VB structures.

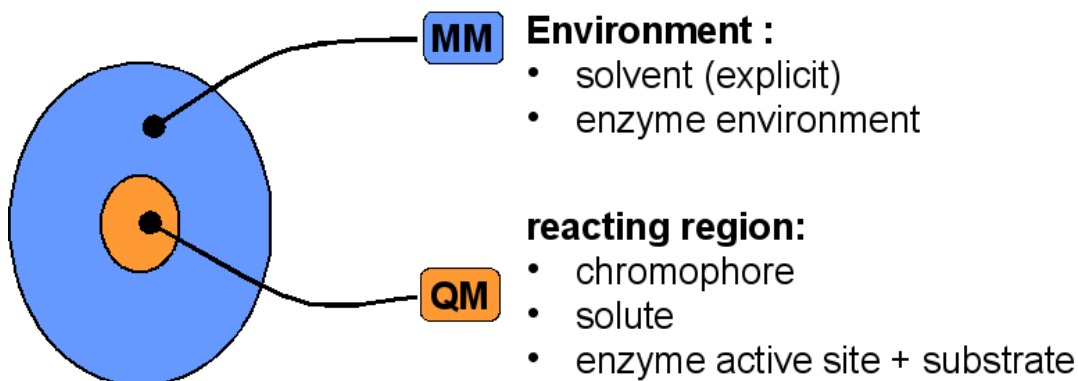
### 2.11 QMMM methods

The study of large molecular systems takes advantage of the use of computational methods based on Molecular Mechanics (MM). These methods simulate atoms almost as rigid charged spheres; the interactions between atoms are modeled on the basis of chemical connectivity, using simple harmonic potentials (or sometimes more complex functions) to describe bonds, angle bending and torsions. Non-bonded electrostatic and Van der Waals interactions are accounted for on the basis of the charge (or dipole) assigned to the atoms using a simple Coulombic potential and by means of a Lennard-Jones potential (or similar), respectively. The analytic functional form of the equations used to compute energies and forces make the MM calculation fast even for large molecules. Anyway, the drawback of these methods is their inability in describing processes involving a change of the “nature” of atoms. Thus, chemical reactions (where bonds are broken and formed, and where connectivity and atom-types are not preserved) cannot be described by MM methods. On the other hand, computational methods based on Quantum Mechanics (QM) have proved to be successful in dealing with such problems because they explicitly treat the electrons (and their couplings) by means of the calculation of the associated wave function.

Unfortunately, the QM methods are much more expensive than MM ones in terms of computational cost and cannot be used to study very large systems. Thus, a problem arises when studying the chemical reactivity of large molecular systems; in response to this need the last three decades were spent on theoretical studies for the development of new computational methods. A promising technique is the partitioning of the whole system (called *real* in the following discussion) in two regions (see figure 2.1): a small part, containing the atoms involved in the chemical process, is described at QM level, while the remaining atoms are treated at MM level, in order to speed-up the calculation and simulate (although at a lower level) the influence of the environment on the reactive core. This hybrid approach is usually called “QM/MM”. Anyway, this method has to overcome a difficult technical problem, which is often a source of significant errors and approximations: the correct description of the boundary region. Thus, the approximations adopted to deal with the interface between the two regions have a fundamental importance for a correct simulation of their reciprocal influence. Consequently, great care is addressed to this problem when developing QM/MM methods because it strongly influences the capability of reproducing the effects of the surrounding regions on the investigated process.

The wide literature on hybrid methods gave us a complete landscape on past attempts to develop a solid and general QM/MM approach.

The QMMM algorithm (implemented by Altoè et al.)<sup>14</sup> is a hybrid approach that acts as an interface between different programs; these programs execute the single phases of the whole calculation and are allowed to communicate each other by the interface, whilst the partial data are gathered and manipulated to give the final result. Modularity is the main feature of this approach, allowing the user to tailor the computational level by selecting (and combining together) specific programs according to the specific requirements of the investigation. The final purpose being to perform the computational task with the more appropriate (and efficient) computational tools available.



**Fig.2.1** Schematic representation of QMMM approach.

### References:

- 1) Foresman, J. B.; Aeleen, F. Exploring Chemistry with Electronic Structure Methods; Gaussian, Inc.: Pittsburgh, 1996.
- 2) Born, M.; Oppenheimer, J. R. *Ann. Physik* 1927, 84, 457.
- 3) Pauli, W. *Z. Physik* 1925, 31, 765
- 4) Szabo, A. & Ostlund, Neil, S. Modern quantum chemistry: an introduction to advanced electronic structure theory, McGraw-Hill, 1989.
- 5) Slater, I. *J. Phys. Rev.* 1930, 36, 57.
- 6) Boys, S. F. Proc. Roy. Soc. (London) 1950, A200, 542.
- 7) Roothaan, C.C.J. *Reviews of Modern Physics* ,1951,23, 69
- 8) Foresman, J. B.; Head-Gordon, M.; Pople, J. A.; Frish, M. J. *J. Chem. Phys.* 1992, 96, 135
- 9) Roos, B. O. *Adv. Chem. Phys.* 1987, 69, 399-446.
- 10) Moller, C.; Plesset, M. S. *Phys. Rev.* 1934, 46, 618.
- 11) a) K. Andersson, P-A Malmqvist, B. O. Ross *J. Chem. Phys.* 1992, 96, 1218.
- 12) MOLCAS, Version 3, K. Andersson, M. R. A. Blomberg, M. Fülcher, V. Kellö, R. Lindh, P.-A. Malmqvist, J. Noga, J. Olsen, B. O. Roos, A. J. Sadlej, P. E. M. Siegbahn, M. Urban, P. O. Widmark, University of Lund, Sweden (1994).
- 13) Finley J, Malqmvist P, B. Roos, *Chemical Physics Letters* 288 (1998) 299–306
- 14) Altoè, P.; Stenta, M.; Bottoni, A.; Garavelli, M. *Theor. Chem. Acc.* 2007, 118, 219.



## -CHAPTER 3-

# Potential Energy Surface

### 3.1 General Concepts

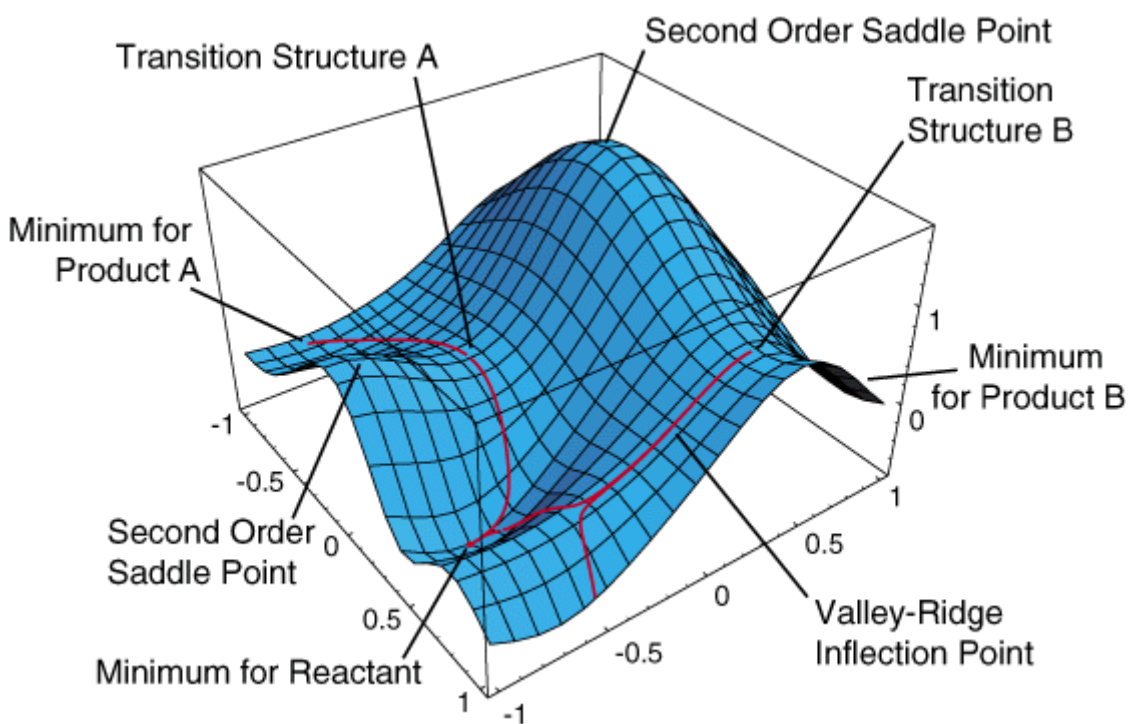
General concepts of energy surfaces for molecular motion, equilibrium geometries, transition structures and reaction paths arises from Born-Oppenheimer approximation. In the previous chapter we have discussed this theory and the eigenvalues equation providing the term  $E_{eff}$  (the potential energy surface value). A molecule or group of molecules comprising  $N$  atoms will possess  $3N-6$  degrees of freedom (or  $3N-5$  in the case of a diatomic). They result from movement of each nucleus in three independent directions, but reducing it by six due to the three translations and rotations of the molecule as a whole, which leave the internal energy unchanged. Further degrees of freedom can sometimes be ignored if they correspond to small changes in potential energy. Thus, a PES can be constructed by plotting potential energy as a function of internal coordinates (or linear combinations of them) such as bond lengths or angles,  $E(R_1, R_2, \dots, R_n)$ . Such a plot is easily obtainable when only one coordinate is needed. However, even for a triatomic system we have  $3 \times 3 - 6 = 3$  degrees of freedom, and thus four variables to plot. The PES must be constructed using a three-dimensional grid of points, each having an energy value  $E(R_1, R_2, R_3)$ . Multi-dimensional hypersurfaces are difficult to visualize but can be analyzed using various graphical tools, such as isosurfaces (surfaces of constant energy). Often it is possible to reduce the problem to two important geometrical variables, and then plot the potential energy as a function of these alone: the axes for a potential energy surface will correspond to changes to the molecules orientation. The third axis, the height corresponds to the energy, of whatever orientation the other two axes specify. In this way each point on the surface corresponds to a different structure. For a complex reaction, this approximation might not be meaningful, but it is still useful to draw a *schematic* potential surface, with reaction paths joining the critical points on the surface, even though the channels are multi-dimensional in reality. This type of visual representation is the most effective way one can see and summarize the pathway of a mechanism<sup>1</sup>. A molecule can be or go to any part of the surface, but logically it will stick to the lowest energy region and the steepest reaction paths. Information about the reaction mechanism and relative barrier heights can be successfully conveyed, and such diagrams are common in the literature.

Particular features on the PES are of chemical interest, since they will determine its topology and so the mechanism by which the molecule is driven along the path and decays to the product (ground state): these are termed critical points<sup>2</sup>. These critical points come in the form of minima, CIs, saddle points (transition states). Minima on the potential energy surface for the nuclei can then be identified with the classical picture of

equilibrium structures of molecules; saddle points can be related to transition states and reaction rates. If the Born-Oppenheimer approximation is not valid, for example in the vicinity of surface crossings, non-adiabatic coupling effects need to be taken in account to correctly describe the evolution of the molecular system. Several algorithms have been implemented to determine where these critical points are and what structures are associated with them. Once a PES has been determined for a molecule the MEP (Minimum Energy Path) can be found which will incorporate a number of these critical points in them. The nuclei and electrons will hypothetically pass through these points, when doing so have that particular isomeric structure.

### 3.2 Critical Points Determination

For a function of several variables, the first derivatives with respect to each of the variables form a vector termed the gradient<sup>3</sup>. At a critical point, the first derivative of the potential energy function with respect to all coordinates will be zero (i.e.  $\partial E / \partial R_i = 0$ , for all  $i$ ). This is true for all the critical points, so for distinguish them different criteria have to be defined. These criteria arise from a requisition relative to the second derivatives of potential energy surface, defining the Hessian matrix, for a function of several variables. If a second derivate of PES is a negative value is called imaginary frequency. In classical mechanics, the first derivative of the potential energy for a particle is minus the force on the particle, and the second derivative (for a quadratic potential) is the force constant. Minima occur if the second derivatives (the force constants) with respect to *all* coordinates are positive, and correspond to the stable products, reactants, or intermediate structures.



**Figure 3.1** There are represented the critical points featuring a reaction path for an hypothetical process  $R \rightarrow A+B$ . There are represented minima for reactant as well for both the products A and B, transition state

(1st order saddle points) for the conversion R→A, and the other for the conversion R→B. The 2th order saddle points have no interest in chemistry.

Graphically correspond to points localized at the bottom of the valley (fig 3.1). If there is one or more eigenvalues negative (imaginary frequencies), the point is a  $n^{\text{th}}$  saddle point (also called a local maximum), depending on the number of imaginary frequencies. The first order saddle point (having only one imaginary frequency) correspond to a transition state, corresponding to a maximum respect to only one coordinate and a minimum to all the rest of coordinates; higher order saddle points on PES have no chemical interest<sup>4</sup>.

At the end conical intersections are critical points characterized from the degenerate value of two electronic state. In reality there should be defined other two coordinates to guarantee the condition of degeneracy, as it will widely discussed next.

### 3.3 Energy minimization

As briefly described in the previous section a minimum is a point where the first derivative of the function  $f$  with respect to its variables is zero and the second derivatives are positive.

$$\frac{\partial E}{\partial R_i} = 0 \qquad \frac{\partial^2 E}{\partial R_i^2} > 0 \qquad (3.1)$$

The variables  $x_i$  for this type of minimization are the internal or Cartesian coordinates of the atom and the function is the total potential energy. The energies are those calculated for a given geometry using the Schrodinger equation if it is requested only a quantum mechanics (QM) calculation (i.e for small systems); for larger system the potential energy may be calculated through molecular mechanics (MM) approach, using one of the several force field MM. At the end the total energy may be also the sum of both contribute , giving the total QMMM potential energy. It does not interest wherever the energy come from for energy minimization algorithms. In order to locate the minima the coordinates are changed in a numerical method to make a configuration with lower energy. This is done in steps until no lower energy structure can be found and until the gradient are *zero*. The criterion can be customized for different types of searches. Geometry optimizations usually attempt to locate minima on the potential energy surface, thereby predicting equilibrium structures of molecular systems. Optimizations can also locate transition states. Since the gradient is the negative of the forces, the forces are also zero at such a point (called *stationary point*). All successful optimizations locate a stationary point, although not always the one that was intended. A geometry optimization begins at the molecular structure specified as its input, and steps along the potential energy surface. It computes the energy and the gradient at that point, and then determines how far and in which direction to make the next step. The gradient indicates the direction along the surface in which the energy decreases most rapidly from the current point as well as the steepness of that slope. Most optimization algorithms also estimate or compute the value of the second derivative of the energy with respect to the molecular coordinates, updating the matrix of force constants (known as the *Hessian*).

### 3.4 The use of derivatives in minimizations

The 1<sup>st</sup> derivative of the energy corresponds to the gradient, its magnitude tells us the steepness of the slope to the minimum and its sign indicates in which direction it lies.

The total energy function is expressed as a Taylor series expansion, truncated at the second order term:

$$\mathcal{V}(x) = \mathcal{V}(x_k) + (x - x_k) \mathcal{V}'(x_k) + (x - x_k)^T \cdot \mathcal{V}''(x_k) \cdot (x - x_k) / 2 + \dots \quad (3.2)$$

In this expression  $\mathcal{V}(x)$  is the total potential energy of system, the vector  $x$  has  $3N$  components corresponding to the  $3N$  Cartesian co-ordinates and  $x_k$  corresponds to the current configuration of the system.  $\mathcal{V}'(x_k)$  is a  $3N \times 1$  dimensional matrix (i.e. a vector), each element of which is the partial derivative of  $\mathcal{V}(x)$  with respect to the appropriate co-ordinate,  $\partial \mathcal{V} / \partial x_i$ . Each element  $(i,j)$  of the matrix  $\mathcal{V}''(x_k)$  is the partial second derivative of the energy function with respect to the two co-ordinates  $x_i$  and  $x_j$ ,  $\partial^2 \mathcal{V} / \partial x_i \partial x_j$ .  $\mathcal{V}''(x_k)$  is the Hessian matrix of dimension  $3N \times 3N$  ( $N$  being the number of atoms in the molecule) containing the second derivatives of the energy with respect to the coordinates and thus also the force constant matrix. This truncated expression defines the quadratic approximation, that using a quadratic potential is enough reliable only for small molecular displacement from equilibrium structure. Nevertheless it is fundamental remember the limitations of the method, being the energy functions not so often fully quadratic. Within the quadratic approximation the curvature of the function is just indicated by the second derivative (force constants) which is used to find out where the function changes direction, providing additional information in order to improve the overall efficiency of the optimization and the rate of convergence to the equilibrium geometry. The optimization methods based on use of the derivatives (called also gradient algorithms) are mathematical algorithms are defined within this approximation, and their application in locating stationary points is very wide. If analytical derivatives are available in the used computational method, these can be significantly efficient and can have good convergence properties.

### 3.5 Newton-Raphson method

Within the quadratic approximation the gradient algorithms approximate the energy surface at step  $k$  by a quadratic expression in terms of the position,  $x_k$ , the computed energy,  $E_k$ , the computed gradient,  $\mathcal{V}'(x)$  (first derivatives), and the approximate Hessian,  $\mathcal{V}''(x_k)$  (second derivatives). One of the most used optimization methods based on the use of the derivatives in the New-Raphson. To explain the origin of the simple mathematical approach is necessary to begin from the quadratic approximation equation (3.2) where the first derivative is:

$$\mathcal{V}'(x) = \mathcal{V}'(x_k) + (x - x_k) \mathcal{V}''(x_k) \quad (3.3)$$

Assuming the function is purely quadratic, the second derivative is the same everywhere, and so

$$\mathcal{V}''(\mathbf{x}) = \mathcal{V}''(\mathbf{x}_k) \quad (3.4)$$

For a stationary point ( $\mathbf{x} = \mathbf{x}^*$ )  $\mathcal{V}'(\mathbf{x}^*) = 0$  and so

$$\mathbf{x}^* = \mathbf{x}_k - \mathcal{V}'(\mathbf{x}_k) / \mathcal{V}''(\mathbf{x}_k) \quad (3.5)$$

For a multidimensional function:  $\mathbf{x}^* = \mathbf{x}_k - \mathcal{V}'(\mathbf{x}_k) / \mathcal{V}''^{-1}(\mathbf{x}_k)$

Where  $\mathbf{x}^* - \mathbf{x}_k = \Delta \mathbf{x}_k$ , and  $\mathcal{V}''^{-1}(\mathbf{x}_k)$  is the inverse Hessian matrix of second derivatives which in the Newton-Raphson method is inverted. The Hessian matrix of second derivatives must be positive definite in a Newton-Raphson minimization meaning all the eigenvalues are positive. This method can find saddle points where the Hessian matrix is not positive definite and where the energy increases i.e a maximum. At the beginning it is convenient to use a guess for the Hessian and during the optimization procedure the Hessian updating is improved, if the surface is quadratic the procedure approaches the true Hessian (or almost quadratic).

An optimization can be divided into the following steps: 1. Start with the geometry  $\mathbf{x}_k$ ,  $k=0$ , and obtain an estimate of the Hessian, 2. For  $\mathbf{x}_k$  calculate the energy,  $E_k$ , and the gradient,  $\mathcal{V}'(\mathbf{x}) = \partial \mathcal{V} / \partial x_i$  3. Update the Hessian so that the model surface fits the current energy and the gradient as well as those from previous steps (omit for the first point,  $k=0$ ). 4 Find the stationary point on the model surface using the gradient and updated Hessian (note that this requires the inverse of the Hessian). If the gradient,  $\mathcal{V}'(\mathbf{x})$ , is small enough and/or the predicted change in the geometry is small enough, stop. 5. Carry out an optimization in the direction of the predicted displacement (line search), i.e. optimize (minimize for a minimization, maximize for a transition state search)  $E(\mathbf{x}_k + \alpha \Delta \mathbf{x}_k)$  with respect to  $\alpha$  (not used in some methods, i.e.  $\alpha=1$  always) 6. Set  $\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha \Delta \mathbf{x}_k$ ,  $k=k+1$ , and return to step 2.

Even for the non-quadratic surfaces encountered in typical geometry optimizations, convergence is usually good, in particular if good estimates of the geometry and the initial Hessian can be made. In practice, as much as an order-of-magnitude increase in efficiency can be obtained with gradient optimization methods compared to energy-only algorithms.

### [3.6 The Steepest Descent and the Quasi-Newton Methods](#)

The overall cost and convergence rate for optimizing a non-quadratic potential depend on the updating scheme for the Hessian, the accuracy of the line searches and the initial estimate of the Hessian (an accurate initial estimate can improve the rate of convergence significantly, but will not effects the final, optimized geometry, since the latter depends only on the gradient - where it goes to zero - and not the Hessian). So among the approximated algorithm (able to guarantee a more efficient procedure), there the one adopted in the simplest gradient methods called the *steepest-descent* algorithm. The minimization procedure is extremely less expensive because the Hessian is take as the unit matrix (or a constant times the unit matrix) and is not updated. Thus the search is along  $\Delta \mathbf{x}_k = -\mathcal{V}'(\mathbf{x}_k)$  i.e. the direction in which the function decreases most rapidly. This method reduces the function value quite rapidly at first, but final convergence is slow.

Closely related to this algorithm is the fixed metric method in which the Hessian is a more general non-diagonal matrix that is not updated.

On the opposite side Newton's method where the Hessian is calculated directly at each step. The most frequently used gradient algorithms fall between these extremes and are termed variable metric or *quasi-Newton* methods. These methods avoid the direct calculation of the Hessian at each step. Instead they start with an estimate of the Hessian and improve it using the gradient information gathered during the course of the optimization. Moreover, for most quasi-Newton methods, the line search does not have to be very accurate and reduction of the directed gradient in the line search by a factor of 0.5 or 0.1 is often sufficient. Many schemes for updating the Hessian have been proposed in literature. Some of the more frequently utilized schemes in geometry minimizations use the DFP, MS, BFGS, or OC algorithms<sup>1</sup>. All these updating formulas have been devised to assure that the Hessian remains symmetric and positive-definite. For non-quadratic functions, the Hessian may have to be reset if it is no longer positive-definite, or if the search direction  $\Delta x_k$  does not lead to a lower value for the energy. The BFGS algorithm appears to be one of the best updating schemes, and it is usually the method of choice during minimizations.

### 3.7 Vibrational normal modes analysis

A non-linear molecule has  $3N-6$  normal modes. The frequencies of these normal modes are calculated from the wavefunction using the Hessian of second derivatives ( $\psi''$ ). The Hessian is calculated previously in a minimization calculation as shown above. The Hessian is converted to an equivalent force constant matrix in mass-weighted coordinates (F), like this:

$$F = M^{-1/2} \psi'' M^{-1/2} \quad (3.6)$$

M is a diagonal matrix of dimension  $3N$  by  $3N$ , containing the atomic masses. All elements of M are zero except those on the diagonal. The non-zero elements of  $M^{-1/2}$  are the inverse square roots of the mass of the atoms. The masses of the atoms are used because a force of a given magnitude will have a different effect on a larger mass than on a smaller one. The secular equation  $|F - \lambda I| = 0$  is solved to give the eigenvalues and eigenvectors of the matrix F. This is achieved using matrix diagonalization. If the Hessian is defined in terms of Cartesian co-ordinates then six of these eigenvalues will be zero as they correspond to translational and rotational motion of the entire system. The frequency of each normal mode is calculated from the eigenvalues using the relationship:

$$\nu_i = \frac{\sqrt{\lambda_i}}{2\pi} \quad (3.7)$$

The calculation of the vibrational frequencies is an essential tool in the determination of the critical points on the PES. A minimum has all positive frequencies corresponding to no degrees of freedom. So a frequency calculation is always run after a critical point is

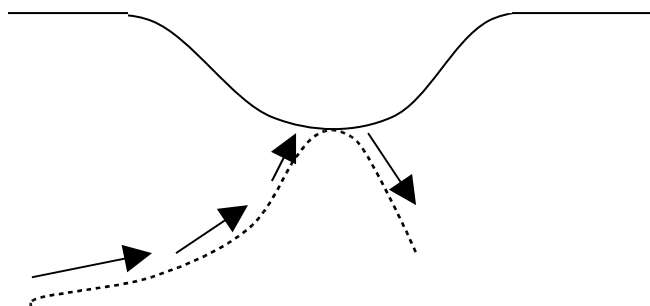
reached to check what it is. A critical point with one imaginary frequency corresponds to a transition state of some sort.

### 3.8 Transition states

Transition structures are sort after to explain how the system changes from one minimum to another, how the energy varies during such a transition and what changes in geometry are involved. When a system moves from one minimum to another, the energy increases to a maximum at the transition structure and then falls, transition structures are critical points which are also referred to as saddle points.

The number of negative eigenvalues in the Hessian matrix indicates what type of saddle point it is, so an  $n$ th order transition or saddle point has  $n$  negative eigenvalues.

Transition structures are first order saddle points, having one negative eigenvalue, they connect two minima by a pathway which passes through an energy maximum (transition structure on one coordinate) and corresponds to a minimum for displacements in all other directions perpendicular to that path. This is shown schematically below:



**Figure 3.2:** The lowest-energy path from one minimum to another passes through a saddle point

The negative eigenvalues of the Hessian matrix are often referred to as imaginary frequencies for motion of the system over the saddle point. The eigenvector corresponding to the single negative eigenvalue of a first-order saddle point is termed the transition vector. Methods for finding transition structures and reaction pathways are often closely related. Some methods start from the transition structure and move down towards a minimum, following the transition vector direction. These methods must be supplied with the transition structure geometry as the starting point. Conversely, some methods for locating transition structures do so by searching along the reaction pathway, or an approximation to it. Other methods require neither transition structure nor the pathway, but can determine both simultaneously from the two minima. It is harder to find transition structures than it is to find a minimum. Methods for locating saddle points are most effective when given as input a geometry that is as close as

possible to the transition structure. It is useful to examine the atomic displacements that correspond to the negative eigenvector, to ensure that it corresponds to the correct motion over the saddle point. As usual there are a number of different methods that can be used to locate a saddle point these can fall into two main categories: those based on interpolation between two minima, and those using only local information. With the TS found from one of these two methods it can be used in locating the reaction path by tracing the Intrinsic Reaction Coordinate which corresponds to a steepest descent path in mass-weighted coordinates, from the TS to the reactant and product. Conversely, it is also possible to use a predetermined reaction pathway to locate the transition structure. In this study we will be first locating the transition structure, which can then be used to find the reaction pathway. Examining the atomic displacements that corresponds to the negative eigenvector with which the species is going to travel over the transition structure will ensure that it corresponds to the correct motion over the saddle point.

### 3.9 Methods based on local information

The most common methods are based on an augmented Hessian Newton-Raphson approach. The quadratic region for a minimum is where all eigenvalues are positive and so the Hessian is positive definite. The quadratic region for a saddle point is the area on the energy surface surrounding the point where the Hessian contains one negative eigenvalue. The Hessian matrix will change from having all positive eigenvalues to including one negative value when moving from a minimum to a saddle point.

Once inside the 'quadratic region', close to the TS, the NR formula will locate the TS rapidly using the negative eigenvalue and the eigenvector indicating the correct direction, along the "reaction coordinate". The augmented Hessian technique, can be used to enlarge the convergence radius of the NR approach, it can locate the saddle point in a region where the Hessian does not have the correct structure. The lowest eigenvector must be in the direction of the saddle point. The NR step near a saddle point maximizes the energy in one direction (along the Hessian TS eigenvector) and minimizes the energy along all other directions. Selecting suitable shift parameters in the augmented Hessian method can improve these steps. This step is similar to the minimization technique described earlier the difference being it's for the unique TS mode. There are various techniques for acquiring the needed guess geometry. In some cases the geometry of the transition structure can be predicted through inspection. A grid search can also be used to help predict the transition structure.

### 3.10 Minimum Energy Path determination

#### Intrinsic Reaction coordinate method

If more than one minimum on a contiguous PES exist, then a family of paths can be constructed that connect one minimum to the other. If the highest-energy point on each path is considered, the transition structure by definition is the structure corresponding to the top of the saddle point, two minima are linked through this energy barrier. The path from one of these minima will lead through a transition structure in order to get to the other minima. If one minima was a reactant and the other a product then this path would be the reaction path, and it could be said that the molecule in question would "travel" along this path. The *Minimum Energy Path* (MEP) is the path of minimum energy linking critical points like these. The *Intrinsic Reaction Coordinate* (IRC)<sup>5</sup> is calculated to represent the MEP, it is defined as the path that would be taken by a classical particle

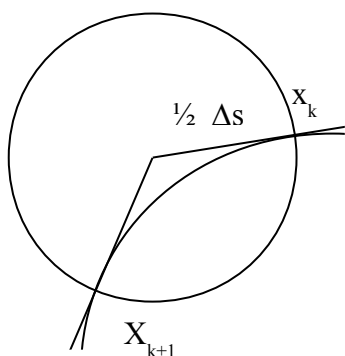


sliding downhill with infinitesimal velocity from the transition state to each of the minima (r.b). In reality a molecule will not necessarily follow such a path in a reaction as they often have translational, vibrational and rotational kinetic energy that would allow them to roam to other parts of the PES, without being restricted to the MEP may in fact deviate quite widely from it such as in the case of an “hot” system (*i.e.* with an high excess vibrational energy) where a dynamics treatment<sup>6</sup> of the motion on the PES would be required . However, the IRC is a convenient measure of the progress of a molecule in a reaction and it is a good approximation of the motion of vibrationally cold systems (*e.g.* as it happens for photochemical reactions in a relatively cold environment where the excited state reactant has a small/controlled amount of vibrational excess energy). It also plays a central role in the calculation of reaction rates by variational transition state theory and with reaction path Hamiltonians. If a transition structure has been found and one is not certain of what it connects then it is possible to use an algorithm to follow the MEP until it reaches the relevant structure(s), and so supplying possible structures of the reaction path. The IRC is the path of steepest descent from the transition state on a surface that is calculated using mass-weighted Cartesian coordinates  $m_\alpha^{1/2}x_\alpha, m_\alpha^{1/2}y_\alpha, m_\alpha^{1/2}z_\alpha$  (15) of the nuclei, where  $m_\alpha$  is the mass of nucleus  $\alpha$ . The steepest-descent path from the saddle point to the minima can be defined easily, but depends on the particular choice of coordinate system. Cartesian coordinates would yield a different path than internal coordinates. Furthermore, internal coordinates are not unique, since a number of different sets of bond lengths, angles and torsions can represent the same structure. An intrinsic reaction path can be defined independently of the coordinate system by appealing to classical mechanics. For a given energy surface, the movement of a classical particle must be the same regardless of whether Cartesian coordinates or any of a number of different sets of internal coordinates are used. The IRC is defined by this differential equation.

$$Dx/ds = g / |g| = v \quad (3.8)$$

Where  $x$  is the (mass-weighted) coordinates,  $s$  is the path length and  $v$  is the (negative) normalized gradient. To solve this the starting geometry needs to be slightly displaced from the TS structure, this can be done by using the normal mode of the imaginary frequency to displace the structure along the coordinate that leads to the associated minima. Once this is done then an optimization technique is used to follow the path of highest negative gradient to a minimum. Several algorithms have been proposed for the calculation of the IRC. One of the most widely used is the one proposed by Schlegel which is useful for following also curved reaction paths (as in many reactions).

(implemented on Gaussian package)<sup>7</sup>. Firstly the gradient is calculated at a point  $\mathbf{x}_k$ , an expansion point is generated from here by taking a step of  $\Delta s/2$  along the direction of the gradient, where  $s$  is a step of predefined size. The point  $\mathbf{x}_{k+1}$  is chosen so that the reaction path between  $\mathbf{x}_k$  and  $\mathbf{x}_{k+1}$  is an arc of a circle and so that the gradients  $\mathbf{g}_k$  and  $\mathbf{g}_{k+1}$  are tangent to this path. This algorithm requires an  $n-1$  dimensional optimization on a hypersphere about the point  $\mathbf{x}_{i+1} = \mathbf{x}_i - \frac{1}{2} s \mathbf{g}_i / |\mathbf{g}_i|$  where  $s$  is the stepsize. In this method quite large step sizes (20-40 bohr amu<sup>1/2</sup>) can be used and also IRC with large curvature are followed quite closely. The energy is then minimized on a hypersphere with radius  $\frac{1}{2} \Delta s$ , located at the expansion point.



**Figure 3.3** : Illustration of the Gonzales-Schegel constrained optimization method for following an IRC.

This is done until a gradient of zero is reached i.e. the minimum. The advantage of this procedure is that the tangent is taken at each point, where the gradient is calculated.

### [3.11 Methods based on interpolation between reactant and product](#)

These types of methods use the reactant and product geometries to locate the TS, by predicting a structure somewhere between the two. These methods are not very accurate and don't find the actual transition structures but rather they locate a point close to it. With this geometry the "local" methods can be used to refine it to the transition structure.

### [3.12 Conical Intersections: no crossing roles](#)

The preceding discussion focused on reactions governed by a single potential energy surface. Two potential surfaces can intersect along an  $n-2$  dimensional hypersurface, where  $n$  is the number of independent variables in the potential function. It is easiest to begin by considering two potential energy curves, intersecting at a single point. Nevertheless in diatomic molecules, where there is only one normal mode involved on the function, the PES of two states will only intersect if the states have a different (spatial or spin) symmetry. However, an analogous statement is *not true* of polyatomic systems: *two PES of a polyatomic molecule can in principle intersect even if they belong to states of the same symmetry and spin multiplicity*.<sup>8</sup>

This sentence leaves open the question whether such intersections ever occur in polyatomic systems. We can try to give a quantitative analysis of this situation<sup>9</sup>. We imagine that all but two of the solutions of the Schrödinger equation for the electronic wavefunction (see section 3.2) have been found, and that  $\phi_1$  and  $\phi_2$  are any two functions which, together with the found solutions, constitute a complete orthonormal set. Of course the two missing solutions correspond to the two states (whose energy is

$E_1$  and  $E_2$ ) whose crossings we are interested on. Then it must be possible to express each of the two remaining electronic eigenfunctions (which describes the states we want to examine) in the form

$$\Psi = c\phi_1 + c\phi_2 \quad (3.10)$$

The well known secular eigenvalue equation is obviously expressed as:

$$\begin{bmatrix} H_{11} - E & H_{12} \\ H_{21} & H_{22} - E \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = 0$$

and after very simple passages we can write down the expressions for the energies  $E_1$  and  $E_2$  of the two states as

$$\begin{aligned} E_1 &= \frac{1}{2} \left[ (H_{11} + H_{22}) - \sqrt{(H_{11} - H_{22})^2 + 4H_{12}^2} \right] \\ E_2 &= \frac{1}{2} \left[ (H_{11} + H_{22}) + \sqrt{(H_{11} - H_{22})^2 + 4H_{12}^2} \right] \end{aligned} \quad (3.11)$$

where, for the matrix elements:

$$\begin{aligned} H_{11} &= \langle \phi_1 | H | \phi_1 \rangle \\ H_{22} &= \langle \phi_2 | H | \phi_2 \rangle \\ H_{12} &= \langle \phi_1 | H | \phi_2 \rangle = H_{21} \end{aligned}$$

Thus, in order to have degenerate solutions (the radican must vanish), it is necessary to satisfy two independent conditions

$$\begin{aligned} H_{11} &= H_{22} \\ H_{12} (= H_{21}) &= 0 \end{aligned} \quad (3.12)$$

and this *requires the existence of at least two independently variable nuclear coordinates*. In a diatomic molecule there is only one variable coordinate - the interatomic distance - so the non-crossing rule follows: for states of different (spatial or spin) symmetry,  $H_{12}$  is always zero and we have to verify the only condition  $H_{11} = H_{22}$ ; this is possible for a suitable value of the single variable coordinate. Otherwise, if the two states have the same symmetry, they will not intersect; but in a system of three or more atoms there are enough degrees of freedom for the rule to break down: the two conditions can be simultaneously satisfied by choosing suitable values for two independent variables, while the other  $n-2$  degrees of freedom ( $n=3N-6$ ) are free to be varied without exiting from the crossing region.

If we denote the two independent coordinates by  $\mathbf{x}_1$  and  $\mathbf{x}_2$ , and take the origin at the point where  $H_{11} = H_{22}$  and  $H_{12} (= H_{21}) = 0$ , the secular equations may be cast in the form

$$\begin{bmatrix} W + h_1 x_1 - E & l x_2 \\ l x_2 & W + h_2 x_1 - E \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = 0 \quad (3.13)$$

or

$$\begin{bmatrix} W + (m + k)x_1 - E & l x_2 \\ l x_2 & W + (m - k)x_1 - E \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = 0$$

where  $m = \frac{1}{2}(h_1 + h_2)$ ,  $k = \frac{1}{2}(h_1 - h_2)$ . The eigenvalues are

$$E = W + m x_1 \pm \sqrt{k^2 x_1^2 + l^2 x_2^2} \quad (3.14)$$

and this is the equation of a *double cone* with vertex at the origin (see fig3.4). For this reason, such crossing points are called *conical intersections*. Indeed, if we plot the energies of the two intersecting states against the two internal coordinates  $x_1$  and  $x_2$  (whose values at the origin satisfy the two conditions  $H_{11} = H_{22}$  and  $H_{12} (= H_{21}) = 0$ ) we obtain a typical double cone shape.

Let's try to have a deeper insight into the physical meaning of the two conditions  $H_{11} = H_{22}$  and  $H_{12} (= H_{21}) = 0$ . If we consider  $\phi_1$  and  $\phi_2$  as the *adiabatic* components of the *adiabatic* electronic eigenfunction (a diabatic function describes the energy of a particular spin-coupling<sup>5a</sup>, while the adiabatic function represents the surface of the real state), the crossing condition (real or avoided) is fulfilled when the two diabatic components  $\phi_1$  and  $\phi_2$  cross each other, and this happens when  $H_{11} = H_{22}$ , *i.e.* the energy of the two diabatic potentials ( $H_{11}$  is the energy for the diabatic function  $\phi_1$  and  $H_{22}$  is the energy for the diabatic function  $\phi_2$ ) is the same.

At the crossing of the diabatic functions ( $H_{11} = H_{22}$ ), the expressions for the energies of the two real states (see equations 3.12) become

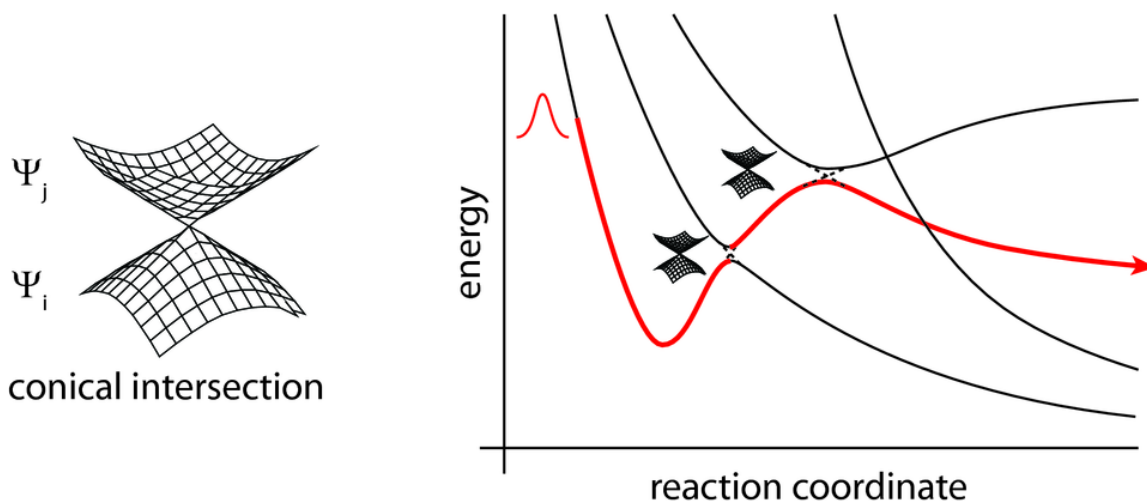
$$\begin{aligned} E_1 &= H_{11} - H_{12} \\ E_2 &= H_{11} + H_{12} \end{aligned} \quad (3.14)$$

and we see that the energy gap between the two real states is

$$E_2 - E_1 = 2 H_{12} \quad (3.15)$$

Thus, if the exchange term  $H_{12}$  is *not zero*, the crossing will be avoided and the potential surfaces of the two real states will “diverge”, being one of the two energies slower and the other higher than the diabatic energy  $H_{11}$ . Moreover, the value of the exchange term  $H_{12}$  determines how deep the avoided crossing minimum is (small values will generate deep minima, big values shallow minima). In general,  $H_{12}$  is *zero* (and the crossings will be real) when the two electronic states have a different (spatial or

spin) symmetry, while it is usually assumed *not zero* for states of the same symmetry (which will generate avoided crossings). Anyway, we have shown that this rule is true only for diatomic molecules: in a polyatomic system we can have real crossings for suitable values of a pair of independent coordinates ( $x_1$  and  $x_2$ ), which will simultaneously satisfy equations 3.12).

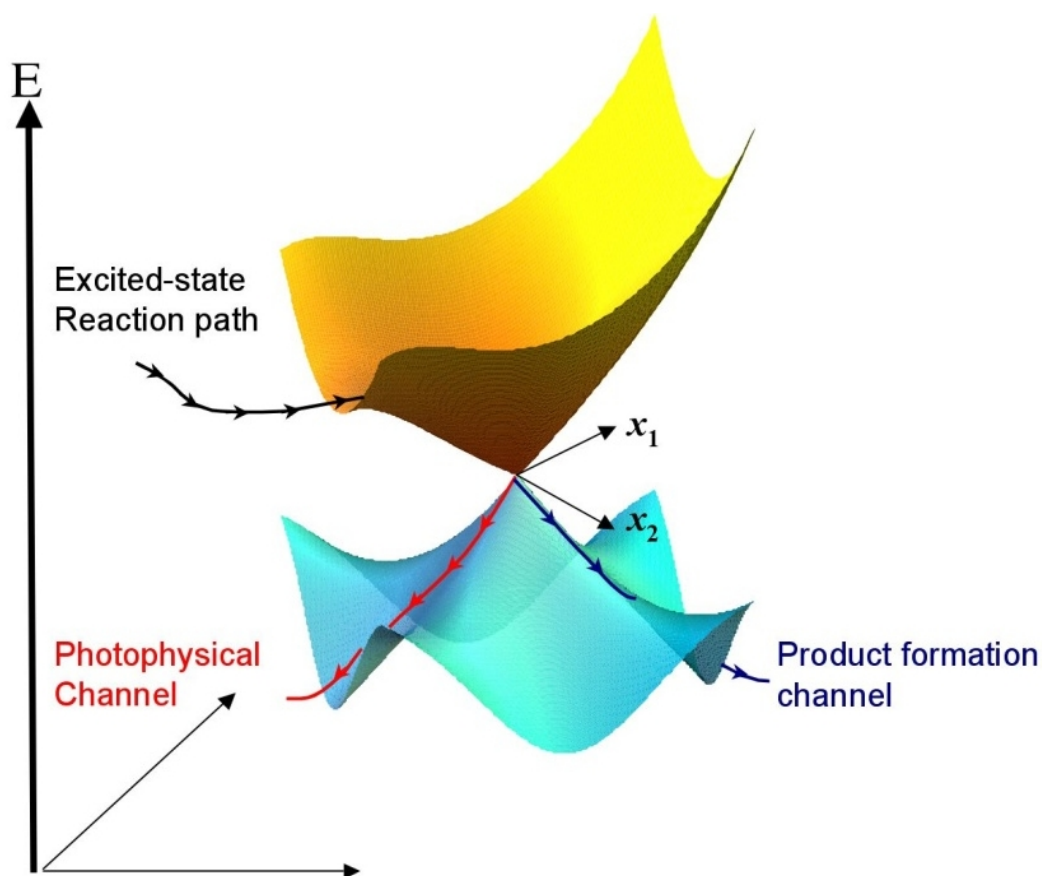


**Figure 3.4** There is represented the double cone conical intersection on the left side, while on the right side a typical view of two surface crossing along the reaction coordinate.

In conclusion the following statement is true: for a polyatomic system two states (even with the same symmetry) will intersect along a  $n-2$  dimensional hyperline as the energy is plotted against the  $n$  internal nuclear coordinates.

### [3.13 The Physical chemistry of conical intersections](#)

To understand the relationship between the surface crossing and photochemical reactivity, it is useful to draw a parallel between the role of a transition state in thermal reactivity and that of a conical intersection in photochemical reactivity<sup>10</sup>. In a thermal reaction, the transition state (TS) forms a bottleneck through which the reaction must pass on its way from reactants (R) to products (P). A transition state separates the reactant and product energy wells along the reaction path. An accessible conical intersection (CI) (fig.3.4) also forms a bottleneck that separates the excited state branch of the reaction path from the ground state branch. The crucial difference between conical intersections and transition states is that, while the transition state must connect the reactant energy well to a *single* product well via a single reaction path, an intersection is a “*spike*” on the ground state energy surface and thus connects the excited state reactant to *two or more* products (P and P’) on the ground state via a branching of the excited reaction path into several ground state relaxation channels. The nature of the products generated following decay at a surface crossing will depend on the ground state valleys (relaxation paths) that can be accessed from that particular structure.



**Fig 3.5** There is represented a “cartoon” of a classic double cone conical intersection, showing how it connects the excited reaction path with the two ground state reaction path.

Theoretical investigations of surface crossings have required new theoretical techniques based upon the “mathematical” description of conical intersections and we now briefly review the central theoretical aspects. The double cone shape of the two intersecting potential energy surfaces can only be seen if the energies are plotted against two special internal geometric co-ordinates of the molecule ( $x_1$  and  $x_2$ , see also section). The co-ordinate  $x_1$  is the gradient difference vector

$$x_1 = \frac{\partial (E_1 - E_2)}{\partial q} \quad (3.16)$$

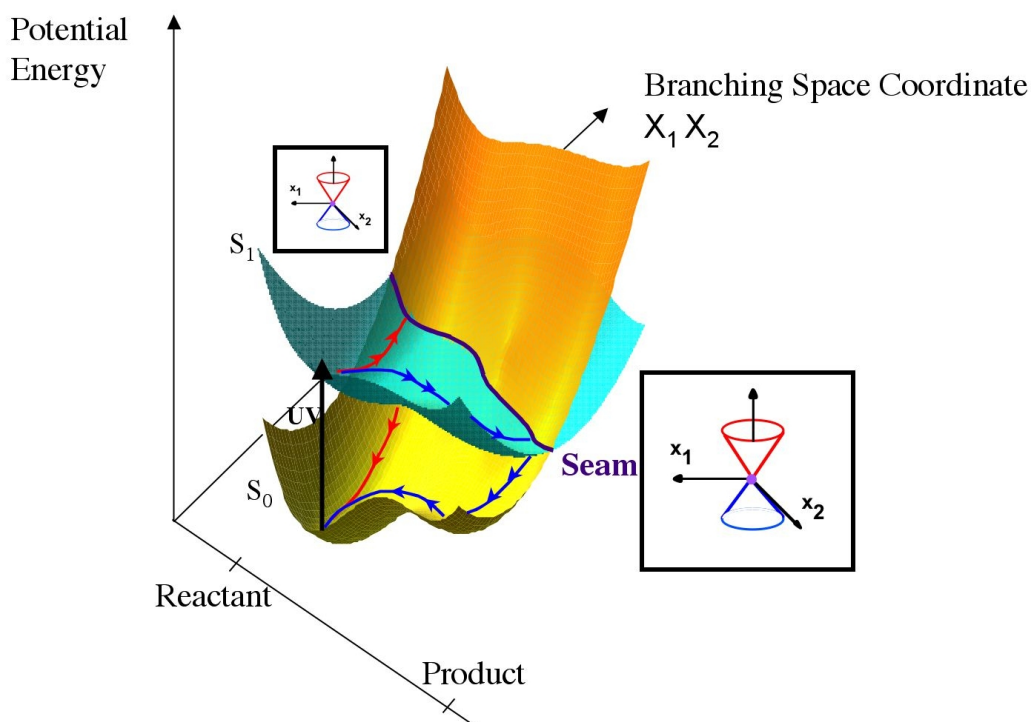
while  $x_2$  is the gradient of the interstate coupling vector

$$x_2 = \left\langle c_1^\dagger \left( \frac{\partial H}{\partial q} \right) c_2 \right\rangle \quad (3.17)$$

where  $\mathbf{C}_1$  and  $\mathbf{C}_2$  are the configuration interaction (CI) eigenvectors in a CI problem and  $\mathbf{H}$  is the CI Hamiltonian. The vector  $\mathbf{x}_2$  is parallel to the non-adiabatic coupling vector:

$$g(q) = \left\langle \Psi_1 \left| \frac{\partial \Psi_2}{\partial q} \right. \right\rangle \quad (3.18)$$

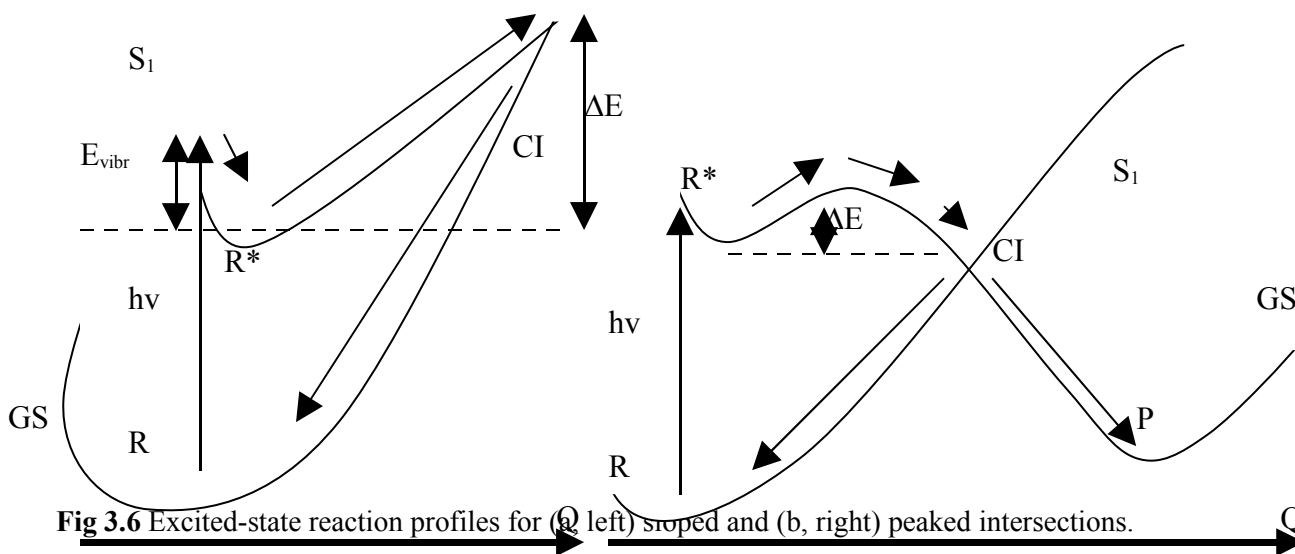
These geometric co-ordinates form the so called “branching space”. The derivative coupling (or nonadiabatic coupling) vector has a special significance when using a trajectory surface hopping algorithm since its magnitude and direction is a factor in determining where the hop occurs. In theory, trajectories can decay at any point on the intersection hyperline. However, if we assume that excited state trajectories will follow the lowest energy paths on the PES, then decay will most likely happen near the lowest energy point on the  $n-2$  dimensional intersection space. As we move in this plane, away from the apex of the cone, the degeneracy is lifted, and ground state valleys must develop on the lower cone. In contrast, if we move from the apex of the cone along any of the remaining  $n-2$  internal co-ordinates (where  $n$  is the number of degrees of freedom of the molecule), the degeneracy is not lifted. This  $n-2$  dimensional space, called “intersection space”, is a hyperline consisting of an infinite number of conical intersection points as it is possible to appreciate in Fig.3.5.



**Fig.3.5** The conical intersection hyperline traced out by a coordinate plotted in a space containing the same coordinate and one coordinate from the degeneracy lifting coordinate.

### 3.14 Conical Intersection Topology

Defining the reaction path through a conical intersection is a more subtle problem. So, as discussed in the previous section potential energy surfaces for excited-state reactions are the combination of several reaction profiles with different co-ordinates; these can be connected by a state crossing. For same state PES's (singlet-singlet crossings and triplet-triplet crossings) these features can be split into two main categories: Peaked crossings and sloped crossings as seen in the following fig.3.6



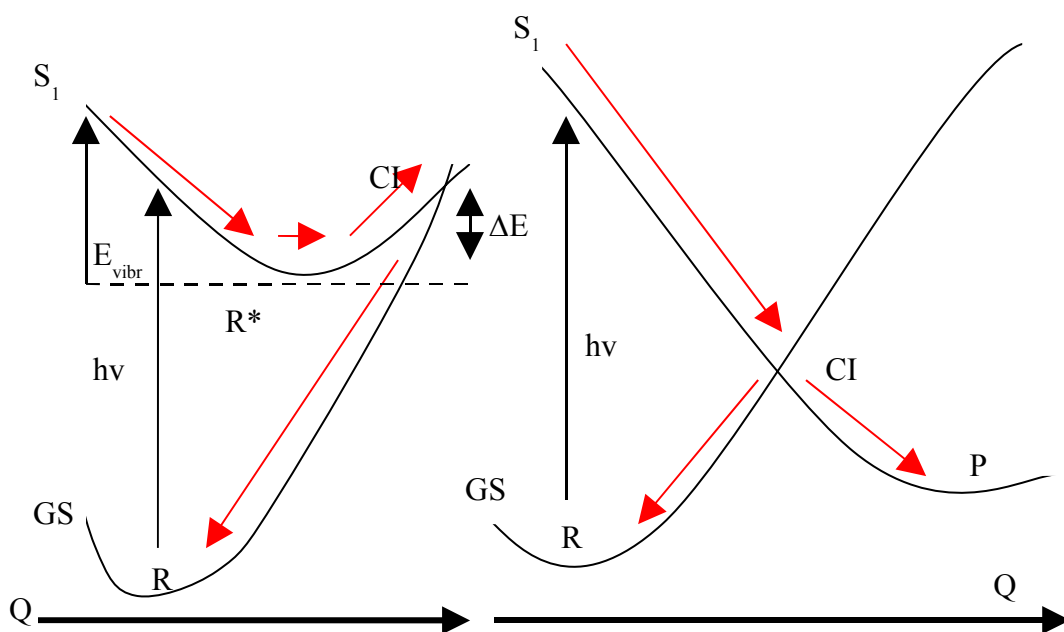
**Fig 3.6** Excited-state reaction profiles for (a, left) sloped and (b, right) peaked intersections.

The characteristics of these features enable us to predict the pathway of the reaction. The approach we take is in accordance with Kasha's rule's which state we consider that photochemical reactivity starts from the minimum of the lowest excited-state surface, and focus on how the products are formed from there. The path from the FC region to the excited state minimum ( $R^*$ ) will be discussed later.

The classification by Ruedenberg of the two categories is based on three relative directions of gradient and the positions of the minima for the crossing states. The peaked crossing has two almost orthogonal gradients and the connected minima lie on the same electronic surface ( $S_0$  in the example). The sloped crossing has gradients approximately parallel to each other and two minima on the same side of the crossing, on different electronic states (the two singlet states  $S_0$  and  $S_1$  in the present case)

A reaction path through a sloped crossing will generally lead back to the reactant,  $R$ , and so be photochemically unreactive, a photophysical process. Peaked crossings have a branched reaction path between  $R$  and  $P$ . The 'branching ratio' is a measure of the photoreactivity of the system, how likely the reaction is to form the product, this can be calculated through dynamics. Aborted CI's are those which lead the species to the reactant geometry, the path to  $P$  is interrupted at the crossing and the trajectory is directed back to  $R$ .

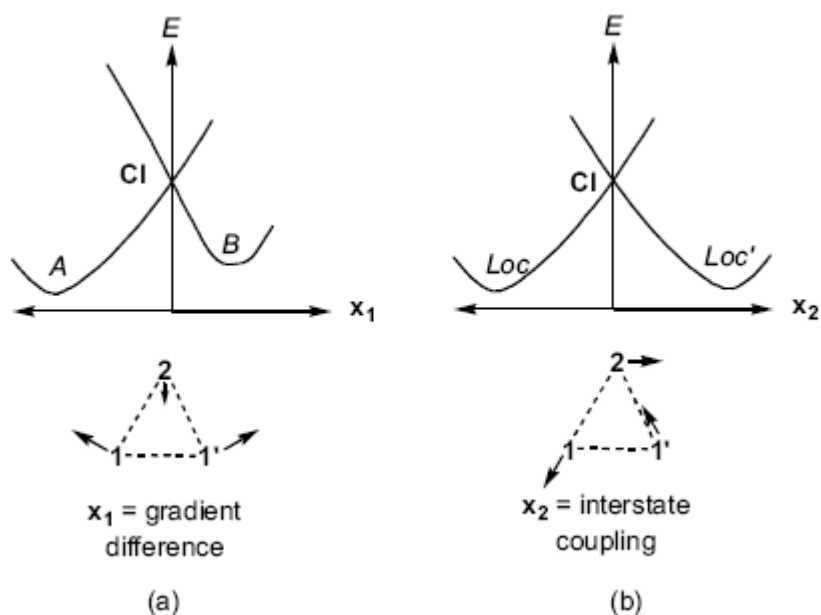




**Figure 3.7:** Reaction profiles for ultrafast excited-state reactions: sloped (a, left) and peaked (b, right) crossings.

In simple cases the reaction is determined by the energy barrier to the crossing region. In the sloped case this is energy difference between the reactant excited-state minimum  $R^*$  and the crossing. In the peaked crossing this is the energy barrier between  $R^*$  and the transition structure to the crossing. In reality the excited state reaction starts from the FC geometry rather than the minimum excited-state geometry. This means the molecule in question would be vibrationally hot, possessing additional vibrational energy ( $E_{\text{vibr}}$  in Fig 3.7(a)). This enables molecules to easily surpass energy barriers and can be the cause of ultrafast reactions. In the example of figure 3.7(a) this would result in an ultrafast reaction as this additional energy will cause the energy barrier to be insignificant. Ultrafast reactions can happen when there is no barrier, like in figure 3.7(b). The reaction profiles shown in figures 3.7(a) and 3.7(b) are only one dimensional, with one reaction co-ordinate.

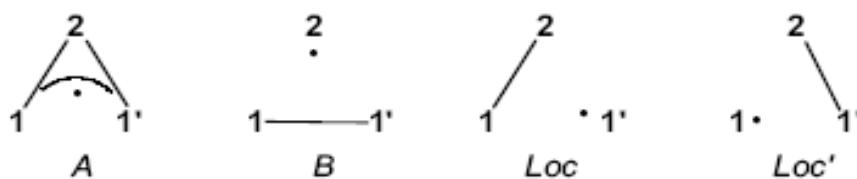
In reality only very few reactions can be represented with only one reaction coordinate. For most reactions the reaction is dependent on two important coordinates which form the branching space of the reaction.



**Fig 3.8** : Energy profiles at a peaked conical intersection along (a) gradient difference and (b) interstate coupling co-ordinates.

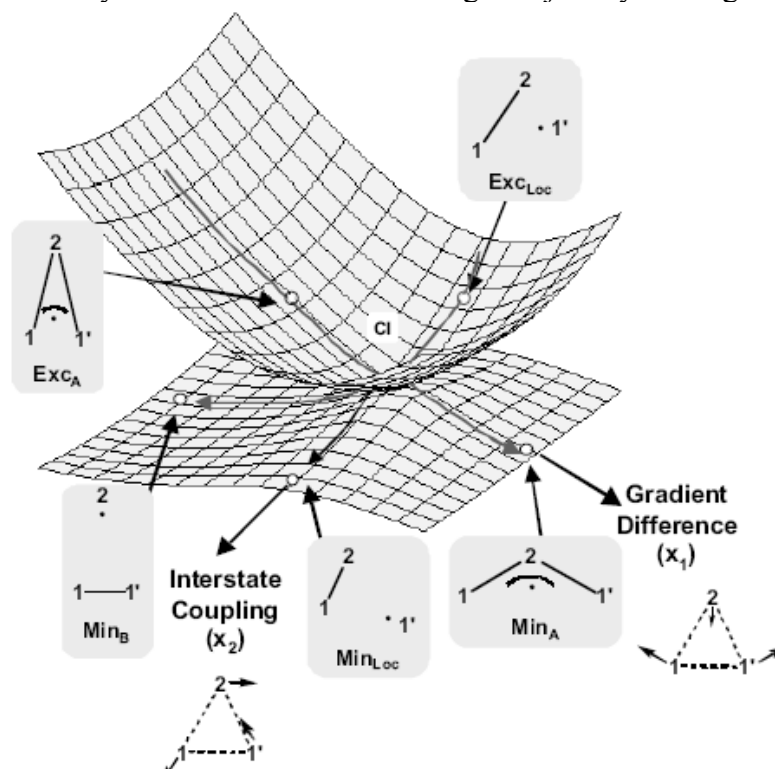
The remaining  $n-2$  coordinates form a hyperline, where displacement along it leaves the states degenerate and there are an infinite number of crossing points. The shape of the cone is important in helping to determine where the reaction path will lead to on the ground state surface, and so what product will be formed. Briefly, if the base of the cone is circular then all directions of decay are equally likely (neglecting the momentum carried from the excited state). However, if the base of the cone is elliptic then two sides automatically become more favorable paths for trajectories. More generally, a number of decay paths may exist at the apex of the cone, and others may develop further down the manifold, all influencing product formation. Prediction of photoproduct distribution depends on this topology, coupled with the effects caused by the kinetic energy possessed by the trajectories. Large amounts of kinetic energy can be gained on the ground state surface, and take trajectories far from the minimum energy paths. In this section there will be analyzed several kinds of crossings and the consequents relaxation channels occurring after the decay from the upper electronic state to the lower one.

To understand the relationship between structures on the excited state and ground state about a CI it is useful to see an example of a CI. We will use the example of an allyl species with 3 atoms and 3 electrons in a simplified version of the classic-three electron problem  $H_3$ . In this example there are two atoms the same labeled 1 and 1' and a third that is different labelled 2. There are two states with valence bond like character, shown below as A and B, with bonds between 2, 1' and 1 forming an allyl-type structure, and bonds between 1 and 1' respectively. The resonant structures of Loc and Loc' are linear combinations of states A and B. States A and B will be degenerate if the sides of the triangle are equal, this will thus form the model CI.



**Scheme 3.1** There are represented the possibilities of electronic distribution

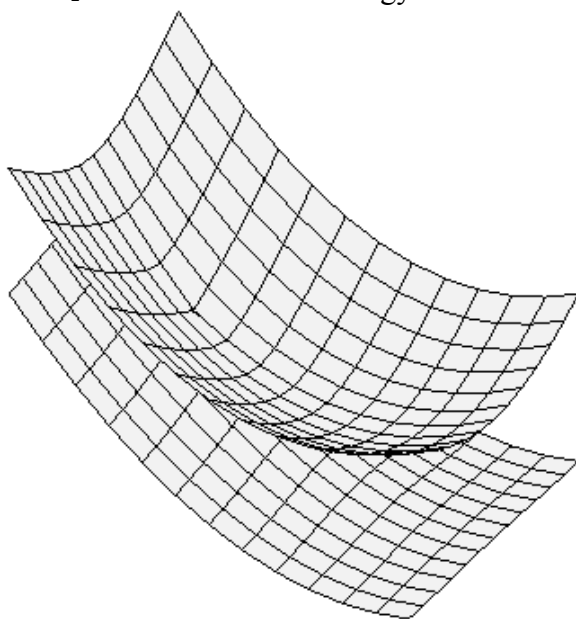
With displacement along the degeneracy lifting coordinate  $x_1$  stabilization of state A will occur in one direction of displacement and of B in the other direction of displacement. Similarly, with displacement in one direction of the degeneracy lifting coordinate  $x_2$  state Loc will be stabilized and in the other direction state Loc' will be stabilized. So plotting the two coordinates of this system gives rise to the conical shape of the CI. There will be four minima corresponding to the four states described, forming a moat about the crossing on the ground state. Since this is a symmetric case there will be four paths leading from the crossing to the four minima. The reaction profiles in figure 4 show cross sections of the CI, in one dimension, either  $x_1$  or  $x_2$ . The structures on the excited state surface are related to those on the ground state, but are electronically unfavorable. These electronic isomers will lead to their ground state counterparts if they are able to travel on a straight trajectory to the ground state.



**Fig 3.9** Potential-energy surface along two degeneracy-lifting co-ordinates  $x_1$  and  $x_2$  for a model peaked intersection.

The analysis of a sloped crossing is very similar to that of a peaked crossing. The difference between the sloped crossing and the peaked is that the former has only one minima (say  $\text{Min}_B$ ) that lies on the lower surface with the other (say  $\text{Min}_A$ ) lying on the upper surface. The sloped crossing possesses the same degeneracy lifting co-ordinates

$X_1$  and  $X_2$ . In this case the energy for the two states decreases along the same coordinate  $X_1$ , but by different amounts. There will be fewer ground state paths with  $\text{Min}_{\text{Loc}}$  and  $\text{Min}_{\text{Loc}}$  potentially disappearing due to a barrierless path to  $\text{Min}_{\text{B}}$ . (9)

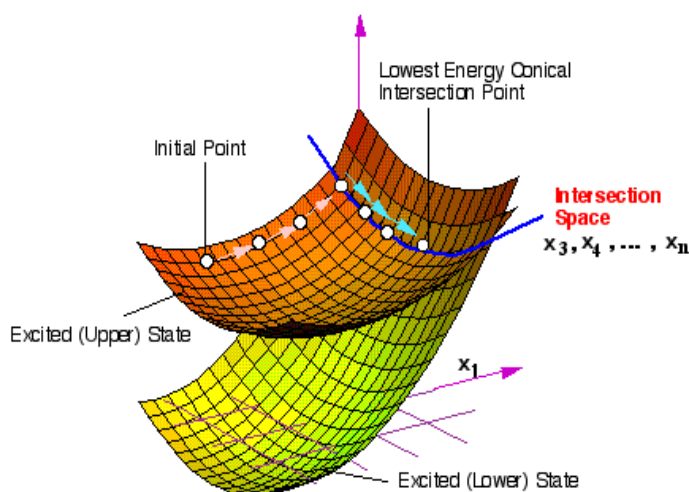


**figure 3.10** : Potential-energy surface along two degeneracy lifting co-ordinates  $x_1$  and  $x_2$  for a steepest sloped conical intersection.

These are an ideal cases not real cases, in the latter, minima may be embedded in the full space of coordinates and so be local minima or saddle points.

### 3.15 Conical Intersection Optimization

In the previous section it has been told that there are situations where there is no transition state connecting an excited state intermediate ( $R^*$ ) to the conical intersection point or where an excited state intermediate on the upper energy surface does not exist. In such situations, mechanistic information must be obtained by locating the lowest lying intersection point<sup>11</sup> along the  $n-2$  intersection space of the molecule (see Fig 3.5). The practical computation of the molecular structure of a conical intersection energy minimum can be illustrated by making an analogy with the optimization of a transition structure. As illustrated in Scheme 4.1a, a transition structure is the highest energy point along the path joining reactants to products and the lowest energy point along all the other  $n-1$  directions orthogonal to it. One can optimize such a structure by minimizing the energy in  $n-1$  orthogonal directions, and maximizing the energy in the remaining direction corresponding to the reaction path.



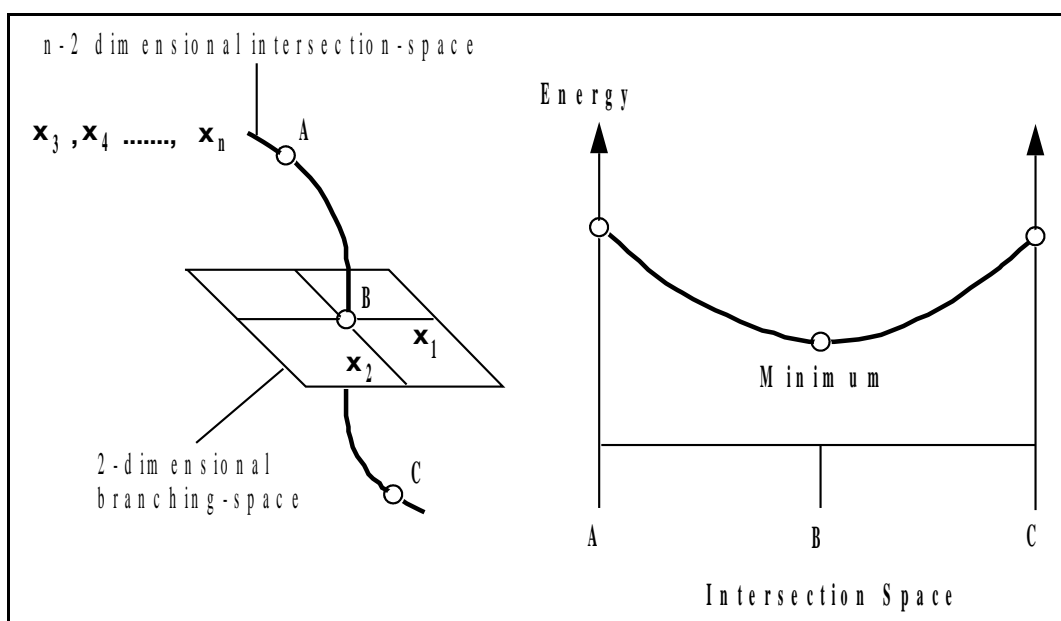
**Figure 3.11** In the figure there is represented a possible reaction path on excited state intersecting the  $S_1/S_0$  degenerate hyperline in a higher point of  $S_1$  Surface. Following the conical intersection optimization it will be located the minimum on the intersection space.

The technique for locating the lowest energy intersection point exploits the fact that the branching space directions  $x_1$  and  $x_2$  play a role analogous to the reaction path at the transition state. Accordingly, the lowest energy point on a conical intersection is obtained by minimizing the energy in the  $n-2$  dimensional intersection space ( $x_3, x_4, \dots, x_n$ ) which preserves the degeneracy (see Scheme 3.5). In practice, in order to properly locate these low-energy stationary points where two potential energy surfaces have the same energy, one must carry out constrained geometry optimizations where the geometry is optimized in directions orthogonal to the two directions  $x_1$  and  $x_2$ . It is important to appreciate that the gradient on the excited state PES will not be zero at an optimized conical intersection point, since it “looks like” the vertex of an inverted cone (see Scheme 3.11). Rather, it is the projection of the gradient of the excited state PES onto the orthogonal complement of  $x_1$  and  $x_2$  (*i.e.* the  $n-2$  dimensional hyperline) that goes to zero when the geometry of the conical intersection is optimized. This situation is distinguished from an “avoided crossing minimum” of two surfaces which is a real minimum where the gradient on the excited state PES would go to zero. Thus, in an optimized conical intersection point, conditions (i) and (ii) have to be fulfilled:

$$(i) E_1 - E_2 = 0$$

$$(ii) \frac{\delta E}{\delta x_3} = \frac{\delta E}{\delta x_4} = \dots = \frac{\delta E}{\delta x_n} = 0$$

**Scheme 3.2**



References:

1) Schlegel, H. B. In *Ab Initio Methods in Quantum Chemistry*; Lawley, K. P., Ed.; John Wiley & Sons Ltd.: New York, 1987;.

- 2) Foresman, J. B.; Aeleen, F. *Exploring Chemistry with Electronic Structure Methods*; Gaussian, Inc.: Pittsburgh, 1996; Chapter 3 and references cited therein.
- 3) Schlegel, H.B. *Adv. chem. Phys.* ,1987,67, 249.
- 4) Helgaker, T. *European Summer School in Quantum Chemistry. Book 2*Lund University, 2005.
- 5) McKee, M. L.; Page, M. in *Reviews in Computational Chemistry*; Lipkowitz, K. B., Boyd, D. B., Eds.; 1993, 4, 35.
- 6) (a) Köppel, H., Domcke, W.; Cederbaum, L. S. in *Adv. Chem. Phys.* 1984, 57, 59. (b) Smith, B. R.; Bearpark, M. J.; Robb, M. A.; Bernardi, F.; Olivucci, M. *Chem. Phys. Letters* 1995, 242, 27.
- 7) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J., J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03, Revision C.02*; Gaussian, Inc.; Wallingford CT; 2004.
- 8) Teller, E. *J. Phys. Chem.* ,1937,41, 109.
- 9) Herzberg, G. & Longuet-Higgins, H.C. *Trans. Faraday Soc.* ,1936,35, 77.
- 10) Bernardi, F.; Olivucci, M. & Robb, M.A. *Chem. Soc. Rev.* ,1996,25, 321-328.
- 11) a) Ragazos, I. N.; Robb, M. A.; Bernardi, F.; Olivucci, M. *Chem. Phys. Letters* 1992, 197, 217. (b) Bearpark, M. J.; Robb, M. A.; Schlegel, H. B. *Chem. Phys. Letters*. 1994, 223, 269.

# -CHAPTER 4-

## Conceptual aspects of photochemistry

### 4.1 What happens when a molecule absorbs light?

The photochemistry is the branch of chemistry studying all the chemical processes initiated from light absorption. In this section there will be briefly described the main concepts featuring the photochemical and photophysical events occurring after the photon absorption<sup>1</sup>.

To understand the photochemistry in terms of the potential energy surface could be useful comparing the photochemical reactions with the thermal reactions. Usually a thermal reaction is initiated by heat and so the reactant is separated from the product by a transition state located on the same and single potential energy surface (adiabatic process). Instead a photochemical reaction, which is initiated by light, is a process governed by the topology of two or more potential energy surfaces<sup>2</sup>.

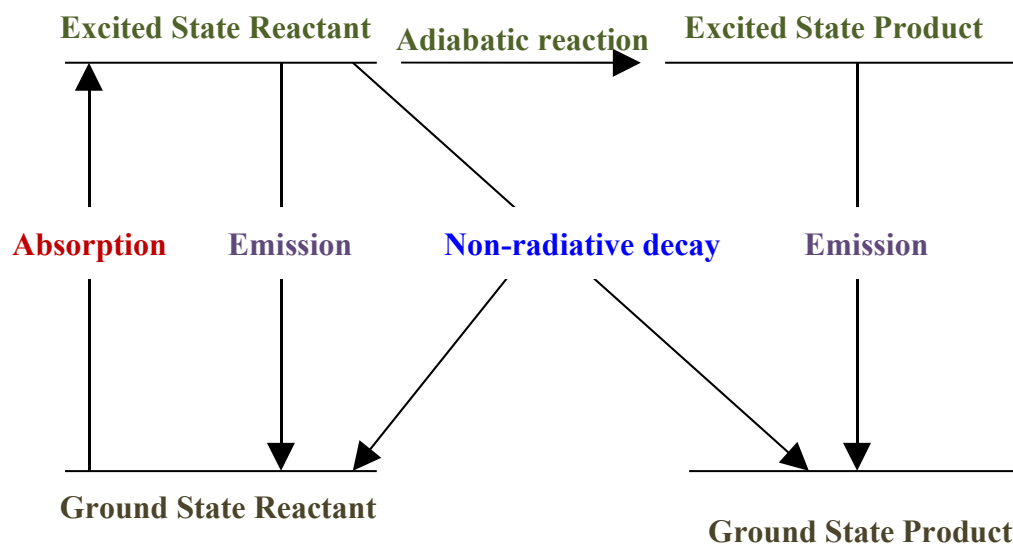
A molecule absorbs energy only in certain quantized amounts called quanta. The wavelength of the photon has to be accurately chosen so that it could correctly match a suitable electronic transition of the molecule: many electronic state are in fact available through suitable absorptions giving rise to different photochemical evolutions.

A molecule before absorption will be in its lowest energy orientation called its ground state ( $S_0$  if it is a singlet state), with all the electrons occupying the low lying orbitals. The unexcited molecule will be residing in a minimum on the potential energy ground state surface. In the local region of the PES the point at the bottom of the minima will be corresponding to the lowest energy structure, a stable structure. Thus being a critical point, the gradient of the surface at this point, is zero. It is possible for a molecule in a minimum to absorb light: the energy absorbed must be stored in some way. According to Franck-Condon principle, because electrons are very light, have negligible mass respect to the nuclei, they move instantaneously on absorption of light. The absorption of a photon, promotes the system from the ground  $S_0$  to an electronically excited state  $S_n$  (singlet states): on  $S_n$  the electrons move into orbitals of higher energy, therefore effectively absorbing and storing the energy. The quanta of light of the correct wavelength cause the molecule to move into an electronic excited state, higher in energy. Due to the heavy masses of the nuclei the structure is the same on the ground state as it is after absorption of the energy. Nevertheless an equilibrium geometry on  $S_0$  is no more a critical point on another state being subjected to a different force field. It

takes a while for nuclei to move compared to electrons due to their respective masses. Since the electronic rearrangement into the orbitals is completely different the system is basically an electronic isomer of the ground state, subjected to a different force field is driven by the new forces, undergoing a relaxation process on the excited state from the Frank-Condon (FC) region towards lower energy configurations.

It is experimentally known that relaxation on the spectroscopic state, from the Frank-Condon region, is an ultrafast process<sup>3</sup> (fs timescale) through which the molecule undergoes a transition to another electronic states. This happens because the system approach a region of degenerate or almost degenerate states (the spectroscopic one, where motion occurs, and a new one which, being higher in energy at the beginning of the relaxation, becomes isoenergetic with the first). Now most of the processes investigated in this thesis are characterized from the electronic transition between a ground state with a *covalent* nature and a *ionic* state featured from a clear charge separation. Thus the spectroscopic state for conjugated  $\pi$ -systems will be characterized by a single electron transition from HOMO to LUMO  $\pi$ -orbitals. For the studied polyenes such a real crossing happens between the spectroscopic ionic and the excited covalent states. It is useful to notice that these states usually have a different (spatial) symmetry and in this case the matrix element  $H_{11}$ , representing the *mixing* or “resonance” term between the two states, is *identically zero* ( $H_{11} = \langle \phi_1 | H | \phi_1 \rangle$ , where  $\phi_1$  and  $\phi_2$  are the electronic wavefunctions of the two states). Thus, due to the absence of mixing between the two states, only condition  $H_{12} = H_{21}$  has to be fulfilled to make the exchange term zero and have a (real) crossing (see formulae 3.10).

Once populated the excited state, after the FC vibrational relaxation, there are several mechanisms through which the molecule decays to the ground state: fluorescence, internal conversion or *radiationless* decay (Fig 3.1).



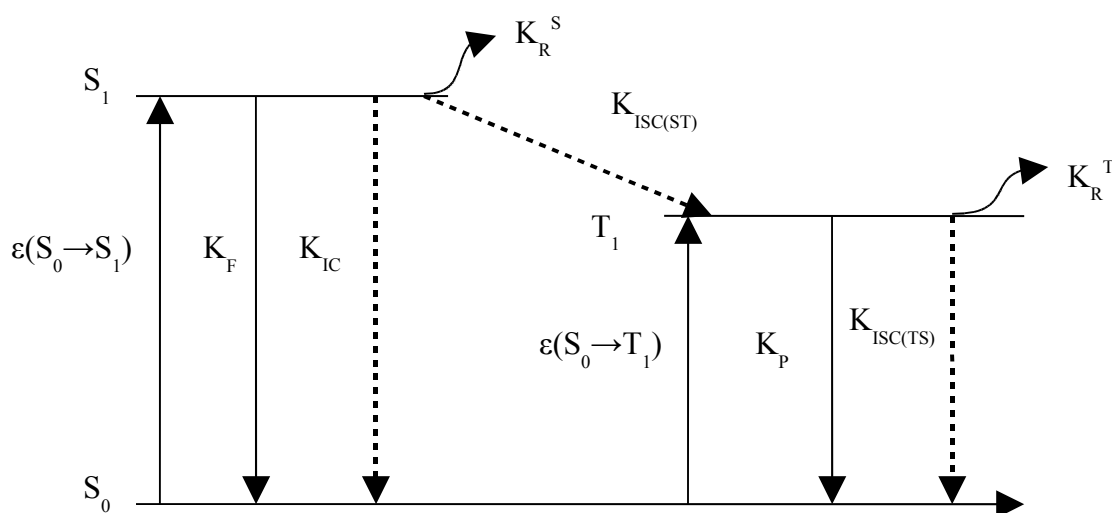
**Figure 4.1:** Photophysical vs photochemical processes

Photochemical processes are those which are initiated by the absorption of light by one of the reactants. Now, primary processes are those where the product is formed directly from the excited state of the reactant. For example the *cis-trans* photoisomerisation of the chromophore in Rhodopsin<sup>5</sup>. In polyenes<sup>6</sup> it is now very



well established that spectroscopic states have a lifetime much smaller than the covalent excited states where they fall. Although the initial motion of the system depends on the force field of the optically active state (ionic), after the ultrafast transition (through intersection point) to the lower covalent state, the motion is driven by the force field of this PES. Here most reactions (whose photoproducts are experimentally observed) happens, and in first approximation it can be stated that chemical reactivity depends on this state: thus this PES will be the most widely studied to understand the mechanisms of a photochemical reaction. Nevertheless, a knowledge of the primary events occurring immediately after the excited state is generated, allows us to have a deeper insight of the changes in bonding and molecular structure providing the initial driving force for the subsequent outcome of the process and the generation of final photoproducts.

Secondary processes are those which form their product directly from the excited state of a reactant. This happens in photosynthesis. In competition with these processes are photophysical processes which de-activate the excited state. The molecule can readily dissipate the energy in the form of light. The emission of light is a photophysical process. The emission can happen instantaneously, the emitted light is not often of the same frequency it is lower in energy. Energy can be dissipated from the molecules, in the form of vibrations. The Jablonski diagram shows the many possible pathways a molecule can take after it absorbs light.



**Figure 4.2:** The Jablonski diagram.

Radiative processes:

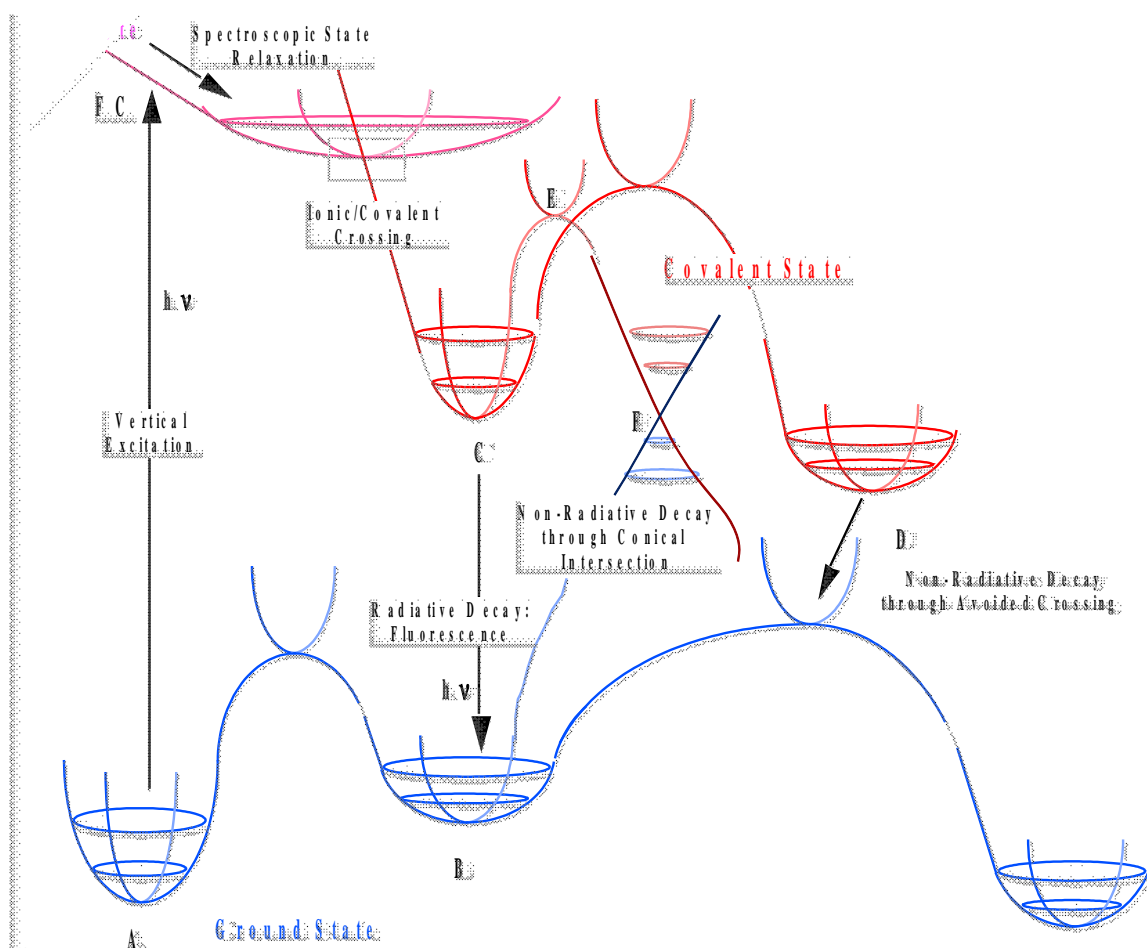
- Singlet-singlet absorption characterized by extinction coefficient  $\epsilon(S_0-S_1)$
- Singlet-triplet (forbidden) absorption characterized by extinction coefficient  $\epsilon(S_0-T_1)$
- Singlet-singlet emission: Fluorescence with radiative rate constant  $K_F$
- Triplet-singlet emission: Phosphorescence with radiative rate constant  $K_P$

Non-Radiative processes:

- $S_1 \rightarrow S_0 + \text{heat}$ : intersystem crossing, between states of same spin, characterized by rate constant  $K_{IC}$
- $T \rightarrow S + \text{heat}$ : intersystem crossing, between states of different spin, characterized by rate constant  $K_{ISC(ST)}$   $K_{ST(TS)}$

Photochemical processes:

- $K_R^T$  yield new structures in triplet state
- $K_R^S$  yield new structure in singlet state.



**Fig.4.3** General evolution of a photochemical reaction (see text for a discussion of these paths).

Despite its interesting to have a complete vision of all possible physical processes occurring upon light irradiation, at this stage it is necessary focus our attention on the specific processes with their electronic aspects<sup>7</sup> mainly investigated with computational photochemistry.

In figure 4.3 descriptions of how a molecule decays back down to the ground state from the excited state, are summarized. These processes can be divided into non-radiative decay and radiative decay. The first occurs through photon emission whose energy represent the energy gap between the two surfaces at the decay point (see point C). This is just right the opposite of the absorption process, following the same rules for vibrational overlaps (Frank-Condon factors) determining absorption probability. Non-

radiative decay processes use up energy in a different way, mostly it can be dissipated to neighboring molecules through rotational, vibrational and translational energy. The energy can be stored in a new structure with electrons into different orbitals. In the case of a photochemical process the energy is used up to form a new structure.

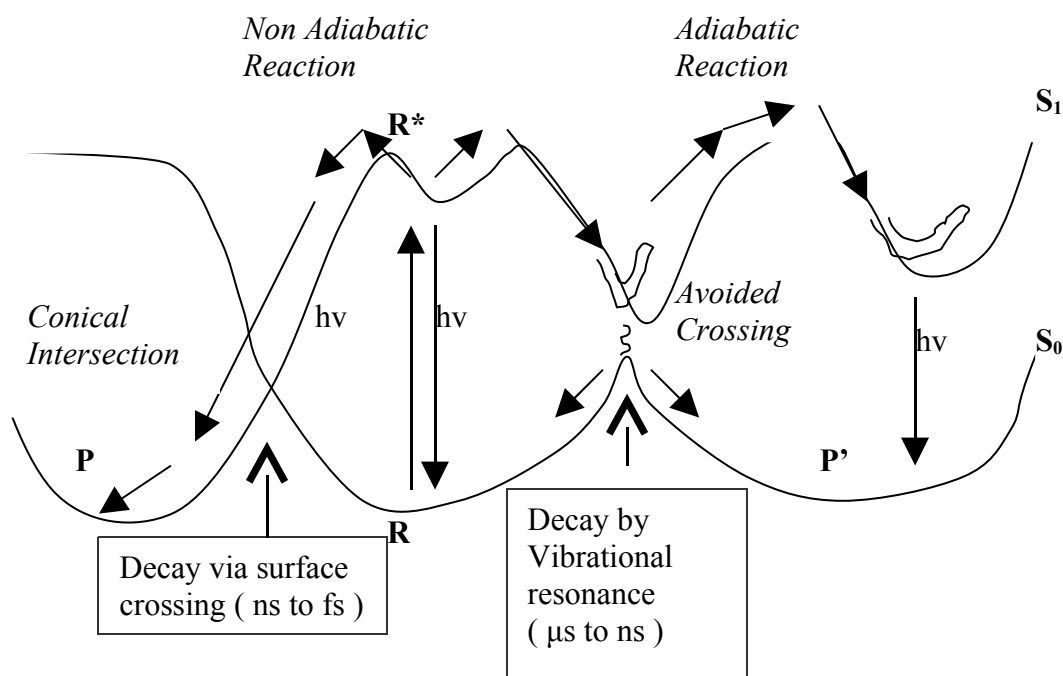
As briefly represented in Jablosky diagram the emission of light involving states of the same spin is called fluorescence, while involving different spin states is called phosphorescence. The latter involve a transition without emission<sup>8</sup> and may occur, for example, at an avoided crossing point (point D). Among the non radiative processes, the efficiency of internal conversion (IC), between states of the same spin multiplicity, and intersystem crossing (ISC), between states of the different spin multiplicity, is usually discussed in terms of the density of states and of the interaction between the vibrational energy levels of the ground and excited state PES using the Fermi Golden Rule. It results that may exist a probability for an *isoenergetic* transition from a vibrationally cold state (of the excited electronic state) to a degenerate (or almost degenerate) vibrationally excited state of the ground state. Then *vibrational relaxation* (VR) follows. Indeed, the vibrational excess energy of the molecule may be redistributed to all the degrees of freedom of the system (*internal vibrational energy redistribution*, IVR) or/and may be dissipated (heat production) by molecular collisions with solvent or other molecules. The decay may take place at an avoided crossing of the excited and ground state PES. At such an avoided crossing, if the energy gap is larger than  $1\text{-}2\text{ kcal mol}^{-1}$ , the intermediate  $R^*$  (on the excited state) will rapidly thermalize and the decay probability will be determined by the Fermi Golden Rule. Accordingly, such processes are supposed to occur on the same timescale (several molecular vibrations) as fluorescence and competition with emission processes may occur. Thus, such a mechanism is usually evoked in photochemistry to account for the observed radiationless decay, even when the energy gap between the involved electronic states is large. Nevertheless, if a real crossing (such as a conical intersection) is accessible to the system (for example by overcoming a small energy barrier, see points E and F in Fig 4.3) it is possible for a very efficient radiationless decay to occur: the probability for such process is very high because it depends on the energy gap between the states, being high when the gap is small (see section 4.3 and equations 4.4, 4.6). This process is similar to the one involved in the initial relaxation on the spectroscopic state, resulting in an ultrafast transition towards the covalent excited state; but in this case the real crossing occurs between states of even the same symmetry (and may occur also when symmetry is completely absent). Obviously, conical intersection search and optimization will be very important to explain all those photochemical processes involving ultrafast radiationless decays or when non-radiative<sup>9</sup> decays cannot be completely understood due to the large energy gap between the states.

## 4.2 Adiabatic and non-adiabatic processes?

With many reactions there is a barrier involved in the reaction co-ordinate of some sort, for instance thermal reactions this is overcome with heat, the bonds broken and reformed. Here the energy is used to push molecules in to configurations they do not want to go in, in order to form the precursor of the desired product molecule.

Photochemical reactions can be split into two main categories adiabatic and non-adiabatic reactions, the former meaning a reaction that happens on the same state (in the

case of photochemical reactions an excited state) and conversely the latter meaning the reaction occurs on a different state. Reactions are taken to be the formation of the product.



**Figure 4.4:** Adiabatic and non-adiabatic reaction pathways from absorption of light to ground state species

The adiabatic processes is usually described by a classical reaction comprising of an initial mechanism leading the photoexcited reactant from FC to an minimum  $R^*$  by overcoming the energy barrier to the transition state  $TS^*$ (Fig.4.5).

If  $R^*$  is a stable minimum on the excited state PES and, provided it has a suitable long lifetime, it will decay to the final ground state photoproduct  $P$  via an emission process (fluorescence). Thus, the overall reaction is a two step process involving first the “thermal” conversion  $FC \rightarrow R^*$ , and second the radiative decay leading to the final photoproduct  $P$ . Notably, all along the path the energy gap between the two PES is always quite large, involving automatically a small probability of radiationless decay. Thus all the reaction paths completing on the excited state PES, followed by radiative decay to the ground state, are called adiabatic.

Nevertheless in this thesis we have mainly interested to the systems are characterized by a non-adiabatic formation of the product since there is reportedly no emission of radiation during the change in conformation. There are two possible ways this can happen. Avoided crossings are where the surfaces don't actually cross but are so close that the structure according to the co-ordinates of the states are so close in energy that it effectively jumps to the state lower in energy. This is achieved in a single vibrational resonance. Conical Intersections (CI) are where the two states actually cross. The molecule is degenerate according to both states.

- In a polyatomic molecule two electronic states of the same symmetry are allowed to cross at a CI (a real surface crossing)
- Radiationless decay from the upper to the lower intersecting state occurs within a vibrational period when the system “travels” in the vicinity of such intersection points.
- CIs provide a common decay mechanism from the lowest excited states of polyatomics.

Thus such kind of photochemical processes, generating the final photoproduct directly from the excited state without emission, are called non-adiabatic.

Competition between adiabatic and non-adiabatic paths may occur in a photochemical process. Barrier heights will determine their efficiency. All this is more deeply discussed in the part regarding systems photochemistry where it is shown that barrierless non-adiabatic paths may account for the lack of fluorescence in fulgide, while the observed fluorescence in dihydronaphthalene may depend on the existence of a significant barrier to the radiationless path; in the part regarding retinal photochemistry it is shown that barrierless non-adiabatic cis-trans isomerization processes exist, thus accounting for the observed ultrafast non-adiabatic photoisomerizations.

### 4.3 The “Physical Chemistry” of Non-Radiative Decay

The precise nature of the molecular mechanism that controls radiationless decay in polyatomics is a problem that has intrigued photochemists and photophysicists for several decades. An excited state species decays non-radiatively via internal conversion (IC) to a state of the same spin multiplicity and via intersystem crossing (ISC) to a state of different spin multiplicity. In textbooks, the efficiency of IC and ISC is usually discussed in terms of the interaction between the vibrational energy levels of the two potential energy surfaces using the Fermi Golden Rule<sup>9</sup>:

$$K_{i \rightarrow f} = \frac{2\pi}{h} \left\langle \Psi_i \left| \hat{H} \right| \Psi_f \right\rangle^2 \rho_z \quad (4.1)$$

where  $K_{i \rightarrow f}$  is the rate of transition from initial (i) to final (f) states,  $\Psi_i$  and  $\Psi_f$  are wavefunctions of initial and final states and  $\rho_z$  is the density of states. Upon simplification this term reduces approximately to:

$$K_{i \rightarrow f} = \left\langle \chi_i \left| \chi_f \right. \right\rangle^2 \beta^{ISC/IC} \quad (4.2)$$

where  $\left\langle \chi_i \left| \chi_f \right. \right\rangle$  are Franck-Condon factors (i.e. vibrational overlaps) and  $\beta^{ISC/IC}$  is an electronic factor. The efficiency of IC and ISC is often discussed in terms of Franck-Condon factors and  $\beta^{ISC/IC}$  is assumed to play a minor role.

The Landau-Zener model provides an alternative semi-classical model for radiationless decay. As shown by Desouter-Lecomte<sup>10</sup> and Lorquet, the probability of radiationless decay is given as:

$$(4.3) \quad P = \exp\left[-\frac{\pi}{4}\xi\right]$$

where  $\xi$  is the Massey parameter given as

$$\xi = \frac{\Delta E(q)}{\frac{h}{2\pi}|q|g(q)} \quad (4.4)$$

where  $\mathbf{q}$  is a vector of nuclear displacement co-ordinates. The term  $g(\mathbf{q})$  is the non adiabatic coupling matrix element defined as

$$g(q) = \left\langle \Psi_1 \left| \frac{\partial \Psi_2}{\partial q} \right. \right\rangle \quad (4.5)$$

while  $|q|$  is the magnitude of the velocity along the reaction path  $q$  and  $\Delta E$  is the energy gap between the two states and . Alternatively, the analogous Landau-Zener formula may be employed

$$P = \exp\left[-\frac{\pi^2 \Delta E^2(q)}{h|q|\Delta s}\right] \quad (4.6)$$

where  $\Delta s$  is the asymptotic slope difference of the two surfaces.

It is possible to make some general estimates on the values of the different parameters in equations 4.3 and 4.6 and we find that, unless  $\Delta E$  is less than 1-2 kcal mol<sup>-1</sup>, the decay probability is vanishing small. In the “classical” photochemical reactions (involving conjugated systems) we are interested on, the energy gap between the excited ( $S_1$ ) and ground ( $S_0$ ) states at an avoided crossing minimum is always quite large (10-30 kcal mol<sup>-1</sup>) and thus it is very unlikely that ultrafast radiationless decays may happen from avoided crossings. However, as we approach a point where the surfaces actually cross (a conical intersection), the decay probability becomes unity ( $P=1$  from equations 4.3 and 4.6 for  $\Delta E=0$ ) and a very efficient non-radiative decay may occur.

To avoid unjustified generalizations, we should mention here a case where one should not infer an easily accessible conical intersection from an ultrafast rate: if two potential surfaces, for example of two excited states, are close in energy and only slightly shifted along a coordinate, their intersection will be clearly above the minimum (i.e.  $\Delta s \sim 0$  in

equation 4.6 and therefore  $P \sim 0$  along this path). But tunneling below this intersection can be ultrafast, as predicted by the matrix-element approach using the Fermi-Golden Rule 4.1, since it is only a small energy which must be converted to vibration<sup>11</sup>. Whereas such nearly parallel surfaces with small distance may be frequent for pairs of excited states, they are not expected for  $S_1$  and  $S_0$  in conjugated systems photochemistry, as it is the case for the “classical” photochemical reactions studied in this thesis.

In previous chapter it has been shown that, for a polyatomic system, two states (even with the same symmetry) intersect each other along an hyperline (a  $n-2$  dimensional intersection space). Competitive paths, driving the system towards such crossing points, may exist so opening a way to the ultrafast radiationless decays often observed in conjugated systems. In the next parts the computational results obtained for such systems will be presented, in conjunction with the available modern experimental evidences. The aim is to provide a new insight for ultrafast photochemical processes, thus increasing our general knowledge of the primary events controlling these reactions.

#### 4. 4 Ultrafast processes

The development of ultrashort laser pulses has made it possible to follow the time evolution of fast photoinduced processes, including excited state relaxation and chemical reactions. Among all the processes here analyzed, the ultrafast (subpicosecond) photochemical processes are recently subject of intensive experimental research. These experimental methods are increasing our general knowledge of the “primary events” which control the final outcome of the absorption of a photon. This modern experimental measurements together with new computational techniques are now providing results that support a new view for understanding organic photochemistry.

Issues such as the efficiency of internal conversion (IC) at a surface crossing, the competition with fluorescence when an excited state barrier is present, and the relationship between the molecular structure at the surface crossing point and the structure of the photoproducts provide the intellectual motivation for the experimental and computational investigation of various photochemical reactions. The rate and the energy thresholds controlling IC can now be experimentally measured by exploiting the advances in laser spectroscopic techniques which have pushed the time resolution of various experimental techniques below the picosecond timescale. Thus very fast IC processes and short excited state lifetimes can now be detected. For example, many detailed experiments are available in the photophysics and photochemistry of conjugated hydrocarbons in solution or isolated conditions. Femtosecond excited state lifetimes have been observed for simple polyenes, and opsin-bound retinal Protonated Schiff Bases (PSB). Experiments on isolated molecules in cold-matrices or expanding-jets have revealed the presence of “thermally activated” fast radiationless decay channels in polyenes and aromatic compounds. Time-resolved studies on such polyenes are of fundamental importance in photochemistry because they unveil the changes in bonding and molecular structure occurring immediately after the excited state is generated and the way in which these changes provide the driving force for the generation of the photoproducts. While laser experiments provide information on the structure and energetics of the excited state potential energy surfaces controlling fast decay, more traditional photochemistry, such as quantum yield measurements, provide information on the molecular structure of the decay channel and on the product formation paths. The detailed characterisation of the stereo- and regio-chemistry of the

primary photoproducts and transient intermediates, their quantum yields and the effect of specifically designed sterically and rotationally hindered reactants on these quantities<sup>12</sup> is now possible.

Molecular dynamics whose consideration is stimulated by modern ultrafast spectroscopy, as well as progress in quantum chemical computations, give new impact to the approach which follows the detailed pathway of the molecule along the potential energy surfaces from the excited to the ground state. This approach is part of a more general way of considering photochemistry which is already employed and it follows the pathway on the surfaces, providing a huge number of details such as slopes, barriers, saddle points, collecting funnels. Thus through this method, called the pathway approach, ultrafast processes are often described very intuitively by the motion of wavepackets on potential surfaces. This has an intimate connection to the pathway approach. Continuous pathways are typical in this method. Although not new, it is receiving new impact by femtosecond spectroscopy and by the advances of quantum chemistry. As opposed to it, we call by matrix-element approach the classical way to describe the jump (tunnelling) between potential surfaces by help of interaction matrix elements (typically between stationary wavefunctions) and by densities of states, using the Fermi Golden Rule, proposed by Fuss et al.<sup>13</sup>.

Moreover, from new quantum chemical computations, with support of modern experimental evidences, we obtain in some cases a very detailed picture of ultrafast photochemical processes and all this is now providing results that challenge the classical models (matrix-element approach) of photochemical reactivity.

We can summarize and further characterize what can be expected for ultrafast photophysical and photochemical processes which involve at least two surfaces (except if they are nearly parallel and energetically close-by):

1. The potential energy surfaces will have a collection funnel in the form of an easily accessible conical intersection where a very efficient radiationless decay may occur (this point connects the excited state part of the photochemical reaction pathway to the ground state part).
2. The molecule moves along these surfaces, guided by the steepest descent path. Local properties of the surfaces (minima, transition states, slopes, curvatures) drive the motion of the system.
3. Many photochemical processes have the initial part of their pathway in common. They branch at the position corresponding to the collection funnel. The relaxation reactions on the ground state lead to the final photoproducts or lead back to the reactant. Thus the characteristic of this approach is that local properties of the surface are considered instead of nonlocal quantities such as integrals over stationary wavefunctions (matrix-element approach). Obviously there is a natural relation between this local approach and the idea of a wavepacket moving on a potential energy surface after excitation<sup>14,15</sup>. Hence, after the recent progress in femtosecond chemistry this approach will probably enjoy increasing attention.



## References:

- 15) Nicholas J. Turro 'Modern Molecular Photochemistry', Yoshizawa, T.; Kuwata, O. In CRC Handbook of Organic Photochemistry and Photobiology;
- 16) Bernardi, F.; Olivucci, M. & Robb, M.A. *Chem. Soc. Rev.*, 1996, 25, 321-328.
- 17) Zewail, A. H. *Femtochemistry: Ultrafast Dynamics of the Chemical Bond*; World Scientific: Singapore 1994; Vols. I and II.
- 18) Zewail, A. H. *J. Phys. Chem.* 1996, 100, 12701-12724
- 19) Palmer, I. J.; Ragazos, I. N; Bernardi, F.; Olivucci, M.; Robb, M. A. *J. Am. Chem. Soc.* 1992, 115, 673.
- 20) a) M. O. Trulson; R. A. Mathies, *J. Phys. Chem.* 1990, 94, 5741. (b) Reid, P. J.; Doig, S. J.; Wickham, S. D.; Mathies, R. A. *J. Am. Chem. Soc.* 1993, 115, 4754 and references cited therein. (b) Pullen, S.; Walker II, L. A.; Donovan; B; Sension, R. J. *Chem. Phys. Lett.* 1995, 242, 415. (c) Cyr, D. R.; Hayden, C. C. *J. Chem. Phys.* 1996, 104, 771.
- 21) Michl, J.; Bonacic-Koutecky, V. *Electronic Aspects of Organic Photochemistry*, Wiley: New York, 1990 and references cited therein.
- 22) Trulson, M. O.; Mathies, R. A. *J. Phys. Chem.* 1990, 94, 5741. (b) Cyr, D. R.; Hayden, C. C. *J. Chem. Phys.* 1996, 104, 771.
- 23) Gilbert; A.; Baggott, J. *Essentials of Molecular Photochemistry*; Blackwell Scientific Publications: Oxford, 1991.
- 24) Desouter-Lecomte, M.; Lorquet, J. C. *J. Chem. Phys.* 1977, 71, 4391
- 25) Bixon, M.; Jortner, J. *J. Chem. Phys.* 1997, 107, 1470.
- 26) W. J. Leigh; A. Postigo *J. Chem. Comm.* 1993, 24, 1836.
- 27) Fuss, W.; Kompa, K. L.; Lochbrunner, S.; Muller, A. M. *Chem. Phys*
- 28) Domcke, W. *Adv. Chem. Phys.* 1997, 100, 1 and references cited therein.
- 29) 18. Manz, J. *Femtochemistry and Femtobiology*; Sundstrom, V., Ed.; World Scientific: Singapore, 1997.

# -CHAPTER 5-

## Fulgides Photochromism

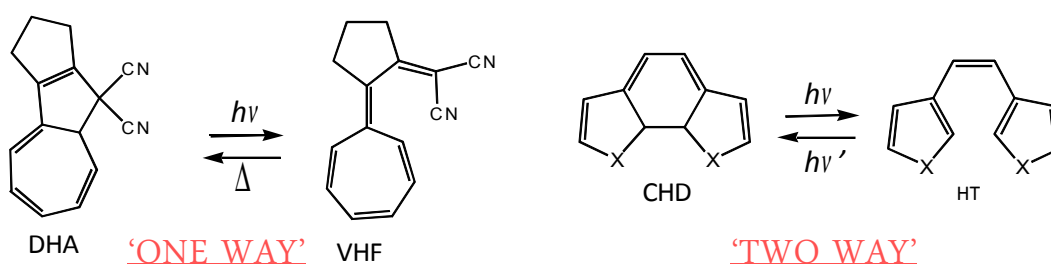
In this section (experimental section) there will be presented and commented the results relative to photochemical properties and light-induced events of the molecules analyzed during the Phd. Starting from this chapter (chapter 5) a computational analysis of the low lying electronic state was computed in order to understand the origin of fulgide photochromism efficiency. The more complicated photochemical behaviour of Stilbene, will be subject of chapter 6. At the end the last two chapters, are dedicated to rhodopins photochemistry: electrostatic effects and “level effect” on the photo-isomerization path were examined respective in chapter 7 and 8.

### 5.1 Introduction

In this chapter 2-5, dimethyl-furyl-fulgide photochromism has been computationally investigated by exploring the surface of the lowest excited state trying to explain its particular efficiency in terms of the mechanistic picture reported.

Photochromism<sup>1</sup> is a reversible, light-induced, transformation of a single chemical species between two states, the absorption spectra of which are clearly different<sup>2</sup>. Several organic compounds<sup>2</sup> show this behaviour after irradiation at appropriate wavelengths: denominated “photochromic materials<sup>2</sup>” (because of the colour changing reversibly) they have been widely studied, due to their huge potential industrial applications such as optical recording and optical switching devices, or data storage<sup>4</sup>. At the microscopic level, photochromism originates from a chemical transformation triggered by a photon absorption, such as ring-closure or ring-opening, the widest and the most important group of photochromic system is that based on pericyclic reactions<sup>5</sup>. There are mainly two photochromic species, called ‘one way’ or ‘two way’<sup>2,6,7</sup>.

In this class of compounds, the photochromic transformation arises from different photochemical process (hydrogen shift, electrocyclyzation, cis-trans or trans-cis isomerization). Of the many pericyclic reactions, electrocyclizations have proved to be especially suitable as a basis for photochromism. So if compounds like fulgides<sup>8</sup> or diarylethenes<sup>9</sup> are ‘two way’ photochromic systems, being photochemically converted both to the open and the closed form when irradiated at suitable different wavelengths, compounds like the dihydroazulene<sup>10</sup> or dihydronaphtalene<sup>11</sup> (Fig.5.1) are examples of ‘one way’ photochromism since the back reaction can occur only thermally. Of the many pericyclic reactions, electrocyclizations have proved to be especially suitable as a basis for photochromism. The  $4n+2$  systems whose photochromism results from the electrocyclic 1,3,5-hexatriene/1,3-cyclohexadiene interconversion<sup>12</sup> comprise a very large number of species.



**Fig.5.1** There are briefly represented two examples of photochromic systems for both the classes: so while on the right side there is a couple (CHD-HT) of diarylethenes thermally irreversible, on the left side there is the couple of dihydroazulene converted to vinyleptafulvene by a process thermally reversible.

The fulgides belong to the second mentioned class of photochromic compounds for which the transformation is largely driven photochemically and consequently the thermal reversal or decoloration of the colored species is forbidden (for this reason they are classified as “two way” photochromic systems). For their features and favorable photochemical properties, fulgides are photochromic systems particularly investigated for the potential industrial application as aberchrome molecules.

The discovery of fulgide systems has to be collocated around one century ago and is due to Hans Stobbe<sup>13</sup>, who observed for the first time triphenyl-fulgide, changing color from yellow-orange to brown, under light exposition, while turning back to the original color under the dark light. Initially Stobbe supposed that the origin of these special properties was correlated with some kind of crystal effect<sup>14</sup>; nevertheless he changed idea while he noticed how such similar properties characterized also similar compounds. At the beginning the first reactions analyzed were electrocyclic reactions converting diphenyl-dimethyl-anydhride-succinic to the 1-8 dihydronaphtalene derivatives<sup>15</sup>.

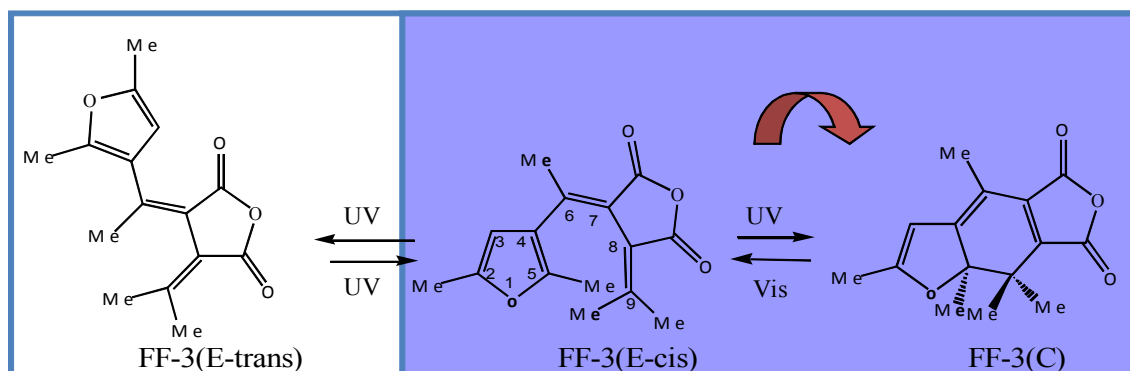
Nevertheless dihydronaphtalene derivatives are characterized from a low quantum yield of photo-conversion and also being subject to (photo-degradation) they are not good photochromic systems.

On the other hand, among the several species of fulgides, the most interesting species are those called “aromatic fulgides” or “heterocyclic-fulgides (as indolyfulgide, furylfulgide), because of its high efficiency as potential photochromic systems. Aromatic fulgides are distinguished depending on the kind of system condensed on the cyclohexadiene-succinic-anidrid moiety, whose photochromism results from the electrocyclic 1,3,5-hexatriene/1,3-cyclohexadiene interconversion: available experimental data are mainly concentrated on the furyl-classes.

In this work computational investigations have been performed on the low lying singlet states, to investigate the efficient photochromism of 2,5-dimethyl-furyl-fulgide system (FF-3), characterized by the light-induced reaction (due to the  $4n+2$  electrocyclization of the hexatriene/cyclohexadiene moiety) between the colorless open forms (the E-form) and the photocyclized colored form (the C-form). Both the processes were examined: the closure-ring reaction taking place from E-isomer as well the C-isomer open-ring reaction. The cis-trans isomerization is associated to another photochemical process (thermally irreversible) not investigated in this thesis (see Fig.5.2).

Despite the analogies with some aromatic polyene related, this system show a different photochemical behaviour and computations reveal that this is due to a different nature

of the excited state  $S_1$  driving its photoreactivity and triggering its deactivation to the ground state.



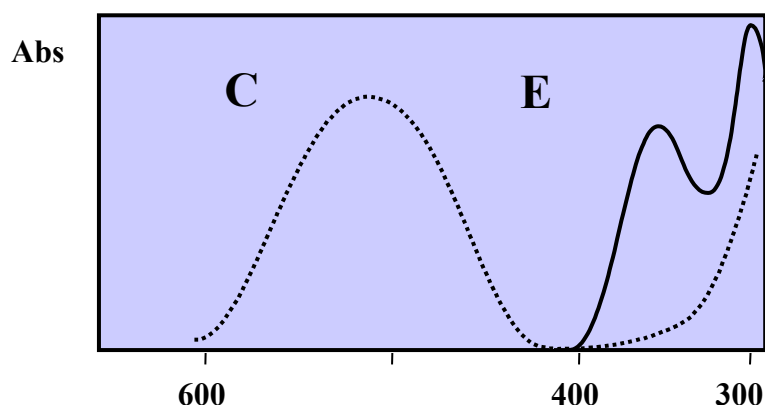
**Figure 5.2** Light-driven electrocyclic ring-closure and ring-opening reactions of the 2,5-dimethyl-furyl-fulgide.

The design of efficient, thermally stable fatigue-resistant photochromic systems was first pursued by Heller with 2,5-dimethyl-3-furyl-fulgide through an analysis for example, of the substituents influence on the electrocyclic reaction<sup>16</sup>. Nevertheless Yokohama was the first one able to conclude how an increment of the substituents dimensions was directly connected with an increment of the photochromic quantum yield hindering the cis-trans isomerization<sup>15</sup>. In fact substitution (e.g. by methyls) of the hydrogen atoms into the ring closing carbons achieves the two-fold effect of eliminating fatigue (as hydrogen-shift side reactions are now prevented) and creating thermal irreversibility (thermal disrotatory ring-opening back reactions are prevented for steric reasons), while employing a heteroaromatic ring allows high conversion efficiency due to both clearly separated absorption bands with high extinction coefficients and high reaction quantum yields (QY). In particular it has to be mentioned that a different heteroatom may influence all the entire photochromism of the species; for example there are experimental evidences proving that with a pyrrole condensed on the hexatriene moiety, its photoreactivity has to be controlled by a state of triplet multiplicity.

Recently the just mentioned fulgide (FF-3) has been subject of primary attraction as possible organic molecules for several kind of applications in the nanoscience field. In particular an experimental paper showing a structuring technique to create a fulgide-based memory device was published the last year (2007) on NanoLetters<sup>17</sup>. It is shown the fabrication of organic nanodots from photoswitchable fulgide molecules, switched between the two isomeric molecular conformations reversibly. The molecule used in these experiments (Phenyl-Tiophen-Fulgide) is very similar to the one we studied computationally (differing only for sulfur instead of oxygen and for the benzene ring connected to the tiophenic ring). Thus through fulgide nanodots it was demonstrated the optical single-access with memory application: for example while the E-isomer is associated with the bright dots storing “1”, the C-isomer is associated with the dark dots storing “0”. The reason of choosing such system is for its extremely high efficiency as potential photochromic systems and thus switch controls (featuring the classes fulgide who belong it).

Despite this last application is probably one of the most interesting, the experimental data do not miss about this system: there is an extremely huge number of papers and reviews documenting the photochemical properties of this system in order to control and maximise its photo-efficiency. Heller<sup>18</sup> was one of the first interested to examine, for example, the substituents influence on the electrocyclic reaction, but Yokohama<sup>19</sup> was the one able to conclude how an increment of the substituents dimensions was directly connect with an increment of the photochromic quantum yield hindering the cis-trans isomerization.

Characterization of the photodynamical behaviour of FF3 has been also pursued by transient spectroscopic measurements<sup>20</sup>. Firstly it has been shown that the reaction's quantum yield of 0.2 ( increasing to 0.6 when isopropyl group are positioned on C6 instead of methyl groups) is not particularly high but it was considered to be independent from solvent effect in which the reaction takes place, interesting feature for using fulgides as aberchrome molecules<sup>21</sup>. However what really makes this photochromic system particularly suitable<sup>22</sup> for several industrial applications (optical data storage media and as molecular functional units such as optical switches) is the thermal stability of both isomers at room temperature. Further, the only one photoproduct observed, full reversibility of the photoreactions without photochemical fatigue and high conversion efficiencies (see fig.5.3) are favorable for several applications.



**Figure 5.3** In this absorption spectrum it is possible distinguish clearly the two absorption bands referred to the two isomers: E and C. On the right side there are the two peaks of 300 and 366 nm featuring the open-ring isomer E absorption: despite both have an high coefficient extension only the 366 nm value seems to be responsible of the photo-conversion to C. On the left side there is the 516 nm peak of the close-ring photoproduct C.

In particular, we were interested to kinetic constants and decay reaction velocities in order to rationalize the origin of such efficient photochromism, by exploring  $S_1$  potential energy surface. From time resolved transient absorption measurements<sup>20</sup> a fast decay of the excited E isomer ( $E^*$ ) absorption (within 600 fs) and the growth of a narrow band, identified with the  $C_0$  ground state absorption, were observed. The rise time of this second absorption band, that indicated the production of C isomer, was 60 ps (in acetonitrile). To explain the reason of the rise time of the C isomer band, differing in a significant way from the decay time of the  $E^*$  transient absorption, the existence of an intermediate (I) on the reaction pathway  $E^* \rightarrow C_0$ , was proposed. The obtained results



**Eq.5.1** There are reported the two main reaction paths  $E^* \rightarrow C_0$ , with the respective population P1 and P2, and its rapport that almost equal to one indicates how both the processes were equally populated.

These last two relaxation paths both conduct to the  $C_0$  isomer: while in the first case  $E^*$  reacts straightforwardly to form the product, in the second one the intermediate I plays the role of a bottleneck reducing the velocity of the  $E^* \rightarrow C_0$  transition and, thus, providing an explanation for the multiexponential increase in the  $C_0$  absorption band. Further from the intermediate I it is possible also to suppose that part of the excited population is directed to the reactant  $E_0$ . Despite the huge experimental evidences available for such system, no real computational model are proposed for explain the FF-3 photoreactivity.

Here for the first time an high-level computational model of the high photo-efficiency of such system, was provided in order to elucidate the origin of the ultrafast mechanism of the photoinduced electrocyclization. This was accomplished through mapping the photochemical reaction paths on the first excited state and the following relaxation channels after the radiationless decay.

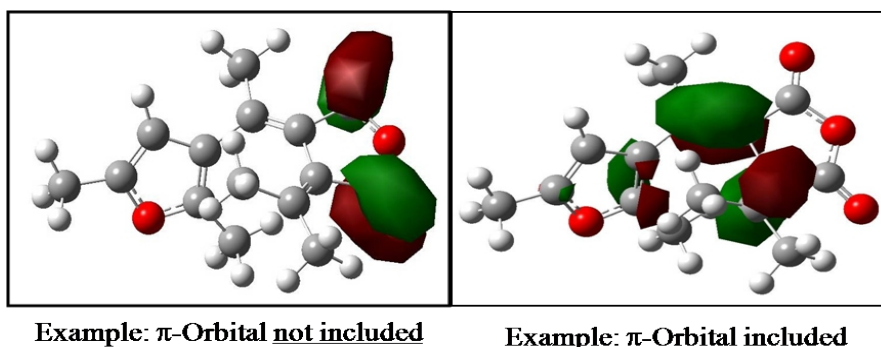
In particular, the aim of our work was to detect stationary points and conical intersections along a minimum energy path that could rationalize the model based on transient-absorption spectroscopy's experiments.

## 5.2 Computational methods

Structure optimization and relaxation path mapping have been carried out using fully unconstrained *ab initio* quantum chemical computations in the framework of the CASPT2//CASSCF/6-31G\* strategy<sup>23-25</sup>. Relaxation paths were computed in terms of minimum energy path (MEP) in mass-weighted coordinates, using the IRD methodology<sup>26</sup>. The CASPT2//CASSCF strategy requires that the reaction coordinate is computed at the complete active space self-consistent field (CASSCF) level and that the corresponding energy profile is re-evaluated on the multiconfigurational second-order Møller-Plesset perturbation theory level (here we used the CASPT2 method<sup>27</sup> implemented in MOLCAS-6<sup>28,29</sup>) to take into account the effect of electron dynamic correlation.

All computations employed the 6-31G\* basis set and an active space comprising 8 electrons in 8 orbitals. To select the optimal active space, in the early stage of our work, a state averaged RASSCF (16,14) was performed on the first three roots. Those orbitals which were mostly involved in the description of these three lowest singlet excited states and whose electronic occupation was lower than 1.8 or higher than 0.2 were considered to be the most important in the description of the photochemical reaction and were, therefore, included in the new reduced active space. These orbitals were mainly localized in the area of the molecule that was mostly involved in the cyclization mechanism, to be more precise, on the six carbon atoms of the hexatrienic system and on the remaining two carbons of the furylic ring. Interestingly MOs localized on the

anhydridic group were not considered to be important in order to describe accurately the molecule reactive path.



**Fig.5.4** There are represented two RASSCF orbitals: the first on the two carbonyl having an occupation number on  $S_1$  very high was not included within the casscf (8,8) ; the second one comprised into the casscf active space (8-8), since it is delocalized on the hexatriene moiety.

All geometry optimizations were carried out using the GAUSSIAN 03 suite of programs<sup>30</sup>.

Energy minima of the electronic ground state were optimized using a single root CASSCF wave function, while, dealing with states that were very close in energy, minima of the electronic first excited state were optimized using a two roots ( $S_0$ ,  $S_1$ ) state average (0.5, 0.5) CASSCF wave function. The  $S_1$  and  $S_0$  potential energy surfaces were connected through Conical Intersections (CI).<sup>31,32</sup>

A two roots ( $S_1$ ,  $S_0$ ) state average (0.5, 0.5) CASSCF wave function  $S_1$  MEP calculation was performed to connect the Franck-Condon point to the  $S_1/S_0$  CI.

As stated before, energies of stationary points, CI and selected points along the MEP have been re-evaluated using single point calculations performed at the CASPT2 level of theory. For each geometry, a five roots ( $S_0$ ,  $S_1$ ,  $S_2$ ,  $S_3$ ,  $S_4$ ) state-average (0.2, 0.2, 0.2, 0.2, 0.2) CASSCF wave function was used as reference function for evaluating the CASPT2 energies.

A further characterization of minima and the first five low-lying excited states was conducted analysing active space occupation numbers of localized molecular orbitals, for both occupied and virtual orbitals, which were obtained through a unitary transformation of the Hartree-Fock canonical molecular orbitals using the Boys's'localization method<sup>33</sup>.

A quantitative analysis of bonding was carried out using the CASSCF implementation of a method based on classic valence bond (VB) theory that calculated the spin exchange density matrix  $P$  with a localized orbital basis<sup>34</sup>. The method was applied to characterize the first four low-lying excited states with resonance structures using ground state as a reference and examining the most significant changes in the spin-exchange density and in the orbital occupations.

Despite structure optimization and relaxation path mapping have been carried out using mainly the CASPT2//CASSCF/6-31G\* strategy, TD-DFT (time-dependent density functional theory) calculations were also performed to get additional investigation and characterization of some structure by using the B3LYP functional, involving the gradient correction of the exchange functional by Becke<sup>35</sup> and the correction functional



by Lee, Yang and Parr<sup>36</sup>, employing a 6-31G\* basis set using the Gaussian 03 program suite<sup>30</sup>. No symmetry or internal coordinate constraints were applied during optimization. At the end energies of vertical and adiabatic transition were calculated using time-dependent density functional theory (TD-DFT) with B3LYP/6-31G\* calculation.

### 5.3 Results and discussion

The origin of the singular photochemical efficiency can be explained in terms of the clear mechanistic picture presented. While the  $S_0$  surface was deeply investigated by Yoshioka et al.<sup>37</sup> (at CASSCF level), through which it was possible to rationalize the thermal irreversibility by locating a transition structure on a barrier of 45 kcal/mol, there is not yet any computational study on the nature and behavior of the lowest excited states.

First in order to reproduce the experimental absorption spectrum, vertical excitations on the Franck Condon geometries were computed to evaluate the  $S_0$ - $S_1$  absorption energy for both the isomers: the open-ring (E) as well the close-one (C).

Relevant information and energetic data are collected in Table 5.1.

#### 5.3.1 Vertical excitations: the open-ring isomer E

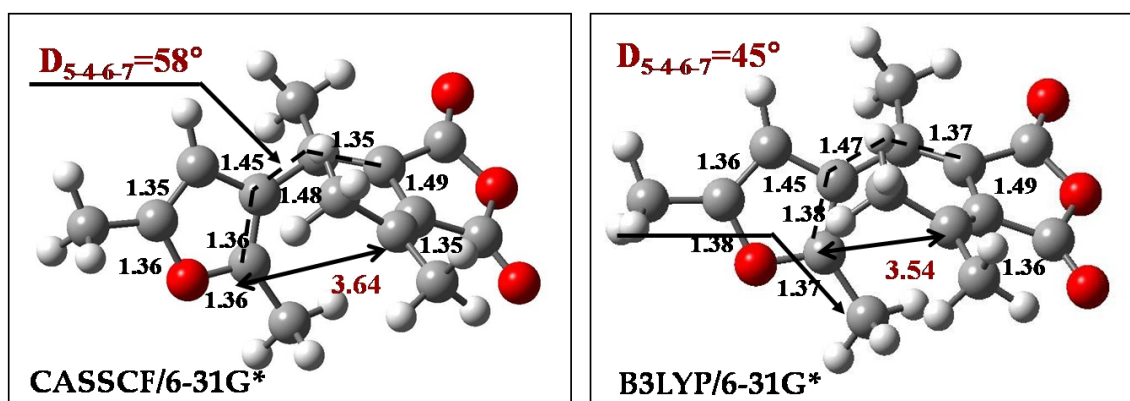
For the open-ring isomer, despite the oscillator strength calculated revealed that  $S_1$  is the spectroscopic state, it is important to point out how the  $S_0$ - $S_1$  absorption, calculated at CASPT2 level at 100 kcal/mol (304 nm) was largely overestimated in comparison to the experimental value (78.12 kcal/mol-366 nm). This was probably due to three main factors: first, the initial conjugated  $\pi$  system, composed by 16 electrons and 14 MOs, was reduced to minimize the computational cost, lowering the quality of the system description. In the second place the broader energy gap could be ascribed to the CASSCF method that was utilized to optimize the  $S_0$  geometry. This method was likely to underestimate the conjugation effect and obtain a less planar optimized structure with an increased value in the vertical excitation energy. At last a third reason, the overestimation could be attributed to the use of 6-31G\* basis set; by using ANO basis, results have been largely improved.

Geometry	State	CASPT2(MS)	CASPT2(MS)	CASPT2(MS)	CASPT2(MS)	TDDFT//
		(8,8)//6-31g*	(12,12)//6-31g*	(8,8)//ANO-S	(12,12)//ANO-S	6-31g*(nm)
		$\Delta E_{S_0-S_n}$	$\Delta E_{S_0-S_n}$	$\Delta E_{S_0-S_n}$	$\Delta E_{S_0-S_n}$	$\Delta E_{S_0-S_n}$
<b>Isomer E</b> (CASSCF Min $S_0$ ) <sup>a</sup>	$S_0$	0.00	0.00	0.00	0.00	0.00
	$S_1$	100.56 (100.57)	119.19	95.35 (95.40)	96.96 95.97	
<b>Isomer E</b> (MP2 Min $S_0$ ) <sup>b</sup>	$S_0$			0.00		
	$S_1$			89.77 (88.29)		
<b>Isomer E</b> (DFT Min $S_0$ ) <sup>c</sup>	$S_0$	0.00	0.00	0.00	0.00	0.00
	$S_1$	93.34		89.84	91.60	81.75= 349.97 nm
	$S_2$	(93.3)		(89.84)	(91.63)	94.26=303.53nm
	$S_3$					97.94(292,15 nm)

**Table 5.1.** Computed CASPT2//6-31g\* and TDDFT//6-31g\* relative energies (kcal/mol), for E ground state geometry and on the two low (or four in the case of TDDFT) lying excited states are reported.<sup>(a)</sup> It is referred to the CASSCF minimum on S<sub>0</sub>.<sup>(b)</sup> It is referred to the MP2 minimum on S<sub>0</sub>.<sup>(c)</sup> It is referred to the DFT minimum on S<sub>0</sub>.

In fact, 6-31g\* basis set was likely to be adequate to describe a covalent state, but insufficient for an optimal characterization of the system reactivity in case that an ionic state was involved. These last two explanations were recognized to be the most consistent, since subsequent calculations performed with larger active space did not lead to significant variations in energy values. The CASPT2//CASSCF(8,8) result of 89,94 kcal/mol, calculated on the DFT geometry (more conjugated) being the lowest value (within CASPT2 range), implies the overestimation arising from a geometrical problem. At the end only TDDFT absorption, with  $\Delta E_{S_0 \rightarrow S_1} = 81$  kcal/mol (350 nm), nicely reproduces the experimental measure of 366 nm.

From a geometrical point of view, the analysis of bond lengths of CASSCF minimum E<sub>0</sub>, reveals an under-estimation conjugation effect on the  $\pi$  hexatrienic system. On the other hand, the DFT S<sub>0</sub> minimum, is less distorted, according with an enhanced planarity, increased the  $\pi$ -electrons delocalization (see Fig.5.5(a)).

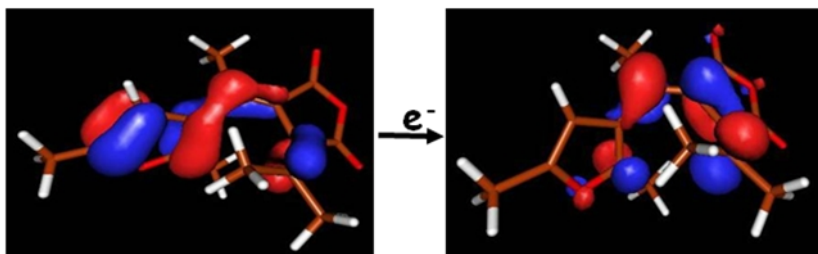


**Fig.5.5(a)** In this figure there are illustrated two minima optimized at two different levels: CASSCF and DFT using the same basis set. Although the two geometries are quite similar, it is emphasized the reduced dihedral angle in the DFT minimum responsible of the more conjugated system.

In these cases, DFT calculations resulted in an overestimated  $\pi$ -electron delocalization effect, yielding shorter single bonds and longer C=C double bonds and, therefore, less rigid and distorted structures. Other evidences of the same tendency, for example, were obtained performing DFT calculations on retinal (containing a carbonyl group as a terminal group) which predicted a smaller bond alternation for the planar structure of the molecule as compared to the experimental values. The increase of  $\pi$ -electrons delocalization and conjugation effect and the resulting gain of stability were reflected by the reduced vertical excitation energy obtained carrying out a TD-DFT single point on the minimum found (see table 5.1).



From the highest oscillator strength value ( $f_{0-1} = 0.141$  in table 5.6) is apparent that the HOMO  $\rightarrow$  LUMO single excitation state S<sub>1</sub>, featuring a ionic state, is the bright state. Notably, while the four states had small dipole moment values and similar Mulliken

charge distribution, a relevant change in its dipole moment reveals a charge transfer character. Remarkable, the charge transfer brings about 0.5 electron density from the heteroaromatic ring to the anhydridic moiety, as reflected also from the electronic transition between the HOMO and LUMO orbitals (see Fig.5.5). This calls for a polar excited state  $S_1$  in agreement with the experiments.

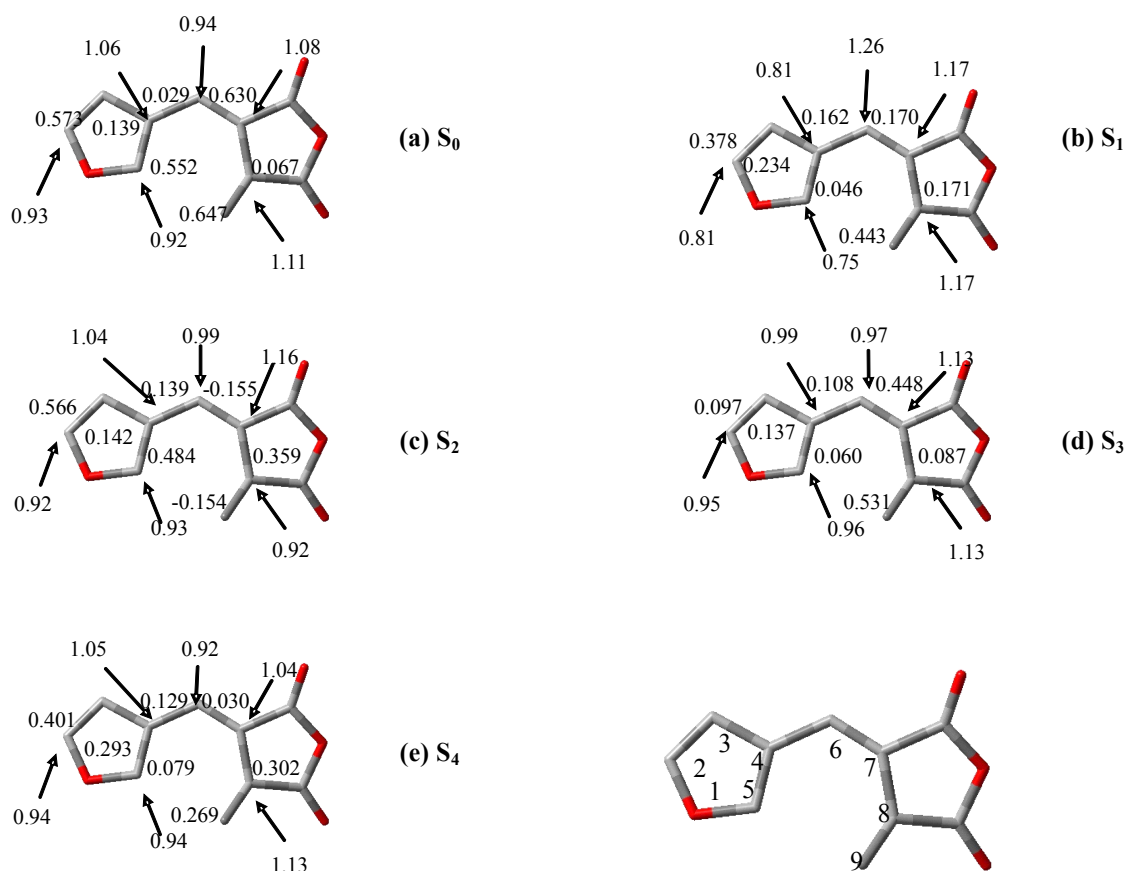


**Fig.5.5** There are represented two frontier orbitals HOMO and LUMO.

To characterize more accurately the nature of these first five states, a state averaged calculation was performed on these 5 roots using Boys localization and printing the spin-exchange density matrix elements  $P_{ij}$  of these states. In addition dipole moments and Mulliken charges values were taken in account to define the nature of these states (Table 5.2 (b)). Notably  $S_1$  and  $S_2$  are almost degenerate in the Franck Condon point.

State	CASSCF(SA5r) <sup>a</sup>	CASPT2(SA5r) <sup>a</sup>	Dipole <sup>b</sup> Moment (Debye)	Mulliken Charges <sup>b</sup>	
	(8,8)//6-31g* Kcal/mol	(8,8)//6-31g* Kcal/mol		A 	B 
$S_0$	0	0	6.06	-0.036	0.036
$S_1$	141.76	101.55	12.82	0.539	-0.539
$S_2$	147.80	127.11	5.79	-0.022	0.022
$S_3$	163.40	129.48	7.73	0.085	-0.085
$S_4$	166.42	142.81	6.04	-0.042	0.042

**Table 5.2** (a) Computed State Average (5 roots) CASSCF and CASPT2 excitation energies ( $\text{kcal}\cdot\text{mol}^{-1}$ ), for the vertical low-lying excited states of Franck Condon E isomer (open-ring). The 6-31G\* basis set was employed throughout (b) Computed dipole moments and Mulliken charges distribution, for the vertical low-lying excited states of the E isomer of 2,5-dimethyl-furyl-fulgide.



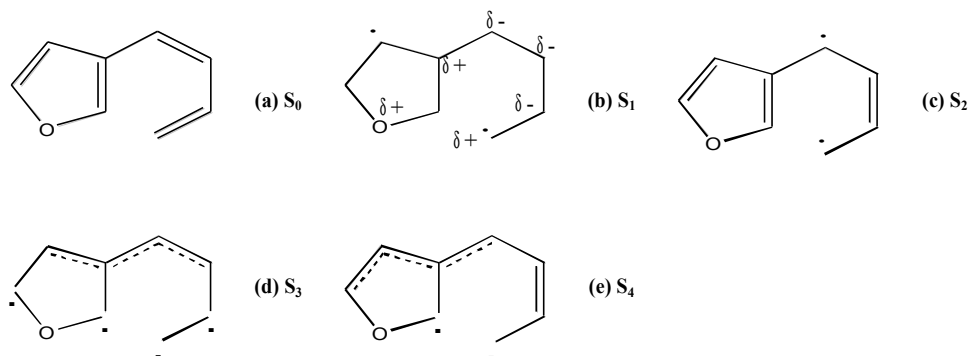
**Fig 5.6** Spin-exchange density matrix elements  $P_{ij}$  ( $\alpha\beta$  terms) for the five lowest singlet excited states of 2,5-dimethyl-furyl-fulgide, CASSCF(8,8)/6-31G\*. The numbers shown with arrows are the occupation numbers of the localized orbitals (diagonal elements of the one-electron matrix,  $D_{ij}$ ) that differ significantly from 1.00.

From the obtained results  $S_0$ ,  $S_2$ ,  $S_3$  and  $S_4$  appeared to be covalent states, while  $S_1$  was labeled as an ionic one. In fact, while these four states had small dipole moment values and similar Mulliken charge distribution, a valuable variation in the moment and a charge displacement from the furanic to the anhydridic ring were observed for  $S_1$ .

This assignment was confirmed by the obtained occupation numbers of localized orbitals and spin-exchange density data. In fact, for the first excited state, the charge-transfer character was clear from the variation of orbital occupations with respect to the ground state and significant variation in some  $P_{ij}$  elements. In particular it was assumed that the charge transfer took place mainly from the  $C_5$ ,  $C_2$  and  $C_4$  carbons (occupation decreased by 0.17, 0.12 and 0.25 electrons in comparison with the ground state, respectively) to  $C_6$ ,  $C_7$  and  $C_8$  carbons. At the same time the  $C_3$ - $C_4$ ,  $C_4$ - $C_5$ ,  $C_3$ - $C_4$  and  $\pi$  bonds character was lost (decrease in  $P_{ij}$  values 0.573, 0.552, 0.630 and 0.647 to 0.378, 0.046, 0.170 and 0.171, respectively) because of the charge transfer.

Moreover an increase in the coupling between  $C_4$  and  $C_6$  ( $P_{ij}$  increased from approximately 0.03 to 0.16) and  $C_4$  and  $C_3$  ( $P_{ij}$  value from 0.14 to 0.23), probably due to stabilizing positive charge delocalization, was detected.

In the same way the other three excited states were examined. For the proposed resonance structures see Figure 5.7.



**Fig 5.7** Proposed resonance structure of 2,5-dimethyl-furyl-fulgide of the first five low lying states, CASSCF(8,8)/6-31G\*.

### 5.3.2 The open→close ring reaction

Once the Franck-Condon point had been characterized, the reaction pathway was studied by computing the Minimum Energy Path (MEP) using the standard intrinsic reaction co-ordinate (IRC) method. A CASSCF(8,8) state averaged IRC calculation on the first two states was performed.

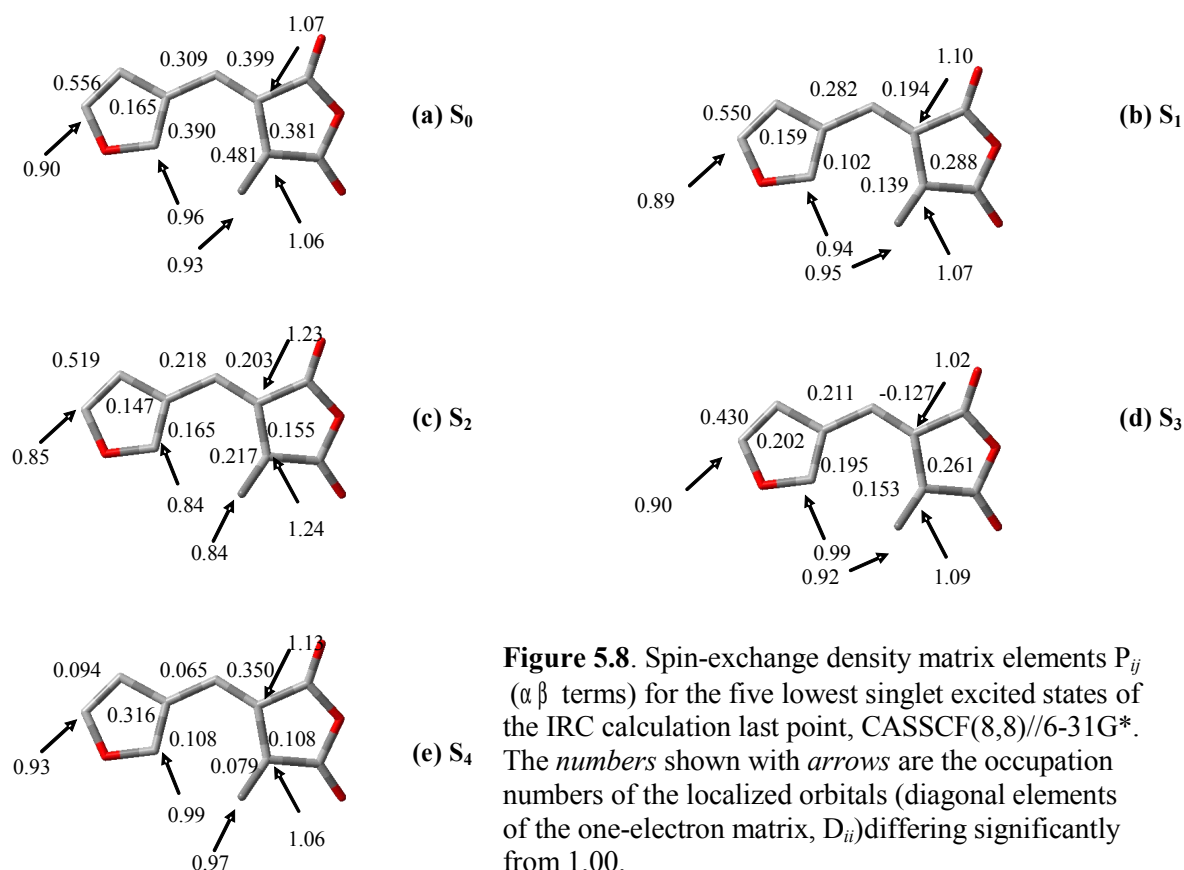
At the last point of the IRC calculation, an energy minimum was reached. Since the nature of such point reveals a different wave function character respect to the starting point (the FC), it was supposed a different CASSCF order of the lowest excited state featuring this critical point. Therefore, in order to analyzing the nature of the reactive state along the  $S_1$  pathway, a series of state averaged single point CASSCF//CASPT2 (8,8) corrections on the lowest 5 roots, were computed along the IRC local minima.

At CASPT2 level the stationary point (IRC last point) was located with an  $S_0$ - $S_1$  energy gap of 21.34 kcal/mol (Table 5.3). Notably, CASPT2 calculations pointed out two important aspects: first of all  $S_1$  and  $S_2$  states were still almost degenerate ( $E = \sim 3.13$  kcal·mol<sup>-1</sup>). Secondly the  $S_0$ - $S_1$  energy gap and, consequently, the  $S_0$ - $S_2$  gap, as well, were significantly reduced with respect to the initial value (decrease from 100.39 to 21.34).

State	CASSCF(SA5r) <sup>a</sup>	CASPT2(SA5r) <sup>a</sup>	Dipole <sup>b</sup> Moment (Debye)	Mulliken Charges <sup>b</sup>	
	(8,8)//6-31g* Kcal/mol	(8,8)//6-31g* Kcal/mol		A	B
$S_0$	0	0	6.47	0.034	-0.034
$S_1$	33.89	21.34	7.07	0.084	-0.084
$S_2$	67.77	24.47	10.45	0.251	-0.251
$S_3$	91.62	76.56	6.47	0.031	-0.031
$S_4$	102.91	85.34	7.47	0.085	-0.085

**Table 5.3.** <sup>(a)</sup> Computed State Average (5 roots) CASSCF and CASPT2 excitation energies (kcal·mol<sup>-1</sup>), for the vertical low-lying excited states of the last IRC point (open-ring). The 6-31G\* basis set was employed throughout. <sup>(b)</sup> Computed dipole moments and Mulliken charges distribution, for the vertical low-lying excited states of the E isomer of 2,5-dimethyl-furyl-fulgide.

Occupation numbers of the localized orbitals, spin-exchange density matrix elements, dipole moment values and Mulliken charges distribution were examined to define the nature of these five lowest excited states (Table 5.3 and Figure 5.8).

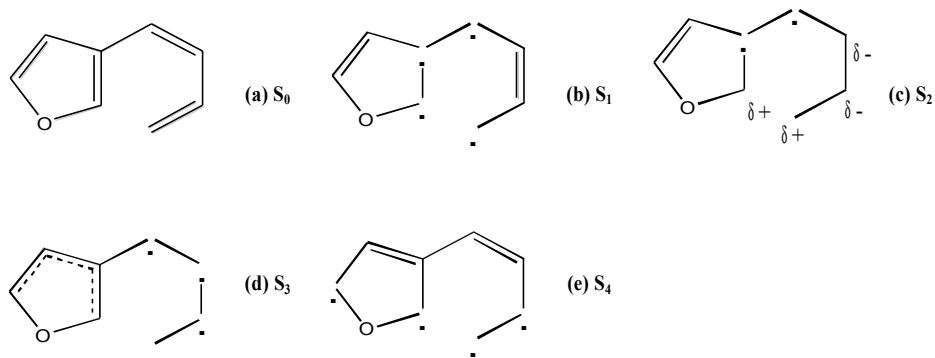


**Figure 5.8.** Spin-exchange density matrix elements  $P_{ij}$  ( $\alpha\beta$  terms) for the five lowest singlet excited states of the IRC calculation last point, CASSCF(8,8)//6-31G\*. The numbers shown with arrows are the occupation numbers of the localized orbitals (diagonal elements of the one-electron matrix,  $D_{ii}$ ) differing significantly from 1.00.

In this case from the dipole moment values and the Mulliken charge distribution it was possible to assert that in this point the  $S_1$ ,  $S_3$  and  $S_4$  excited states had the typical features of a covalent state, while, with its significant variation of dipole moment and displacement of charge,  $S_2$  appeared to be an ionic state.

Occupation numbers of localized orbitals confirmed this assumption: in the second excited state the charge transfer took place from  $C_5$  and  $C_9$  to  $C_7$  and  $C_8$  carbon atoms (occupation decreased by 0.12 and 0.9 electrons and increased by 0.16 and 0.18 electrons, respectively, in comparison with ground state values). Furthermore the  $C_4$ - $C_5$ ,  $C_6$ - $C_7$  and  $C_8$ - $C_9$  bond character was lost and a decrease in coupling between  $C_7$  and  $C_8$  carbons was observed, probably due to charge displacement.

The remaining three excited states occupation numbers and spin-exchange density matrix elements were analyzed as well and it was clear that all these states had typical features of covalent states. The proposed resonance structures are reported in Figure 5.9



**Figure 5.9.** Proposed resonance structure of 2,5-dimethyl-furyl-fulgide of the last point of the IRC calculation, CASSCF(8,8)/6-31G\*.

After wave function analysis, it is remarkable to notice also, how the MEP calculated on the reactive state follow the same reaction coordinate, clearly dominated by the vibrational normal mode between the two reactive carbons C<sub>5</sub>-C<sub>9</sub> (those involved in the  $\sigma$ -bond making reaction). According with this, geometrical analysis of the IRC last point reveals how the C<sub>5</sub>-C<sub>9</sub> distance was reduced from the initial value of 3.64 Å ( of the Franck Condon) to 2.54 Å. Furthermore a decrease in single bonds length as well as an increase in double bonds length were observed, showing an abrogation in bond alternation and a slight inversion of the bond-order of the hexatrienic system in respect to the ground state, due to an increase in  $\pi$  electrons delocalization.

Since an IRC calculation often terminates if the PES is sufficiently flat, not necessarily individuating a real minimum, a CASSCF (SA2roots) geometry optimization of last point, was performed on S<sub>1</sub>. The system, moving along the S<sub>1</sub> potential, reached an energy minimum 11.30 kcal/mol lower in energy. The wave function character of the real minimum located on S<sub>1</sub> is again changed; further single point calculations were again necessary to obtain a more accurate information (Table 5.4).

At this stage, it is necessary to be careful and do not make confusion between CASPT2 SA 5roots and SA 2roots calculations. Either, the CASSCF/CASPT2 SA 5roots and SA 2roots results, are collected in the table 5.4. Besides, as for the previously examined two points, dipole moment values and Mulliken charges distribution data were examined for a more accurate characterization of the state.

State	CASSCF(SA5r) <sup>a</sup>	CASPT2(SA5r) <sup>a</sup>	Dipole <sup>b</sup> Moment (Debye)	Mulliken Charges <sup>b</sup>	
	(8,8)//6-31g* Kcal/mol	(8,8)//6-31g* Kcal/mol		A	B
S <sub>0</sub>	0	0	6.48	0.027	-0.027
S <sub>1</sub>	26.36(18.11)	7.53(3.84)	10.52	0.209	-0.209
S <sub>2</sub>	65.26	33.89	8.52	0.127	-0.127
S <sub>3</sub>	86.60	75.30	6.54	0.023	-0.023
S <sub>4</sub>	103.54	83.46	6.77	0.037	-0.037

**Table 5.4** .SA 5 roots computed CASSCF and CASPT2 calculations, in italic SA 2roots calculations, computed dipole moments and Mulliken charges distribution, for the vertical low-lying excited states of the minimum E\*. The 6-31G\* basis set was employed throughout.

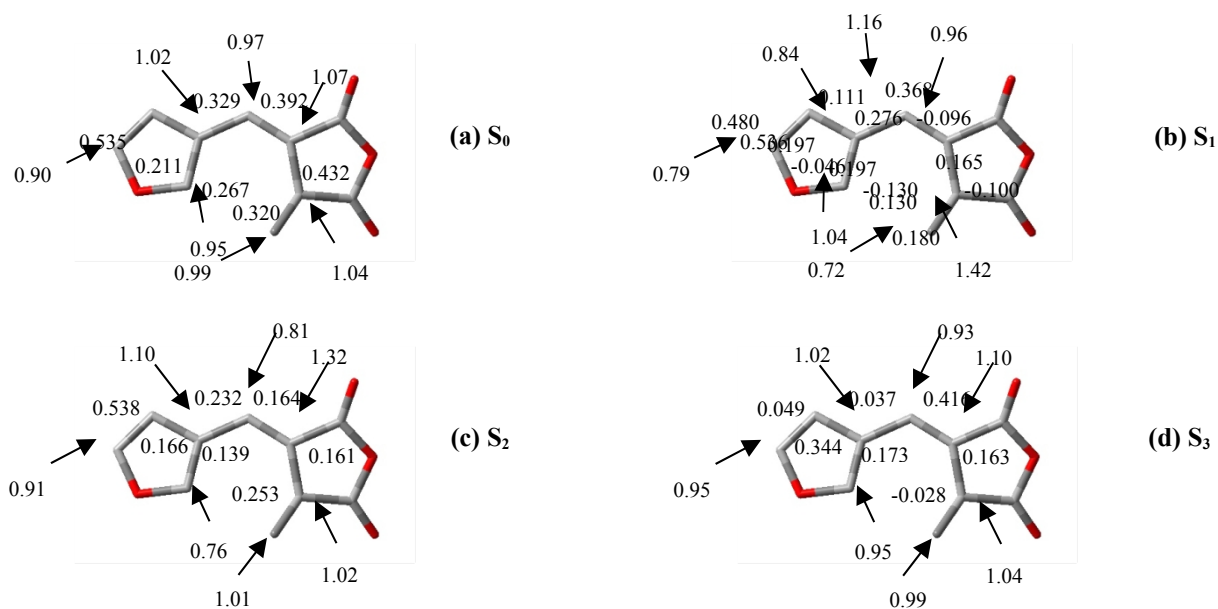
Since  $S_1$  is a ionic state, due to its CASSCF overestimation, the computed CASSCF  $\Delta E(S_0-S_1)$  of 18.00 kcal/mol, was further reduced by CASPT2 correction to less than 4 kcal/mol (both SA2roots), taking into account the effect of electron dynamic correlation, showing a remarkable approaching between the two states ( 0.87 kcal/mol at Multistate level ).

Therefore the  $S_1/S_0$  radiationless decay, leading to photoproducts on  $S_0$ , is explained through such crossing ,lying 54 kcal·mol<sup>-1</sup> lower than FC, that is an “avoided crossing” at CASPT2 level.

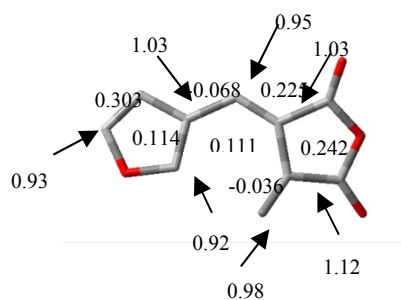
From the dipole moment collected in table 5.4, if the first excited state is ionic, the fourth and the fifth, instead, appeared to be covalent states not showing remarkable differences in charge position and in dipole moment values as compared with the ground state. Regarding  $S_2$ , instead, it was difficult to define its nature, simply observing dipole moment and charge distribution data.

In regard to the first excited state the charge transfer character was confirmed by an increase by 0.38 electrons in the occupation number of the orbital localized on C8 and a decrease by 0.27 electrons in the occupation of the orbital on C9. In addition a consequent decrease in the coupling between these two carbons ( $P_{ij}$  decrease from 0.320 to 0.130) was observed (fig.5.10).

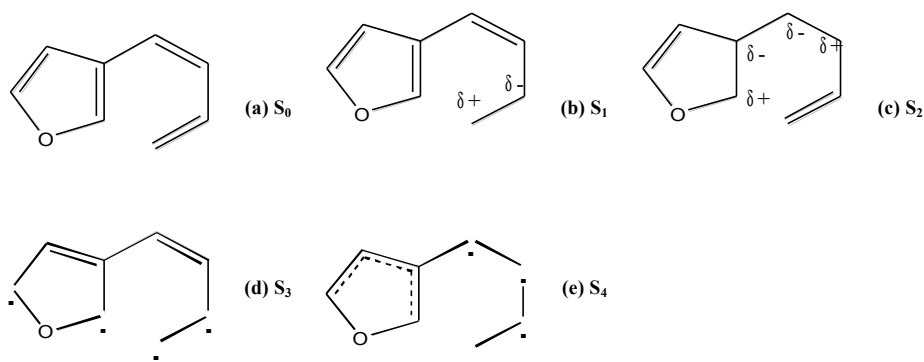
Once again the values referred to the other two states were examined and resonance structures for the four excited states were proposed (Figure 5.11).





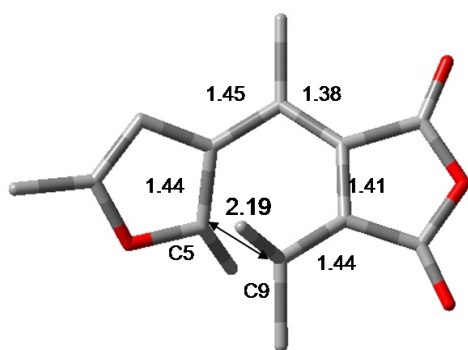


**Figure 5.10** Spin-exchange density matrix elements  $P_{ij}$  ( $\alpha\beta$  terms) for the five lowest singlet excited states of the crossing  $E^*$ . CASSCF(8,8)/6-31G\*. The numbers shown with arrows are the occupation numbers of the localized orbitals (diagonal elements of the matrix one-electron  $D_{ii}$ ) that differ significantly from 1.00.



**Figure 5.11.** Proposed resonance structure of 2,5-dimethyl-furyl-fulgide of the minimum  $E^*$ , CASSCF(8,8)/6-31G\*.

Moreover, looking at the crossing geometry, it was remarkable the further reduced C5-C9 distance of 2.19 Å (Figure 5.12).



Concerning to the bond-order, no significant variations in comparison with the previously reached minimum were observed. It was therefore possible to state that this  $S_1$  relaxation pathway, has well described the ring closure reaction starting from FC. In fact, moving along the path, the distance between C5-C9 forming the  $\sigma$  bond in C conformer, gradually decreased together with redistribution of electrons of the conjugated  $\pi$  hexatrienic system.

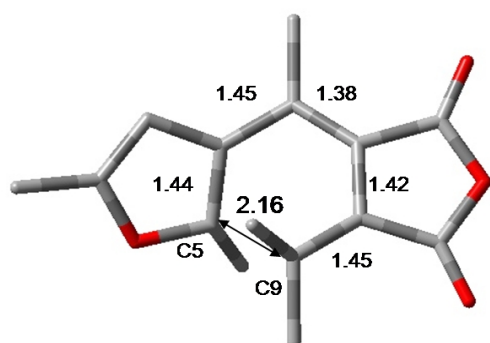
**Figure 5.12** Optimized structure on the potential energy surface of the  $E^*$  minimum at CASSCF(8,8)/6-31G\* level ( $S_1/S_0$  conical intersection at CASPT2 level). Distances are in angstroms.

Notably, no transition states were detected on the cyclization reaction coordinate. The existence of such  $S_1/S_n$  crossing and a barrierless profile could provide both an explanation of the absence of fluorescence from the excited open ring isomer  $E^*$ .

In order to confirm the  $S_1$  reaction pathway and the  $S_1/S_0$  crossing calculated with CASPT2, further investigations on the reached minimum were carried out utilizing TD-DFT/6-31G\* level of theory. This was accomplished by a  $E^*$  geometry optimization on the first excited state, employing B3LYP hybrid functional.

A new equilibrium structure was found, in which the two lowest lying singlet states were practically degenerate (2.17 kcal/mol).

Therefore it was concluded TDDFT calculations are in agreement with CASPT2 potential: through the TDDFT  $S_1$  optimization a similar  $S_1/S_0$  crossing was detected, confirming the possibility of a radiationless decay to the ground state via such crossing. It is confirmed not only by the  $S_1-S_0$  energy gap (even lower with TDDFT) but also by the electronic and geometric nature of such critical point. Thus TDDFT calculations not only permit to well reproduce the experimental absorption spectrum, but also its potential reflects the same barrierless  $S_1$  profile: after the photon absorption, from the FC the photoexcited population reaches directly the  $S_1/S_0$  degenerate region (CI  $S_1/S_0$  confirmed with a computed  $\Delta E(S_1-S_0)=0,41$  kcal/mol) through which it is possible to rationalize an efficient radiationless decay to  $S_0$ . Thus if the wavefunction analysis

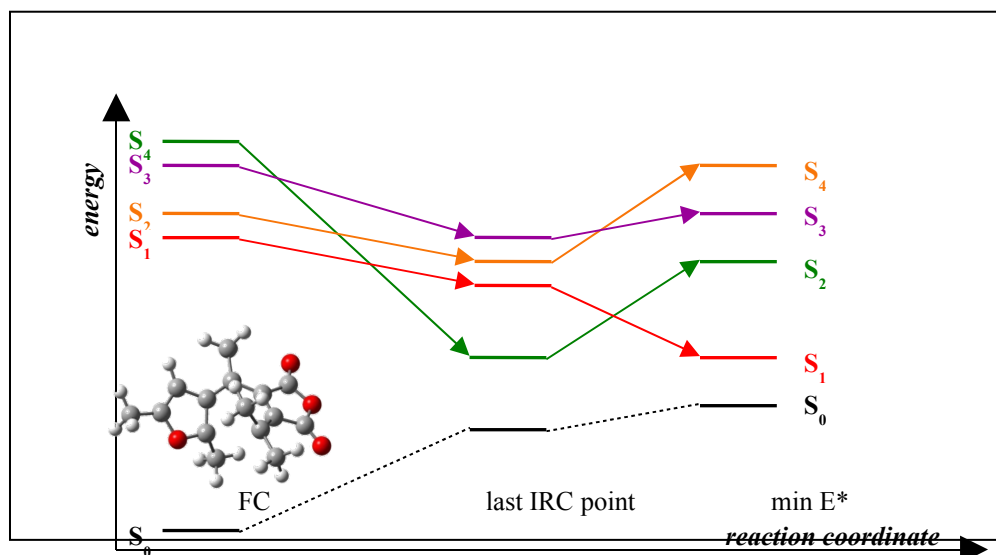


reveals how the almost degenerate states are the covalent  $S_0$  and the ionic  $S_1$  geometrically an ulterior approaching of C5 and C9 carbon atoms was detected (decrease in the atom distance by 0.03 Å), while no significant differences in the bond length values were observed.

**Figure 5.13.** Optimized structure on the  $S_1$  potential energy surface of the  $E^*$  minimum at B3LYP/6-31G\* level. Distances are in angstroms.

Once defined the nature of the five low-lying states in the detected minima, a precise description of the potential energy surfaces and a detection of possible crossings between them were aimed.

Therefore dipole moment and charge displacement data, proposed resonance structures and, in term of molecular orbital excitations, the weights of configurations that composed the states were examined to outline the states trend at CASSCF/6-31G\* level of theory. At this level three crossings were detected between the initial FC point and the last point of the IRC calculation: the fourth excited state crossed  $S_3$ ,  $S_2$  and  $S_1$  in sequence. All the crossings resulted from a decrease in  $S_4$  energy along the relaxation path. Between the last IRC point and the minimum  $E^*$ ,  $S_3$  and  $S_4$  swapped as well as the two lowest singlet excited states. The outlined state level trend was summarized in Scheme 5.



**Scheme 5.2.** Proposed model for the five lowest singlet excited states energy levels in the ring closure photochemical reaction of the 2,5-dimethyl-furyl-fulgide at CASSCF(8,8)/6-31G\* level of theory.

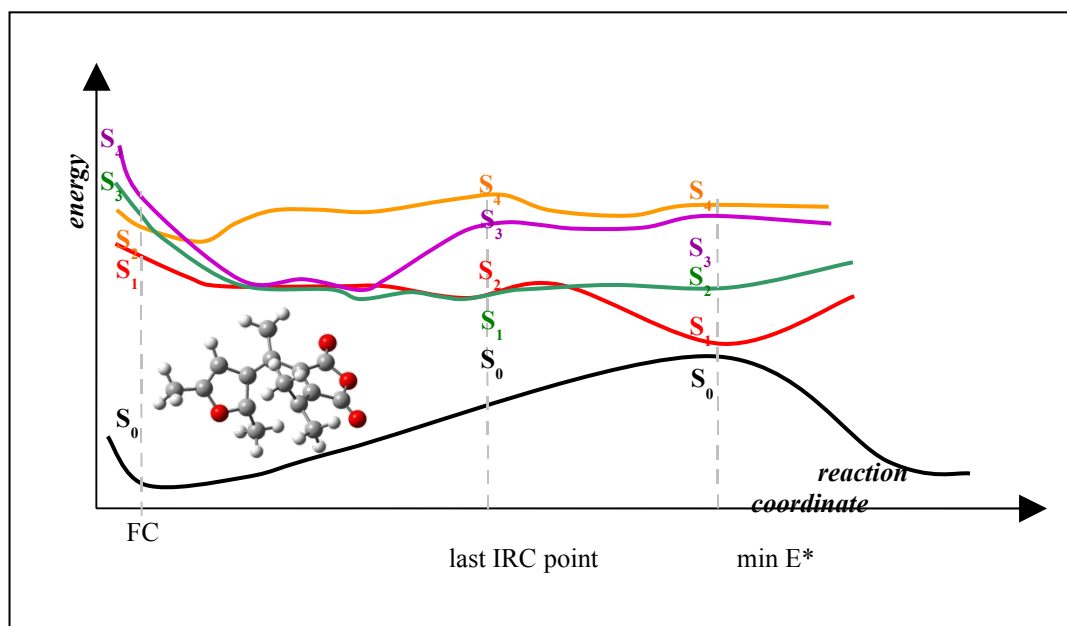
Furthermore CASPT2 single point calculations provided some important information concerning the nature of local intermediate points of the IRC calculation and the minimum E\*. In each point, a five roots state-average CASSCF wave function was used as reference function. Combining the obtained data for the five low-lying states in these intermediate points with all the information we have already acquired relating to the three previously characterized points, it was possible to outline the states trend at CASPT2/6-31G\* level of theory. Examining the dominant configurations in the description of the five roots and the energy variations along the MEP, some differences in the excited states order were observed in comparison to the CASSCF proposed outline (Scheme 5.3).

In fact, comparing the CASSCF excited state order and main configurations involved in the description of the states with that resulted from CASPT2 calculations, several observations could be advanced: firstly, in the starting FC point, at CASPT2 level, taking into account the effect of electron dynamic correlation, the CASSCF state S<sub>4</sub> energy was reduced. At CASPT2 level, the S<sub>3</sub> states corresponded to the CASSCF S<sub>4</sub> state and vice versa.

Second of all, during the initial relaxation path, the starting S<sub>1</sub>, S<sub>3</sub> and S<sub>4</sub> were quite close in energy. Then, an increase in the S<sub>4</sub> energy led this state distant from the other two states and closer to S<sub>2</sub>. As seen from the CASSCF results, in addition, the FC S<sub>1</sub> and S<sub>3</sub> states crossed and in the last point of the IRC calculation the covalent state was lower in energy and became the first excited state. Moreover, in this point, at CASPT2 level, the order of the two highest excited states, S<sub>3</sub> and S<sub>4</sub>, was inverted in respect to CASSCF evidences.

Thirdly, relaxing from the last IRC point, there was another crossing between the starting S<sub>1</sub> and S<sub>3</sub> states and the ionic state changed back to the first excited state, as detected at CASSCF level.

Moreover no further crossing between the considered two highest singlet excited states were observed, while at CASSCF level these two roots were inverted in the minimum E\* in respect to the last point examined.



**Scheme 5.3.** Schematic outline (based on the corrected IRC points) of the five lowest singlet excited states energy levels in the ring closure photochemical reaction of the 2,5-dimethyl-furyl-fulgide at CASPT2/6-31G\* level of theory.

To summarize the state of theoretical evidence, a radiationless decay path for the photoinduced electrocyclic ring closure reaction of 2,5-dimethyl-furyl-fulgide was proposed. After photoexcitation the charge transfer state  $S_1$  was populated. The spectroscopic state was rapidly depopulated by an internal conversion process that involved a covalent state  $S_2$  that was degenerate and strongly coupled with the pumped  $S_1$  state in the FC point. During the relaxation in the covalent state, the system evolved to a structure that gradually resembled that of the close ring isomer and, at a certain point, the relaxation path reached a crossing point between the first two excited state and the system decay on the initial charge transfer state.

Moving along the path, the  $S_0$ - $S_1$  energy gap gradually reduced up until a point in which the two surfaces were practically degenerate and the fulgide system decay to the ground state through a conical intersection. The system was, then, free to return to the initial open ring isomer E or turn into the photoproduct.

In the next section it will be briefly rationalized the close-open ring reaction taking place on the opposite side of the potential energy surface: the close-ring isomer region (C). Similarly to the just discussed one, this process is initiated upon light irradiation of C conformer and once the spectroscopic state  $S_1$  is populated, the open-ring process is completed on the same state.

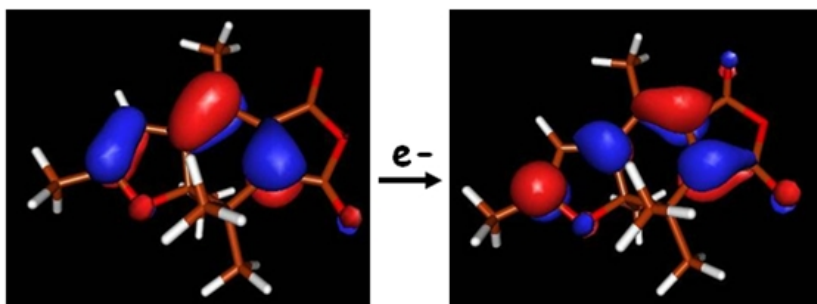
### [5.3.2 The close-ring isomer and the open ring reaction](#)

Respect to  $C^*$  vertical excitations, rather than a geometry problem, the overestimation of  $S_0 \rightarrow S_1$  energy absorption is mainly due to the restricted active-space: in reality increasing the active space size by using the CASPT2 (12,12) wave function, the error on the  $S_0 \rightarrow S_1$  vertical energy, was reduced, recovering more than 10 kcal/mol. Thus, although the TDDFT result is even better ( $\Delta E(S_0 \rightarrow S_1) = 58.57 \text{ 487 nm}$ , the CASPT2(12,12)  $S_0 \rightarrow S_1$  absorption of 62 kcal/mol (438 nm) sufficiently matches with the experimental one (516 nm). All the data are collected in the Table 5.5

Geometry	State	CASPT2(MS)	CASPT2(MS)	CASPT2(MS)	TDDFT//
		(8,8)//6-31g*	(12,12)//6-31g*	(8,8)//ANO-S	6-31g*(nm)
		$\Delta E_{S_0-S_n}$	$\Delta E_{S_0-S_n}$	$\Delta E_{S_0-S_n}$	$\Delta E_{S_0-S_n}$
Isomer C	S <sub>0</sub>	0.00	0.00		
(CASSCF MinS <sub>0</sub> )	S <sub>1</sub>	70.06(SS) 77.86 (67.70)	66.16(SS) 82.53(SA4r) (64.02 MS4r)		
Isomer C	S <sub>0</sub>			0.00	0.00
(DFT MinS <sub>0</sub> )	S <sub>1</sub>			<b>59.95</b> <b>(61.85)</b>	<b>58.57</b> <b>(487.29 nm)</b>
	S <sub>2</sub>				81.93 (349.22 nm)
	S <sub>3</sub>				96.53 (296.38 nm)

**Table 5.5.** Computed CASPT2//6-31g\* and TDDFT//6-31g\* relative energies (kcal/mol), for C ground state geometry and on the two low (or four in the case of TDDFT) lying excited states are reported.

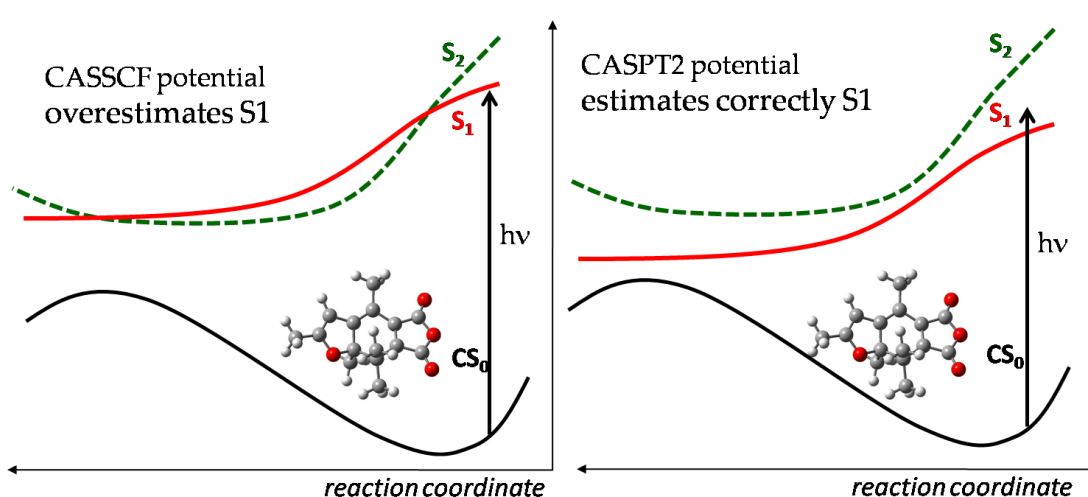
The highest oscillator strength value ( $f_{0-1} = 0.246$ , see table 5.6) confirms the S<sub>1</sub> state is the optically active state. C\* wavefunction, featured by a high charge separation, has an ionic character (also reflected from orbitals, refigured in fig.5.14) and (moving) along the S<sub>1</sub> reaction path for the opening reaction, its nature will be preserved (until the S<sub>1</sub>/S<sub>0</sub> degenerate point).



**Fig.5.14** There are represented two frontier orbitals HOMO and LUMO

Also for the open-ring reaction the S<sub>1</sub> MEP was studied through a CASSCF (SA2r) IRC calculation, that lead the molecule directly from the FC to a minimum located in a flat region of the S<sub>1</sub> potential energy surface.

In reality, due to the CASSCF overestimation of the (S<sub>1</sub>) ionic state, the CASSCF ordering of lowest excited states along the reaction coordinate behaviour, is quite different from the CASPT2 one; if the CASSCF analysis of the electronic nature reveals mainly two S<sub>2</sub>/S<sub>1</sub> (covalent/ionic) crossing along the MEP S<sub>1</sub>, at CASPT2 level the first excited state S<sub>1</sub> is always dominated by an ionic nature (it is the ionic state, the charge separation state the lowest along the reaction path mapped by the calculation, as it is clear from the scheme 5.4).



**Scheme 5.4.** Schematic outline of the  $S_1, S_2$  energy profile along the ring opening photochemical reaction of the FF-3: comparing the CASSCF and CASPT2 potentials. Due to CASSCF  $S_1$  overestimation  $S_1$  cross  $S_2$  along the  $C^*$  relaxation path, this is not verified for CASPT2 potential.

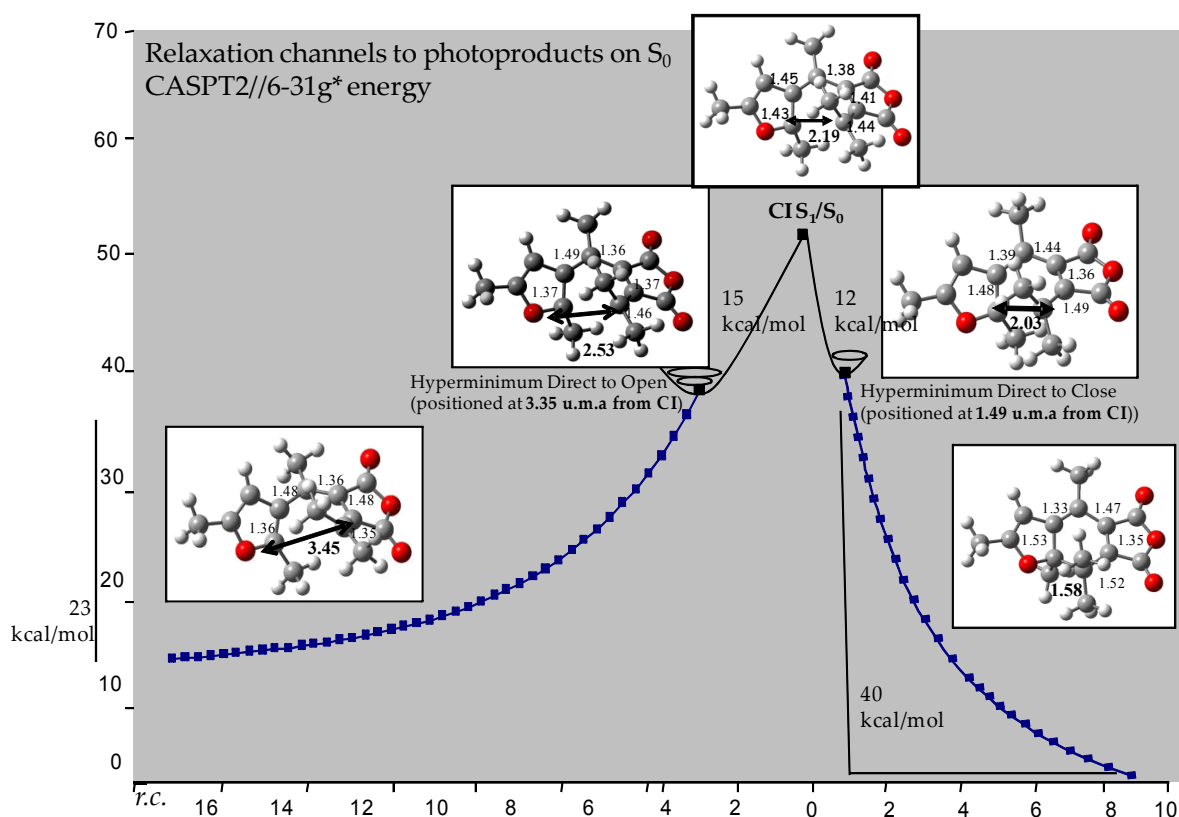
However, also for this branch, the vibrational relaxation process on  $S_1$  causes a complete  $\pi$ -system re-organization together with  $\sigma$ -bond elongation. This suggests how  $C^* \rightarrow E$  conversion has been started and completed on the same  $S_1$ . Experimentally  $C^*$  emission was not observed, confirmed by a barrierless reaction path on  $S_1$ ; from the FC point the system relaxes directly toward a flat region of  $S_1$  potential energy surface, until a minimum located about 22 kcal/mol lower in energy than the FC, is populated (reached). Also for C isomer is correct to affirm  $S_1$  corresponds to the reactive state, responsible of  $\sigma$ -bond breaking reaction: this will be confirmed by comparing the FC geometry with the  $C^*$  equilibrium geometry. Such point is characterized by a geometry where the  $C_1-C_6$  bond distance is increased from 1.58 Å (FC) to 1.67 Å ( $C^*$ ), with an  $S_1-S_0$  energy gap of still 30 kcal/mol. The high energy gap at such point is not agree with a radiationless decay (experimentally observed also  $C^*$  population). In addition, the transient-absorption experiments support the evidence of an intermediate I, to be located in the middle of the route decay  $E^* \rightarrow C_0$ : this may be associated with the  $C^*$  relaxed structure, calculated on  $S_1$ .

However, if  $C^*$  population is trapped in this valley, how may occur radiationless decay on  $S_0$ ? It is, therefore, necessary to know if the  $S_1/S_0$  crossing (CI), easily reached by  $E^*$  population, can be accessible also from  $C^*$  population. For this reason a reaction path, connecting the  $C^*$  with CI, with a smooth barrier of 5 kcal/mol was computed. This energy profile was calculated through CASPT2 single-point calculations on the intermediate geometries generated from a linear interpolation between the two structures. Although there were not estimated precisely the lifetime for ring-opening reaction, a barrier of 5 kcal/mol should be in agreement with ps lifetime.

Further the kinetic energy accumulated coming down directly from the FC point is able to explain how the photoexcited population is able to overcome such barrier of 5 kcal/mol.

Finally in order to validate the photoproducts formation after the radiationless decay, the relaxation channels, connecting the  $S_1/S_0$  CI with the two isomers, were calculated on the ground state. Also for such case an IRC calculation was employed by using a  $S_0$  Single State CASSCF wave function, but if for calculating the  $S_1$  reaction path (the

close to open and the opposite) the starting point was for both the cases the Franck Condon structure (on which the gradient necessary for launching the IRC calculations are not zero), for the relaxation channels the starting point being a CASSCF minimum have the gradient null. Thus for E isomer (as well for C), the necessary forces have been calculated on an hyperminimum calculated by projecting the  $S_0$  E-minimum on the  $S_1$  CASSCF  $E^*$  minimum (the CASPT2  $S_1/S_0$  crossing) geometry. Remarkably, no relaxation channel, leading the system to an eventual bicycle photoproduct (through a diradicalic intermediate), was located. Furthermore, according with the evidence of  $\sigma$ -bond almost formed at the  $S_1/S_0$  CI, the open-ring hyperminimum was difficult to locate because particularly far (in u.m.a) from the CI geometry.



**Fig.5.15** Relaxation channels to both the photoproducts on  $S_0$  starting from  $S_1/S_0$  degenerate point trough the Hyperminima located .

At the end, CASSCF calculations (scaled with CASPT2 energy correction as plotted in fig.5.15 ) reveal that on both side the reaction paths are barrierless, no transition states were located on this channels, according with the short lifetime of 3 ps trough which it was detected the C isomer formation upon light irradiation on the E reactant. Nevertheless it is fundamental to notice how the shape of both relaxation channels, where the deeper is the one leading to the close-ring isomer, are not in agreement with the quantum yield of 0.2, that seems to suggest the primary photoproduct should be the reactant open-ring.



## 5.4 Interpretation of the results

If in the previous section the collected results, relative to the light drive electrocyclicization, have been presented, in this section, they will be discussed and commented in order to give a complete interpretation and a refined picture of 2,5-dimethyl-furyl-fulgide photochromism, trying to rationalize its photoefficiency.

The relative energy of the main critical points, located on the reactive  $S_1$  state are collected in a unique table 5.6 and plotted respect to the reaction coordinate in the fig.5.16: there are summarized the  $S_1$  reaction path for all the processes, the ring-opening and closing reaction as well as the relaxation channels to both the photoproducts.

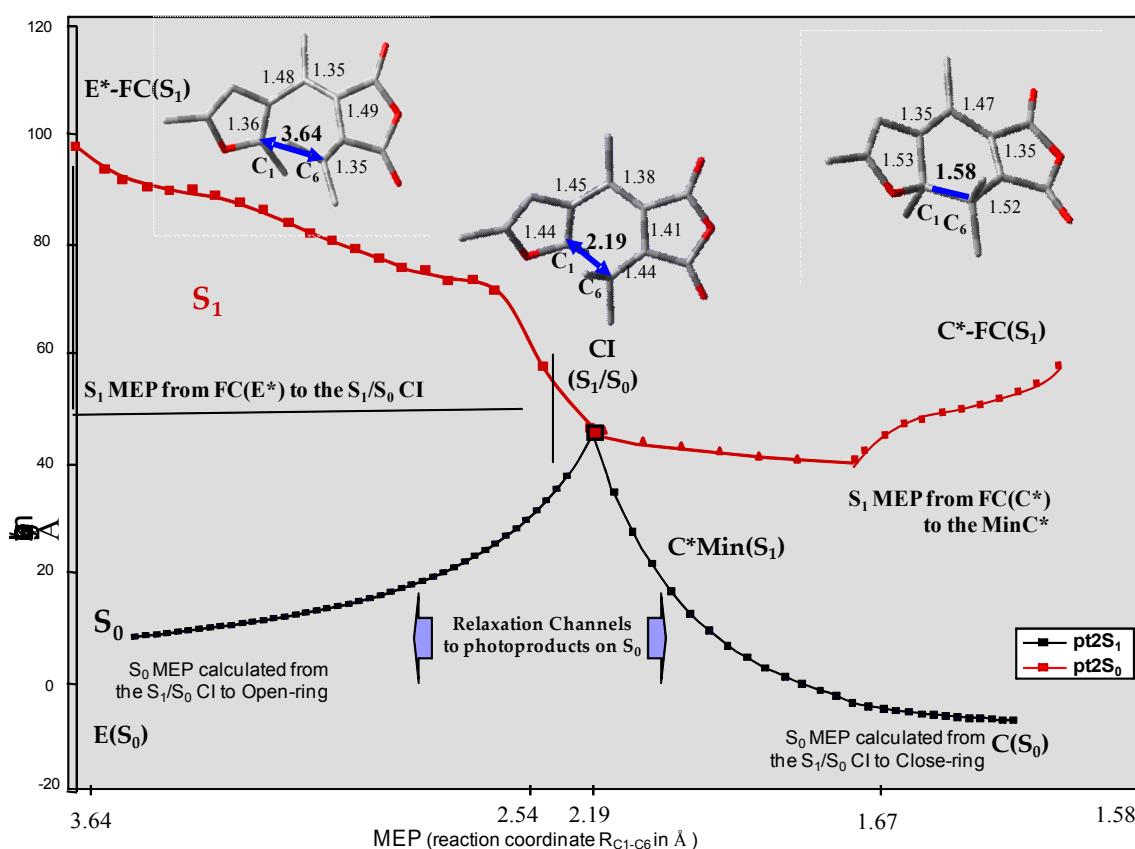
Figure 5.16 displays the MEPs disclosing the photochemical electrocyclic process. Starting on the open-ring side the results have provided evidence that immediately following  $E(S_0 \rightarrow S_1)$ FC excitation, molecules get funnelled along a barrierless path that develops entirely on  $S_1$ , and drives the system directly into a peaked  $S_1/S_0$  conical intersection (CI), whose geometry displaying a highly shortened (2.19 Å) C5-C9 bond with respect to the FC structure (3.64 Å), reveal how at this stage the ring-closure reaction is going to be completed. The analysis of the CI wave function and of the charges distribution, point out how the structure does not reflect the classical tetraradicalic character, typical of the polyene conical intersections. The charge distribution confirms its ionic nature being a crossing between the charge transfer state  $S_1$  and the covalent state  $S_0$ . Thus if the electronic re-coupling typical of a tetraradicalic CI is able to explain the several photoproducts observed on  $S_0$  (like several others photochromic systems), the electronic rearrangement of this ionic CI is completely different according with the only one photoproduct observed: the ring-closed isomer.

**Table 5.6** Relative ( $\Delta E$  in kcal mol<sup>-1</sup>) and absorption ( $\lambda$  in nm) energies, corresponding oscillator strength ( $f$ ), wavefunction characterization, dipole moment ( $\mu$  in debyes) and Mulliken charges analysis for the computed structures. [a] CASPT2 and TD-DFT values (in italic). Imaginary level-shift 0.2 au. [b] The state is 51.1 kcal mol<sup>-1</sup> above the  $S_0$  minimum (FC structure). [c] Referred to  $S_2$

Geometry		$\Delta E^{[a]}$ (kcal/mol)	$\lambda_{\text{calc}}^{[a]}$ ( $\lambda_{\text{exp}}$ ) (nm)	$f_{0>n}^{[a]}$	Config weight <sup>[a]</sup> casscf/ <i>td</i> dft	$\mu$ debye	Fragment charge <sup>[b]</sup>	
								
<b>Isomer E</b> ( $S_0$ )	$S_0$	0.00	exp(366)	0.00	(4a) <sup>2</sup> (5a) <sup>0</sup> 0.856	6.06	-0.036	0.036
	$S_1$	89.84 <i>81.75</i>	318.48 <i>349.97</i>	0.141 0.100	(4a) <sup>1</sup> (5a) <sup>1</sup> 0.786/0.644	12.82	0.539	-0.539
<b>Isomer C</b> ( $S_0$ )	$S_0$	0.00	exp(516)	0.00	(4a) <sup>2</sup> (5a) <sup>0</sup> 0.859	5.56	-0.039	0.039
	$S_1$	59.95 <i>58.57</i>	477.26 <i>487.29</i>	0.246 0.156	(4a) <sup>1</sup> (5a) <sup>1</sup> 0.882/0.619	11.58	-0.371	0.370
<b>CI</b> ( $S_1/S_0$ )	$S_0$	0.00	0.00	0.00	(4a) <sup>2</sup> (5a) <sup>0</sup> 0.683	6.48	0.027	-0.027
	$S_1$	3.85 <i>0.41</i>	7431.6 <i>69785.4</i>	0.002 0.	(4a) <sup>1</sup> (5a) <sup>1</sup> 0.820	10.86	0.209	-0.209
<b>Isomer C*</b> ( $S_1$ )	$S_0$	0.00	0.00	0.00	(4a) <sup>2</sup> (5a) <sup>0</sup> 0.715	7.24	-0.049	0.046
	$S_1$	28.51	1003.58	0.100	(4a) <sup>1</sup> (5a) <sup>1</sup> 0.764	12.92	-0.362	0.362



Despite here we discuss and comment only the  $S_1$  minimum energy path, whose ionic character is preserved along the pathway, in reality the real CASPT2 energy order of the states is more complicated along the path (see Scheme 5.3): several crossing, along the reaction path occur between the first singlet excited state,  $S_1$  (the charge transfer state) and the dark covalent state  $S_2$  or the higher singlet excited states. The mechanistic features (shortening of the bond formed upon ring-closing) and energetic (i.e. lack of excited state intermediates and barriers) nicely account for a photoinitiated (yet not finalized on  $S_1$ ) cyclization process and an ultrafast (subpicosecond) IC that prevents  $E^*$  light emission. The ring-closing reaction occurring on a barrierless surface explained the empirically time demand of the photocyclization on the order of picosecond, as the lifetime  $\tau_1$  of 0.3 ps relative to the  $E_0$  formation and as the  $\tau_2$  of 3.3 ps relative to the  $C_0$  formation.



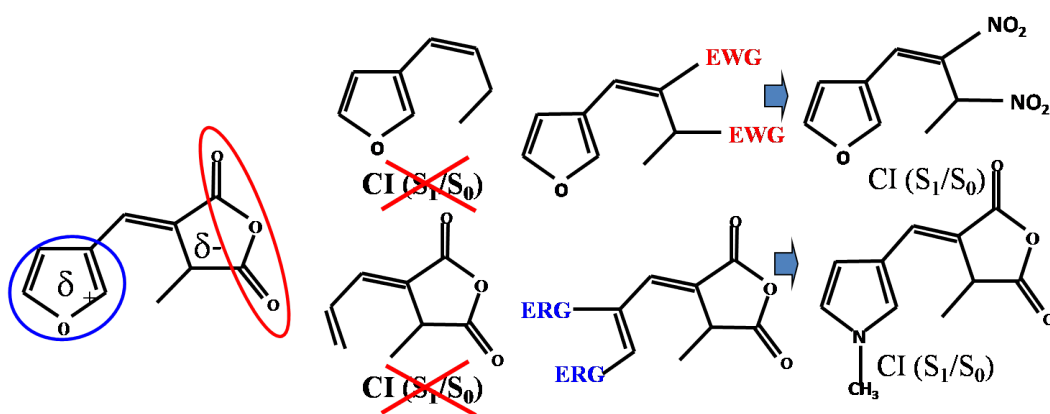
**Figure 5.16**  $S_1$  MEP-computed at CASSCF level and corrected at CASPT2//6-31g\* level-are quantitatively plotted in this scheme, to represent both the ultrafast processes: the ring-closing and the ring-opening reactions. In addition there are represented also the relaxation channels connecting the  $S_1/S_0$  CI to the photoproducts.

Mapping of the relaxation channels on  $S_0$  following decay through the CI reveals that only two paths exist: one leading back to  $E$ , the other to the photocyclization product  $C$ . The former appears more favoured than the latter as it starts closer to the CI and is steeper, matching with a  $E \rightarrow C$  quantum yield that is lower than 50% (ca. 20%). Notably, other possible processes (e.g. C5-C8 bond formation that is seen in many other less efficient photocyclizing photochromic compounds) do not occur here, thus ruling out unwanted side reactions that would cause intrinsic fatigue. This is again in nice

agreement with the high photochromic performance and fatigue resistance found in such particular fulgide (FF-3).

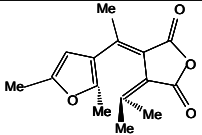
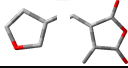
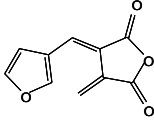
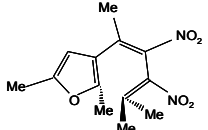
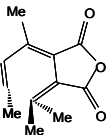
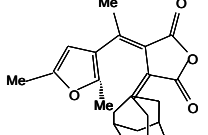
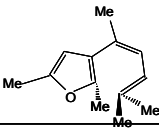
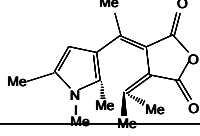

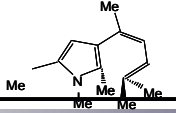
A somewhat different behaviour is recognized for the photoinduced process prompted by  $C(S_0 \rightarrow S_1)$ FC excitation. Interestingly, the  $S_1$  MEP drives the system into an excited state intermediate ( $C^*_{\text{Min}}$ ) where the  $S_1$ - $S_0$  energy gap is still large (ca. 30 kcal/mol) and that displays a stretched (although not yet broken) C5-C9 bond (1.67 Å). By proceeding along the C5-C9 bond-elongating coordinate, and by overcoming a small 5 kcal/mol barrier, the same  $S_1/S_0$  CI described before is assessed, although through a sloped path (see fig.5.16). Notably, the existence of a minimum on  $S_1$ , the barrier to the crossing region and the sloped CI well accounts for the observed weak signal of fluorescence and the smaller photocyclization quantum yield (ca. 8%) of the closed ring form that can still undergo (although in a longer timescale) a non-radiative ring opening process through the CI due to the small barrier. Additionally, it is suggested here that  $C^*_{\text{Min}}$  is the intermediate (I) associated to the slow C formation path: in fact, part of the photoexcited  $E^*$  molecules can escape the direct  $E^* \rightarrow \text{CI}$  route and be trapped in  $C^*_{\text{Min}}$ , thus following the fate of  $C^*$  (that is indeed involved in a slower IC process).

The reversibility of the photocyclization, the high ring-closing rate and quantum yield stability, the lack of unwanted photoinduced processes (e.g. formation of bicyclic products) and, more generally, all those remarkable properties that do characterize the *two-way* photochromism of furyl-fulgides, do all depend on the electronic nature of the photochemically relevant excited state  $S_1$  (and the characteristics of the CI involved), namely its ionic charge-transfer character (and the ionic/covalent nature of the corresponding crossing with  $S_0$ ): this marks and distinguishes the photoreactivity of furyl-fulgides to that of other (less efficient) 6-electrons systems (e.g. cyclohexadiene, DHN, azulene, diaryl-ethenes, etc) whose photochemistry is driven by a covalent  $S_1$  state. Thus, for instance, while the latter state triggers the formation of a biradical minimum and a barrier is necessary to access the  $S_1/S_0$  CI for the photocyclization, the former does lead to a direct barrierless path from  $E^*$  to the CI. Additionally, remarkable differences appear in the geometrical/electronic properties of the CI involved that prompt different photochemical outcomes.



**Scheme 5.5** In this scheme there is briefly summarized the effect of removing electron withdrawing or electron releasing groups at the CI geometry: without the stabilizing effect on the charge transfer state of these groups the  $S_1$ - $S_0$  degeneracy is removed and the conical intersection is not a crossing anymore. The energy gap is again reduced by introducing other EWG (as nitro groups) or ERG (methyl-pyrrole group) groups instead of the original.

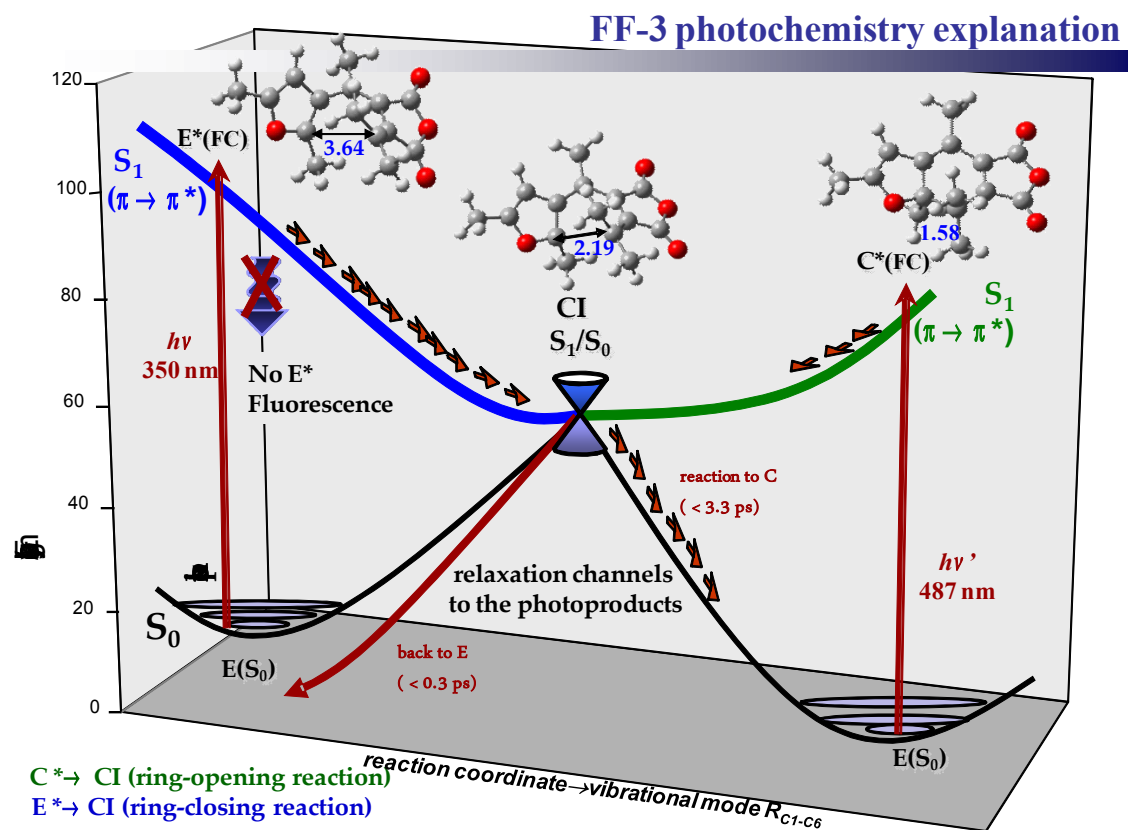
The reason for the intrinsic high stability of the ionic state in furyl-fulgides stands in the high electron withdrawing (EW) and electron releasing (ER) ability of the anhydride and heteroaromatic groups, respectively, that stabilize this state and its underlying charge transfer. Very remarkably, this understanding opens unprecedented scenarios for the *de novo* design of efficient related photochroms. In fact, it may be assumed that other substituents that act as good EW/ER groups can indeed stabilize this state and disclose an ‘ionic’ photochemistry that is different (and more efficient) than the covalent one seen in the other photochroms. Thus, provided no other side effects are generated, we can recreate fulgides-likes conditions by using different groups. Scheme 5.5 depicts this reasoning by displaying the  $S_1$ - $S_0$  energy gap at the CI structure of fugide for different substitutions. Very remarkably, it is shown that when we do not have such a stabilization effect on the ionic state, the crossing is removed and the covalent excited state becomes  $S_1$ . On the other hand, when we have strong EW/ED groups, than the ionic state turns back to be  $S_1$  and the ionic/covalent CI is recovered. Interestingly, EW groups such as the nitro groups, cyano groups, and ED groups such as methyl-pyrrole one seem to be promising replacements as they turn out to produce an ionic state and a CI of similar (or even better) stability. With this work we warmly suggest to look at this direction as a new potentially attracting route for the design of novel efficient photochromic systems.

Geometry Model	CASPT2 $\Delta E(S_0-S_1)$ kcal/mol	MS CASPT2 $\Delta E(S_0-S_1)$ kcal/mol	Substituted Groups	CASPT2 $\Delta E(S_0-S_1)$ kcal/mol	MS CASPT2 $\Delta E(S_0-S_1)$ kcal/mol
	3.85	2.01		12.05	12.07
	9.13	9.41		4.91	10.61
	19.59	20.25		5.60	5.61
	17.65	8.99		17.75	20.43
	29.15	25.16		29.15	25.16

**Table 5.5** Substituents modification effects on the CI. On the first column there are reported how much the  $S_1$ - $S_0$  energy gap increase removing the anhydride group or the furyl group, and on the second column how the same  $S_1$ - $S_0$  energy gap is again reduced with the introduction of EW and ER groups.

## 5.5 Conclusions

The main purpose of this work was to provide for the first time, a detailed high level computational model with the aim of rationalize such efficient photochromism due to 2-5-dimethyl-furyl-fulgide electrocyclization, such largely experimentally studied. This was accomplished through identifying the most efficient decay and isomerization path for the photoinduced open-ring (and reverse) reaction mechanism. The highest-level ab initio computational methods were employed to adopt an adequate quantum chemical strategy, through which we have identified the minimum energy path of both the process (the close-ring as well as the open-ring reactions) on the  $S_1$  PES and a radiationless decay point, extremely relevant for the photophysical and photochemical properties of the system. In both the relaxation channels the MEP was found to follow the reaction coordinate dominated by the same vibrational mode causing the progressive decrease or increase in the distance of the two reactive carbon atoms.



The barrierless profile was perfectly in agreement with the absence of fluorescence and the lifetime (ps) of the ring-closing reaction suggested as an ultrafast process. The ionic nature of the reactive state is able to rationalize how the electronic re-coupling after the radiationless decay through  $S_0/S_1$  CI, permits only one photoproduct, according with the only  $C_0$  formation and its high efficiency. Finally it is proposed in fig.5.17 a schematic outline of both the photoinduced electro-cyclization featuring this particular efficient photochromism.

### References:

- 1) a). Irie, M. *Chem. Rev.* 2000, 100, 1685. b) Bouas-Laurent, H. Bouas-Laurent, H. ; Durr, H. *Pure Appl. Chem.* 2001, 73, 639
- 2) Alfimov M., Fedorova O, Gromov S, *Jour. Of Photochemistry and Photobiology A*, 2003, 158, 183-198.
- 3) Heller, H. G. and Szewczyk M., "Photoreactions of photochromic  $\alpha$ -phenylethylidene succinyl anhydrides" *J. Chem. Soc. Perkin*, 1, 1321, 1972.
- 4) Yokohama Y. "Fugides for Memories and Switches" *Chem. Rev.* 2000, 100, 1717-1739.
- 5) Santiago H., Die Farbe der Fulgensauren und Fulgides, *Cgem. Ber.*, 38, 3673, 1905
- 6) Crano, J. C.; Guglielmetti, R. J., *Organic Photochromic and Thermochromic Compounds*. Plenum Press: New York, 1999; Vol. 1.
- 7) Crano, J. C.; Guglielmetti, R. J., *Organic Photochromic and Thermochromic Compounds*. Kluwer Academic: New York 1999; Vol. 2.
- 8) (a) Heller, H.G.; Szewczyk, M. *J. Chem. Soc. Perkin Trans. 1* 1974, 1487-1492 (b) Heller, H.G.; Oliver, S. *J. Chem. Soc. Perkin Trans 1* 1981, 197-201.
- 9) Boggio-Pasqua, Marcella Ravaglia, Michael Bearpark, Marco Garavelli and Mike Robb. *J.A.C.S.* 2003, 107, 11139-11152
- 10) (a) J Ern, Petermann, T. Mrozek, J. Daub, K. Kuldova', Kryschi, C. *Chem, Phys*, 2000, 259, 331-337 (b) Martial-Boggio-Pasqua, ecc-- *J. Phys. Chem A* 2001
- 11) (a) Helmut Gerner, Thomas Mrozek, and Jorg Daub. *Chem. Eur.* 2002, 8, 174008-4016 (b) G. Tomasello, Francois Ogliaro, Michael J. Bearpark, Michael A. Robb, and Marco Garavelli *J. Phys. Chem. A*; 2008; 112(41) pp 10096 – 10107.
- 12) . Marco Garavelli Phs Thesis 1997.

- 13) Stobbe, H. Die Fulgide, *Annalen*, 380, 1, 1911.
- 14) Stobbe, H. Phototropiererscheinungen bei Fulgiden und Anderen Stoffen, *Annalen*, 359, 1, 1908.
- 15) Yokoyama, Y. *Chem. Rev.* 2000, 100, 1717.
- 16) Heller, H. G. and Szewczyk M., Photoreactions of photochromic  $\alpha$ -phenylethylidene succinici anhydrides, *J. Chem. Soc. Perkin*, 1, 1321, 1972
- 17) Rath, Helling, Port et al. *Nanoletters* 2007, 7, 12,
- 18) Heller, H. G.; Hughes, D. S.; Hursthouse, M. B.; Koh, K. V. S. *J. Chem. Soc., Chem. Commun.* 1994, 2713.
- 19) (a) Yokoyama, Y.; Goto, T.; Inonue, T. Yokohama, M.; Y. *Chem. Lett.* 1988 1049-1052 (b) Yokohama, Y., Kurita Y. *Mol. Cryst. Sect. A* 1984, 246, 87-94
- 20) . M. Handschuh, M. Seibold, H. Port and H. C. Wolf. *J. Phys. Chem. A* 1997, 101, 502-506.
- 21) Rappon, Syvitski, *Journ. of Photochemistry and Photobiology A. Chemistry A* 1996, 243-247.
- 22) Kurita, Y. Yokoyama, Y. Goto, T. Inoue, T. and Yokoyama, M., Fulgides as efficient photochromica compunds. *Chem. Letters*, 1049, 1988.
- 23) Gonzalez-Luque, R.; Garavelli, M.; Bernardi, F.; Merchan, M.; Robb, M. A.; Olivucci, M. *Proc. Natl. Acad. Sci. U.S.A.* 2000, 97, 9379.
- 24) Garavelli, M.; Negri, F.; Olivucci, M. *J. Am. Chem. Soc.* 1999, 121 (1), 1023.
- 25) De Vico, L.; Page, C. S.; Garavelli, M.; Bernardi, F.; Basosi, R.; Olivucci, M. *J. Am. Chem. Soc.* 2002, 124, 4124.
- 26) Celani, O.; Robb, M. A.; Garavelli, M.; Bernardi, F.; Olivucci, M. *Chem. Phys. Lett.* 1995, 243, 1.
- 27) Andersson, K.; Malmqvist, P.-Å.; Roos, B. O. *J. Chem. Phys.* 1992, 96, 1218.
- 28) Andersson, K.; Barysz, M.; Bernhardsson, A.; Blomberg, M. R. A.; Carissan, Y.; Cooper, D. L.; Fülcher, M. P.; Gagliardi, L.; de Graaf, C.; Hess, B. A.; Karlström, G.; Lindh, R.; Malmqvist, P.-Å.; Nakajima, T.; Neogrady, P.; Olsen, J.; Roos, B. O.; Schimmelpfenning, B.; Schutz, M.; Seijo, L.; Serrano-Andrés, L.; Siegbahn, P. E. M.; Stålring, J.; Thorsteinsson, T.; Veryazov, V.; Widmark, P.-O. *MOLCAS, Version 6.0; Lund University: Lund, Sweden, 2003.*

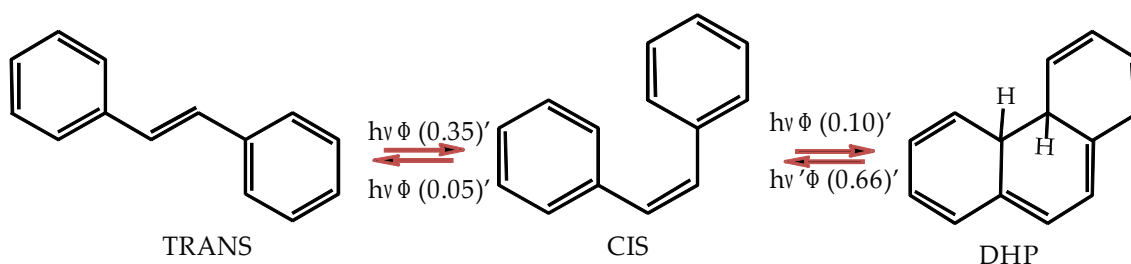
- 29) Andersson, K.; Barysz, M.; Bernhardsson, A.; Blomberg, M. R. A.; Carissan, Y.; Cooper, D. L.; Fülischer, M. P.; Gagliardi, L.; de Graaf, C.; Hess, B. A.; Karlström, G.; Lindh, R.; Malmqvist, P.-Å.; Nakajima, T.; Neogrady, P.; Olsen, J.; Roos, B. O.; Schimmelpfenning, B.; Schutz, M.; Seijo, L.; Serrano-Andrés, L.; Siegbahn, P. E. M.; Stålring, J.; Thorsteinsson, T.; Veryazov, V.; Widmark, P.-O. MOLCAS Version 6.2; Dept. of Theor. Chem.: Lund, Sweden, 2005.
- 30) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. Gaussian 03, Gaussian, Inc.: Pittsburgh, PA, 2003.
- 31) Olivucci, M.; Robb, M. A.; Bernardi, F. In *Conformational Analysis of Molecules in Excited States*; Waluk, J., Ed.; Wiley-VCH: John Wiley and Sons Inc.: New York, 2000; p 297.
- 32) Robb, M. A.; Garavelli, M.; Olivucci, M.; Bernardi, F. In *Reviews in Computational Chemistry*; Lipkowitz, K. B.; Boyd, D. B., Eds.; Wiley-VCH: John Wiley and Sons Inc.: New York, 2000; p 87.
- 33) Boys, S. F. *Rev. Mod. Phys.* 1960, 32, 296.
- 34) Blancafort, L.; Celani, P.; Bearpark, M. J.; Robb, M. A. *Theo. Chem. Acc.* 2003, 110 (2), 92.
- 35) a) Becke, A. D. *Phys. Rev. A* 1988, 38, 3098. b) Becke, A. D. *J. Chem. Phys.* 1993, 98, 1372.
- 36) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* 1998, 37, 785.
- 37) Yasunori, Yoshioka, M. Usami, M. Watanabe, Yamaguchi. *Journal of Molecular Structure(Theochem)* 623 2003 167-178.

## -CHAPTER 6-

### Stilbene photo-physical properties

#### 6.1 Introduction

In the field of spectroscopy, stilbene<sup>1</sup> is certainly one of the most extensively investigated organic compounds. One of the reasons for the continuing interest in this molecule is its leading role in the investigation of photoinduced isomerizations<sup>2</sup>. These processes are not only crucial for the understanding of many photochemical transformations and their solvent dependence but also for our own view of the external world, as they form a key step in the visual process. Stilbenes and stilbenoid compounds are indeed frequently textbook examples of photochemical cis-trans isomerization<sup>2</sup>. Their reactivity has an important role as model compound of biological phototropic systems. They, also serve as building blocks for organic materials, whose properties could be used in optical and electro-optical applications such as optical data storage, laser dyes, nonlinear optics, or photochemically, cross-linked polymers<sup>3</sup>.



**Figure 6.1** There is a summary of stilbene photochemistry. The trans-cis photoisomerization and the reverse one is represented on the left side with the quantum yields reaction reported on the parenthesis. On the right side it is represented the electrocyclic reaction, initiated and completed on the dark covalent side, leading to the close-ring product (dihydrophenantrene-DHP) with also the relative quantum yields of the reactions.

Stilbene systems react in presence of light by undergoing a change in their electronic and geometric structures (fig.6.1), in particular there are two possible photoreactions: starting from cis-isomer the electrocyclic process leading to the dihydrophenantrene (DHP) is less favorite ( $hv\Phi=0.10$ ) than the cis-trans isomerization ( $hv\Phi=0.35$ )



We have mainly focused our attention to modeling the photoisomerization, with its reaction coordinate defined by the torsion around the central C=C bond, and despite the remarkable advances in our understanding of such processes<sup>4</sup>, they have not been satisfactorily interpreted.

There are still several aspects of stilbene photoreactivity that need to be clarified, since the huge amount of experimental data<sup>5</sup> are not completely explained (rationalized).

For example there is still lacking a commonly accepted description of main steps defining the photo-induced isomerization and if it is sufficient only one normal mode defining the reaction coordinate. If for fulgide photoreactivity (chapter 5) the reaction coordinate was clearly dominated by a vibrational mode between the reactive carbons, here the central torsional mode, responsible of the cis-trans isomerization, may be not sufficient to guarantee the radiationless decay to S<sub>0</sub> will occur. Two main models were proposed through which the isomerization mechanism may occur: whereas in the conventional one-bond flip (OBF) mechanism the substituents of the C=C bond move along the surface of a cone, in the hula-twist (HT)<sup>6</sup> mechanism only the CH group of the double bond rotates out of plane and the other parts reorient within the plane.

In particular S<sub>0</sub> and S<sub>1</sub> potential energy surfaces, with special emphasis to the potential energy curves along the torsion around the ethylenic C=C bond, have been crucial for a complete understanding studied by many authors and by different theoretical approaches. Following this line guide, summarized by the OBF (One Bond Flip) model, the conventional picture of this process has been a two-state, one-dimensional model, stressing torsional motion as the primary nuclear coordinate. Nevertheless with the accumulation of ultrafast data for several hydrocarbons some theoretical proposal for a more complicated model have began to be considered for the smallest unsaturated hydrocarbon<sup>7,8</sup>.

After excitation of cis- and trans-stilbene in the first UV band, the ethylenic double bond twists by 90°, and an S<sub>1</sub> minimum at perpendicular geometry (p\*) is reached from both isomers<sup>9</sup>. For leaving the fluorescent region (cis\* and trans\*) towards p\*, the molecules must overcome a small barrier which is, however, barely detectable for cis\*.

In the gas phase, the lifetime for S<sub>1</sub> trans population (trans\*) extends from 10 to 100 ps depending on the excess energy<sup>10</sup> and 300 fs for cis photo-excited population<sup>11-12</sup>.

The lifetime in the perpendicular minimum (p\*) has only recently been measured and found to be 135–160 fs<sup>11</sup>. This short lifetime implies that the path leads through an easily accessible CI, as pointed out previously<sup>12-15</sup>. After the C=C twist, the path hence changes direction. The new coordinate either involves a twist of the adjacent single bond (leading to a CI of Hula-Twist type) or pyramidalization<sup>16</sup> of one ethylenic C atom (OBF-type CI).

Nevertheless, some evidence of the inadequacy of a one-dimensional model is available for the trans isomer<sup>17</sup>, and even more for the cis-isomer<sup>18,19</sup>, partially expected since steric repulsion forces the ground-state geometry of the cis isomer to be non-planar, and both photoisomerization and photocyclization pathways can be accessed from the cis\*.

The cis-isomer upon  $\pi \rightarrow \pi^*$  excitation have two main decay route: following the first one from the lowest covalent 1B state the dihydrophenantrene derivative is observed (photocyclization product)<sup>20</sup>, while the second one is responsible of the cis-trans isomerization. For example among the several unsolved problems there is the possibility of an adiabatic cis-trans isomerization process, especially observed for the larger systems as stiff-stilbene and others. Further several experiments have recovered a weak

fluorescence from  $S_1$  cis-isomer population according with a flat topology of  $S_1$  the potential energy surface along the torsional coordinate.

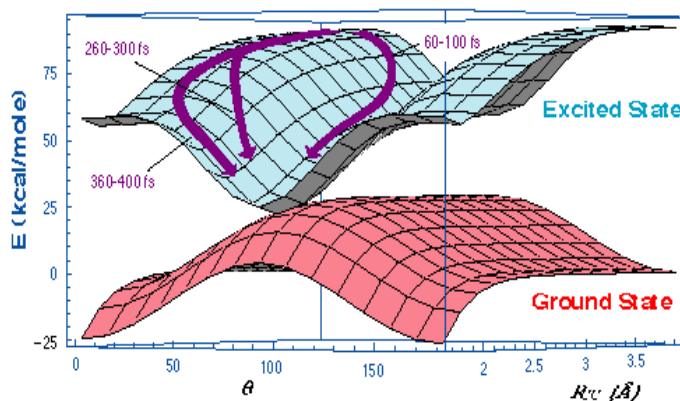
If no barrier to torsion in cis-isomer was detected, on the opposite side such barrier is clearly established for trans-isomer. On the opposite side the commonly accepted picture suggests the mechanism of photoisomerization of trans-stilbene is based on the following concepts. The  $S_1$  state (belonging to B symmetry), as indicated by intensity of the lowest energy absorption band, it has large transition oscillator strength. According also with the lifetime and fluorescence spectra,  $S_1$  is the optically active state that, characterized by the dominant HOMO-LUMO excitation, generates a torsional barrier of about  $1250\text{ cm}^{-1}$  (3 kcal/mol) connecting the planar relaxed intermediate (trans\*) with the perpendicular one (p\*). This barrier has been interpreted as arising from either the interaction of an optically dark A state with the optically accessed 2B state, or the lowest 1B state with the higher-lying 2B state. In this chapter the calculations computed will reproduce such torsional barrier. The doubly excited (HOMO-LUMO)<sup>2</sup> singlet state, of A symmetry, observed in the two-photon spectroscopy<sup>21</sup>, may become the lowest excited state at the twisted geometry and favor the photoisomerization.

In this chapter such torsional barrier, connecting the planar  $S_1$  excited minimum (optimized from FC) with the twisted region, has been reproduced through only one potential. Although in the past, to calculate vertical excitations and the correct state order was also a big challenge, recently most of the computational effort was concentrated not on the Franck Condon region, but rather on the intersection region. The wavefunction nature of the state approaching each others, is still not clearly identified. The so called “phantom state” is the supposed lowest excited state featuring the twisted intermediate (p\*), and any accurate *ab initio* calculations<sup>22</sup> seem to have provided a sufficiently clear wavefunction analysis of such state. A quantitative assessment of the ordering of the lowest electronic states and of the shapes of this region of potential energy surfaces and thus of possible this mechanism to the decay of  $S_1$  is still lacking.

**Fig6.2** There is represented the  $S_1$  topology at the twisted geometry.

An important contribution in this is provided by a recent paper by Martinez et al<sup>16</sup> in which for the first time conical intersection geometries and possible reaction path leading the molecule to the decay, have been provided.

In particular, they have been useful for the characterization of pyramidalization coordinate and in addition the assignment of the charge-transfer character to the wavefunction state to such structure. In reality the two things are strictly correlated and it is just the particular distortion of the geometry with the  $Csp^2$ - $Csp^3$  pyramidalization (rehybridization) of the “anionic” carbon, having a stabilizing effect on its energetic. Until now this is the best description of stilbene photoreactivity, and it was done through less accurate level of CASSCF with a limited active space, of two electrons in two orbitals (2,2). Just for this reason the aim of the presented work here



was to improve the one-dimensional model (inadequate for stilbene photoisomerization) through a major level of computational accuracy. Several combined CASSCF//CASPT2 strategies were adopted, trying to improve the model nowadays known, in order to get a more accurate classification of electronic states involved in the  $S_1$  reaction path, for providing a more detailed mechanism of vibrational relaxation and the geometries of the decay to  $S_0$  still not well established. In reality the major part of tentative we tried did not give us reliable results, but a series of discrepancies not comparable with the experiments. This is due to the CASSCF limitation in treating the ionic reactive state whose nature request a particular description<sup>23</sup>. For computing these states, correlated<sup>24</sup> and multireference methods are usually necessary for a qualitatively correct representation of the wave function. Furthermore, obtaining accurate electronic excitation energies usually requires an additional high-level treatment of dynamic electron correlation effects. Unfortunately, this treatment is likely to be computationally expensive, lacking also the analytical energy gradients for geometry optimizations. This is a severe limitation, as the gradient has to be computed numerically, which becomes more expensive and less accurate as the number of atoms increases. Among the method that usually combine reasonable accuracy with speed sufficient there is CASSCF<sup>25</sup>, that usually offer a poor description of the dynamic electron correlation so important for ionic excited states. In particular CASSCF includes only the static (long-range, non-dynamic) correlation by default.

Despite mono-determinantal, the particular influence of the  $\sigma$ - $\pi$  excitation on the  $S_1$  stilbene wave function, dominated by a  $\pi \rightarrow \pi^*$  character, reveals how the non-dynamical correlation energy was important to correctly describe the excited state behavior and the topological aspect of its potential energy surface. Remarkable CASSCF method focuses on just a few “active” electrons, in active orbitals chosen to give multiple configurations describing the required electronic state, but does not take in account the influent. Unfortunately, correlating only the active electrons often gives energies that are quantitatively incorrect. For reliable calculations of ionic excited states, one must go beyond the CASSCF level and include dynamic correlation between active and inactive electrons. In fact once chosen the active space, all orbitals spectator and active are fully optimized in the calculation, but correlation energy is only recovered for the active electrons, thus only the  $\pi$ -component of the nondynamic correlation is recovered. Therefore if with CASSCF still use the same core  $\sigma$  orbitals for all configurations, preventing the polarization of the core needed to stabilize ionic configurations further.

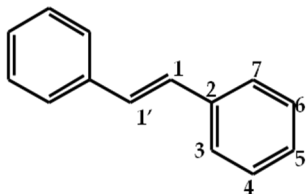
To remedy this flaw, the strategy that seems to best describe the  $S_1$  reaction path for the photoisomerizations of both isomers giving energy profiles in agreement with the experimental data, consists of using a RASSCF<sup>26</sup> wavefunction comprising of a sub-active space (denominated Ras2) of 2 electrons into 2 orbitals within them all the excitations are considered (the Martinez method) plus the contribution of only single excitations between other two subspaces called Ras1 and Ras3. Through this method the contribute of single  $\sigma$ - $\pi$  excitation is taken in account and is determinant for stabilizing ionic state and its energy profile along the isomerization coordinate

## 6.2 Computational strategies adopted and consequent results.

### 6.2.1 Applying CASSCF//CASPT2 methods: Vertical excitations.

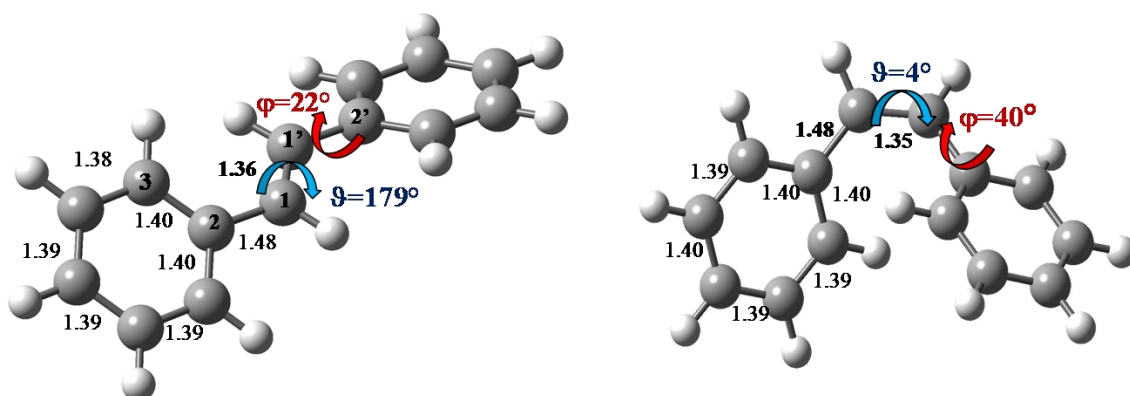
In order to improve the simple two-state model proposed by Martinez with its active space treatment of 2 electrons in 2 orbitals, we, first carried out several CASSCF//CASPT2<sup>24</sup> strategies of calculations by including a higher number of electrons into the orbitals, trying to find a reliable description of the S<sub>1</sub> state dominated by the HOMO-LUMO excitation.

The first choice was a combined CASSCF//CASPT2(10,10) approach, thus with 10 electrons into 10 orbitals, both series of calculations were performed with and without preserving symmetry. Two choices of basis set were mainly adopted: the first one consisted of treated all the atoms with the same 6-31g\* basis set, while the second one consisted of adding diffuse functions to only the ethylenic central C1 and C1' (as numbered on the left) carbons, in order to get a more reliable description of the ionic 'charge-transfer'



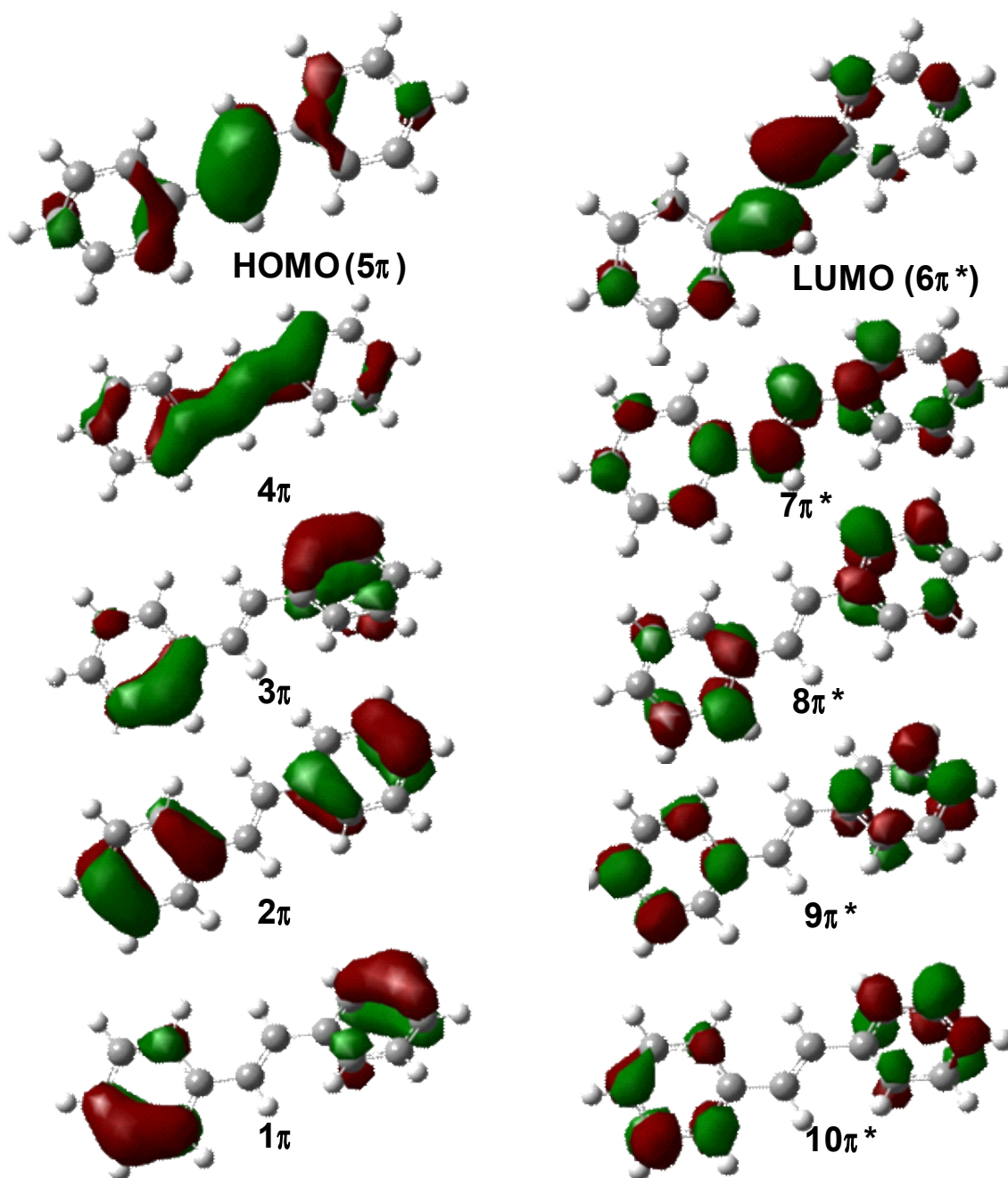
character of the S<sub>1</sub> state (dominated by the HOMO-LUMO transition). The additional basis function are often added in order to stabilize the ionic configurations by recovering the  $\pi$ -component of dynamic valence-virtual correlation as well, which can be done by extending the  $\pi$  active space. Nonetheless additional diffuse functions on the ethylenic carbons have only some moderate effect on the excitation energy and on the geometrical parameters. Most of the energy values discussed will be at the CASPT2 level.

On the ground state two minima were located on the opposite sides, the cis and trans isomers (the last one more stable of about 2 kcal/mol) separated from a twisted transition state (90°) located on a barrier of 42 kcal/mol (43 kcal/mol with the diffuse). Both minima were optimized by using CASSCF(10,10) potential, and the geometries found belong to C<sub>2</sub> symmetry: the trans isomer is not planar<sup>27</sup>, due probably to the steric hindrance of two hydrogens (absent into planar azobenzene S<sub>0</sub> minimum) that make the  $\varphi$  dihedral angle to be 22°. On the opposite side the S<sub>0</sub> cis minimum is characterized from a geometry twisted with dihedral angles  $\theta=4^\circ$  and  $\varphi=40^\circ$  (see figure 6.3).



**Fig6.3** There are represented the two minima geometries of both isomers: on the left side there is the trans isomer whose structure assigned to C<sub>2</sub> representation and not C<sub>2h</sub> being not planar, due to the hydrogens hydrance making the dihedral angle  $\varphi$  to be 22°, while the central angle  $\theta$  is almost 180°. On the right side the cis geometry (minimum on S<sub>0</sub>) belong to C<sub>2</sub> symmetry, being twisted of only 4° respect to the central dihedral angle  $\theta$ , while the two benzene ring are twisted ( $\varphi$ ) respect to the double bond C=C of 40°.

This work reexamines the ordering of the lowest excited states<sup>28</sup> at the  $S_0$  equilibrium structures of both the conformers, using this more advanced technique. However, before the beginning of discussion, one should remember that vertical excitation energy is a theoretical concept and can be estimated by experiment only qualitatively. With the Franck-Condon spectra available, a direct comparison of the 0-0 transitions and the vibronic progressions becomes possible. Nevertheless in this work, the  $S_0$ - $S_n$  vertical transitions were computed, thus the comparison between theoretical results and experimental measures will be directly limited by approximations adopted.



**Figure 6. 4** There are represented the orbitals comprised into the active space of CASSCF(10-10)/6-31g\* for the Trans isomer optimized on  $S_0$ . As it possible to observe they are all  $\pi$  orbitals with the HOMO and LUMO featured by an ethylenic character.

The figure 6.4 collects all the CASSCF  $\pi$  orbitals included in the active space at trans geometry. A CASSCF wave function is built by distributing active electrons in all consistent ways among the active orbitals: the active space selection is a crucial step in a CASSCF calculation. The stilbene molecule with its 14 valence  $\pi$  MO's and 14  $\pi$  electrons is a problematic case. For most of the calculations the choice adopted was a compromise between accuracy and computational cost, with the aim of guarantee an active space balanced and stable: the choice of 10  $\pi$  electrons in 10  $\pi$  orbitals seem to respond to such requirements.

Tables 6.2<sup>(a)</sup> and 6.2<sup>(b)</sup> summarized the results obtained for the vertical excitations of both isomers. The data collected into the table 6.2 reveal a clear discrepancy between CASSCF and CASPT2 behavior in the order of the lowest excited states. Part of the discrepancy in the vertical excitation energy is certainly due to the lack of dynamical electron correlation in the CASSCF method. We have carried out several calculations which include dynamic correlation in order to further characterize this discrepancy, and it appears that part of the problem is closely associated with the question of planarity in the  $S_0$  equilibrium geometry, especially for the trans isomer. Indeed, Roos and co-workers<sup>29</sup> have achieved good agreement with experimental vertical transition energies for the trans isomer by assuming a planar  $S_0$  geometry and using multireference perturbation theory (CASPT2). However, the planarity issue has been a source of some controversy,<sup>30-32</sup> and here we only comment that vertical excitation energy is sensitive to the degree of phenyl rotation and the trans global minimum we has the phenyl rings 22° rotated out of plane (as shown in figure 6.3).

However, in our work, although CASSCF//CASPT2(10,10) potential will fail on reproducing a correct  $S_1$  reaction pathway, the vibronic structure of absorption spectra have been determined in agreement with the experimental data with also a reliable CASPT2 ordering of the lowest excited state for both conformers.

All these calculations were performed in  $C_2$  symmetry, being the trans  $S_0$  minimum not planar. By placing the molecule in the such symmetry group the orbitals belong to the A and B irreducible  $C_2$  representations.

From the table 6.1<sup>(a)</sup> at CASSCF level for trans isomer only  $S_4$  is the optically active space (symmetry B) with an oscillator strength value of 1.48. The spectroscopic state, dominated by the (HOMO-LUMO)<sup>1</sup> single excitation, if corrected by the correlated method CASPT2, became the second root  $S_1$ , with  $S_0 \rightarrow S_1$  absorption energy calculated to be 102 kcal/mol (103 kcal/mol with the diffuse basis functions), only 3 kcal/mol underestimated respect to the experimental measure of 4.13 eV (105 kcal/mol). Following the CASPT2 order, among the others excited states, two of them, both of benzenic-nature are degenerate while the dark state, bi-excited (HOMO-LUMO)<sup>2</sup> state, was individuated as the highest root ( $S_4$ ).

If for the trans isomer the bright state is the first excited state (CASPT2 level) (according with the experimental data), on the opposite side for the cis isomer the electronic state associated with the most intense transition is only  $S_2$ , with an oscillator strength value of 0.59, and described by one-electron (HOMO-LUMO)<sup>1</sup> promotion. The  $\Delta E(S_0 \rightarrow S_2)$  calculated to be 111 kcal/mol nicely reproduces the experimental value of 113 kcal/mol.

State <sup>(a)</sup> (Trans Isomer)	$\Delta$ E_CASSCF <sup>a</sup> 6-31g*/ (6-31g(+)*C1-C1') Kcal/mol	$\Delta$ E_CASPT2 <sup>a</sup> 6-31g*/ (6-31g(+)*C1-C1') Kcal/mol	Oscillator Strength $f_{0-n}$ <sup>(c)</sup>	Configuration wavefunction weight	Symmetry
S <sub>0</sub>	0	0	---	(5 $\pi$ ) <sup>2</sup> (6 $\pi$ *) <sup>0</sup> 0.823	A
S <sub>1</sub>	135.37 (136.21)	113.24 (103.81)	0.00016	(5 $\pi$ ) <sup>1</sup> (9 $\pi$ *) <sup>1</sup> 0.118 (3 $\pi$ ) <sup>1</sup> (6 $\pi$ *) <sup>1</sup> 0.152	A
S <sub>2</sub>	135.86 (137.04)	113.47 (112.10)	0.03404	(5 $\pi$ ) <sup>1</sup> (10 $\pi$ *) <sup>1</sup> 0.175 (1 $\pi$ ) <sup>1</sup> (6 $\pi$ *) <sup>1</sup> 0.256	B
S <sub>3</sub>	146.22 (149.13)	122.93 (120.75)	0.00002	(2 $\pi$ ) <sup>1</sup> (6 $\pi$ *) <sup>1</sup> 0.178 (5 $\pi$ ) <sup>0</sup> (6 $\pi$ *) <sup>2</sup> 0.227	A
S <sub>4</sub>	<u>151.15</u> (148.64)	<u>101.55</u> (102.60)	<u>1.46816</u>	<u>(5<math>\pi</math>)<sup>1</sup> (6<math>\pi</math>*)<sup>1</sup> 0.826</u>	<u>B</u>

State <sup>(b)</sup> (Cis Isomer)	$\Delta$ E_CASSCF 6-31g*/(6- 31g(+)*C1-C1') Kcal/mol	$\Delta$ E_CASPT2 6-31g*/ (6-31g(+)*C1-C1') Kcal/mol	Oscillator Strength $f_{0-n}$ <sup>(c)</sup>	Configuration wavefunction weight	Symmetry
S <sub>0</sub>	0	0	---	(5 $\pi$ ) <sup>2</sup> (6 $\pi$ *) <sup>0</sup> 0.822	A
S <sub>1</sub>	109.72 (138.66)	106.13 (115.92)	0.00017	(5 $\pi$ ) <sup>1</sup> (8 $\pi$ *) <sup>1</sup> 0.308 (3 $\pi$ ) <sup>1</sup> (6 $\pi$ *) <sup>1</sup> 0.295	B
S <sub>2</sub>	154.69 (142.15)	139.27 (122.16)	0.02283	(5 $\pi$ ) <sup>1</sup> (8 $\pi$ *) <sup>1</sup> 0.154 (1 $\pi$ ) <sup>1</sup> (6 $\pi$ *) <sup>1</sup> 0.153	B
S <sub>3</sub>	<u>159.84</u> (159.98)	<u>111.25</u> (108.82)	<u>0.58837</u>	<u>(5<math>\pi</math>)<sup>1</sup> (6<math>\pi</math>*)<sup>1</sup> 0.806</u>	<u>B</u>
S <sub>4</sub>	174.34 (161.77)	145.50 (134.25)	0.00209	(2 $\pi$ ) <sup>1</sup> (6 $\pi$ *) <sup>1</sup> 0.121 (5 $\pi$ ) <sup>0</sup> (6 $\pi$ *) <sup>2</sup> 0.131	A

**Table 6.1** There are collected the S<sub>0</sub>-S<sub>n</sub> vertical excitations for the lowest five roots, for both the isomer (trans table 6.1<sup>(a)</sup> and cis table 6.2<sup>(b)</sup>), following the CASSCF order. The active space used was of 10 electrons into 10 $\pi$  orbitals, both the basis sets were used in parenthesis and in italic there is reported the energy value adding the diffuse functions on C1-C1'. In order to correct the energetic, correlated method CASPT2 was employed. <sup>(c)</sup>The oscillator strength values are referred to the  $\Delta$ E\_CASSCF. For both the isomer there is emphasized the bright state, the only excited state dominated by a single configuration. For the wavefunction analysis the orbitals represented into the Figure 6.4 may be helpful.

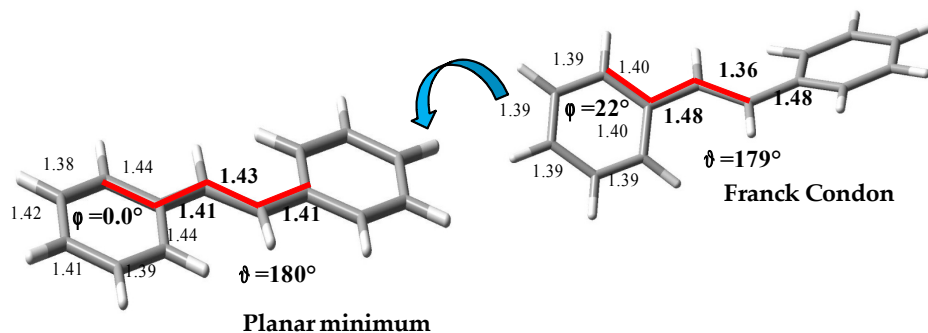
Despite in this chapter only the lowest five roots are discussed, the energy calculations were performed at the CASSCF//CASPT2(10,10) level with a State Average extended within the lowest 10 roots. The highest occupied and the lowest unoccupied molecular orbitals (HOMO and LUMO) are  $\pi$  and  $\pi^*$ -type. Therefore the active space consisted of the  $\pi^2$ ,  $\pi\pi^*$ , and  $\pi^{*2}$  configurations.

### 6.2.2 Applying CASSCF//CASPT2 and CC2 methods: trans\* relaxation.

If CASSCF//CASPT2(10,10) calculations correctly reproduce the ordering of lowest excited states for both the conformers, most of the results following presented and discussed do not have an experimental counterpart and are not supported by a reasonable explanation (rationalization). Once computed the trans  $S_0$ - $S_1$  excitation,  $S_1$  vibrational relaxation on  $S_1$  was investigated by using SS (Single State) and SA (State Average) CASSCF potential for geometry minimization. The geometries obtained from such calculations reveal how the CASSCF gradient direction was inadequate and in contrast with the wavefunction character. In particular for such problem, the choice of a proper active space was a crucial point and surprisingly the dependence of the geometry on the active space selection has an unexpected impact. In order to find a balanced active space for a correct gradient evaluation, we have explored several possibilities, going from a minimal choice of 10 active electrons in 10 active MO's (10/10), up to the largest calculation including all the 14 active electrons in the 12 active valence  $\pi$  MO's (14/12). At the end, considered the failure of the CASSCF gradient calculations and its inadequate potential for treatment of an HOMO-LUMO electronic state, a correlated method was requested. Since CASPT2 gradient (not yet available in MOLCAS<sup>33</sup>) are computationally too expensive, Coupled Cluster CC2 geometry minimizations were performed (using TURBOMOLE<sup>34</sup> package).

Nevertheless, although the single-reference configuration (with the ethylenic orbitals clearly involved into the electronic transition) is preserved along the  $S_1$  relaxation pathway, all the tentative converge to a common minimum planar. The relaxed geometry obtained was different for each of the different active space and characterized by a series of geometrical parameters not coherent with such wavefunction nature. The bond lengths collected in the table 6.2 provide a clear picture of the inadequacy of CASSCF potential for ionic state, that for its nature, request a correlated method. Because of single reference determinant nature of  $S_1$ , CC2 method, at a qualitative level, correctly predicts the structural relaxation leading the molecule toward a planar structure before twisting around the ethylenic C=C bond. We wish to point out that, although the  $\pi \rightarrow \pi^*$  excited state is well described by CC2, this method may not offer a reliable treatment of the excited state where the reference determinant changes to  $(\pi^*)^2$ . The  $S_1$  vibrational relaxed structure located trough CC2 potential was the only one reliable with the wavefunction of the examined electronically excited state.





Methods	$\phi$	C1-C1'	C1-C2	C2-C3	C3-C4	C4-C5	C5-C6	C6-C7	C7-C2
Casscf(10,10)/6-31g*	180°	1.406	1.405	1.441	1.381	1.372	1.411	1.378	1.435
Casscf(10,10)/diff	180°	1.405	1.408	1.438	1.355	1.416	1.391	1.387	1.433
Casscf(10,12)/diff	168°	1.439	1.411	1.418	1.374	1.419	1.389	1.445	1.449
Casscf(12,12)/diff	180°	1.418	1.404	1.435	1.392	1.411	1.405	1.390	1.436
Casscf(14,12)/diff	180°	1.393	1.440	1.398	1.422	1.423	1.388	1.438	1.408
CC2/def SV(P)	180°	<b>1.429</b>	<b>1.407</b>	<b>1.445</b>	<b>1.392</b>	<b>1.409</b>	<b>1.416</b>	<b>1.387</b>	<b>1.440</b>

All the bond lengths refer to the planar geometry optimized on  $S_1$  state.

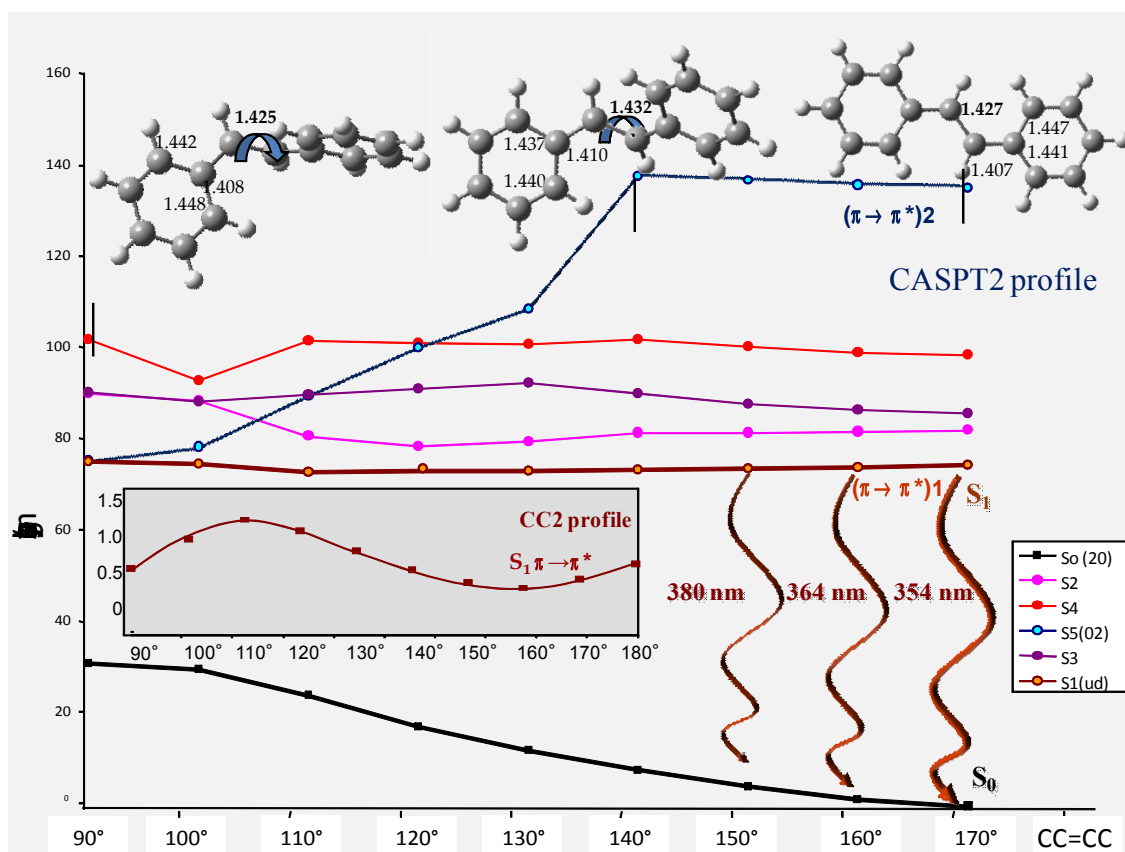
**Fig6.5** There are collected all the bond lengths of the minima obtained through the different combination of CASSCF active space and basis set. There is no coherence between similar active space such as the (14,12) and (10,12). The  $^1B$  only reliable equilibrium geometry was located at CC2 level.

Using CASSCF approach was not expected to increase the accuracy, unless considerably larger active space (including correlation arising by  $\sigma \rightarrow \pi^*$  contribution) than those used in this work are employed. Since geometry optimization were performed by using CASSCF potential, it was decided to preserve symmetry in order to include the state interested, the  $2B$  at CASSCF level. The electronic state associated to the lowest intense transition (with the highest oscillator strength value) is described mainly by a single one-electron (HOMO-LUMO) promotion, for all the active space combination taken in account. If we have a look to the HOMO and LUMO orbitals, we may appreciate the clear ethylene character slightly mixing with benzene ring, through which we may expect a precise gradient direction towards an equilibrium region. For example in basis of the wavefunction nature we may expect an extending of C1-C1' length bond and consequently a shortening of C1-C2 bond length, according with the single electron transition from a  $\pi$ -bonding orbital(HOMO) and the corresponding to a  $\pi^*$ -bonding orbital (LUMO) between C1-C1'. Surprisingly, any of the CASSCF equilibrium geometry obtained seem to reflect such characteristic. For example, despite the similarity of (10,12) and (14,12) combinations of active space, the relaxed trans\* obtained were completely different: if for the first choice there is an overestimation of such behavior (with a C1-C1' distance of 1.439 and C1-C2 distance of 1.411), for the second one there is an underestimation and a completely opposite geometrical parameters (with a C1-C1' distance of 1.390 and C1-C2 distance of 1.440). Now, although in the next section we will see how CC2 potential fail on  $S_1$  energy profile treatment along the torsional coordinate, in this case it was the only method capable to correctly describe the vibrational relaxation on the ionic state leading to an accurate equilibrium structure, whose geometry reflects the electronic transition character. Such evidence is clear if we analyze the bond lengths reported into the table 6.5: in the middle there is a sort of allylic character due to the combined elongation of the central C1-C1' distance (1.425) with the shortening of the adjacent C1-C2 bond length (1.407).

The phenyl rings are characterized by geometrical rearrangement whose bond length values are directly the consequence of de-populating  $\pi$ -bonding orbitals and populating of  $\pi^*$ -anti-bonding orbitals.

Following, we focus on aspects of the excited electronic surfaces important to electronic state population quenching for cis- and trans-stilbene along the isomerization reaction coordinates. We do not specifically address the barriers to ethylenic torsion which play a role in trans-stilbene photochemistry, nor the DHP reaction coordinate, which is a minor product channel for the cis isomer<sup>18-19</sup>. The torsional barrier from the trans isomer is very important to the many studies which have used photo-induced isomerization of trans-stilbene in gas phase, cluster, and condensed phase environments as a benchmark system for studying statistical rate theories and intramolecular vibrational energy redistribution. For calculating the minimum energy path along the torsional barrier connecting the planar trans\* with the perpendicular intermediate p\*, the IRC method cannot be employed, being difficult to estimate a barrier with such method. Thus it was necessary to get starting points (accurate geometries) along which the energy profile evaluation may be completed. Therefore, starting from CC2 planar trans\*, a series of optimized scan were computed by using CC2, along the torsional mode, keeping frozen the central dihedral angle  $\theta$  at the requested value. Once completed the CC2 energy profile, CASPT2 corrections (preserving C2 symmetry) were performed on the lowest six roots (until the dark bi-excited state is included) along the CC2 scan optimized points. The figure 6.6 summarizes the CASPT2 energy profile along the torsional normal mode for the relevant states. According with the commonly theoretical picture the bi-excited state (4A) is correctly reproduced giving a notable energy stabilization along the pathway (until 90° where is degenerate with 1B), while no barrier but a flat potential was detected on the mono-excited state along the twisting coordinate. Therefore, the presence (or absence) of the barrier seems to be mainly an intrinsic feature of the HOMO-LUMO (B symmetry) state, although its height can of course be modulated by the interaction with other states. The HOMO-LUMO state always corresponds to the lowest energy transition, and its nature does not change significantly before reaching the saddle point.

Although the lacking of a barrier, at least the  $\Delta E(S_1 \rightarrow S_0)$  of about 350 nm corresponding to 10° or 20° twisting, nicely reproduce the experimental emission (fluorescence) spectrum (Fig.6.7). The low fluorescence quantum yield (0.05) matches with the small isomerization barrier, not enough to let the reactive  $S_1$  being a trap state. On the other hand despite the use of CC2 gradient, if we compare the  $S_1$  isomerization profile calculated at CC2 level with the CASPT2 one, only a slightly improved outline may be appreciate: the CC2 barrier located at a structure 60° twisted, was estimated to be only 0,5 kcal/mol.



**Figure.6.6** There is reported the CASPT2 energy profile along the torsional normal mode: the isomerization coordinate. The  $S_1$  (ud) indicates the mono-excited  $(\pi \rightarrow \pi^*)^1$  state (symmetry B), while the  $S_5$  (02) is the  $(\pi \rightarrow \pi^*)^2$  bi-excited state.

### [6.2.3 Applying CASSCF//CASPT2 and CC2 methods: twisted region.](#)

Once the geometry distortion reaches the perpendicular structure, the choice of preserving the symmetry or not, is fundamental.

First at the rigid twisted ( $\theta=90^\circ$ ) geometry ( the transition state optimized on  $S_0$ ) if it is preserved the symmetry there are two almost degenerate, lowest excited states (at CASPT2 level). Nevertheless for both these states (1B and 2A, emphasizes in the Table 6.2<sup>(a)</sup>) the  $\Delta E(S_0-S_1)$  energy gap of 50 kcal/mol is too high for explaining an ultrafast depopulation process.

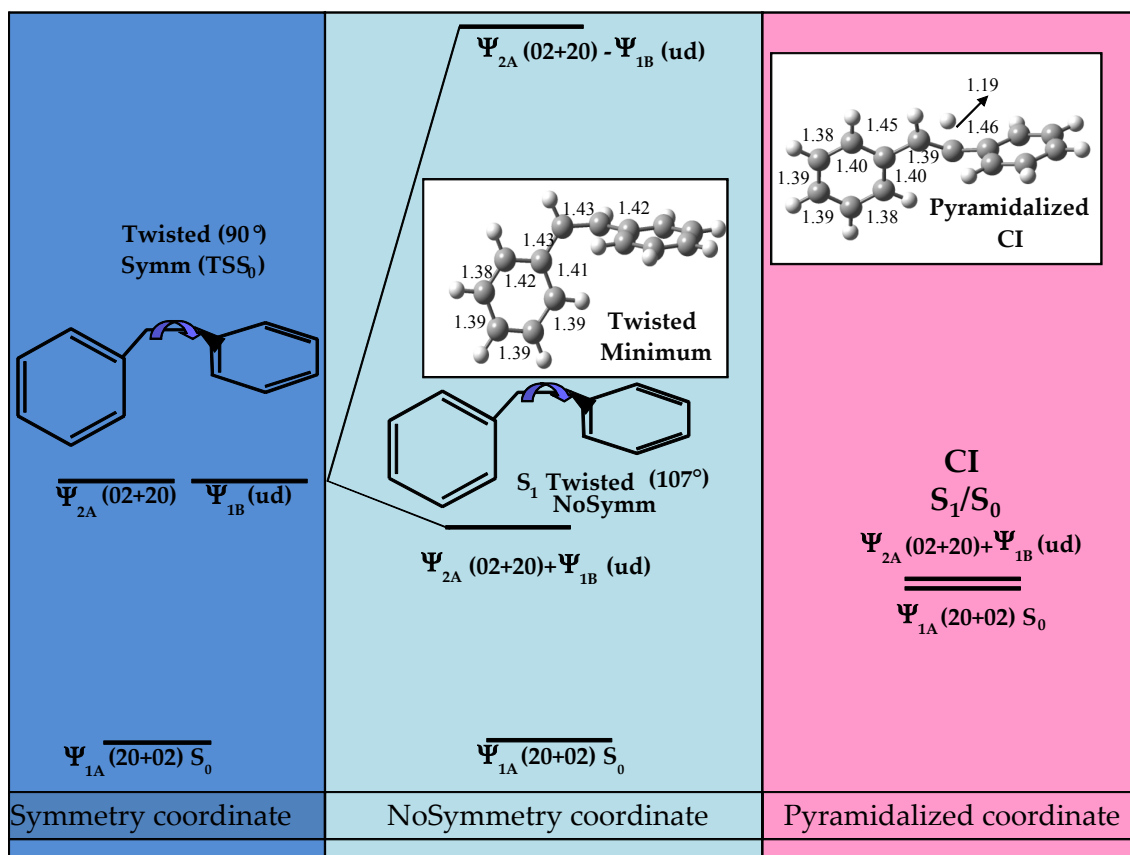
At this stage the lacking of an adequate correlated method able to correctly treat the HOMO-LUMO state has began to cause several difficulties on finding a reliable picture of  $S_1$  topology. At the twisted geometry ( $90^\circ$ ) the optimization procedure calculated on the state of B symmetry failed to find a stable intermediate at around  $90^\circ$ , it seems that the CASSCF potential is characterized by a singular topology with a minimum at about  $48^\circ$  twisted (the central dihedral angle  $\theta=48^\circ$ ) suggesting a possible trap state towards the cis-isomer region. A stable intermediate was located on 2A state at  $\theta=107^\circ$  twisted, according with some hypothesis affirming the perpendicular minimum presents a covalent character.

State <sup>(a)</sup> (Symmetric Twisted)	$\Delta$ E CASSCF 6-31g* Kcal/mol	$\Delta$ E CASPT2 6-31g*/ Kcal/mol	Oscillator Strength $f_{0-n}^{(c)}$	Configuration wavefunction <i>weight</i>	Symmetry
<b>S<sub>0</sub></b>	0	0	---	$(5\pi)^2(6\pi^*)^0$ 0.415 $(5\pi)^0(6\pi^*)^2$ 0.377	A
<b>S<sub>1</sub></b>	76.97	65.99	0.00326	$(4\pi)^1(6\pi^*)^1$ 0.190 $(3\pi)^1(10\pi^*)^1$ 0.147 $(2\pi)^0(5\pi)^0(6\pi^*)^2$ 0.199	B
<b>S<sub>2</sub></b>	77.46	65.72	0.00093	$(6\pi)^1(9\pi^*)^1$ 0.141 $(2\pi)^1(6\pi^*)^1$ 0.196 $(3\pi)^0(4\pi)^0(6\pi^*)^2$ 0.164	B
<b>S<sub>3</sub></b>	92.42	74.89	0.00085	$(2\pi)^1(6\pi^*)^1$ 0.145 $(3\pi)^1(4\pi^*)^1$ 0.132 $(2\pi)^1(6\pi^*)^1$ 0.109	B
<b>S<sub>4</sub></b>	93.21	80.69	0.00155	$(4\pi)^1(6\pi^*)^1$ 0.189 $(2\pi)^0(5\pi)^0(6\pi^*)^2$ 0.172	A
<b>S<sub>5</sub></b>	<u>99.57</u>	<u>48.05</u>	<u>0.00104</u>	<u><math>(5\pi)^1(6\pi^*)^1</math> 0.853</u>	<u>B</u>
<b>S<sub>6</sub></b>	<u>100.90</u>	<u>51.22</u>	<u>0.00015</u>	<u><math>(5\pi)^2(6\pi^*)^0</math> 0.389</u>	<u>A</u>

State <sup>(b)</sup> (No symmetric Twisted)	$\Delta$ E CASSCF 6-31g* Kcal/mol	$\Delta$ E CASPT2 6-31g* Kcal/mol	Oscillator Strength $f_{0-n}^{(c)}$	Configuration wavefunction <i>weight</i>	Symmetry
<b>S<sub>0</sub></b>	0	0	---	$(5\pi)^1(6\pi^*)^1$ 0.805	A
<b>S<sub>1</sub></b>	79.89	55.61	0.004628	$(5\pi)^2(6\pi^*)^0$ 0.369 $(1\pi)^0(6\pi^*)^1$ 0.152 $(5\pi)^0(6\pi)^1(7\pi^*)^1$ 0.127	B
<b>S<sub>2</sub></b>	<u>87.01</u>	<u>54.58</u>	<u>0.000315</u>	<u><math>(5\pi)^2(6\pi^*)^0</math> 0.469</u>	<u>B</u>
<b>S<sub>3</sub></b>	88.54	64.47	0.000362	$(5\pi)^1(9\pi^*)^1$ 0.229 $(2\pi)^0(5\pi)^0(6\pi^*)^2$ 0.293	<u>B</u>
<b>S<sub>4</sub></b>	89.96	74.24	0.001380	$(2\pi)^0(5\pi)^0(6\pi^*)^2$ 0.303 $(5\pi)^1(9\pi)^1$ 0.197	A
<b>S<sub>5</sub></b>	95.44	73.52	0.001878	$(5\pi)^2(6\pi^*)^0$ 0.215 $(4\pi)^1(6\pi)^1$ 0.286	A
<b>S<sub>6</sub></b>	<u>109.25</u>	<u>73.07</u>	<u>0.003429</u>	<u><math>(5\pi)^0(6\pi^*)^2</math> 0.860</u>	<u>A</u>

**Table 6.2** There are collected the  $S_0$ - $S_n$  vertical excitations for the lowest seven roots, for both the perpendicular intermediate (the symmetric table 6.2<sup>(a)</sup> and no-symmetric table 6.2<sup>(b)</sup>), following the CASSCF order. The active space used was of 10 electrons into  $10\pi$  orbitals. In order to correct the energetic the correlated method CASPT2 was employed. <sup>(c)</sup>The oscillator strength values are referred to the  $\Delta E$ \_CASSCF. For both the structures there are emphasized the relevant states, of ionic or covalent character. For the wavefunction analysis the orbitals represented into the Figure 6.4 may be helpful.

On the other hand by removing symmetry, the ordering of the lowest excited states changes: as predicted by ethylene theory, either the degenerate lowest states are splitted, into two linear combinations of both the states, corresponding to two other states covalent and ionic, each of one subject to energy stabilization and de-stabilization (as summarized in the scheme 6.1). It is quite simple to recognize both these states, (table 6.2<sup>(a)</sup>) being dominated by a single configuration (HOMO-LUMO mono or bi-excited transition) while the state, correlated with the ionic one (symmetry B) is kept at almost the same energy a notable de-stabilization of the covalent state (symmetry A), as refined in the scheme 6.1. Significantly if we have a look at the table 6.2<sup>(b)</sup>, referred to the non-symmetric perpendicular minimum the configuration dominating the  $S_0$  wavefunction is seems a mono-electron transition, while a bi-excited dominates the ionic  $S_1$ . This is not in contrast with the covalent nature of the ground state and the ionic nature of  $S_1$ , and it is possible to explain through the active space orbitals, being localized on each atom, when the symmetry is removed.



**Scheme 6.1** This scheme reports the comparison between three different structures. On the left side there are the two perpendicular intermediates, removing the symmetry the degeneracy is lost and both the state

are splitted into two different energy levels. Nevertheless an asymmetric twisted intermediate does not guarantee the degeneracy  $S_1-S_0$ , for which a pyramidalization of one ethylene carbons is requested.

Remarkably, the  $\Delta E(S_1-S_0)$  is still too high for explaining a non-emissive decay trough this intermediate, it will be necessary a geometrical deformation: in particular, the ultrafast decay of photoexcited cis-stilbene motivates a search for minimal energy conical intersections of  $S_0$  and  $S_1$  (the ground and first excited-state potential surfaces) which lead to product formation. Bearpark and co-workers<sup>35</sup> reported several such intersections, obtained using molecular mechanics valence bond (MMVB) theory, because the MMVB method does not include ionic states, only a subset of the relevant (covalent) intersections can be described with this method. Amatatsu<sup>15</sup> reported a conical intersection obtained with limited basis sets and *ab initio* quantum chemistry, but gave very little information on its geometrical structure. Martinez searched for conical intersections using the method, which simultaneously minimizes the energy of the upper electronic state and the electronic energy gap, implemented in GAUSSIAN package<sup>36</sup>.

**Table 6.4** There are collected the energy value for the pyramidalized conical intersection (Martinez), found with CASSCF(2,2), and here recalculated with CASSCF//CASPT2(10,10)

State <sup>(a)</sup> (PyrCI)	$\Delta E$ CASSCF 6-31g*/ (6-31g(+)*CI-CI') Kcal/mol	$\Delta E$ CASPT2 6-31g*/ (6-31g(+)*CI-CI') Kcal/mol	Oscillator Strength $f_{0-n}$ <sup>(c)</sup>	Configuration wavefunction weight	Symmetry
$S_0$	0	0		$(5\pi)^2(6\pi^*)^0 0.472$	A
				$(5\pi)^1(6\pi^*)^1 0.383$	
$S_1$	0.017 (0.022)	0.573 (-2.83)		$(5\pi)^2(6\pi^*)^0 0.395$	B
				$(5\pi)^1(6\pi^*)^1 0.462$	

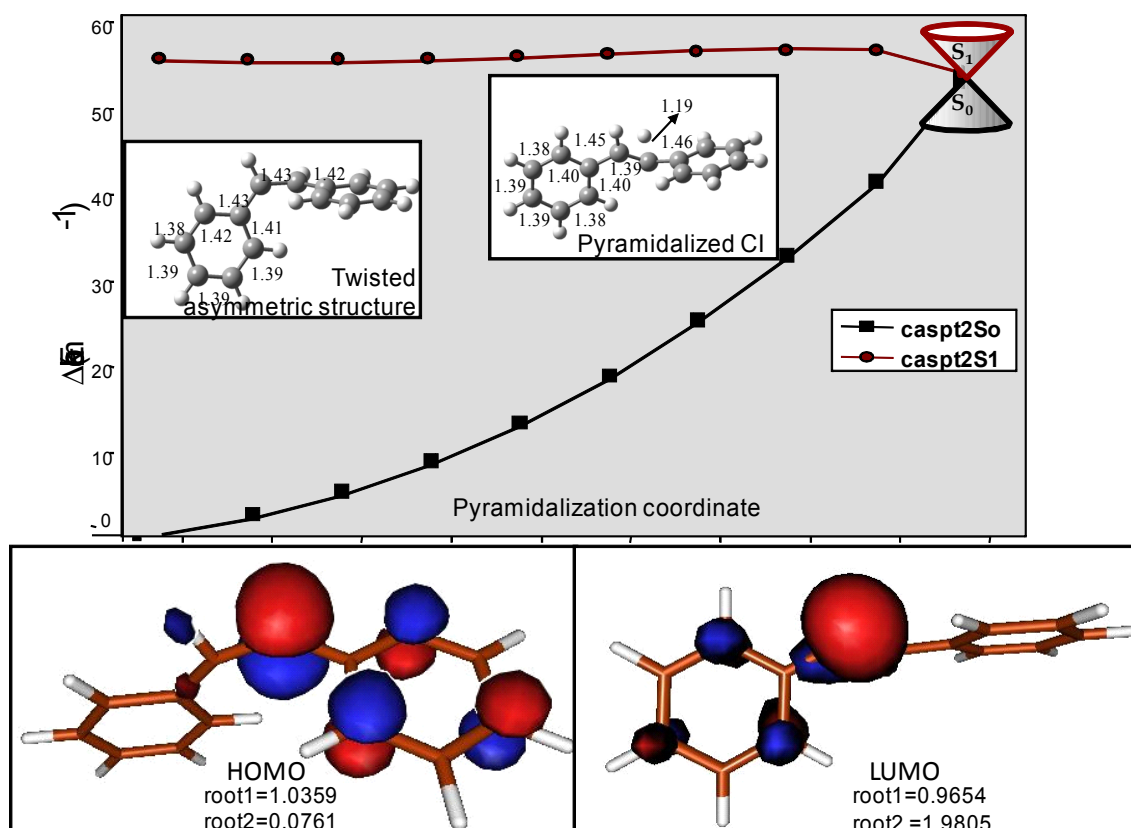
The lowest energy  $S_0/S_1$  intersection that is pictured in Figure 6.7, is twisted by  $90^\circ$  and has significant pyramidalization of one of the ethylenic carbon atoms, defining also a sort of zwitterionic character. This geometry is strikingly similar to the minimal energy conical intersection of ethylene therefore we use the same label, Pyr-CI (Pyramidalized Conical Intersection). According with ethylene model, the pyramidalization of one of the two central carbons refine the following picture: if the ground state is described by the radical character with each one of the two electrons localized on each of the carbons, the first excited state is characterized by a charge-transfer character, with both the electrons localized on a single carbon atom. This is reflected by the wavefunction and the orbital symmetry: all the active space included, in a 10-10 space, localized orbitals on each of the benzene rings. In particular the LUMO has only one localized orbital into one of the two ethylene carbons. This is more expanded featured by a negative charge, due to the presence of both the electrons on  $S_1$  configuration. Interestingly the nature of almost an anion is reflected by the extended C-H bond length, leading to a possible hydrogen migration. The natural occupation numbers reflect this behavior, as indicated into parenthesis of figure 6.7.

Despite several tentative of optimizing such conical intersection, CASSCF(10,10) potential has revealed its inadequacy to treat such “charge-transfer” state. In particular for such system, the algorithm for minimizing CI, implemented in GAUSSIAN, which simultaneously minimizes the energy of the upper electronic state and the electronic

energy gap, was totally unable to locate a lower crossing along the  $S_1/S_0$  degenerate hyperline.

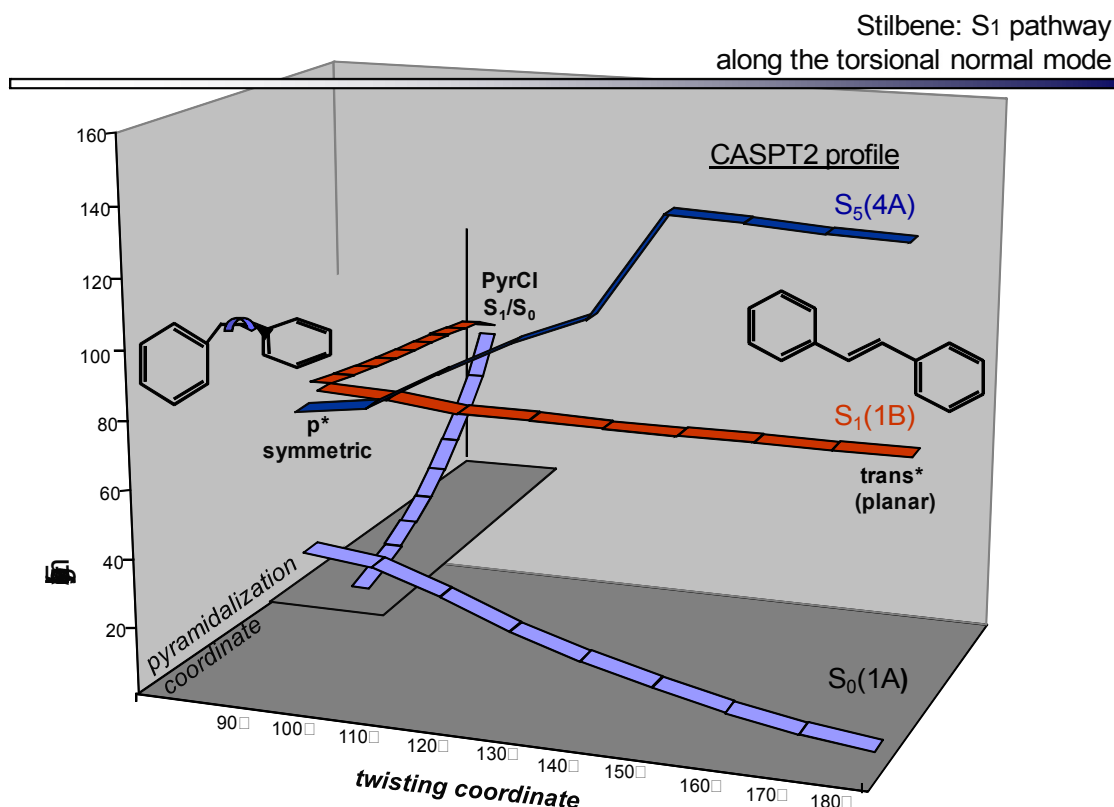
Therefore we have concluded that the pyramidized CI (PyrCI), found by Martinez, was the most accessible crossing at both levels CASSCF(2,2) and CASSCF(10,10).

Interestingly the minimum energy path connecting the twisted asymmetric structure  $p^*$  with PyrCI was calculated through the CASSCF(10,10)//CASPT2: despite the barrierless pathway according with the lifetime of ps, no real energy stabilization was observed for PyrCI compared with  $p^*$ .



**Figure 6.7** On the top of the figure is reported the almost barrierless  $S_1$  reaction path leading the photo-excited population to the accessible CI. The pyramidized CI, according with the OBF model is featured by an anionic character, and by a wavefunction described by localized orbitals on ethylenic carbons. The occupation number of 1.98 on LUMO and 0.07 on HOMO reflect a negative charge on C1 (LUMO), due to both the electrons located on the same orbital and a positive charge on C1, due to absence of electrons.

Finally it is fundamental to emphasize how the mono-dimensional model is not sufficient for explaining the efficient radiationless process of internal conversion trough which stilbene photoexcited population is relaxed on the ground state. At beginning the reaction coordinate is only dominated by the ethylene torsional mode and moving along the twisting reaction pathway lead the photo-excited molecule from  $trans^*$  planar to  $p^*$  (where the  $S_1-S_0$  energy gap  $S_1-S_0$  being 50 kcal/mol). Nevertheless for reducing the  $S_1-S_0$  energy gap it is necessary a geometrical distortion (no symmetry preserved) as pyramilization according with the charge–transfer character on  $S_1$  until it will be degenerate with the covalent (radicalic)  $S_0$ . The figure 6.8 summarizes the two-dimensional mechanism proposed.



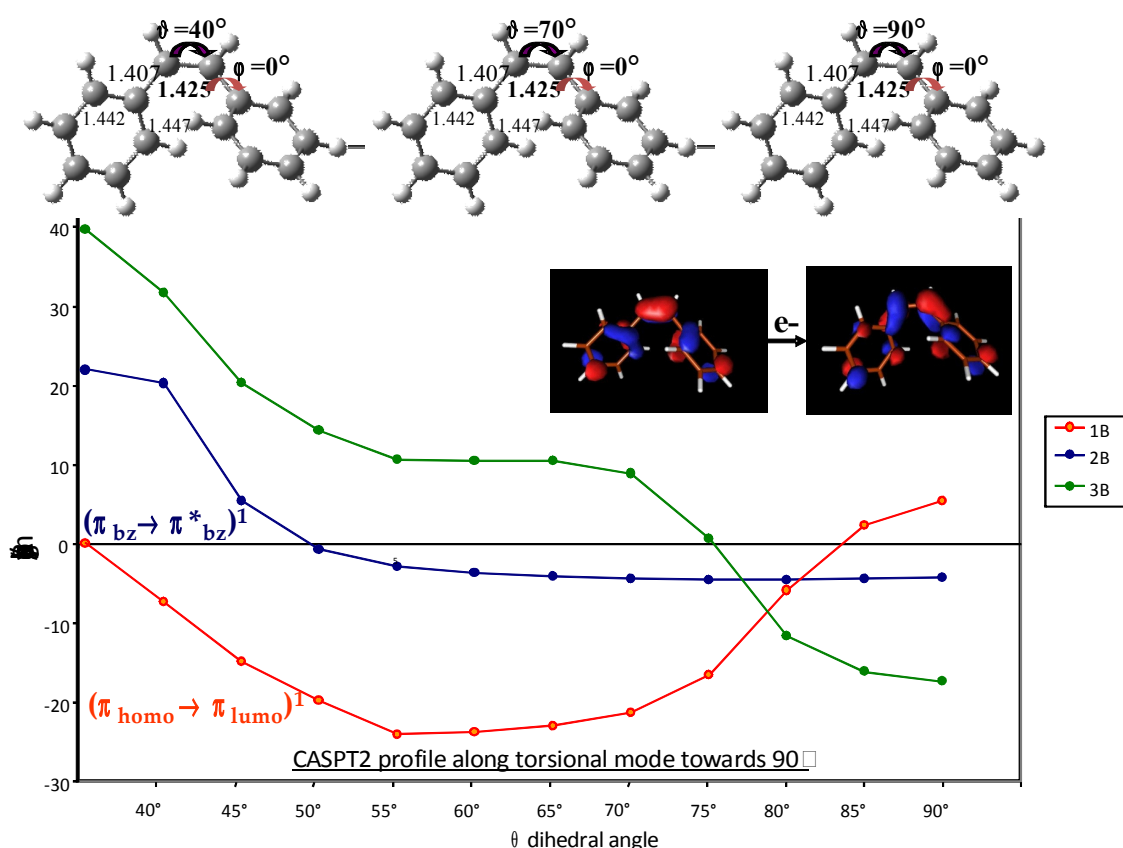
**Figure 6.8** There is summarized the multi-dimensional scheme for rationalizing the complete Stilbene photoreactivity, from trans\* fluorescent region to the PyrCI reached, populating an orthogonal coordinate to the torsional mode.

#### 6.2.4 Applying CASSCF/CASPT2 and CC2 methods: cis\* relaxation.

On one hand the trans isomer fluorescence detected upon HOMO-LUMO S<sub>1</sub> excitation and the subpicosecond lifetime measured for this state suggest that S<sub>1</sub> is associated with an almost flat potential energy surface along torsion, with a small barrier of 3 kcal, making isomerization feasible on the same excited state. On the other hand the cis\* population is (almost) not fluorescent and has two decay routes: the geometrical cis-trans isomerization process occurring through a barrierless reaction pathway and the electronic rearrangement to yield dihydrophenantrene on S<sub>0</sub>. Upon light irradiation, vibrational relaxation from FC proceeds towards to two different minima depending on which state the gradient are calculated. Either the potentials, CASSCF and CC2, were employed for performing geometry optimization. Despite the almost equivalent structures, CC2 results were mainly considered, in order to compare the cis-trans reaction with the previous one, discussed before. The ordering of the states characterizing both the equilibrium structures is different, between CC2 and CASPT2 potential. Vibrational relaxation on the bright state (2B) proceeds following the isomerization reaction coordinate and lead the molecule to a flat region where the minimum on the HOMO-LUMO state was detected, at 40° of torsion. Therefore, after crossing a small energy barrier, it can further twist the  $\theta$  angle and reach pseudo-perpendicular arrangements (p\*) where it accesses the S<sub>1</sub>/S<sub>0</sub> conical intersection responsible for the cis-trans photoisomerizations. To gain more insight into the motions



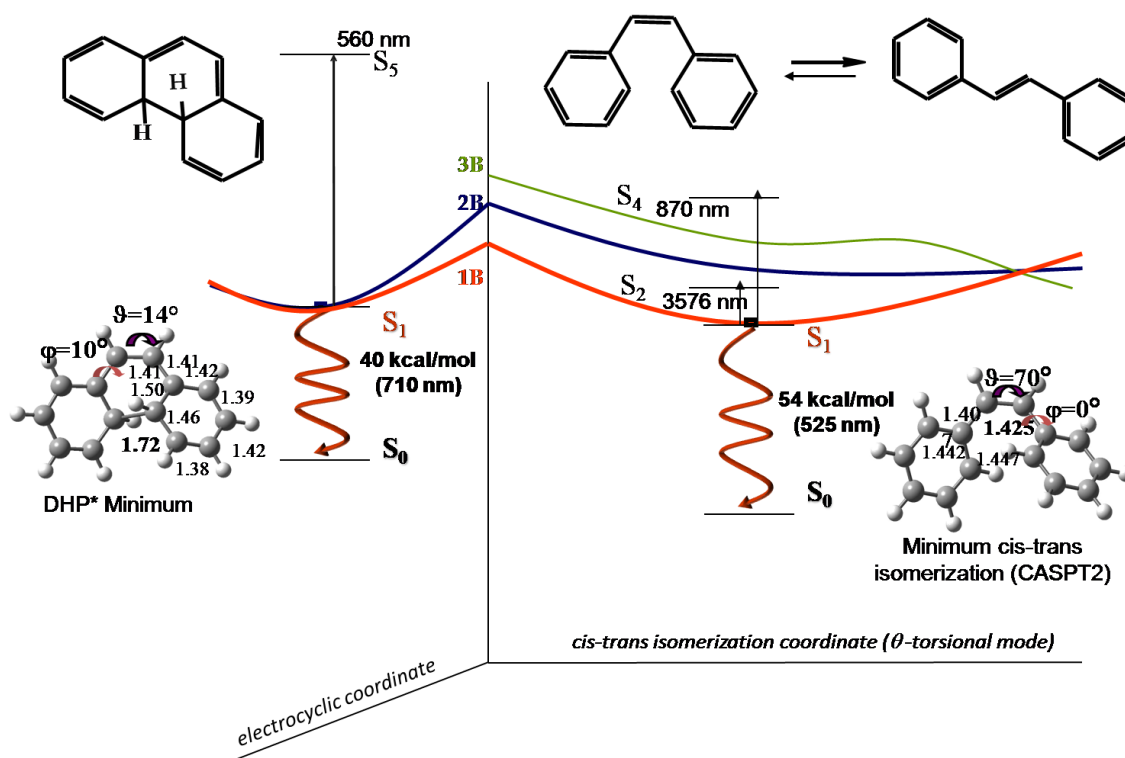
involved in the initial cis\* relaxation, we computed CASPT2 rigid scans extended to the lowest B states. In the figure 6.9 there is represented the behavior of such states along the torsion: surprisingly the CASPT2 minimum (located at about 60°) is displaced respect to the CASSCF, with the 1B, the lowest state corresponding to the ionic HOMO-LUMO state. Proceeding towards perpendicular region, the HOMO-LUMO state increases in energy but there are not reported the energy profile of A states along. However since measurements indicate that the wavepacket oscillates and cross the barrier in the  $\theta$ -twisting direction in 300 fs, the barrier on the cis-isomer should be even lower than the one detected on the trans side where the lifetime of  $S_1$  photo-excited population of ps order.



**Figure 6.9** Energy profiles for B states are computed for rigid scans geometries connecting cis\* with p\*.

Once completed the discussion about the cis-trans isomerization channel, it is important to mention the electrocyclization channel responsible of dihydrophenantrene (DHP) formation observed on the ground state with a quantum yield of 0.1. The commonly accepted vision of the electrocyclic process affirms the existence of a conical intersection trough which an ultrafast decay with the DHP formation may occur. Here we did not search any  $S_1/S_0$  crossing, we have limited our calculations to a geometry optimization on the benzenic covalent state 1B. Both the potentials CC2 and CASSCF adopted provide almost the same equilibrium geometry, characterized by  $sp^2 \rightarrow sp^3$  hybridization of the carbons (C3-C3') involved in the  $\sigma$ -bond forming reaction and a shorter distance between them (C3-C3'=1.72 at CC2 level, C3-C3'=1.82 at CASSCF level). Further a same ordering of the states was observed for both the level of calculations (preserving along this path the benzenic state as the lowest B state).

Nevertheless introducing CASPT2 corrections a different ordering of the states may be appreciated, being 1B and 2B degenerate at the CC2 optimized equilibrium geometry (see fig.6.10). The figure 6.10 summarizes the two possible cis\* pathways: along the  $\theta$ -twisting coordinate the cis-trans isomerization may occur with an experimental quantum yield of 0.35, while following the electrocyclic coordinate dihydrophenanthrene photo-product is observed on  $S_0$ . Since recently two-photon experiments production and its influence are increasing, we have also calculated the eventual transient absorptions from the minima located on the 1B and 2B states with the higher absorbing states (characterized by an high oscillator strength value  $f_{l-n}$ , coupling the lowest excited state, such as 1B and 2B, with higher roots), and we have collected the results in the figure6.10.



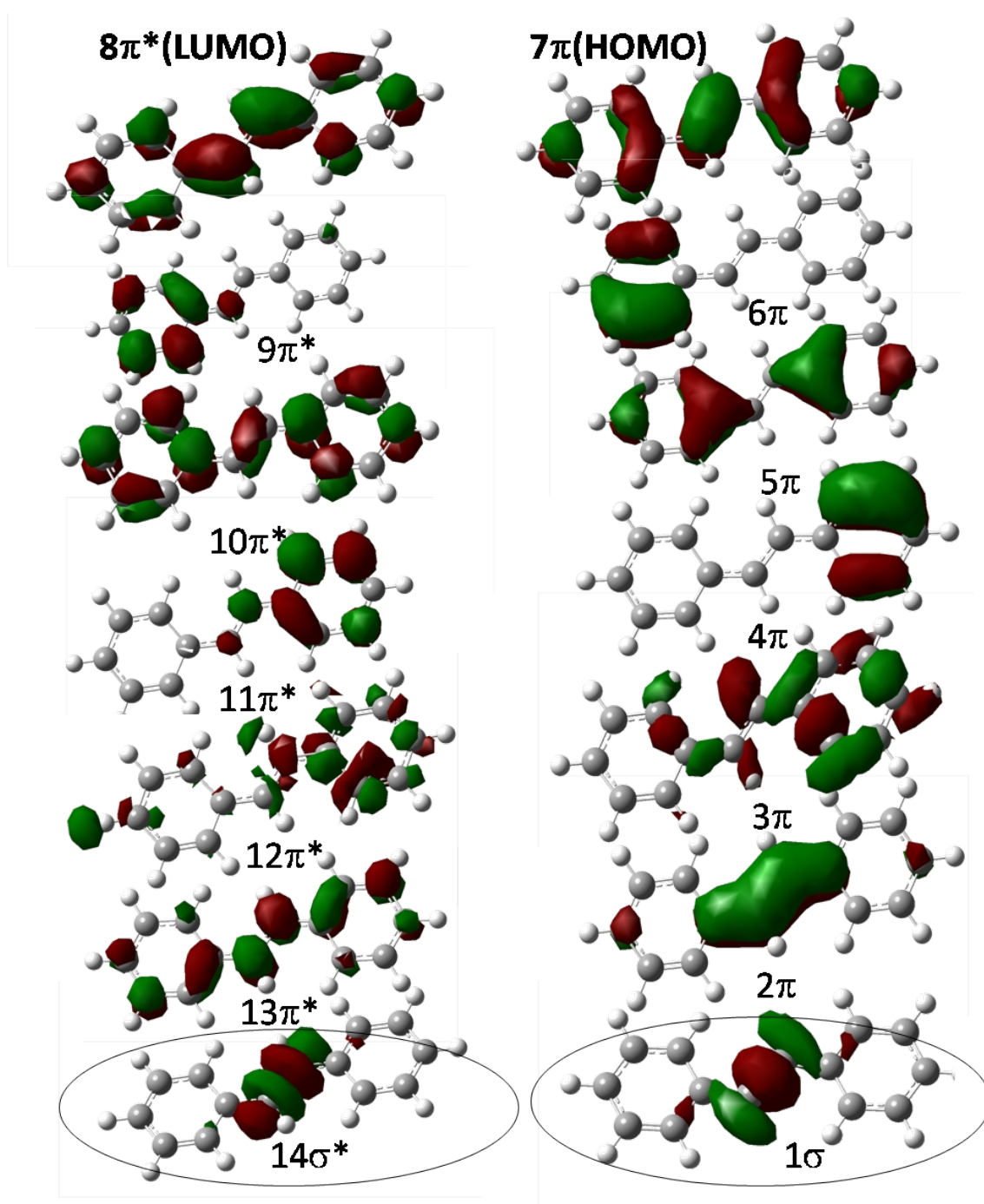
**Figure6.10** There is here reported a schematic outline of the two main decay routes for cis\* photo-excited population. After relaxation to both the minima, following different coordinate transient absorption  $S_1$ - $S_n$  may be calculated by including higher A roots, coupling with minima B roots.

### 6.3 Appling RASSCF method.

Once we have collected a series of failures of CASSCF//CASPT2 method on reproducing the experimental data, we have individuated the origin of such insufficient treatment to the CASSCF wavefunction and its limit. The problem of finding an adequate method for the HOMO-LUMO state arises from its particular nature, and the singular correlation requested for a reasonable treatment. Interestingly if for other systems the combined CASSCF//CASPT2 method for locating critical points and estimating the relative energy was an excellent strategy, stilbene may not be treated with CASSCF potential because of its limitation on treating the ionic state. In particular in

tilbene, the reactive state, although dominated by a single  $\pi \rightarrow \pi^*$  configuration, is particularly affected by  $\sigma \rightarrow \pi^*$  configuration contributions. The CASSCF calculations described above still use the same  $\sigma$  core orbitals for all configurations, preventing the polarization of the core, needed to stabilize ionic configurations further. To remedy this flaw, the second part of our strategy was performing  $\sigma$ - $\pi$  RASSCF (Restricted Active Space Self Consistent Field) calculations. Typically, only one or two electrons populate the orbitals added from outside the valence space in any important electron configuration. Many-electron configurations (which would be included in a full CASSCF calculations) are therefore largely redundant. Recognizing this, the spirit of the RASSCF method is to subdivide the active space into three categories: orbitals with a limited number of vacancies, called the Ras1 space, a fully active orbital set (Ras2), and orbitals with a limited number of electrons (Ras3). By eliminating the redundant configurations, the size of the configuration interaction problem can be greatly reduced with RASSCF compared to CASSCF without compromising accuracy

Thus for introducing the effect of  $\sigma$ - $\pi^*$  transitions, RASSCF calculations were employed with including all the 14  $\pi$  orbitals. Several choices were made imposing the Ras2 equal to zero and considering until quadruple excitations from Ras1 to Ras3. But any satisfactory result was really appreciated, and often problems with the wavefunction convergence were detected. Therefore, considering the incredible contribution provided by Martinez model, we have used a RASSCF wavefunction on the basis of the most simple case, the CASSCF(2,2), trying to follow this example without introducing too many configurations: the choice converged towards a RASSCF wavefunction with a Ras2 (full active space) space of two electrons into two orbitals. The dynamic correlation between  $\sigma$  core and  $\pi^*$  electrons can be included by adding single  $\sigma$  excitations to all of the  $\pi^*$  configurations. This enables the  $\pi$  system to polarize the  $\sigma$  framework in the ionic state, thus stabilizing this state relative to the covalent states. Hence only one vacancy and one electron have been allowed in Ras1 and Ras3, respectively, defining only single excitations Ras1 $\rightarrow$ Ras3. This method adopted choice have suggested to finally proceeds towards a correct direction providing a stable wavefunction description and a more reliable topological picture. At this stage it is fundamental to emphasize how for the moment only preliminary results will be next presented and discussed.



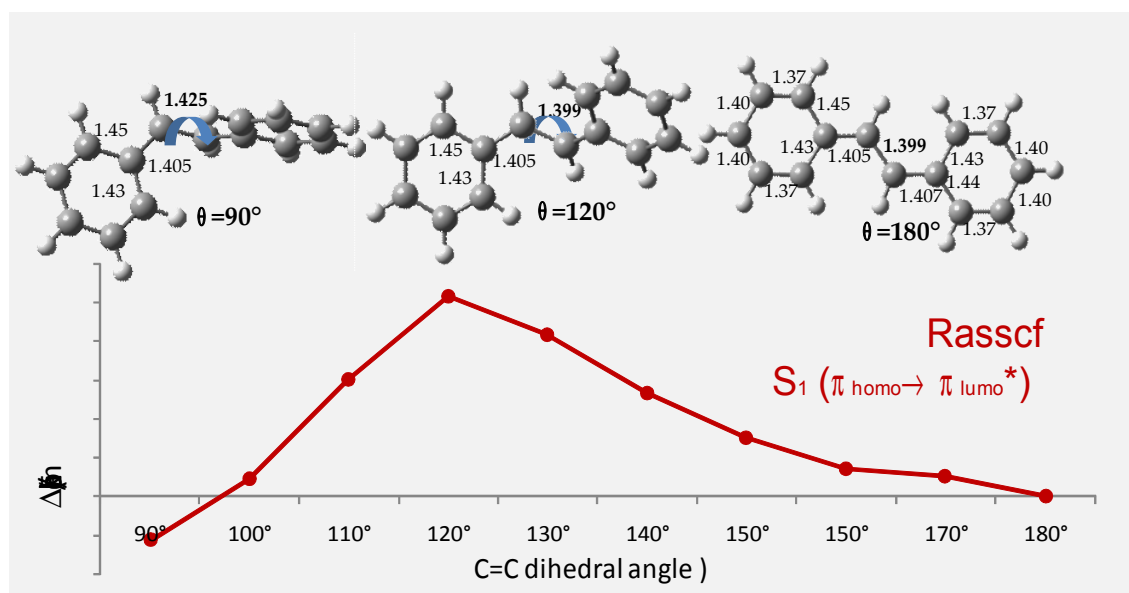
**Figure 6.11** There are collected the 14 RASSCF orbitals at the trans isomer. Interestingly within the 14,14 active space there is included the  $\sigma$  bonding and anti-bonding orbitals between the two ethylenic carbons.

First of all, being considered only two-electrons into two orbitals, the ionic (HOMO-LUMO) is the lowest excited state at RASSCF level, and its energetic is relatively correct being explicitly considered the influence of a  $\sigma \rightarrow \pi^*$  excitation. In fact if we have a look to the complete active space we may appreciate the presence of the  $\sigma$

bonding orbital (C1-C') on the two ethylene carbons (among the others six  $\pi$ -bonding orbitals), and the corresponding  $\sigma^*$ -anti-bonding (among the others six  $\pi$ -bonding orbitals)(see fig6.11).

Although we did not have yet a complete RASSCF picture of stilbene photo-reactivity, the preliminary results seem to suggest that this potential is the most appropriate for modeling its photochemistry, trough an adequate evaluation of the ionic state topology.

Therefore, the  $\theta$ -twisting barrier for trans-cis isomerization, along the pathway connecting the planar with pseudo-perpendicular intermediate, was estimated to be 5 kcal/mol. In such way the underestimation made by CASPT2 was corrected. At beginning using RASSCF wavefunction we have optimized the geometry on the optically active space (the ionic state) in order to get reliable geometrical parameters for a correct evaluation of the barrier. The trans\* equilibrium geometry found, has a reasonable structure reflecting the mono-determinantal wavefunction; consequently these series of geometrical parameters were used to evaluate the torsional barrier. In reality, being a series of non-optimized scan (rigid structures), the energy barrier may be a little bit overestimated. We believe that once completed the geometry optimization with the  $\theta$  dihedral kept frozen during the procedure, an improved evaluation of the barrier (closer to the experimental value /3 kcal/mol) will be obtained.



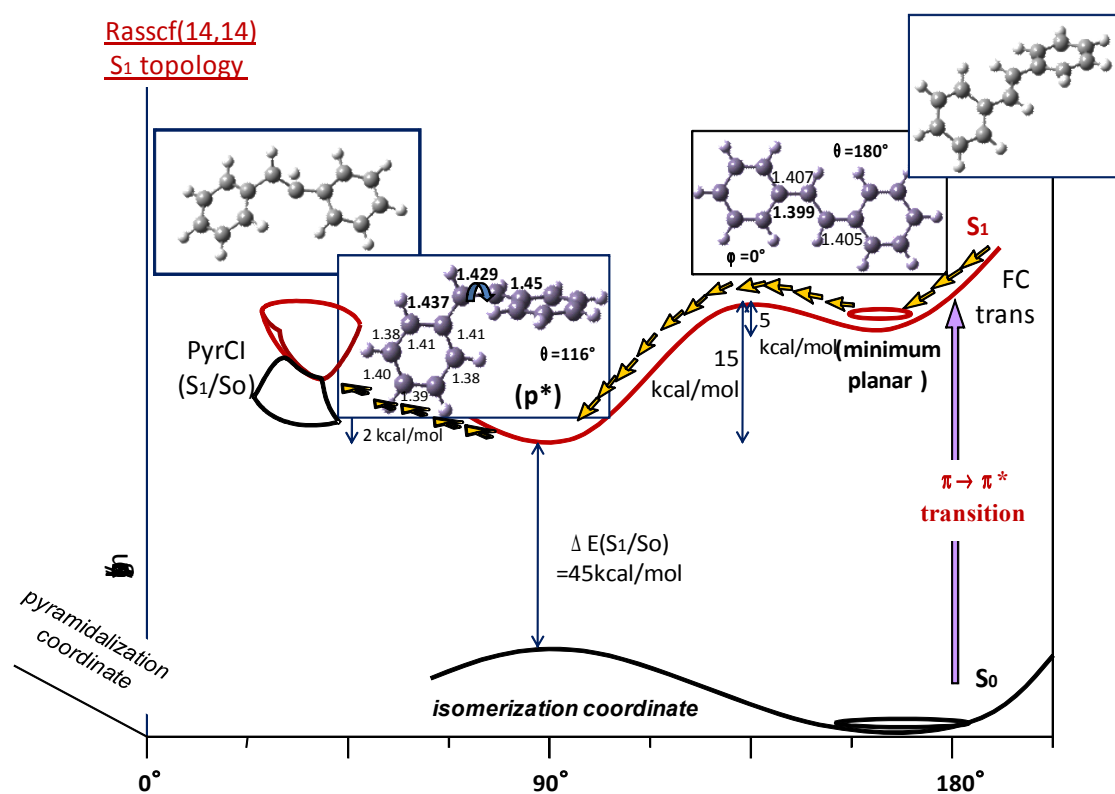
**Fig.6.12** Once there considered the  $\sigma$ - $\pi^*$  excitations the torsional barrier, connecting trans\* with p\* in finally well reproduced.

Further we have searched for a perpendicular intermediate lower in energy and able to trigger the photo-excited  $S_1$  population towards an  $S_1/S_0$  degenerate region moving along another coordinate. The geometry optimization lead to a perpendicular intermediate (p\*) on the ionic state, found 15 kcal/mol more stable than the planar equilibrium geometry, accounting for a reactive trap state at a pseudo-perpendicular region, non-emissive because able to drive the system towards  $S_1/S_0$  crossing.

Nevertheless, the  $S_1$ - $S_0$  energy gap of 45 kcal/mol, found at the optimized p\*, implying a long excited-state lifetime, is not accounting for radiationless decay from such point. However, (experimentally) no appreciable fluorescence was observed and the short

lifetime, recently measured of about 150 fs, is in agreement with a fast radiationless  $S_1 \rightarrow S_0$  deactivation (decay), which is most easily explained by regions of small or vanishing electronic energy gap.

Therefore, an easily accessible conical intersection should may be reached from perpendicular minimum if a different coordinate, orthogonal to the C=C twist, can be populated. Through the RASSCF potential we will investigate both the new coordinates, either involving a twist of the adjacent single bond, leading to Hula-Twist CI or pyramidalization of one ethylenic C atom, leading to the OBF-type CI, as the PyrCI (Martinez). For the moment the calculations we carried out were limited to the PyrCI, trying to minimizing it. The RASSCF optimized PyrCI is almost degenerate (2 kcal/mol higher) with the twisted minimum, and may be easily populated if it is considered the kinetic energy accumulated after passing the C=C barrier (located almost 20 kcal/mol higher), and once IVR (Internal Vibrational Redistribution) into the necessary orthogonal reaction coordinate, has been completed.



**Fig6.13** There is pictured a schematic outline of the RASSCF  $S_1$  topology, in much more agreement with the experimental data.

## 6.4 Conclusion

In this chapter we have examined the complex stilbene photoreactivity, in order gain more insight the radiationless event origin from the phantom state  $S_1$ . Due to its ionic nature, affected by  $\sigma \rightarrow \pi$  excitations, the dynamic correlation plays a crucial role in the description of such electronic state. The CASSCF//CASPT2 level of calculation gives usually reliable results for vertical excitations, but it fails to describe accurately the

ionic/covalent conical intersection geometries. This approach does not provide the proper topology of the excited  $S_1$  potential energy surface for stilbene. The inclusion of  $\sigma\pi$  correlation was shown to be necessary, and RASSCF has proved to be a successful approach to achieve this. The valence–core correlation energy is recovered by adding single  $\sigma$  excitations to all the  $\sigma$  configurations. In such way it is avoided that CASSCF still use the same core  $\sigma$  orbitals for all configurations, preventing the polarization of the core needed to stabilize ionic configurations further. This gives a balanced description of the differential electron correlation and greatly improved results for the  $S_1$  topology in stilbene.

### References:

- 38) a) Syage, J. A.; Felker, P. M.; Zewail, A. H. *Chem Phys* 1981, 11, 4685; b) Syage, J. A.; Lambert, W. R.; Felker, P. M.; Felker, P. M.; Zewail, J. *Phys. Chem.*, 1985, 89, 5402. c) Orlandi, G.; Siebrand, W. *Chem. Phys. Lett.* 1975, 30, 352
- 39) “Photochromism, Molecules and Systems in Studies in Organic Chemistry” (Eds.; H. Durr, H. Bouasri-Laurent) Elsevier, Amsterdam, 1990.
- 40) Evans, J.; Reynisson, J. K.; Geirgson, A.; Kravan, W. P.; McGimpsey, P. *Photochem. Photobiology* 1998, 115, 57-61
- 41) Fuss, Schmid, Kosmidis, *Angew. Chemie IntEd.* 2004, 43, 4178.
- 42) a) Malkin, S.; Fischer, E. *J. Phys. Chem.* 1964, 68, 1153. b) Saltiel, J.; D’Agostino, J.; Megarity, D. E.; Metts, L.; Neuberger, K. R.; Wrighton, M.; Zafarion, O. C. *Org. Photochem.* 1971, 3, 1. Saltiel, J.; c) Sun, Y.-P. In “Photochromism Molecules and Systems” Durr, H., Bouasri-Laurent, H., Eds.; Elsevier: New York, 1990; p 64. d) Saltiel, J.; Sun, Y.-P. *J. Phys. Chem.* 1989, 93, 8310. (e) Hochstrasser, R. M. *Pure Applied Chemistry* 1980, 52, 2683. f) Sension, R. S.; Repinec, S. T.; Szarka, A. Z.; Champagne, B. B.; Pfansfiel, J. P.; Plusquellic, F.; Pratt, D. W.; van Harpen, W. M.; Meerts, W. L. *J. Phys. Chem.* 1990, g) Meier, H. *Angew. Chem.* 1992, 31, 1399. (h) Arai, T.; Tokumaru, K. *Chem. Rev.* 1993, 93, 23. Whitten, D. G.
- 43) Fuss, Kosmidis, Schmid, Truskin, *Angew. Chem.* 2004, 4178-4182
- 44) (12) Farmanara, P.; Stert, V.; Radloff, W. *Chem. Phys. Lett.* 1998, 288
- 45) Mestdagh, J. M.; Visticot, J. P.; Elhanine, M.; Soep, B. J. *Chem. Phys.* 2000, 113, 237.
- 46) a) D. H. Waldeck, *Chem. Rev.* 1991, 91, 415 – 496; b) H. GÖrner, H. J.

- Kuhn, *Adv. Photochem.* 1995, 19, 1 – 117
- 47) a) J. S. Baskin, L. BaQares, S. Pedersen, A. H. Zewail, *J. Phys. Chem.* 1996, 100, 11920 – 11 933; b) J. A. Syage, W. R. Lambert, P. M. Felker, A. H. Zewail, *Chem. Phys. Lett.* 1982, 88, 266 – 270.
- 48) W. Fuß, C. Kosmidis, W. E. Schmid, S. A. Trushin, *Chem. Phys. Lett.* 2004, 385, 423 – 430.
- 49) M. J. Bearpark, F. Bernardi, S. Clifford, M. Olivucci, M. A. Robb, T. Vreven, *J. Phys. Chem. A* 1997, 101, 3841 – 3845.
- 50) W. Fuß, S. Lochbrunner, A. M. Miller, T. Schikarski, W. E.
- 51) Schmid, S. A. Trushin, *Chem. Phys.* 1998, 232, 161 – 174.
- 52) Y. Amatatsu, *Chem. Phys. Lett.* 1999, 314, 364 – 368.
- 53) J. Quenneville, T. J. MartRnez, *J. Phys. Chem. A* 2003, 107, 829 – 837.
- 54) a) Park, N. S.; Waldeck, D. H. *Chem. Phys. Lett.* 1990, 168, 379. b) Lee, M.; Haseltine, J. N.; Smith, A. B., III; c) Hochstrasser, R. M. *J. Am. Chem. Soc.* 1989, 111, 5044.
- 55) a) Todd, D. C.; Jean, J. M.; Rosenthal, S. J.; Ruggiero, A. J.; Yang, D.; Fleming, G. R. *J. Chem. Phys.* 1990, 93, 8658. 836 J Abrash, S.; Repinec, S.; b) Hochstrasser, R. M. *J. Chem. Phys.* 1990, 93, 1041.
- 56) Rodier, J.-M.; Myers, A. B. *J. Am. Chem. Soc.* 1993, 115, 10791.
- 57) Hochstrasser, R. M. *J. Chem. Phys.* 1993, 98, 6291.
- 58) Georg Hohlneicher\* and Ralph Wrzal, *J. Phys. Chem. A* 1999, 103, 8969-8975
- 59) a) Amatatsu, *Journ. of Molec. Struc. (Theochem)* 461-462 1999 311-316  
b) Amatatsu *Chem. Phys. Letters* 314 (1999) 364-368
- 60) Martial Boggio-Pasqua, Michael J. Bearpark, Michael Klene, and Michael A. Robb, *Journ. of Chem. Physics* 120, 17, 7849
- 61) . K. Andersson, P-Å Malmqvist, B.O. Roos. *J. Chem. Phys.* 1992 92, 1218-1226
- 62) Roos, B. O. *Adv. Chem. Phys.* 1987, 69, 399-446.
- 63) Szabo, A. & Ostlund, Neil, S. “Modern quantum chemistry: an introduction to advanced electronic structure theory,” McGraw-Hill, 1989.
- 64) Catàlan, *Chemical Physics Letters* 421 (2006) 134–137



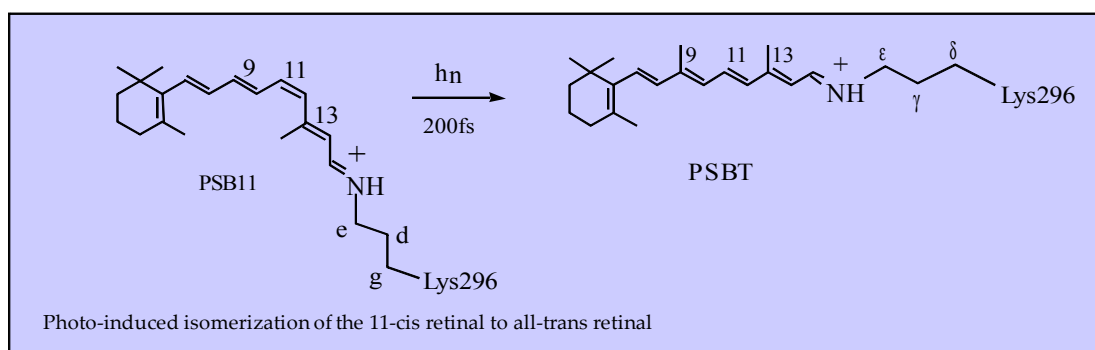
- 65) Gagliardi L,Orlandi G,Molina V,Roos B.*J.Phys.Chem.A* 2002,106, 7355-7361
- 66) Molina, V.; Merchan, M.; Roos, B. O. *J. Phys. Chem. A* 1997,101, 3478.
- 67) Warshel, A. *J. Chem. Phys.* 1975, 62, 214.
- 68) Suzuki, T.; Mikami, N.; Ito, M. *J. Phys. Chem.* 1986, 90, 6431.
- 69) Spangler, L. H.; van Zee, R. D.; Blankespoor, S. C.; Zwier, T. S.*J. Phys. Chem.* 1987, 91, 6077.
- 70) Andersson, K.; Barysz, M.; Bernhardsson, A.; Blomberg, M. R. A.; Carissan, Y.; Cooper, D. L.; Fülcher, M. P.; Gagliardi, L.; de Graaf, C.; Hess, B. A.; Karlström, G.; Lindh, R.; Malmqvist, P.-Å.; Nakajima, T.; Neogrady, P.; Olsen, J.; Roos, B. O.; Schimmelpfenning, B.; Schutz, M.; Seijo, L.; Serrano-Andrés, L.; Siegbahn, P. E. M.; Stålring, J.; Thorsteinsson, T.; Veryazov, V.; Widmark, P.-O. MOLCAS Version 6.2; Dept. of Theor. Chem.: Lund, Sweden, 2005.
- 71) Ahlrichs, R.; Bar, M.; Haser, M.; Horn, H. & Kolmel ,C. *Chemical Physics Letters* ,1989,162, 165-169.
- 72) M.Bearpark.Bernardi,Cliffors,Olivucci, Robb M. Vreven T. *The Journ.of Phys.Chem.A* 1997, 101, 21,3841
- 73) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03, Gaussian, Inc.: Pittsburgh, PA, 2003.*

## -CHAPTER 7-

# Electrostatic control of the photoisomerization efficiency in Rhodopsin

### 7.1. Introduction

The protonated Schiff base of the 11-*cis* retinal (PSB11) is the chromophore of visual pigments.<sup>1-6</sup> These include rhodopsin (Rh) that peaks at 498 nm<sup>7</sup> and is used in twilight vision, and the human three color vision pigments that peak at 425 nm (S-cone or blue), 530 nm (M-cone or green), and 560 nm (L-cone or red).<sup>7,8</sup> The biological activity of rhodopsins is triggered by the ultrafast (200 fs in Rh, see Eq. 1) light-induced *cis-trans* isomerization of the corresponding retinal chromophores that, in turn, induces a conformational change in the protein on longer timescales.<sup>1,5</sup> This ultrashort photochemical step is usually referred to as the primary event of the protein photocycle.



**Equation 7.1**

Garavelli et al.<sup>9-14</sup> have previously reported the results of *ab initio* CASPT2//CASSCF minimum energy path (MEP) mapping for the photoisomerization *in vacuo* of reduced models of the retinal chromophore and, very recently, the results for the real (i.e. unreduced) PSB11 chromophore have been presented.<sup>15</sup> It has been shown that the reaction mechanism is characterized by a *two-mode* photoisomerization path (first stretching then one-bond-flip (OBF)<sup>16,17</sup> torsion of the reacting double bond) that develops entirely on the spectroscopic (charge transfer) state  $S_1$ . This path is essentially barrierless and drives the system into a central double bond twisted  $S_1/S_0$  conical

intersection (CI) funnel whose geometry and electronic structure is consistent with that of a twisted intramolecular charge transfer (TICT) state.<sup>10,12,13</sup> This point triggers the ultrafast decay to the ground state and prompts an efficient photoproduct formation.

While the aforementioned results disclose the intrinsic photoisomerization ability of the retinal chromophore, environmental factors (such as the ones involved in the solvent or protein) are known to affect both its spectroscopy and photochemistry.<sup>18</sup> For instance, the absorption maximum in Rh, 498 nm, appears to be red-shifted with respect to the one observed in solution (440 nm), and the protein structure itself plays a role in the spectral tuning of the pigments as revealed by the three (S, M and L) color vision pigments (in this context, the spectral effects of the substitution of one or more amino acids have been widely investigated experimentally).<sup>8,19,20</sup> Secondly, while in the protein the photoreaction is finalized in 200 fs with a high (ca. 67%) quantum yield (QY)<sup>21,22</sup> under stereoselective control (i.e. only the *all-trans* (PSBT) photoproduct is produced), the isolated chromophore in solution (e.g. methanol or hexane) features a bi-exponential excited state decay dynamics with a dominant (almost 20 fold longer) 2 ps component<sup>21,23,24</sup>, the stereoselectivity is lost, and the QY decreases (ca. 25%).<sup>21,25</sup> Consistently, an excited state energy barrier has been observed for PSBT in solvent<sup>23</sup>, while no barriers are expected for Rh due to the sub-picosecond nature of the process,<sup>9,11</sup> arising questions on the origin of the protein catalytic effect that makes this event as one of the fastest photochemical reactions observed so far in nature.

It is apparent that both steric and electrostatic interactions between the chromophore and the surrounding protein pocket must play a key role in tuning/controlling the photochemical and photophysical behavior of the system. The reasons behind the optical properties and catalytic effect of Rh have been widely investigated in the past both experimentally and computationally. Those works reveal that different cooperative effects are responsible for the increased efficiency and photoisomerization rate observed in the visual pigment. For instance, previous experimental works by Mathies and coworkers<sup>26</sup> have shown that intramolecular steric interactions on between the retinal C<sub>10</sub>-hydrogen and the C<sub>13</sub>-methyl do significantly affect the efficiency (*i.e.* QY and rate) of the photoisomerization process. Consistently with this view, Buss and coworkers<sup>27</sup> have recently shown that the ground state twists of the C<sub>11</sub>=C<sub>12</sub> and the C<sub>12</sub>-C<sub>13</sub> bonds (which are mainly due to interaction with the Rh protein pocket) help to effectively and rapidly transfer stretch energy into torsion energy and thus populate efficiently the isomerization coordinate. Furthermore, recent reports on crystal structures of Rh<sup>28,29</sup> have pointed out the close proximity of Cys-187 with the retinal C<sub>12</sub>-hydrogen, suggesting how such an interaction could lead to enhanced isomerization upon photoexcitation. Besides steric interactions, also electrostatic effects are expected to play a major role in photoisomerization catalysis as the photoactive state S<sub>1</sub> has a charge transfer nature and its energy is expected to depend on interactions with surrounding charged residues and polar groups. The same holds (for the same reasoning) for the TICT S<sub>1</sub>/S<sub>0</sub> CI funnel and, more generally, for the S<sub>1</sub>/S<sub>0</sub> crossing seam. In fact, its position is dictated by external charges that may displace the S<sub>1</sub>/S<sub>0</sub> crossing seam out of the ideal photoisomerization channel or can even remove it.<sup>18,30</sup> Anyway, while steric effects on the photoisomerization have been deeply analyzed, studies on electrostatic effects appear to be mainly focused on static spectral properties (*i.e.* vertical excitations) rather than on the efficiency of the photochemical reaction, and to the best of our knowledge no precise and systematic analysis of these effects on the photoisomerization ability of the chromophore in the visual pigments is available up to date. For example, a

static as well as dynamical molecular level description of the primary photoisomerization event in Rh has been recently reported by Olivucci and coworkers<sup>31,32,33</sup> using Quantum Mechanics/Molecular Mechanics (QM/MM)<sup>34</sup> computations at the multiconfigurational(CASSCF)/correlated energy(CASPT2) resolution. This is the reference simulation up to date on Rh photoreaction. Anyway, that study does not focus or analyze the reasons behind Rh photoisomerization efficiency, such as environment electrostatic effects. Among these, the interaction with the formal E113 counterion and nearby charged/polar residues must be carefully analyzed as they are likely to be the principal contributors of these effects. Interestingly, the protonation state of the E181 residue (that is located right above the photoactive central double bond of the chromophore)<sup>28,29,35</sup> has been recently reconsidered: while previous two-photon absorption<sup>36</sup> and mutation experiments<sup>37,38</sup> seemed to suggest a neutral binding site and assigned it to a neutral (i.e. protonated) species, recent experimental<sup>39-41</sup> and computational studies<sup>42</sup> suggest a deprotonated (i.e. negatively charged) 181 residue. Implications of this nearby negative charge on the photoisomerization efficiency and stereoselectivity have been suggested.<sup>18,30,41</sup> Among others, these still unresolved issues represent a stimulus for the present QM/MM investigation, although here a more systematic study of all (both close and far) residues playing a role on these effects is considered.

As mentioned above, Rh absorption peaks at 498 nm.<sup>7</sup> This is a well established result, whereas uncertainties exist for the absorption values of higher energy states. The Rh absorption spectrum recorded by Ebrey and Honig<sup>43</sup> displays, besides the dominant 498 nm component, a second smaller maximum at shorter wavelength peaking around 340 nm. This should refer to a less absorbing (still not forbidden) higher energy state with an oscillator strength that is roughly one third of  $S_1$ , as the inspection of the spectrum approximately reveals. Two-photon absorption experiments by Birge and coworkers<sup>36</sup> on Rh (containing a 11-cis-locked retinal analog to prevent bleaching) suggest a  $S_2$  (covalent dark  $A_g$ -like) state that is only 2000  $\text{cm}^{-1}$  above  $S_1$  (which would lead to a vertical 63 kcal/mol (i.e. 453 nm) energy value for this state). If this were the case, the higher energy 340 nm band observed by Ebrey and Honig would originate from a state higher than  $S_2$ . Anyway, based on the absorption values recorded for isolated retinal chromophores in the gas phase (that place the signatures of the  $S_2$  state at much higher energies), Andersen and coworkers<sup>44</sup> have very recently reassigned the observed 2000  $\text{cm}^{-1}$  two-photon absorption band to a vibrational feature of the bright lowest energy  $S_1$  state spectrum, thus suggesting that  $S_2$  does correspond indeed to the 340 nm band.

In recent computational investigations, both Buss<sup>45</sup> and Olivucci<sup>31</sup> successfully reproduced Rh absorption by applying the same approach (with the former value in remarkable good agreement with the experimental data): multireference perturbation level (CASPT2/CASSCF(12,12)) calculations on the chromophore put in the bath of the protein point charges. Anyway, the reasons leading to reproduce the correct absorption value are contrasting and striking different in the two cases. In the former, the Rh electrostatic environment is shown to have only a very minor effect on the absorption properties of the chromophore/counterion (E113) couple (that alone does almost reproduce the correct  $S_1$  energy found in the protein), while in the latter this effect is major as it quenches (i.e. counterbalance) almost fully the electrostatic effect of the counterion, basically recovering the absorption energy computed for the isolated distorted (i.e. optimized within the protein) chromophore (that alone does already almost reproduce the value observed in the protein). Additionally, Buss reproduces

fairly well (66.7 kcal/mol)<sup>45</sup> the lower energy value reported by Birge for S<sub>2</sub> (63 kcal/mol),<sup>36</sup> while for the same (i.e. covalent A<sub>g</sub>-like) S<sub>2</sub> state Olivucci replicates satisfactorily (87.5 kcal/mol)<sup>31</sup> the 340 nm (84 kcal/mol) absorption feature observed by Ebrey and Honig.<sup>43</sup> While those results were both obtained considering a neutral Glu181 residue, more recently Bravaya et al. (yet applying another perturbative QM/MM approach) successfully reproduced Rh absorption (as well as the gas and solution values)<sup>46</sup> using a charged E181; shielding of the counterion by the protein field was also predicted in this case. Additionally, recent SAC-CI QM/MM computations<sup>47</sup> also nicely predicted S<sub>1</sub> excitation (using a neutral E181), but no shielding emerged. Finally, Altun et al.<sup>48</sup> have very recently reproduced Rh absorption as well (and that of its mutants) employing a DFT/TD-DFT QM/MM procedure: E181 was still considered neutral, and the counterion was predicted to produce a much smaller blue shift (from 4 to 6 kcal/mol only) than in all previous studies, and to be unquenched by the protein (i.e. the net effect of the other amino acids is negligible). It is apparent that all those contrasting findings call for a major reinvestigation of Rh spectroscopic properties. This is far from being just a simple academic issue, as it allows to disclose the role played by protein electrostatic effects in spectral tuning.

In this chapter, the primary photochemical event of Rh and the singlet manifold characterizing this process are explored by *ab initio* multiconfigurational perturbation theory calculations (CASPT2//CASSCF) via adopting a novel (recently developed by our group) hybrid QM/MM procedure accounting for electrostatic embedding at the QM level.<sup>49</sup> The most recent and best-resolved crystallographic structure available for Rh is employed<sup>29</sup> and the protein embedded photoisomerization path is mapped. While the structure of the photoisomerization coordinate in Rh is not the focus of this study (as it has been documented and discussed in details elsewhere),<sup>31,32,33</sup> herein we provide information on the effects that the protein electrostatic environment may exercise on the (i) optical (i.e. spectral tuning) and (ii) photoreactivity properties of rhodopsins. For this purpose, the effect of both close and distant residues is explicitly accounted for and analyzed. While we do not want to reject or question out other effects that have been previously recognized in photoisomerization catalysis, here we provide computational evidences for a mechanism of electrostatic control in the photoactivity of visual pigments that have not been recognized before. A thoughtfully discussion of all the most recent experimental/computational findings compared to the presented results is shown, which supports this new view. Finally, a unified model is drawn that discloses the relationship between spectroscopic and mechanistic properties in Rh and related (blue, green, and red) color vision pigments. This leads to formulate a solid mechanism for spectral tuning in color vision pigments. It is anticipated here that these findings may open novel scenarios in the rationale of mutations-dependent vision deficiencies.

## 7.2. Computational Details

All computations are performed using a new hybrid QM/MM<sup>34</sup> potential recently developed by our group.<sup>49</sup> Details of this hybrid approach are here not presented (see appendix A). Briefly, the method is based on a hydrogen link-atom scheme.<sup>50</sup> QM and MM layers interact in the following way: (i) all QM atoms feel the electrostatic potential of all MM point charges (i.e. an electrostatic embedding scheme has been adopted), (ii) all the bonding (i.e. stretching, bending, and torsion) and the Van der Waals QM/MM cross terms are described by the standard MM potential, and (iii)

electrostatic interactions involving MM atoms are all accounted for classically. In this case, the QM-MM frontier is placed at the C $\gamma$ -C $\delta$  bond of the Lys-296 side chain (see eq. 1). The QM region (i.e. the chromophore atoms) is treated by ab initio CASSCF/6-31G\* computations using the Gaussian03<sup>51</sup> suite of programs (a complete active space of 12  $\pi$ -electrons in 12  $\pi$ -orbitals (12,12) is employed to describe the CASSCF wavefunction), while the AMBER8<sup>52</sup> software and the *ff99* force field<sup>52</sup> is used for the MM region (i.e. the opsin) and its charges. CASSCF/AMBER QM/MM computations are accomplished using the COBRAMM interface,<sup>49</sup> implemented by our group, that links Gaussian03<sup>51</sup> and AMBER8<sup>52</sup> packages.

Protein chain (the monomer A) of the recently crystallized and best-resolved X-ray structure available for bovine Rh (pdb code: 1U19, resolution = 2.2 Å),<sup>29</sup> which has no missing amino acids, is used for the protein framework. Hydrogen atoms are added by means of the H++ software<sup>53</sup>. In general, all titratable groups are carefully reconsidered and through an accurate analysis the following choice is adopted: the Glu-113 counterion is taken charged, Glu-122 is neutral, Glu-181 is considered both charged and neutral in two different protein set ups (Rh<sub>181(-)</sub> and Rh<sub>181H</sub>, respectively), all other glutamate residues are charged, lysine residues are protonated, Asp-83 is protonated, and three (Hip-65, Hip-100, Hip-278) of the six histidine residues are protonated.

The protein framework is kept fixed during all QM/MM geometry optimizations (we assumed that it has no time to change during the ultrafast (200 fs) timescale of the primary photo-induced event), while QM chromophore atoms and MM atoms comprising the Lys-296 side chain and the two water molecules (W1 and W2) are left free to relax.

A S<sub>0</sub>-S<sub>n</sub> n-root state average CASSCF wavefunction (equally weighting all the n roots) is always used for geometry optimizations on the S<sub>n</sub> state, while a single-state wavefunction is used for ground state optimizations. A relaxed scan (with the central dihedral angle fixed at specified values) is used to trace the photoisomerization channel in the protein, while fully unconstrained optimizations are employed to locate all the critical points.

To account for correlation energy, both CASPT2<sup>54</sup> and MS-CASPT2 single point computations are performed on the optimized points in the bath of the protein (AMBER) point charges, by using the MOLCAS-6.0 package.<sup>55</sup> Unless otherwise stated, a three-root state average wavefunction is employed. To minimize the influence of weakly interacting intruder states at the second-order level, the so-called imaginary level shift technique is used (imaginary shift value=0.2).<sup>10</sup> QM(CASPT2)/MM energies are taken as the reference values throughout the work, while MS-CASPT2 corrections are considered only when a multi-state problem appears (this happens only in one specific situation as will be discussed in the text below), see also supporting information for details. The CASSCF state interaction (CASSI) method<sup>56</sup> is used (as implemented in MOLCAS-6.0<sup>55</sup>) to calculate the transition dipole moments. Oscillator strengths (*f*) are computed using CASSCF transition moments and CASPT2 corrected energies.

The charge distribution along the chromophore chain is evaluated according to state-averaged CASSCF wavefunctions by Mulliken-charges analysis and is used to characterize the electronic nature (*i.e.* ionic *vs.* covalent) of the investigated states.

A Reverse Fingerprint (RFP) analysis is employed on the investigated residues to highlight their electrostatic effects on the excitation energy: this is done by switching off specifically the charges of the investigated residue (while leaving on all the other charges of the protein environment) and recording the CASPT2 (blue or red) shift in the

chromophore  $S_1$  excitation energy with respect to the reference Rh value (that is represented by the zero line in the corresponding figure, see below).

## 7.3 Result and Discussion

### 7.3.1 The Glu181 issue: neutral or protonated?

Inspired by the recent works<sup>39-42</sup> that have readdressed the charge state of the Glu-181 residue in Rh (considered neutral so far), the chromophore structure has been optimized into both a neutral ( $Rh_{181H}$ ) and a negatively charged ( $Rh_{181(-)}$ ) protein binding pocket, corresponding to a neutral (i.e. protonated) or a charged (i.e. deprotonated) Glu-181 residue, respectively. Table 7.1 reports CASPT2 vertical energies computed for the  $S_1$  (ionic/charge-transfer  $B_u$ -like) and  $S_2$  (covalent  $A_g$ -like) states of the chromophore optimized in the ground state of the two different set-ups together with their oscillator strengths and wavefunction characterization. Interestingly, while  $Rh_{181H}$  leads to an  $S_1$  energy (53.6 kcal/mol) that is slightly underestimated with respect to the observed absorption value (peaking at 57.4 kcal/mol, i.e. 498 nm),  $Rh_{181(-)}$  reproduces remarkably well this energy (57.6 kcal/mol). On the other hand,  $S_2$  appears unaffected by the electrostatic environment, leading to an energy value that is ca. 82 kcal/mol in both cases. This is consistent with  $S_2$  possessing the same electronic nature as  $S_0$  (covalent  $A_g$ -like, i.e. they have a similar electronic density distribution), leading to a  $S_0$ - $S_2$  energy gap that is barely affected by external electrostatic effects. This value and its computed oscillator strength (0.3, see Table 7.1) nicely agree with the secondary (less intense) higher energy absorption feature observed in Rh at 340 nm (84 kcal/mol) by Ebrey and Honig.<sup>43</sup> Anyway, it must also be said that it greatly differs from the other experimental value (by two-photon absorption) observed by Birge and coworkers<sup>36</sup> that is much closer to the value computed by Buss.<sup>45</sup> Nonetheless, this assignment has been recently questioned out<sup>44</sup> so that uncertainties still exist on the  $S_2$  absorption energy in Rh. It is worth to note that this would be a very valuable information since it would provide an outstanding test for QM/MM Rh approaches as it appears that very different energies are predicted for this state according to different computational approaches employed.

Although this spectroscopic scenario is in qualitative agreement with the one reported by Olivucci and coworkers,<sup>31</sup> the quantitative agreement with the experiments appear to be significantly improved. Furthermore, an electrostatic environment (with a charged Glu-181) that is different than the one conventionally used, has been considered here (as also done in a more recent work by Bravaya et al.<sup>46</sup>).

Protein structure	Chromophore State (electronic nature)	$\Delta E$	$f_{S_0 \rightarrow S_n}$	Wavefunction <sup>a</sup> (coefficient)
$Rh_{181(-)}$	$S_0$ ( $A_g$ -like)	0.0	0.00	$(6a)^2 (7a)^0$ (0.77)
	$S_1$ ( $B_u$ -like)	57.6	0.74	$(6a)^1 (7a)^1$ (0.63)
	$S_2$ ( $A_g$ -like)	82.0	0.27	$(6a)^0 (7a)^2$ (0.11) $(6a)^1 (7a)^1$ (0.19)
	$S_0$ ( $A_g$ -like)	0.0	0.00	$(6a)^2 (7a)^0$

Rh <sub>181H</sub>				(0.77)
	S <sub>1</sub> (B <sub>u</sub> -like)	53.6	0.75	(6a) <sup>1</sup> (7a) <sup>1</sup> (0.64)
	S <sub>2</sub> (A <sub>g</sub> -like)	81.6	0.28	(6a) <sup>0</sup> (7a) <sup>2</sup> (0.10) (6a) <sup>1</sup> (7a) <sup>1</sup> (0.18)

<sup>a</sup>6a and 7a represent the HOMO and LUMO  $\pi$  orbitals, respectively.

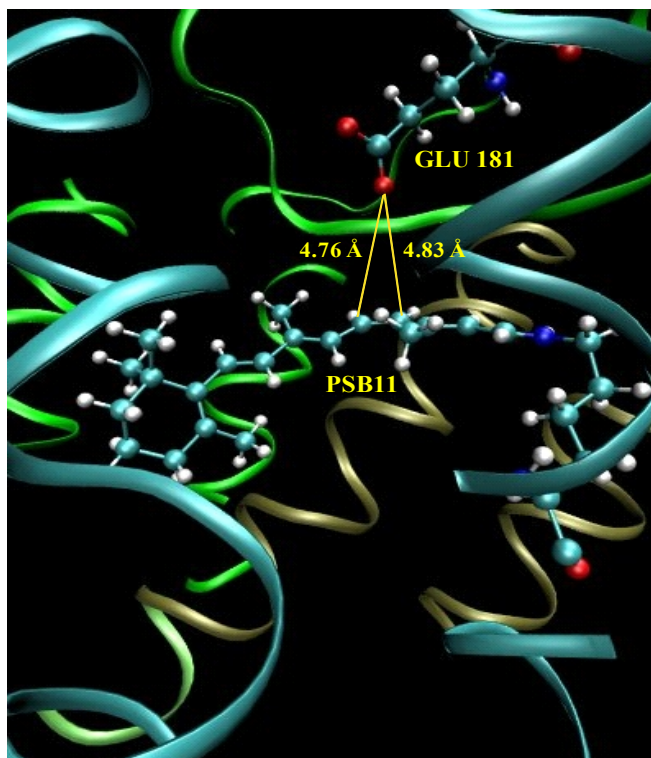
**Table 7.1.** CASPT2 relative energies ( $\Delta E$ , kcal/mol) for the lowest three singlet states of two different protein set-ups (Rh<sub>181(-)</sub> and Rh<sub>181H</sub>). Oscillator strengths for ground state excitations ( $f_{S_0 \rightarrow S_n}$ ) and wavefunction characterization are also reported.

While a charged Glu-181 residue is consistent with the changes observed in the spectrum of some 181-mutated Rh (e.g. E181Q and E181D),<sup>39</sup> it is in contrast with other mutations<sup>37,38</sup> that revealed a very minor effect on the S<sub>1</sub> absorption energy due to 181 substitution and that were taken as a proof for a neutral Glu-181. It is apparent that absorbance measurements on Rh mutants do not give a clear-cut answer about whether E181 is ionized in the dark. This is not surprising as this residue is placed right above the central C<sub>11</sub>=C<sub>12</sub> double bond of retinal (i.e. in a symmetric position with respect to the PSB head and the  $\beta$ -ionone tail of the chromophore that supports charge translocation upon photoexcitation to S<sub>1</sub>, see Figure 7.1), which makes S<sub>1</sub> rather insensitive to its protonation state, as recent computational investigations indeed reveal<sup>18,30,46</sup>. This has also been pointed out recently by Kliger and coworkers.<sup>39</sup> Consistently, it must be noted that the computed effect of Glu-181 ionization on S<sub>1</sub> absorption, although not negligible (ca. 3 kcal/mol), is very much less than the one expected in general for a nearby charged residue (e.g. the effect of the Glu-113 counterion residue (ca. 0.6 eV) reported and discussed below!). As its protonation state has much less impact on Rh absorbance than generally assumed, mutation experiments must be carefully analyzed. Anyway, many other experimental/computational evidences have been very recently collected that suggest a charged Glu-181, specifically: i) Glu-181 acts as the counterion in invertebrate rhodopsins;<sup>40,57</sup> ii) Glu-181 must be ionized to stabilize the late photointermediates (e.g. MI) and make efficient the photoactivation cycle of visual pigments (the so called Glu-113/Glu-181 *counterion switch model*); iii) Glu-181 is deprotonated (together with the primary counterion) in the mouse short-wavelength sensitive visual pigment (MUV);<sup>58,59</sup> iv) Glu-181 is the counterion in parapinopsin that shares with Rh a great sequence identity;<sup>57</sup> v) FTIR spectroscopy studies<sup>41</sup> and MD simulations<sup>42</sup> of Rh are consistent with a negatively charged Glu-181 in the ground state. Taken all together, and jointly with the present results, these findings suggest a deprotonated 181 residue and an overall negatively charged binding pocket.



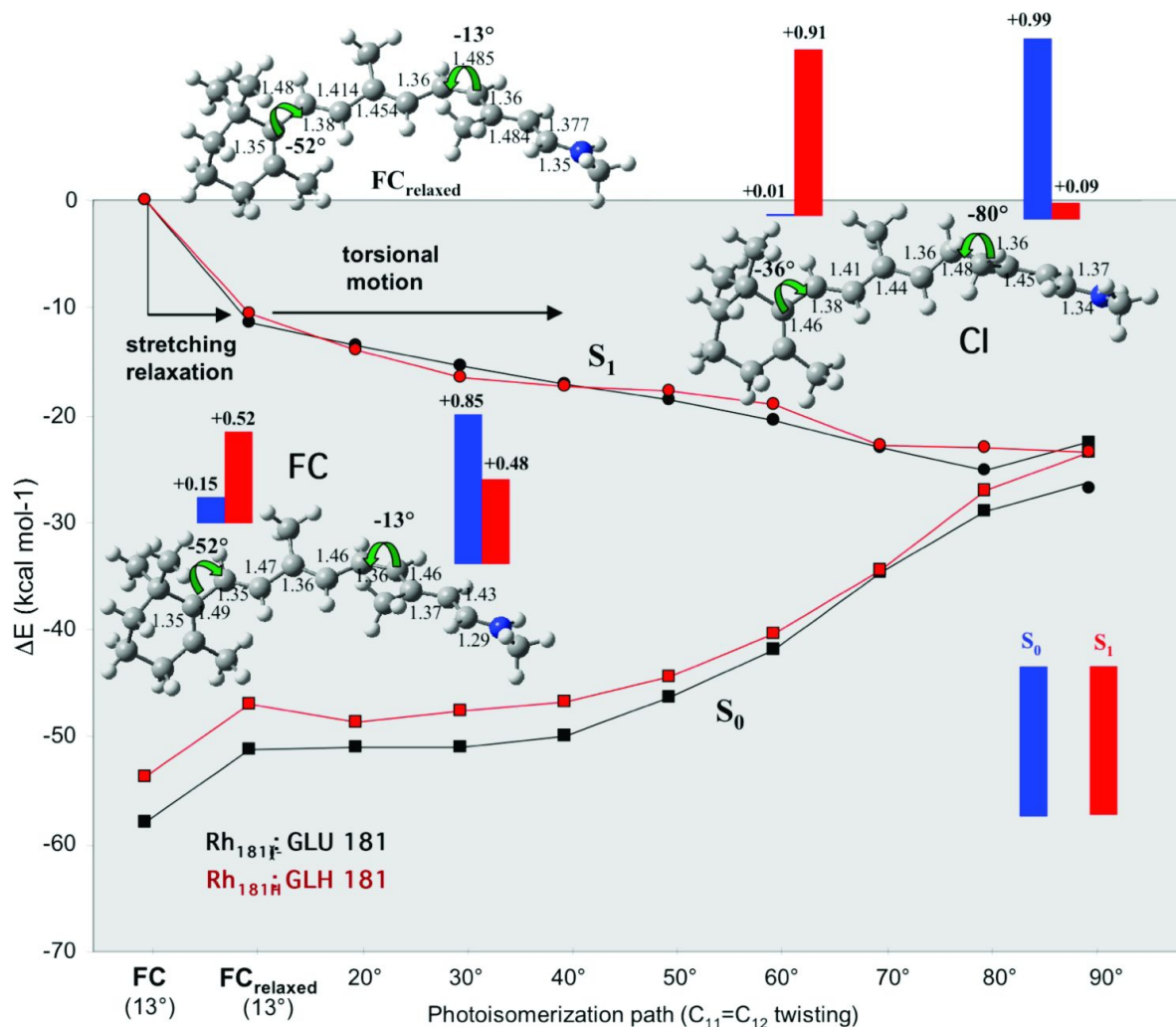
**Figure 7.1.** Retinal binding pocket in Rh drawn from its crystallographic structure.<sup>29</sup> It is apparent that Glu-181 is placed right above the central C<sub>11</sub>=C<sub>12</sub> double bond of the chromophore, in a symmetric position with respect to its  $\beta$ -ionone tail and protonated Schiff base head.

To disclose the possible effects that the protonation state of Glu-181 may have on the chromophore photoisomerization ability, the photoisomerization channel is computed for both Rh<sub>181H</sub> and Rh<sub>181(-)</sub>, and the two paths are reported and compared in Figure 7.2. Since Glu-181 is positioned near the center of the chromophore, no significant effects are foreseen on the photoisomerization process for the same reasons discussed above. Consistently, Figure 7.2 shows that



the photoreaction efficiency remains the same whatever is the 181 residue charge: both the slope and the CI funnel are unaffected. These results seem to question out a possible involvement of the Glu-181 residue in assisting the primary photochemical event, invoking its role “only” in the stabilization of the later photocycle intermediates and the promotion of Rh photoactivation (as according to the counterion switch model). Under this respect, both Rh<sub>181H</sub> and Rh<sub>181(-)</sub> can be safely taken to properly study the photoisomerization mechanism; anyway, since Rh<sub>181(-)</sub> does reproduce remarkably well the spectroscopic properties observed in Rh and agrees with the most recent experimental findings on this topic (as discussed above), we will consider this setup thereafter for Rh.

It is worth noting that previous QM/MM studies on Rh photoisomerization<sup>31,33</sup> surprisingly showed the existence of an intermediate (i.e. a minimum) on S<sub>1</sub> that seems apparently to contrast with the ultrafast photoisomerization observed in Rh. Although this fact does not appear to disfavor retinal sub-picosecond photoisomerization dynamics (but this result is based on just a single MD simulation at 0 K),<sup>33</sup> in this work (starting from the best-resolved X-ray structure of Rh)<sup>29</sup> we partially revise this picture as we find a fully barrierless path with no excited state intermediate and get a static description of the mechanism whose energies are more coherent with experimental observations.



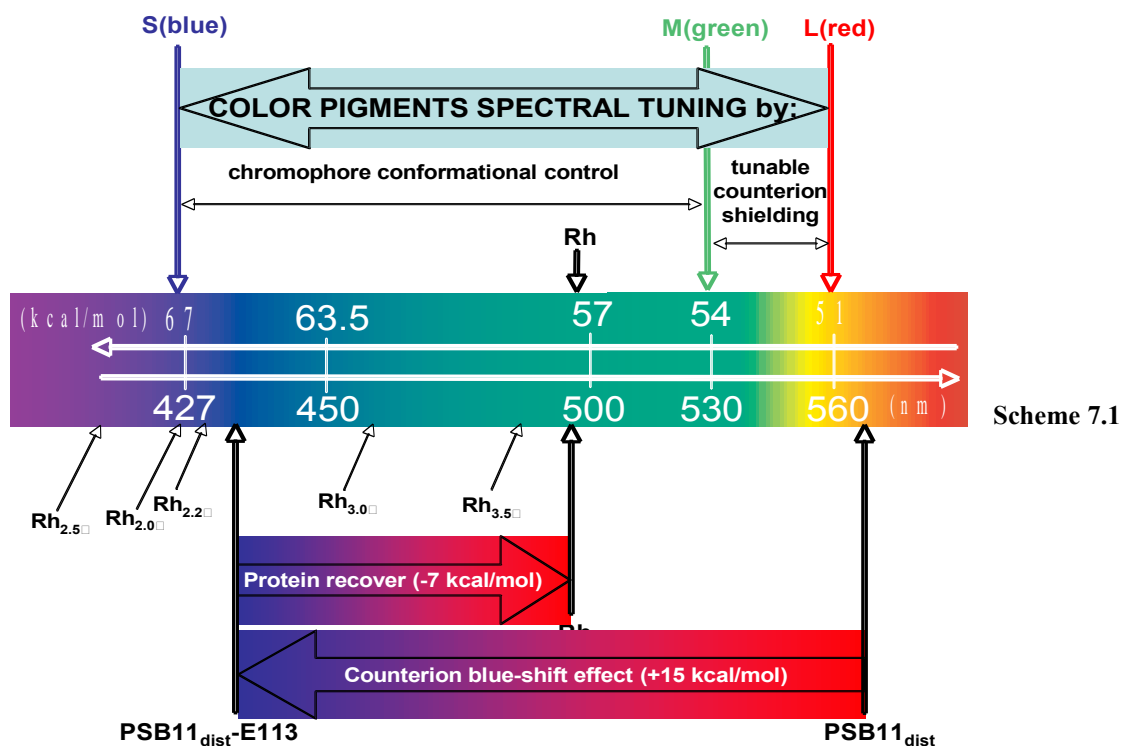
**Figure 7.2.** CASPT2 corrected S<sub>0</sub> and S<sub>1</sub> energy profiles along the QM(CASSCF)/MM relaxed scan on S<sub>1</sub> tracing the photoisomerization process in the two different Rh set-ups, from the Franck-Condon (FC) point to the twisted conical intersection funnel (CI), through a point (FC<sub>relaxed</sub>) where only skeletal bond stretchings have been relaxed (i.e. dihedral angles have been frozen here). Bar diagrams display the fractional positive charge on each chromophore half with respect to the rotating central bond. Molecular parameters are reported in Å and degrees.

Obviously, the photoisomerization mechanism is the same as documented before,<sup>9,11</sup> i.e. a complex reaction coordinate<sup>29</sup> is involved, differing from the simple OBF mode: that is, cooperative/concurrent twisting motions of adjacent C-C bonds result into a space saving mechanism that resembles (although aborted and developed only asynchronously and transiently) the bicycle pedal model proposed few decades ago by Warshel in his pioneering computational studies,<sup>60,61</sup> that already successfully predicted, by semi-empirical molecular dynamics simulations, the correct 200 fs timescale for the photoisomerization process in Rh. Notably, the computed direction of twisting is in agreement with the sign of the initial C<sub>11</sub>=C<sub>12</sub> twist which has been previously derived<sup>27,31,62</sup> in Rh.

### 7.3.2 The counterion and the other polar/charged residues: counterion quenching by the protein pocket.

In this section we will decompose the protein electrostatic effects eventually building up the absorption energy observed in Rh. This analysis appears to be relevant for spectral tuning, but turns out to be crucial also for the photoisomerization efficiency, as it will be discussed further below.

**Optical properties.** Table 7.2 shows how the  $S_1$  (ionic, single excitation  $B_u$ -like) and  $S_2$  (covalent, double excitation  $A_g$ -like) absorption properties change upon increasing the complexity of the system, i.e. from the isolated chromophore ( $PSB11_{dist}$ ) to the full protein ( $Rh_{181(-)}$ ), passing through the ion-pair ( $PSB11_{dist}$ -E113) system and shells of increasing size (i.e. comprising all the residues within a specified radius around the chromophore:  $Rh_{r=2.0\text{\AA}}$ ,  $Rh_{r=2.2\text{\AA}}$ ,  $Rh_{r=2.5\text{\AA}}$ ,  $Rh_{r=3.0\text{\AA}}$ ,  $Rh_{r=3.5\text{\AA}}$ ). Note that the chromophore structure is always the distorted one ( $PSB11_{dist}$ ) optimized in the protein. This analysis allows us to progressively identify the protein residues that finally lead to the recorded Rh absorption. Scheme 1 summarizes these results graphically. As expected (see the discussion in the previous section),  $S_2$  is much less affected by external charges and its energy remains substantially unchanged, while  $S_1$  is very much sensitive to the electrostatic environment. Not surprisingly, the major effect on  $S_1$  comes from the counterion (E113): an  $S_1$  vertical absorption energy of 65.7 kcal/mol is predicted in the  $PSB11_{dist}$ -E113 couple that is blue shifted by ca. 15 kcal/mol with respect to the value calculated in the bare chromophore  $PSB11_{dist}$  (50.3 kcal/mol). This effect is in quantitative agreement with that calculated by Buss and coworkers<sup>45</sup> (that is, the computational reference on this topic up to now), while Olivucci,<sup>31</sup> Bravaya,<sup>46</sup> and Nakatsuji<sup>47</sup> all predict a larger value ( $> 1$  eV). It is worth noting that multi-state CASPT2 (MS-CASPT2) corrections are needed in our case due to the degenerate  $S_1$  and  $S_2$  CASPT2 energies found in the  $PSB11_{dist}$ -E113 couple.



Note that without MS-CASPT2 we do also predict an higher  $> 1$  eV blue shift, see Table 7.2; on the other hand, Buss and coworkers do not need MS corrections as  $S_1$  and  $S_2$  CASPT2 energies are quite different in his system.<sup>45</sup> Thus, we will consider the MS-CASPT2 values as the reference ones (i.e. the most accurate result) in this case, while standard CASPT2 results will be considered in all the other situations as the  $S_1$  and  $S_2$  states are not degenerate anymore (see computational details and supporting information for a more in depth discussion on this topic). Quite surprisingly, DFT/TD-DFT ONIOM-based QM/MM computations on Rh very recently reported by Altun et al.<sup>48</sup> predict a much smaller blue shift effect ( $< 0.5$  eV) for the counterion that goes against all the other reported values.

Structure	Chromophore State (electronic nature)	$\Delta E$	$f_{S_0 \rightarrow S_n}$	Wavefunction <sup>a</sup> (coefficient)	Residues
PSB11 <sub>dist</sub>	$S_0$	0.0	0.00	(6a) <sup>2</sup> (7a) <sup>0</sup> (0.78)	
	$S_1$	50.3	0.78	(6a) <sup>1</sup> (7a) <sup>1</sup> (0.65)	
	$S_2$	79.3	0.23	(6a) <sup>0</sup> (7a) <sup>2</sup> (0.15) (6a) <sup>1</sup> (7a) <sup>1</sup> (0.13)	
PSB11 <sub>dist</sub> -E113	$S_0$	0.0 <sup>b</sup> (0.0)	0.00	(6a) <sup>2</sup> (7a) <sup>0</sup> (0.75)	Glu-113
	$S_1$	65.7 <sup>b</sup> (75.7)	0.86	(6a) <sup>1</sup> (7a) <sup>1</sup> (0.54)	
	$S_2$	95.2 <sup>b</sup> (74.7)	0.11	(6a) <sup>0</sup> (7a) <sup>2</sup> (0.23) (6a) <sup>1</sup> (7a) <sup>1</sup> (0.18)	
Rh <sub>r=2.0Å</sub>	$S_0$	0.0	0.00	(6a) <sup>2</sup> (7a) <sup>0</sup> (0.75)	Glu-113, Ala-295, Lys-296, Thr-297
	$S_1$	67.2	0.62	(6a) <sup>1</sup> (7a) <sup>1</sup> (0.47)	
	$S_2$	77.4	0.35	(6a) <sup>0</sup> (7a) <sup>2</sup> (0.19) (6a) <sup>1</sup> (7a) <sup>1</sup> (0.40)	
Rh <sub>r=2.2Å</sub>	$S_0$	0.0	0.0	(6a) <sup>2</sup> (7a) <sup>0</sup> (0.75)	Rh <sub>r=2.0Å</sub> + Trp-265, Tyr-268, Met-207, Phe-212, Gly-121, Glh-122, Phe-293
	$S_1$	66.4	0.60	(6a) <sup>1</sup> (7a) <sup>1</sup> (0.49)	
	$S_2$	77.6	0.37	(6a) <sup>0</sup> (7a) <sup>2</sup> (0.20) (6a) <sup>1</sup> (7a) <sup>1</sup> (0.39)	
Rh <sub>r=2.5Å</sub>	$S_0$	0.0	0.00	(6a) <sup>2</sup> (7a) <sup>0</sup> (0.75)	Rh <sub>r=2.2Å</sub> +Phe-91, Gly-114, Cyx-187, Ile-189, Phe-261, Ala-292, Ser-298
	$S_1$	69.0	0.66	(6a) <sup>1</sup> (7a) <sup>1</sup> (0.49)	
	$S_2$	77.0	0.31	(6a) <sup>0</sup> (7a) <sup>2</sup> (0.20) (6a) <sup>1</sup> (7a) <sup>1</sup> (0.39)	
	$S_0$	0.0	0.00	(6a) <sup>2</sup> (7a) <sup>0</sup>	Rh <sub>r=2.5Å</sub> +Ala-299,

Rh <sub>r=3.0Å</sub>				(0.76)	Met-44, Leu-47, Thr-94, Ala-117, Thr-118, Leu-125, Gly-188, Tyr-191, Phe-208, His-211, Phe-294
	S <sub>1</sub>	62.2	0.43	(6a) <sup>1</sup> (7a) <sup>1</sup> (0.57)	
	S <sub>2</sub>	79.6	0.46	(6a) <sup>0</sup> (7a) <sup>2</sup> (0.18) (6a) <sup>1</sup> (7a) <sup>1</sup> (0.32)	
Rh <sub>r=3.5Å</sub>	S <sub>0</sub>	0.0	0.00	(6a) <sup>2</sup> (7a) <sup>0</sup> (0.76)	Rh <sub>r=3.0Å</sub> +Wat-358, Tyr-43, Tyr-178, Ser-186, Wat-354
	S <sub>1</sub>	58.7	0.49	(6a) <sup>1</sup> (7a) <sup>1</sup> (0.61)	
	S <sub>2</sub>	80.9	0.48	(6a) <sup>0</sup> (7a) <sup>2</sup> (0.17) (6a) <sup>1</sup> (7a) <sup>1</sup> (0.27)	
Rh	S <sub>0</sub>	0.0	0.00	(6a) <sup>2</sup> (7a) <sup>0</sup> (0.77)	Full protein (all residues)
	S <sub>1</sub>	57.6	0.74	(6a) <sup>1</sup> (7a) <sup>1</sup> (0.63)	
	S <sub>2</sub>	82.0	0.27	(6a) <sup>0</sup> (7a) <sup>2</sup> (0.11) (6a) <sup>1</sup> (7a) <sup>1</sup> (0.19)	

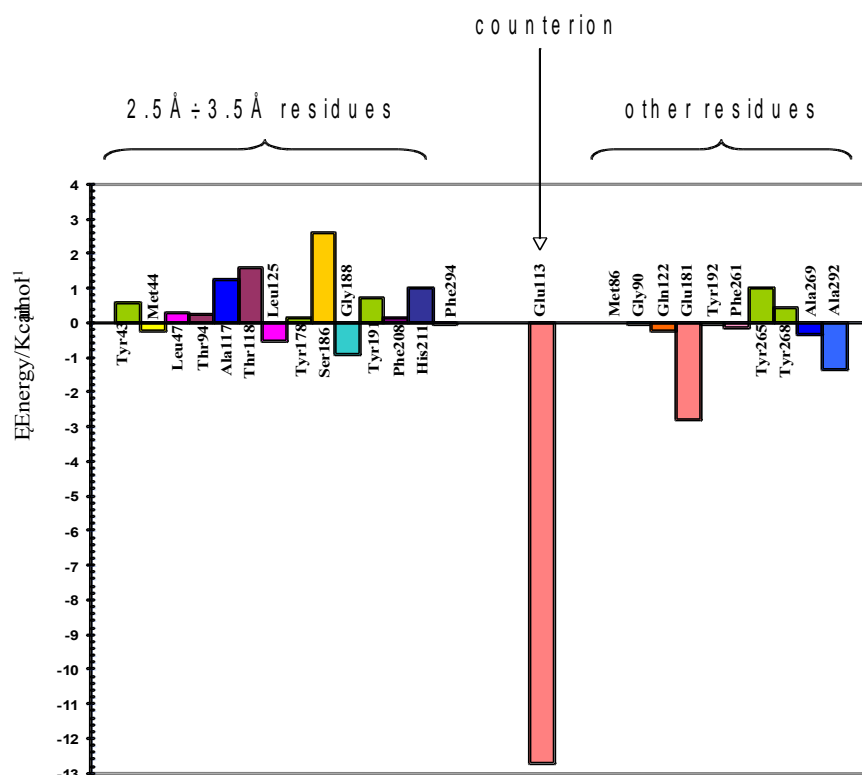
<sup>a</sup>6a and 7a represent the HOMO and LUMO  $\pi$  orbitals, respectively.

<sup>b</sup>Multi-State CASPT2 results: S<sub>1</sub> and S<sub>2</sub> are degenerate at the CASPT2 level (values in brackets).

**Table 7.2.** CASPT2 relative energies ( $\Delta E$ , kcal/mol) for the lowest three singlet states, oscillator strengths for ground state excitations ( $f_{S_0 \rightarrow S_n}$ ) and structure characterization (list of comprised residues) for systems of increasing size and complexity: the isolated chromophore (PSB11<sub>dist</sub>), the ion-pair (PSB11<sub>dist</sub>E113), proteinshells of increasing size (i.e. comprising all the residues within a specified radius  $r$  around the chromophore: Rh<sub>r=2.0Å</sub>, Rh<sub>r=2.2Å</sub>, Rh<sub>r=2.5Å</sub>, Rh<sub>r=3.0Å</sub>, Rh<sub>r=3.5Å</sub>), the full protein (Rh). The chromophore structure is always the distorted one (PSB11<sub>dist</sub>) optimized in the protein.

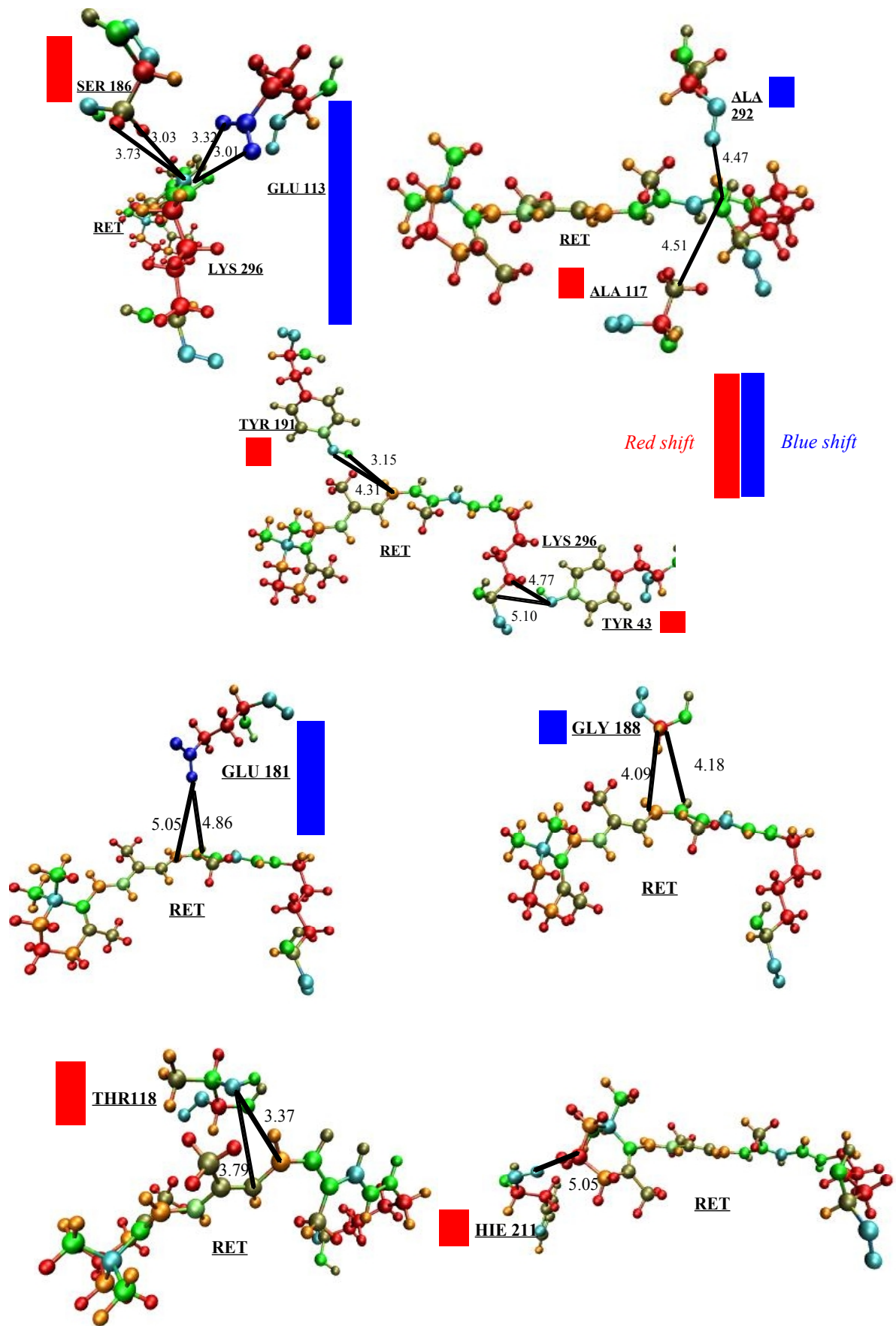
Interestingly, both the CASPT2 (74.7 kcal/mol) and MS-CASPT2 corrected (65.7 kcal/mol) values for S<sub>1</sub> in PSB11<sub>dist</sub>-E113 (see Table 7.2) are far higher than the one simulated/observed in Rh (57.6/57.4 kcal/mol). It is apparent that a quenching by the protein of the electrostatic effect due to the counterion is needed to recover the correct absorption energy. Anyway, the shielding effect that the protein dielectric exercises on E113 must not be full, as we need to recover back only ca. 50% (i.e. ca. 7 kcal/mol) of the counterion blue shift, see Scheme 7.1. Very remarkably, Table 7.2 shows that most of the residues responsible for this effect must be found nearby the retinal chromophore between the 2.5 and 3.5 Å shell, as it is on going from 2.5 Å to 3.5 Å that the counterion is mostly quenched and Rh absorption is recovered (see Scheme 7.1). Thus, these results reveal that only a limited number of actors (14 amino acids overall, see Table 7.2) is potentially involved in counterion quenching. To highlight the effect of these residues, a RFP analysis (see computational details) is performed on each of them and, for completeness, also on the Glu-113 counterion. The blue or red shift in the chromophore S<sub>1</sub> excitation energies with respect to the reference Rh value (that is represented by the zero line) is displayed in Figure 7.3 for each investigated residue.<sup>63</sup> Very interestingly, Figure 7.3 shows that there is not a specific residue that is responsible alone of the 113 quenching, as this appears more as a cooperative effect of a rather limited number of polar residues surrounding the chromophore, namely Ser-186 (2.6), Thr-118 (1.6), Ala-117 (1.3), His-211 (1.0), Tyr-191 (0.7), Tyr-43 (0.6) (the shielding effect in kcal/mol is reported in brackets), see Figure 7.3. These are the only

residues (that are comprised between the 2.5 and 3.5 Å shell) that produce a significant backward red shift of the absorption maximum and *provide alone almost all the shielding power of the protein.*

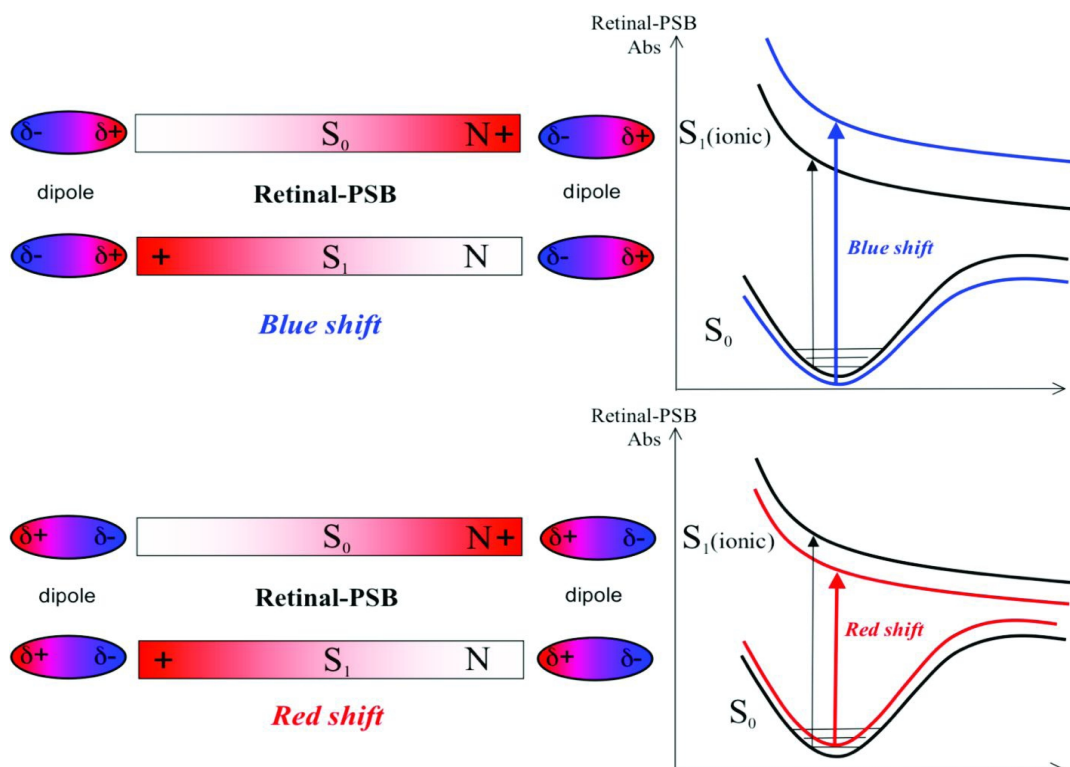


**Figure 7.3.** Reverse Finger Print (RFP) analysis of the 14 rhodopsin residues comprised within 2.5Å-3.5Å radius around the chromophore, the counterion, and other residues recognized relevant for spectral tuning in the visual pigments. All the reported values refer to CASPT2 corrected energies.

Figure 7.4 shows the position of these residues around the chromophore and the associated charges (in a colors code: from blue/negative to red/positive), and it provides a rationale for their effects. It is apparent that all the shielding residues display a significant positive (negative) charge close to the N-head (C-tail) of the chromophore that therefore produces a red shift in the absorption by destabilizing the ground state and stabilizing the excited state (the other way round occurs for blue shifting residues), see Scheme 7.2. Thus, the protein is designed here to produce a net effect that partially shields the primary Glu-113 counterion.



**Figure 7.4.** Displacement around the chromophore of the residues mostly relevant for counterion shielding and corresponding charges analysis. The colors spectrum from blue to red displays negative to positive partial charges on the atoms, respectively. Bar blue/red diagrams display the blue/red shift effect for each residue as shown also in Figure 7.3.



**Scheme 7.2.** Blue/red shift effect on the absorption of the chromophore due to the nearby dipoles of the residues. Positive/negative charges are displayed in red/blue, while the linear polyenic chain of the chromophore is represented as a horizontal bar.

Enough fairly, it must be said that other surrounding polar groups (that are not comprised in the residues analyzed above) affect spectral properties, as they lead to a significant red/blue shift absorption as well. For instance, Glu-181, Glu-122, Tyr-192, Tyr-268, 292, 90, 261, 265, 269 have been already recognized important in dictating/affecting spectral tuning and optical properties. Specifically, Glu-181 and Glu-122 are charged/highly-polar nearby amino acids; Tyr-192 and 268 are very close to the chromophore<sup>29</sup> and intervene in the H-bonded network;<sup>64</sup> amino acid change at site 292 in the C-terminus (transmembrane VII) causes a variable level of  $\lambda_{\max}$ -absorption shift in different visual pigments<sup>65</sup> (it has been recently found that far distant amino acids are also able to modulate the absorption maximum of rhodopsins too);<sup>66</sup> sites 90, 261, 265 and 269 correspond to “key residues” for spectral tuning in color vision (i.e. cone pigments).<sup>67,68</sup> Although some of these residues produce a significant change in the absorption energy (see again Figure 7.3 for a RFP analysis of these groups), anyway their net electrostatic effects is negligible: *it is only for the residues in the 2.5Å-3.5Å shell that electrostatic effects add-up in a constructive manner, so that the primary counterion get shielded.* This is the first time such an effect has been recognized and experimentalists are now encouraged to target these residues in their studies (e.g. by



mutation experiments) to prove and quantify their effects. More important, mutations involving residues that are crucial for spectral modulation and shielding efficiency could be responsible for vision pathologies related to Rh malfunctions, which is a further stimulus for experimental efforts in this direction. This issue will be discussed and studied in depth in section 7.4 below.

In conclusion, we predict for Rh a situation of partial (ca. 50%) quenching of the 113 counterion, that falls right on between the two extreme scenarios presented by Buss/Morokuma/Nakatsuji<sup>45, 47, 48</sup> (no quenching, i.e. Rh absorbs as the ion-pair system) and Olivucci<sup>31</sup> (complete shielding, i.e. Rh absorbs as the distorted retinal in vacuo). This appears in line with recent QM/MM studies by Bravaya et al.<sup>46</sup>

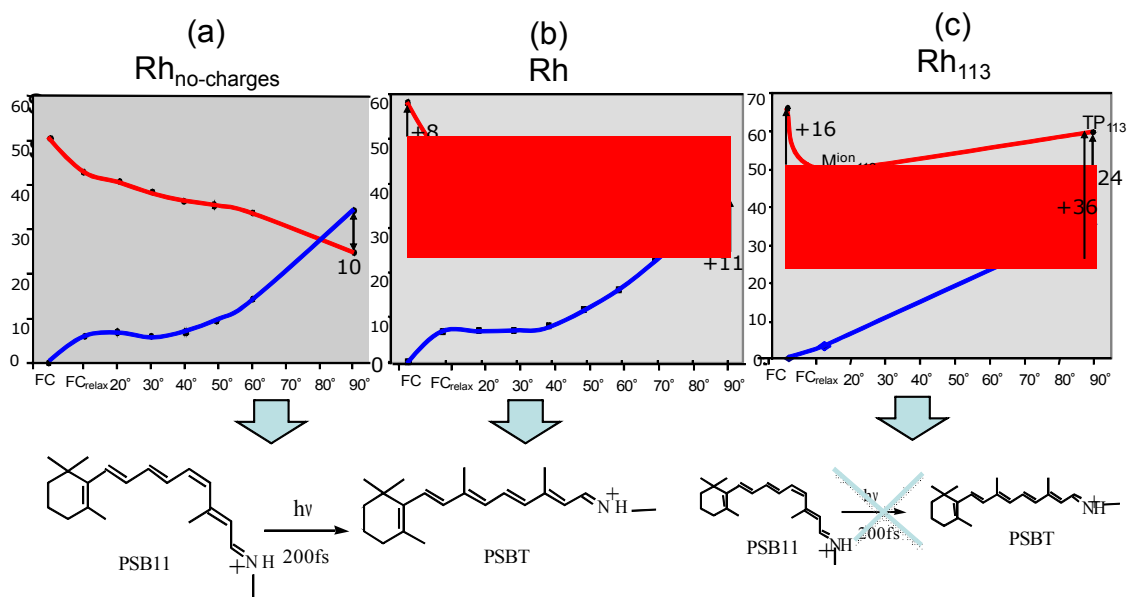
#### Photoreactivity effects.

**Tuning the slope of the photoisomerization channel.** The counterion dependent destabilization of the  $S_1$  energy that has been recognized at the FC point is foreseen all along the points of the  $S_1$  path. Moreover, we expect this effect not to be identical, but to increase steadily along the path as the charge transfer character of the  $S_1$  ( $B_u$ -like) state is increasing upon rotation,<sup>15</sup> reaching a maximum at the twisted funnel point where a net charge transfer occurs (see bar diagrams in Figure 7.2): this is the point where external electrostatic effects on the  $S_1$  energy are expected to be the higher. Thus, it is apparent that the electrostatic effects can also play a role in the photoisomerization process, as the slope of the channel and the location/existence of the twisted conical intersection funnel and crossing seam can change. To quantify these effects the energy of the MEP computed in Rh has been reevaluated in two different electrostatic conditions, i.e. into a protein environment where all external charges have been switched off ( $Rh_{no-charges}$ ) and where only the counterion charge has been explicitly considered ( $Rh_{113}$ ). So doing, all protein steric constraints have been preserved, although electrostatic effects have been zeroed ( $Rh_{no-charges}$ ) or limited only to Glu-113 ( $Rh_{113}$ ), respectively. Thus, while the former reproduces the scenario for a fully quenched counterion, the latter displays the limiting situation of a fully unquenched counterion.

Figure 7.5 summarizes these results by comparing the path and key points calculated for  $Rh_{no-charges}$  (see Figure 7.5a), Rh (see Figure 7.5b) and  $Rh_{113}$  (see Figure 7.5c), while Table 7.3 reports the corresponding energies. The shielding effect that the protein dielectric exercises on the counterion (that results into a redshift in the  $S_1$  energy with respect to the ion-pair) weights ca. 7 kcal/mol at the FC point that, as shown in the previous section, is ca 50% of the overall counterion blue-shift effect at FC (that ranges from 14.4 to 15.4 kcal/mol depending on the calculation level adopted, see Scheme 7.1 and Figure 7.5c). On the other hand, the overall counterion blue-shift effect is estimated to be ca. 34 kcal/mol at the twisted structure (see Table 7.3 and Figure 7.5c). Notably, this is roughly double the value calculated at the FC point as we have a net (half) charge transfer at the CI (FC). As anticipated above, this is the trend expected for  $S_1$ : counterion destabilization on  $S_1$  energies increases upon increasing the charge transfer character of  $S_1$ , i.e. upon rotation. Calculations show (see Table 7.3 and Figures 7.5b and 7.5c) that counterion shielding in Rh recovers ca. 23 kcal/mol at the  $CI_{Rh}$ .

**Figure 7. 5.** The  $S_1$  and  $S_0$  energy profiles (kcal/mol) along the  $S_1$  photoisomerization MEP computed in Rh (b) have been reevaluated in the protein without charges (a) or only with the counterion charge (c). The profile displayed in (a) is also reported as a shaded line in (b) and (c). CASPT2 energies are reported in (a) and (b), while MS-CASPT2 values are reported in (c).

## Mechanicistic scheme: Comparison of different S<sub>1</sub> MEP



**Table 7.3.** CASPT2 relative energies ( $\Delta E$ , kcal/mol) for the lowest three singlet states, oscillator strengths for ground state excitations ( $f_{S_0 \rightarrow S_n}$ ) and wavefunction characterization at the Franck-Condon (FC) and twisted structures of the rhodopsin photoisomerization path computed in the protein with three different electrostatic environments: no charges (Rh<sub>no-charges</sub>), all charges (Rh), only the counterion charge (Rh<sub>113</sub>).

Environment	Structure	Chromophore State (electronic nature)	$\Delta E$	$f_{S_0 \rightarrow S_n}$	Wavefunction <sup>a</sup>
Rh <sub>no-charges</sub>	FC point	S <sub>0</sub>	0.0	0.00	(6a) <sup>2</sup> (7a) <sup>0</sup> (0.78)
		S <sub>1</sub>	50.3	0.78	(6a) <sup>1</sup> (7a) <sup>1</sup> (0.65)
		S <sub>2</sub>	79.3	0.23	(6a) <sup>0</sup> (7a) <sup>2</sup> (0.15) (6a) <sup>1</sup> (7a) <sup>1</sup> (0.13)
	Twisted structure (twisted PSB11 found in Cl <sub>Rh</sub> )	S <sub>0</sub>	23.1	0.00	(6a) <sup>2</sup> (7a) <sup>0</sup> (0.81)
		S <sub>1</sub>	35.1	0.003562	(6a) <sup>1</sup> (7a) <sup>1</sup> (0.74)
		S <sub>2</sub>	86.9	0.000623	(6a) <sup>1</sup> (8a) <sup>1</sup> (0.30) (5a) <sup>1</sup> (6a) <sup>1</sup> (8a) <sup>2</sup> (0.74)
Rh	FC point	S <sub>0</sub>	0.0	0.00	(6a) <sup>2</sup> (7a) <sup>0</sup> (0.77)
		S <sub>1</sub>	57.6	0.74	(6a) <sup>1</sup> (7a) <sup>1</sup> (0.63)
		S <sub>2</sub>	82.0	0.27	(6a) <sup>0</sup> (7a) <sup>2</sup> (0.11) (6a) <sup>1</sup> (7a) <sup>1</sup> (0.19)
	Twisted structure (Cl <sub>Rh</sub> )	S <sub>0</sub>	32.5	0.00	(6a) <sup>1</sup> (7a) <sup>1</sup> (0.73)
		S <sub>1</sub>	34.5	0.000013 16	(6a) <sup>2</sup> (7a) <sup>0</sup> (0.82)
		S <sub>2</sub>	84.8	0.000008 75	(6a) <sup>1</sup> (8a) <sup>1</sup> (0.31) (5a) <sup>1</sup> (6a) <sup>1</sup> (0.34)
		S <sub>0</sub>	0.0 <sup>b</sup> (0.0)	0.0	(6a) <sup>2</sup> (7a) <sup>0</sup> (0.75)

Rh <sub>113</sub>	FC point	S <sub>1</sub>	65.7 <sup>b</sup> (75.7)	0.13566	(6a) <sup>1</sup> (7a) <sup>1</sup> (0.54)
		S <sub>2</sub>	95.2 <sup>b</sup> (74.7)	0.74604	(6a) <sup>0</sup> (7a) <sup>2</sup> (0.23) (6a) <sup>1</sup> (7a) <sup>1</sup> (0.18)
	Twisted structure (TP <sub>113</sub> )	S <sub>0</sub>	35.7 <sup>b</sup> (32.6)	0.000127	(6a) <sup>1</sup> (7a) <sup>1</sup> (0.70)
		S <sub>1</sub>	59.8 <sup>b</sup> (57.1)	0.000578	(6a) <sup>2</sup> (7a) <sup>0</sup> (0.79)
		S <sub>2</sub>	88.3 <sup>b</sup> (84.7)	0.03199	(5a) <sup>1</sup> (6a) <sup>1</sup> (7a) <sup>2</sup> (0.33) (6a) <sup>1</sup> (8a) <sup>1</sup> (0.12) (6a) <sup>1</sup> (9a) <sup>1</sup> (0.15)
	M <sup>ion</sup> <sub>113</sub>	S <sub>0</sub>	4.3 <sup>b</sup>	0.00	(6a) <sup>2</sup> (7a) <sup>0</sup> (0.64)
		S <sub>1</sub>	51.2 <sup>b</sup>	0.7182	(6a) <sup>1</sup> (7a) <sup>1</sup> (0.53)
		S <sub>2</sub>	78.4 <sup>b</sup>	0.006818	(6a) <sup>0</sup> (7a) <sup>2</sup> (0.21) (6a) <sup>1</sup> (7a) <sup>1</sup> (0.14)

<sup>a</sup>6a and 7a represent the delocalized HOMO and LUMO  $\pi$  molecular orbitals at the FC point, respectively. At the twisted structure  $\pi$  orbitals localize on each semi-fragment, with 6a that is localized on the N containing half, and 7a that is localized on the  $\beta$ -ionone containing half. Consequently, at the twisted structure the closed shell (6a)<sup>2</sup> (7a)<sup>0</sup> configuration describes the charge transfer state (that is diabatically connected to S<sub>1</sub>/(6a)<sup>1</sup> (7a)<sup>1</sup> in the FC point), while the open shell (6a)<sup>1</sup> (7a)<sup>1</sup> configuration describes a diradical state (that is diabatically connected to S<sub>0</sub>/(6a)<sup>2</sup> (7a)<sup>0</sup> in the FC point).

<sup>b</sup>Multi-State CASPT2 results: corresponding CASPT2 values are in brackets, when reported.

Another consequence of these effects is that the slope of the photoisomerization channel gets dramatically decreased in Rh<sub>113</sub> (as it increases the destabilization of the counterion along the path, see Figure 7.5c). Eventually, a barrier emerges along the path and an excited state intermediate appears (M<sup>ion</sup><sub>113</sub>) preventing an efficient twisting. More interesting, the CI funnel disappears and a twisted point (TP<sub>113</sub>) with a large S<sub>1</sub>/S<sub>0</sub> energy gap separation (ca. 21 kcal/mol, see Figure 7.5c) replaces it. Thus, it is apparent that the crossing seam has been removed (or displaced to much higher energies, that are not of (photo)chemical relevance). To further investigate this issue and unambiguously define this mechanistic picture, we have fully optimized (into the protein and in the field of the counterion charges alone, i.e. Rh<sub>113</sub>) both the excited state intermediate (M<sup>ion</sup><sub>113</sub>) and the twisted structure of the chromophore (TP<sub>113</sub>). These results (that are shown in the supporting information) confirm the mechanistic picture shown in Figure 7.5c. It is apparent that the 21 kcal/mol S<sub>1</sub>/S<sub>0</sub> energy gap found at TP<sub>113</sub> prevents any efficient (and ultrafast) internal conversion process in Rh<sub>113</sub>. Interestingly, this energy gap is the *minimum* energy that the protein dielectric must recover (by shielding the counterion) in order to generate a conical intersection funnel and trigger an efficient (i.e. ultrafast) photoisomerization process. Although this quenching does not appear full, anyway it is significant and is estimated to be ca. 60% (i.e. 21 kcal/mol) of the overall counterion blue-shift effect at the twisted structure (i.e. 34 kcal/mol): *this is the least shielding that we need in order to have photoactive visual pigments*. Very remarkably, this is almost identical to the shielding that is estimated in Rh at the twisted point (23 kcal/mol) and that is needed to recover S<sub>1</sub>/S<sub>0</sub> degeneration and generate a twisted conical intersection.

It can be concluded that *the quenching of the counterion by the protein (by at least a 60% shielding factor in the twisted structure) is a prerequisite to have an efficient and ultrafast photoisomerization in visual pigments*: it recovers a barrierless path and a twisted CI funnel. Thus, besides providing a tool for spectral tuning that regulates absorption energies, counterion quenching is also mechanistically fundamental

since it promotes photoisomerization catalysis that would be otherwise depressed if the counterion were unshielded.

Interestingly, a consequence of the above reasoning is that if we produced a mutated Rh with an altered (i.e. smaller) counterion shielding power, the photoreactivity of the pigment would be impaired and fluorescence (e.g. from the  $M^{ion}_{113}$  intermediate) may become a competitive process. Experimental investigations in this direction (by site directed mutations of those residues recognized here as responsible of the shielding efficiency such as Ser-186, Thr-118, Ala-117, Hie-211, Tyr-191, Tyr-43, Tyr-265) would be welcome as this would provide the way to switch from a photoisomerizable to a fluorescent device.

**Spectral tuning mechanism in color vision.** Notably, the 435 to 570 nm range of  $S_1$  excitation wavelengths spanned by the vacuo(PSB11<sub>dist</sub>)/ion-pair(PSB11<sub>dist</sub>-113E) systems covers the appropriate range for the color pigments S (425nm), M (530 nm), and L (560 nm), see Scheme 7.1. Thus, such a spectral modulation could be achieved in principle by tuning the shielding effect of the 113 counterion from a fully quenched (vacuo-like) system (that reproduces the red absorption of the L-cone pigment), to the unquenched ion-pair (that absorbs in the blue as the S-cone pigment), through an intermediate state absorbing in the green (M-cone pigment) where the counterion is only partially quenched. As a result, a tunable shielding of the counterion could appear as a suitable tool for regulating color absorptions. Anyway, under the light of the mechanistic results presented above, it is apparent that *the absorption of the blue (S-cone) pigment cannot be achieved by a fully unquenched 113 residue*: this would produce the correct blue absorption found in the ion-pair, but it would lead to an inefficient photoreaction. In fact, we have proved that the counterion must be significantly shielded to have a photoactive pigment: at least a 60% shielding factor must be operative at the twisted structure. This means that also other mechanisms must cooperate to recover the ca. 10 kcal/mol needed to reach the high absorption energy of the blue pigment (from 57 kcal/mol in Rh to 67 kcal/mol in S), as the counterion shielding cannot be further reduced. We have suggested in previous works<sup>69</sup> that the  $\beta$ -ionone handle may be also used as a regulatory tool for absorption energies. In fact, a fully twisted  $\beta$ -ionone ring should produce a blue-shifted retinal absorption, as the chromophore conjugation has been reduced. Interestingly, recent experimental investigations on a Rh with an acyclic retinal analog (i.e. lacking the  $\beta$ -ionone ring) show that this absorbs at 460 nm,<sup>70</sup> corresponding to a ca. 5 kcal/mol blue shift of the Rh wild-type absorption. This is also the shift that is expected in a chromophore with a fully deconjugated  $\beta$ -ionone ring. Anyway, this is only half the energy that we need to recover S-pigment absorption. Thus, it is apparent that this effect is not enough alone, and that a further deconjugation of the chromophore (triggered by steric interactions with the protein pocket) is necessary. That is achieved for example by twisting about one (or more) of its internal double bonds (such as the  $C_8$ - $C_9$  single bond), thus leading to a higher energy absorbing chromophore. Very remarkably, recent experimental works on a Rh containing a retinal analog with only 4 conjugated double bonds (i.e. a 7,8-dhydro retinal has been used)<sup>71</sup> show that this chromophore already absorbs at blue energies (435 nm), as found in the S-cone pigment. Thus, if the protein binding pocket of this pigment is designed to force the retinal chromophore into a conformation with a highly distorted structure (e.g. fully twisted about the  $C_8$ - $C_9$  single bond), a blue absorption should be already observed without the need of other effects. On the other hand, a complete (partial) shielding of the counterion by the protein dielectric appears

as the most viable way to produce alone the lower energy absorbing L (M) pigment: this would preserve the photoactivity of the pigment itself as counterion shielding would be higher than 60% anyway. Scheme 7.1 summarizes these results.

Interestingly, this reasoning provides an explanation for the similarity in the structure of the L and M opsins (they share over 96% amino acid sequence identity) and the different structure observed in the S pigment (only 43% identity with L and M),<sup>72,73</sup> as here the binding pocket must be differently designed so that a highly twisted chromophore is hosted (which causes the very different sequence identity). Additionally, this spectral tuning mechanism validates the tentative structures for the L, M and S pigments that have been very recently generated using homology modeling techniques:<sup>74</sup> while the chromophore structure appears very similar in both the L and M pigments, its conformational twisting plays a major role in the S-pigment as retinal appears much more distorted in this environment.

## [7.4 Mutations and Vision Deficiencies.](#)

### [7.4.1 Rh related vision deficiencies.](#)

It has been shown above that several Rh residues appear as major players in spectral modulation (by a regulatory mechanism of the shielding effect) and their effect has been quantified in Table 7.2 and Figure 7.3. It is likely that natural mutations involving those groups may compromise more than others the optical properties and photoactivity of the photoreceptor, and may be the source of vision pathologies related to Rh malfunctions. More specifically, we suggest here that *many of the mutations impairing the photoactivity of Rh do that by reducing its counterion shielding efficiency*: indeed we have seen that at least a 60% shielding power must be operative at the twisted structure to preserve the photoefficiency of the pigment. If it is reduced below that threshold, an inefficient photoisomerization that impairs Rh photoactivity should result.

More than 40 natural mutations have been recognized until now to cause retinitis pigmentosa 4 (in the genetic classification of retinitis pigmentosa, the form that is due to mutations in the rhodopsin gene is referred to as retinitis pigmentosa 4).<sup>75</sup> If we exclude those mutations that inactivate Rh by disrupting its structure and stability (or that of its photocycle intermediates) or its ability to bind G-protein, it is conceivable that the remaining ones harm Rh because they prejudice its counterion shielding power. Although the systematic computational analysis of those genetic changes is out of the scope of this work, the collected data on Rh make this hypothesis computationally sound. To the best of our knowledge, this is new paradigm in Rh related vision pathologies that experimentalists are now encouraged to investigate and address to (e.g. by novel mutation experiments). If positive, these experimental proofs would provide a validation of the above model as well as open a new scenario in the rationale of night blindness.

### [7.4.2 Color pigments related pathologies.](#)

It is worth noting that mutations related to color blindness in humans are known to be mostly located in the red (L) and green (M) cone-pigments, while the blue (S) pigment is much less affected.<sup>73</sup> We suggest this is also a clue for a different color tuning mechanism in the formers with respect to the latter, which is based on a tunable counterion quenching and a chromophore conformational control, respectively (as shown in section 7.3 and Scheme 7.1 above). The first is by definition very much

sensitive to changes in the electrostatic environment (such as the ones due to mutations) that could alter the optical response of the pigment and can involve potentially all the surrounding polar residues that are found to be important in spectral tuning. We have seen here (see Figure 7.3) that many residues may play such a role in Rh and all appears as good candidates for mutations leading to color vision deficiencies: less than 3 kcal/mol separates the absorption energy of the L and M pigments, and most of these residues produce spectral modulations falling in the 1-3 kcal/mol range. On the other hand, a much larger energy gap (ca. 13 kcal/mol) separates the green (M) from the blue (S) pigment absorptions and, once the retinal is accommodated into the protein pocket, it is much harder for a single point mutation to alter its optical properties and photoactivity efficiency so strongly, as these are mostly regulated by the steric dependent conformational twisting of the chromophore. It is worth to stress here that this reasoning focuses on those residues and mutations that are involved in spectral tuning only, while it neglects all those genetic changes that inactivate the photopigment by disrupting its structure or stability (which may be another source of vision deficiency).

### 7.5. Conclusions

We have provided computational evidence that electrostatic effects do strongly control both spectral and photoreactivity properties in visual pigments, and that the two concepts are strictly related to each other: a unified model has been produced that discloses the key role played by counterion quenching in assisting retinal photoisomerization and setting up its optical features. Based on that, a sound mechanism for spectral tuning in color vision pigments emerges that is also validated by recent experimental and molecular dynamics investigations, and that provides a rationale for the different structures found in the M/L pigments as compared to the S one. These results suggest a new paradigm in vision deficiencies related to Rh mutations, as night blindness may also result by an impaired counterion shielding efficiency in the mutated pigment. It is also shown that the easiness in M/L malfunctions with respect to the ones found in S can be easily accounted for by considering the different mechanism for color modulation found in those pigments.

In conclusion, these results contribute to the understanding of the factors controlling spectral modulation and photoisomerization catalysis in rhodopsins, and suggest that vision pathologies may be also produced by mutations that harm the counterion shielding power of the visual pigment. To the best of our knowledge, this is a novel concept in vision deficiencies that experimentalists are now encouraged to investigate and address. Finally, it has been pointed out that an intriguing consequence of the presented results is that it would be possible in principle to design modified (i.e. mutated) rhodopsins with a smaller counterion shielding power that can become fluorescent. That is, playing with the protein electrostatic field would be the way to switch from photoisomerizable devices to fluorescent dyes, and the rationale for such a molecular design has been provided here.

Details of the QM/MM scheme adopted in this chapter (section A.1 and Figure A.1), validation of MS-CASPT2 results (section A.2, Tables A.1-A.6 and Figure A.2), complete Refs. 51 and 55 (section A.3), energies and coordinates of all the structures discussed in the paper (section A.4) are given in the Appendix A.

## References:

- (1) Kandori, H.; Shichida, Y.; Yoshizawa, T. *Biochemistry-Moscow* 2001, 66, 1197.
- (2) Needleman, R. Bacteriorhodopsin and Rhodopsin. In *CRC Handbook of Organic Photochemistry and Photobiology*; ed.; Horspool, W. M., Song, P.-S., Eds.; CRC Press: Boca Raton, FL, 1995; Vol. pp 1508-1515.
- (3) Ottolenghi, M.; Sheves, M. *Isr. J. Chem.* 1995, 35, U3.
- (4) Wald, G. *Science* 1968, 162, 230.
- (5) Mathies, R.; Lugtenburg, J. The Primary Photoreaction of Rhodopsin. In *Molecular Mechanism of Vision*; Stavenga, D. G., DeGrip, W. J., Pugh, E. N. J., Eds.; Elsevier Science Press: New York, 2000; Vol. 3, pp 55-90.
- (6) Yoshizawa, T.; Kuwata, O. Vision: Photochemistry. In *CRC Handbook of Organic Photochemistry and Photobiology*; ed.; Horspool, W. M., Song, P.-S., Eds.; Boca Raton, FL, 1995; Vol. pp 1493-1499.
- (7) Wald, G.; Brown, P. K. *Science* 1958, 127, 222.
- (8) McBee, J. K.; Palczewski, K.; Baehr, W.; Pepperberg, D. R. Confronting Complexity: The Interlink of Phototransduction and Retinoid Metabolism in Vertebrate Retina. In *Progress in Retinal and Eye Research*; Eds.; Elsevier: Great Britain, 2001; Vol. 20, pp 469.
- (9) Garavelli, M.; Bernardi, F.; Olivucci, M.; Vreven, T.; Klein, S.; Celani, P.; Robb, M. A. *Faraday Discuss.* 1998, 110, 51.
- (10) González-Luque, R.; Garavelli, M.; Bernardi, F.; Merchán, M.; Robb, M. A.; Olivucci, M. *Proc. Natl. Acad. Sci. U. S. A.* 2000, 97, 9379.
- (11) Cembran, A.; Bernardi, F.; Olivucci, M.; Garavelli, M. *J. Am. Chem. Soc.* 2003, 125, 12509.
- (12) De Vico, L.; Page, C. S.; Garavelli, M.; Bernardi, F.; Basosi, R.; Olivucci, M. *J. Am. Chem. Soc.* 2002, 124, 4124.
- (13) Garavelli, M.; Vreven, T.; Celani, P.; Bernardi, F.; Robb, M. A.; Olivucci, M. *J. Am. Chem. Soc.* 1998, 120, 1285.
- (14) Garavelli, M.; Celani, P.; Bernardi, F.; Robb, M. A.; Olivucci, M. *J. Am. Chem. Soc.* 1997, 119, 6891.
- (15) Cembran, A.; Gonzalez-Luque, R.; Serrano-Andres, L.; Merchan, M.; Garavelli, M. *Theor. Chem. Acc.* 2007, 118, 173.
- (16) Turro, N. J. *Modern Molecular Photochemistry*; Benjamin-Cummings: Menlo Park, CA, 1991; Vol. p.
- (17) Gilbert, A.; Baggott, J. *Essentials of Molecular Photochemistry*; Blackwell Science: Oxford, 1991; Vol. p.
- (18) Cembran, A.; Bernardi, F.; Olivucci, M.; Garavelli, M. *J. Am. Chem. Soc.* 2004, 126, 16018.
- (19) Nathans, J.; Piantanida, T. P.; Eddy, R. L.; Shows, T. B.; Hogness, D. S. *Science* 1986, 232, 203.
- (20) Kochendoerfer, G. G.; Lin, S. W.; Sakmar, T. P.; Mathies, R. A. *Trends Biochem. Sci.* 1999, 24, 300.
- (21) Hamm, P.; Zurek, M.; Roschinger, T.; Patzelt, H.; Oesterhelt, D.; Zinth, W. *Chem. Phys. Lett.* 1996, 263, 613.
- (22) Kandori, H.; Sasabe, H.; Nakanishi, K.; Yoshizawa, T.; Mizukami, T.; Shichida, Y. *J. Am. Chem. Soc.* 1996, 118, 1002.

- (23) Logunov, S. L.; Song, L.; ElSayed, M. A. *J. Phys. Chem.* 1996, *100*, 18586.
- (24) Kandori, H.; Katsuta, Y.; Ito, M.; Sasabe, H. *J. Am. Chem. Soc.* 1995, *117*, 2669.
- (25) Freedman, K. A.; Becker, R. S. *J. Am. Chem. Soc.* 1986, *108*, 1245.
- (26) Kochendoerfer, G. G.; Verdegem, P. J. E.; vanderHoef, I.; Lugtenburg, J.; Mathies, R. A. *Biochemistry* 1996, *35*, 16230.
- (27) Sugihara, M.; Hufen, J.; Buss, V. *Biochemistry* 2006, *45*, 801.
- (28) Teller, D. C.; Okada, T.; Behnke, C. A.; Palczewski, K.; Stenkamp, R. E. *Biochemistry* 2001, *40*, 7761.
- (29) Okada, T.; Sugihara, M.; Bondar, A. N.; Elstner, M.; Buss, V. *J. Mol. Biol.* 2004, *342*, 571.
- (30) Cembran, A.; Bernardi, F.; Olivucci, M.; Garavelli, M. *Proc. Natl. Acad. Sci. U. S. A.* 2005, *102*, 6255.
- (31) Andruniów, T.; Ferré, N.; Olivucci, M. *Proc. Natl. Acad. Sci. U. S. A.* 2004, *101*, 17908.
- (32) Migani, A.; Sinicropi, A.; Ferre, N.; Cembran, A.; Garavelli, M.; Olivucci, M. *Faraday Discuss.* 2004, *127*, 179.
- (33) Frutos, L. M.; Andruniow, T.; Santoro, F.; Ferre, N.; Olivucci, M. *Proc. Natl. Acad. Sci. U. S. A.* 2007, *104*, 7764.
- (34) Lin, H.; Truhlar, D. G. *Theor. Chem. Acc.* 2006, *117*, 185.
- (35) Palczewski, K.; Kumasaka, T.; Hori, T.; Behnke, C. A.; Motoshima, H.; Fox, B. A.; Le Trong, I.; Teller, D. C.; Okada, T.; Stenkamp, R. E.; Yamamoto, M.; Miyano, M. *Science* 2000, *289*, 739.
- (36) Birge, R. R.; Murray, P. P.; Pierce, B. M.; Akita, H.; Balogh-Nair, V.; Findsen, L. A.; Nakanishi, K. *Proc. Natl. Acad. Sci. U. S. A.* 1985, *82*, 4117.
- (37) Yan, E. C. Y.; Kazmi, M. A.; De, S.; Chang, B. S. W.; Seibert, C.; Marin, E. P.; Mathies, R. A.; Sakmar, T. P. *Biochemistry* 2002, *41*, 3620.
- (38) Yan, E. C. Y.; Kazmi, M. A.; Ganim, Z.; Hou, J. M.; Pan, D. H.; Chang, B. S. W.; Sakmar, T. P.; Mathies, R. A. *Proc. Natl. Acad. Sci. U. S. A.* 2003, *100*, 9262.
- (39) Lewis, J. W.; Szundi, I.; Kazmi, M. A.; Sakmar, T. P.; Klinger, D. S. *Biochemistry* 2004, *43*, 12614.
- (40) Terakita, A.; Yamashita, T.; Shichida, Y. *Proc. Natl. Acad. Sci. U. S. A.* 2000, *97*, 14263.
- (41) Ludeke, S.; Beck, R.; Yan, E. C. Y.; Sakmar, T. P.; Siebert, F.; Vogel, R. *J. Mol. Biol.* 2005, *353*, 345.
- (42) Rohrig, U. F.; Guidoni, L.; Rothlisberger, U. *Biophys. J.* 2002, *82*, 223A.
- (43) Ebrey, T. G.; Honig, B. *Proc. Natl. Acad. Sci. U. S. A.* 1972, *69*, 1897.
- (44) Nielsen, I. B.; Lammich, L.; Andersen, L. H. *Phys. Rev. Lett.* 2006, *96*, 018304.
- (45) Sekharan, S.; Sugihara, M.; Buss, V. *Angew. Chem., Int. Ed.* 2007, *46*, 269.
- (46) Bravaya, K.; Bochenkova, A.; Granovsky, A.; Nemulkhin, A. *J. Am. Chem. Soc.* 2007, *129*, 13035.
- (47) Fujimoto, K.; Hayashi, S.; Hasegawa, J.; Nakatsuji, H. *Journal of Chemical Theory and Computation* 2007, *3*, 605.
- (48) Altun, A.; Yokoyama, S.; Morokuma, K. *J. Phys. Chem. B* 2008.
- (49) Altoè, P.; Stenta, M.; Bottoni, A.; Garavelli, M. *Theor. Chem. Acc.* 2007, *118*, 219.
- (50) Singh, U. C.; Kollman, P. A. *J. Comput. Chem.* 1986, *7*, 718.
- (51) Frisch, M. J. et al. *Gaussian 03, Revision C.02; Gaussian, Inc., Wallingford CT*, 2004.



- (52) Case, D. A.; Cheatham, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. *J. Comput. Chem.* 2005, 26, 1668.
- (53) Gordon, J. C.; Myers, J. B.; Folta, T.; Shoja, V.; Heath, L. S.; Onufriev, A. *Nucleic Acids Res.* 2005, 33, 368.
- (54) Andersson, K.; Malmqvist, P. A.; Roos, B. O. *J. Chem. Phys.* 1992, 96, 1218.
- (55) Andersson, K. et al. *MOLCAS*, 6.0; Department of Theoretical Chemistry, Chemical Centre, University of Lund: Lund, 2004.
- (56) Malmqvist, P.-Å.; Roos, B. O. *Chem. Phys. Lett.* 1989, 155, 189.
- (57) Terakita, A.; Koyanagi, M.; Tsukamoto, H.; Yamashita, T.; Miyata, T.; Shichida, Y. *Nat. Struct. Mol. Biol.* 2004, 11, 284.
- (58) Kusnetzow, A. K.; Dukkupati, A.; Babu, K. R.; Ramos, L.; Knox, B. E.; Birge, R. R. *Proc. Natl. Acad. Sci. U. S. A.* 2004, 101, 941.
- (59) Ramos, L. S.; Chen, M. H.; Knox, B. E.; Birge, R. R. *Biochemistry* 2007, 46, 5330.
- (60) Warshel, A. *Nature* 1976, 260, 679.
- (61) Warshel, A. *Proc. Natl. Acad. Sci. U. S. A.* 1978, 75, 5250.
- (62) Buss, V.; Kolster, K.; Terstegen, F.; Vahrenhorst, R. *Angew. Chem., Int. Ed.* 1998, 37, 1893.
- (63) Notably, both the reverse (Figure 7.3) and the direct (Table 7.2) analysis do roughly agree for the E113 effect, the difference in the two results being due to the different polarization that the QM environment experiences in the two approaches: in the first we do have the full electrostatic environment of the protein but the residue under investigation, while in the second we have only the electrostatic effects of the investigated residues.
- (64) Yan, E. C. Y.; Ganim, Z.; Kazmi, M. A.; Chang, B. S. W.; Sakmar, T. P.; Mathies, R. A. *Biochemistry* 2004, 43, 10867.
- (65) Fasick, J. I.; Lee, N.; Oprian, D. D. *Biochemistry* 1999, 38, 11593.
- (66) Yoshitsugu, M.; Shibata, M.; Ikeda, D.; Furutani, Y.; Kandori, H. *Angew. Chem., Int. Ed.* 2008, 47, 3923.
- (67) Lin, S. W.; Kochendoerfer, G. G.; Carroll, H. S.; Wang, D.; Mathies, R. A.; Sakmar, T. P. *J. Biol. Chem.* 1998, 273, 24583.
- (68) Coto, P. B.; Strambi, A.; Ferre, N.; Olivucci, M. *Proc. Natl. Acad. Sci. U. S. A.* 2006, 103, 17154.
- (69) Cembran, A.; González-Luque, R.; Altoè, P.; Merchán, M.; Bernardi, F.; Olivucci, M.; Garavelli, M. *J. Phys. Chem. A* 2005, 109, 6597.
- (70) Bartl, F. J.; Fritze, O.; Ritter, E.; Herrmann, R.; Kuksa, V.; Palczewski, K.; Hofmann, K. P.; Ernst, O. P. *J. Mol. Biol.* 2005, 280, 34259.
- (71) DeGrip, W. J.; Bovee-Geurts, P. H. M.; van der Hoef, I.; Lugtenburg, J. *J. Am. Chem. Soc.* 2007, 129, 13265.
- (72) Nathans, J.; Thomas, D.; Hogness, D. S. *Science* 1986, 232, 193.
- (73) Neitz, M.; Neitz, J., Color Vision Defects. In *ENCYCLOPEDIA OF LIFE SCIENCES*, John Wiley & Sons, Ttd: 2005.
- (74) Trabanino, R. J.; Vaidehi, N.; Goddard, W. A. *J. Phys. Chem. B* 2006, 110, 17230.
- (75) McKusick, V. A. et al. RHODOPSIN; RHO. In *OMIM: Online Mendelian Inheritance in Man*, Johns Hopkins University:

## - CHAPTER 8-

# Steric control of photoisomerizations efficiency and spectroscopic properties in Rhodopsin

### 8.1 Introduction

The rhodopsin photoreactivity, as already widely discussed in chapter 7, is one of the most interesting challenges in the chemistry of the vision consists in understanding the mechanism responsible of regulating the absorbance of the visual photoreceptors. In particular, absorption of a photon by rhodopsin initiates the primary photochemical reaction in the chromophore<sup>1-6</sup>, implying a deep conformational change in the protein. Despite all the experimental<sup>7</sup> and computational<sup>8-14</sup> progress in the field, details of the excited state geometrical and electronic evolution of the protein embedded chromophore are still partially unknown: this is where computational photochemistry can give considerable insights. Concerning the mechanism of cis-trans isomerization there is yet an incredible high interest on trying to individuate and rationalize the details of such mechanism and the factors controlling and determining its efficiency. For instance, among the unsolved problems, environmental effects on spectral shifts, photoreaction mechanism for retinal chromophore have been investigated using several techniques.

Recently the crystal structure of bovine rhodopsin has been resolved<sup>15-16</sup>, providing in such way a deeper knowledge of the binding site structure, whose influence on the isomerization mechanism is still unclear. In fact the chromophore, covalently linked to a lysine residue of the protein, may be subject to steric hindrance during the torsional motion occurring while cis-trans isomerization proceeds. It was supposed a crucial role of  $\beta$ -ionone ring<sup>17</sup>, probably characterized by a non negligible stabilizing effect when conjugates significantly with the rest of the conjugated chain. In this chapter the so called “lever effect” has been analyzed: the influence of the large conjugated  $\pi$ -system, extended on the entire chromophore, which probably may affect the  $S_1$  topology, has been quantitatively estimated by comparison with the topology of a modified less conjugated chromophore.

Further transient absorption (pump-probe) experiment have revealed that the photoisomerization in rhodopsin is a vibrationally coherent process. Femtosecond fluorescence spectroscopy directly captured excited state dynamics of rhodopsin, so that

both coherent reaction process and unreacted excited state were observed<sup>18</sup>. In particular probing at near-IR wavelengths allowed detection of stimulated emission, the rise and decay of which was too fast to be resolved with our 70 fs pulses. The fast dynamics of the emission suggest a multidimensional potential energy surface for isomerization, possibly involving different reaction coordinates dominated by carbon double-bond stretching in addition to bond torsion, as previously invoked for model systems by quantum mechanical calculations. Further non-exponential fluorescence dynamics of bovine rhodopsin were measured by Kandori et al. at various wavelengths from 530 to 780 nm. This wide range of wavelength have point out how fluorescence measures depend on the excited state populated, suggesting the second excite state  $S_2$  to be involved in the photon absorption and following events. Remarkable the experiments have revealed the contribution of  $S_2$  to transient absorption and fluorescence, suggesting even *a possible trapping-mechanism for the photoexcited population involving reasonably the  $S_2$  state*, being a trap state. Therefore,  $S_2$  topology has been explored, in order to explain and rationalize the trap state effect, the calculations have located an absorbing second excited state, that featured by a minimum, behaves as a trap and which may be possibly involved in the formation of the unreactive population recently observed.

## 8.2 Computational Methods

Computational methods employed for such problematic are the same of those already described in the previous chapter: for reason they will be here briefly summarized.

Computations have been performed through our recently developed Quantum Mechanics Molecular Mechanics (QM/MM) tool<sup>19</sup>, in a new implementation of a general hybrid approach with a modular structure called COBRAMM (Computations at Bologna Relating Ab-initio and Molecular Mechanic Methods).

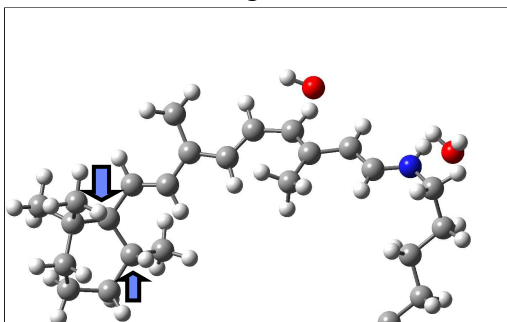
The recent refinement down to 2.2 Å resolution for the bovine Rh crystal structure<sup>16</sup> (1U19.pdb in the protein Data Bank archive) was used as a basis for the protein framework. In particular we selected the A monomer, as for the electrostatic control evaluation.

In order to quantitatively estimate the steric catalytic effect on the  $S_1$  reaction pathway, calculations were performed on two different set-up, differing only for quantum mechanical systems. The comparison was in fact between two different minimum energy path calculated for the 11-cis-retinal (already presented in the previous chapter) and for a modified chromophore, the 5,6 dihydroretinal, where the double bond on the  $\beta$ -ionone ring is saturated. Therefore the protein set-up is the same for both the quantum-mechanical systems: since its absorption spectrum best reproduces the experimental data we have chosen to use the set-up including the GLU181 charged. For locating the  $S_2$  minimum geometry optimization algorithm were performed without searching the minimum energy path leading to equilibrium structure.

The *ab initio* QM calculations are based on a CASPT2//CASSCF approach with a 6-31g\* basis set. The active space comprises the full  $\pi$ -system of the chromophore (i.e. 12 electrons in 12  $\pi$  orbitals for the standard retinal and 10 electrons in 10  $\pi$  orbitals for dihydroretinal). To minimize the influence of weakly interacting intruder states at the second-order level, the so-called imaginary level shift technique was employed<sup>20</sup>. The

CASSCF state interaction (CASSI)<sup>21</sup> method was used to calculate the transition dipole moments. In the expression for the oscillator strength, the CASSCF transition moment and the energy difference obtained in the CASPT2 computation were employed. All calculations were performed with the tools available in the MOLCAS-6<sup>22</sup>, Gaussian03 and the COBRAMM<sup>19</sup> software packages. To account for correlation energy, both CASPT2<sup>23</sup> and MS-CASPT2 single point computations are performed on the optimized points in the bath of the protein (AMBER) point charges, by using the MOLCAS-6.0 package. Concerning the transient absorption calculations computed from such minimum or from local minima located all along the S<sub>1</sub> reaction pathway, a ten-root state average wavefunction is employed, correcting the energy eigenvalues obtained with both the correlated methods, CASPT2 and MSCASPT2 methods. To minimize the influence of weakly interacting intruder states at the second-order level, the so-called imaginary level shift technique is used (imaginary shift value=0.2).

The protein framework is kept fixed during all QM/MM geometry optimizations (we assumed that it has no time to change during the ultrafast (200 fs) timescale of the primary photo-induced event), while QM chromophore atoms and MM atoms comprising the Lys-296 side chain (C<sub>8</sub>H<sub>2</sub>, C<sub>7</sub>H<sub>2</sub>, C<sub>β</sub>H<sub>2</sub>) and the two water molecules (W1 and W2) are left free to relax. A S<sub>0</sub>-S<sub>n</sub> n-root state average CASSCF wavefunction (equally weighting all the n roots) is always used for geometry optimizations on the S<sub>n</sub> state, while a single-state wavefunction is used for ground state optimizations. A



relaxed scan (with the central dihedral angle fixed at specified values) is used to trace the photoisomerization channel in the protein, while fully unconstrained optimizations are employed to locate all the critical points.

**Figure 8.2** 5,6 dihydroretinal chromophore used to study the  $\beta$ -ionone ring effect on the steric catalysis.

## 8.3 Results and Discussion

### 8.3.1 Steric catalysis

The  $\beta$ -ionone ring effect implying a “lever effect” has been proposed already in previous theoretical works<sup>17</sup>, where concurrent shortening of the single bond (adjacent to the ring double bond) of chromophore along the S<sub>1</sub> reaction path calculated *in vacuo*, have caused the simultaneous planarization of the  $\beta$ -ionone ring with its twisting angle remarkably reducing passing from the FC region to CI. The planarization motion which progressively turns the chromophore from a five to a six conjugated double bond system, may turn out to contribute to the efficient photoisomerization in the protein. In particular, according to the crystallographic structure, the residue 265 is very close to the  $\beta$ -ionone ring (less than 3 Å) and, by twisting the C<sub>6</sub>-C<sub>7</sub> bond (red arrow), the distance between the two groups dramatically shortens leading to steric interactions. Therefore, it is likely that the  $\beta$ -ionone ring acts like a lever upon planarization when the chromophore is surrounded by the protein, forcing the central bond to twist earlier in a way to guarantee an increased distance between the ring and the residue 265 and steric interactions are reduced. Remarkably, this direction is in agreement with the sign of the initial C<sub>11</sub>=C<sub>12</sub> twist which has been previously derived in Rhodopsin. Together

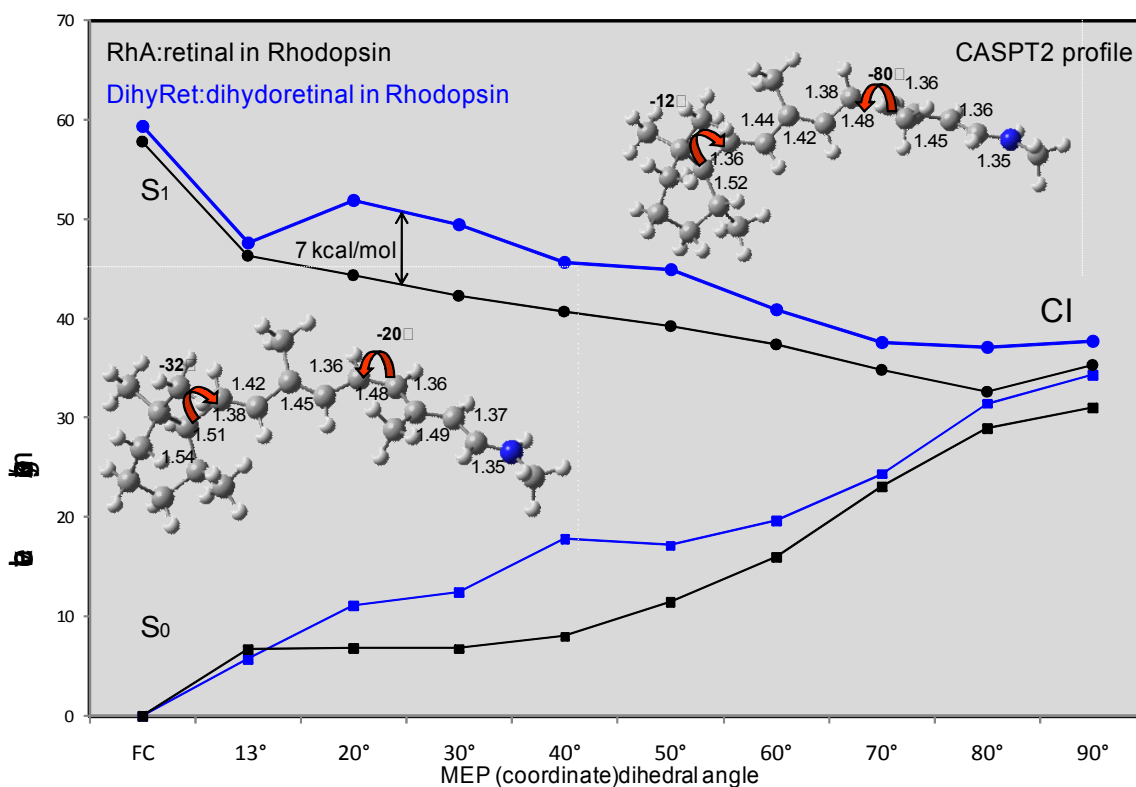
with the electrostatic catalysis and the intramolecular/intermolecular interactions discussed in previous chapter this provides a possible additional reason for the faster photoisomerization in the protein. That is, a lever effect would contribute in tilting the  $S_1$  potential energy surface in such a way that the wavepacket would be more efficiently funneled into the rotational mode. Therefore if it is real the  $\beta$ -ionone ring catalytic effect by switching off it should be noticed a slower isomerization process because of the changed topology of potential energy surface featuring the modified chromophore.

In fact we have estimated quantitatively the  $\beta$ -ionone ring effect by comparing two  $S_1$  isomerization pathway calculated by using two different chromophores. For evaluating the effect we have to choose between two possibility: removing the entire  $\beta$ -ionone ring or keeping it, but hydrogenating the double bond (C5-C6). In order to ensure a more reliable comparison with the real chromophore, the (*trans*)5,6-dihydro-retinal was chosen as system on which evaluating the cis-trans isomerization efficiency. Although Rhodopsin is yet activated, no information about the timescale and efficiency of the ultrafast primary event (i.e. the initial photoisomerization) is provided for such system. Furthermore, Rh activation is not surprising since it primarily depends on the cis  $\rightarrow$  trans isomerization of the central double bond of the chromophore, which in fact occurs also in other media in a longer (ps) timescale, in presence as well as in absence of the  $\beta$ -ionone ring.

Once the ground state, optimization was computed by using the CASSCF(10,10) wavefunction. On such equilibrium structure,  $S_0$ - $S_1$  absorption energy was calculated trough QMMM potential, for including the proteic environment. According with a less conjugated system, the  $\Delta E(S_0-S_1)$  with a value of 62 kcal/mol is blue shifted of few kcal/mol, and it is relatively close to the experimental measure (60 kcal/mol). Despite the hypothetic catalytic lever effect, the experimental evidence for dihydroretinals and for acyclic retinals (without the  $\beta$ -ionone ring) establish the cis-trans isomerization is completed and photoproducts are however collected on the ground state.

In order to directly compare the two reaction pathways for both the models (the classical retinal and dihydro-retinal), in the figure 8.3 there is represented the comparison between the two minimum energy path calculated on the two chromophore starting from Franck Condon region towards  $S_1/S_0$  conical intersection.

These aspects were already investigated through calculations carried out on the isolated chromophore: in this chapter the QMMM method elsewhere presented, was used to the lever effect due to  $\beta$ -ionone ring, and to study the  $S_2$  behavior of retinal surrounded by protein. Since the  $S_0$ - $S_1$  absorption is in an excellent agreement with the experimental measure, when introducing the charged Glu181 in the protein, this set-up is still adopted.



**Figure 8.2** There is represented the comparison between the two S<sub>1</sub> MEP calculated for both the systems: the retinal and dihydoretinal in rhodopsin.

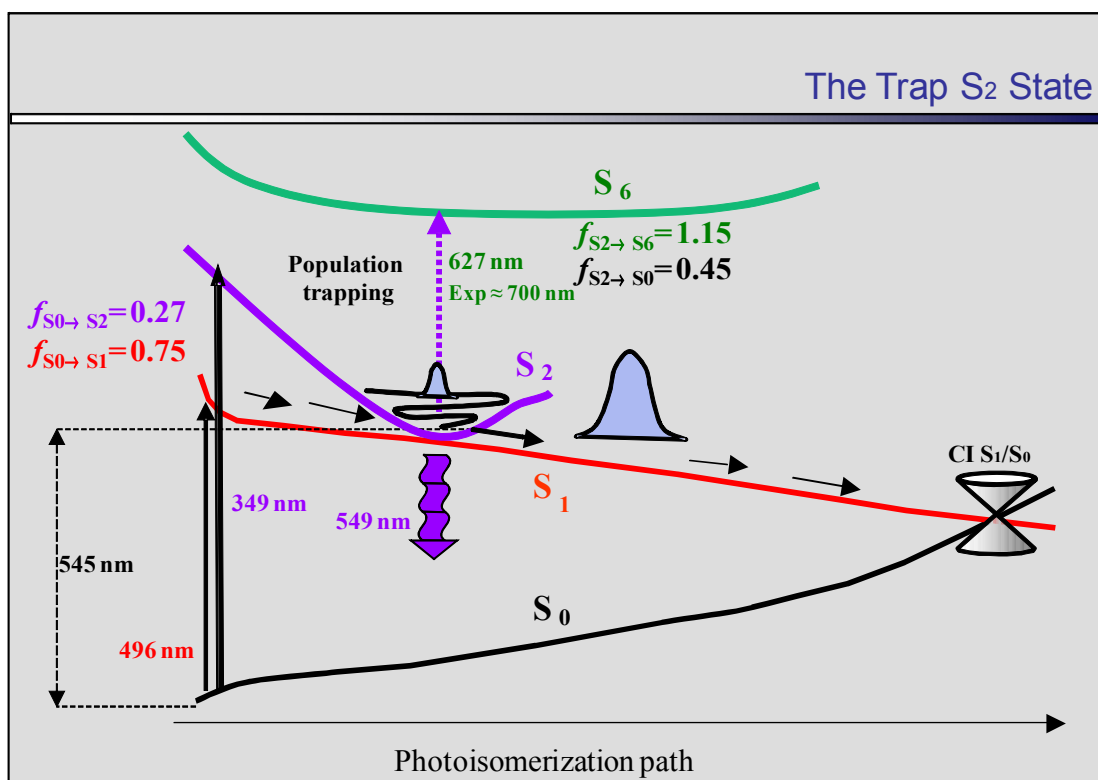
From the figure 8.2 it is clear that on S<sub>1</sub> of dihydoretinal, a sort of small barrier, of about 5 kcal/mol, was located between the relaxed stationary point (13°) and the 20° twisted point. However it is possible to conclude how the S<sub>1</sub> PES for the model is flatter than full retinal, making the isomerization less feasible. On the basis of precedent calculations (comparing the retinal and an acyclic retinal, with lacking completely of  $\beta$ -ionone ring) where the lever effect of the retinal in vacuo was quantitatively evaluated of 7 kcal/mol on S<sub>1</sub> we were expected to find a significant difference on the S<sub>1</sub> reaction path of dihydoretinal. Remarkable S<sub>1</sub> energy profile of dihydoretinal, being a less conjugated system, is less stable than the full system, and through a quantitative evaluation of such difference it was found the  $\Delta E$  between the two profile corresponds to just 7 kcal/mol.

### 8.3.2 The trap S<sub>2</sub> state

Furthermore, a rationale is proposed for the unreactive population recently observed in protein, which is here addressed to the accessible S<sub>2</sub> state, behaving as a trap. The experimental transient absorption and (absorption-wavelength dependent) emission are discussed and interpreted under the light of the results following discussed. In the last decade a huge amount of new experimental techniques have been applied to model new aspects of the Rhodopsin photoreactivity<sup>24,25</sup>. Recently the isomerization reaction of retinal in bovine rhodopsin was studied by femtosecond optical techniques. If the S<sub>1</sub> photoisomerization and its mechanism have been widely investigated and rationalized,

the  $S_2$  vibronic structure of absorption and emissive spectra have been studied and compared with the available experimental data. Among the several experiments there are evidences about an aspect of rhodopsin photoreactivity, consisting of detection of unreactive photo-excited population, suggesting a crucial role of some higher excited state. No dynamic Stokes shift was detected but the fastest part of the fluorescence at both blue and red sides of the spectrum decayed faster than that at the center, which has led to the conclusion that the Franck–Condon state→fluorescence state conversion occurs within time-resolution of apparatus ( $<100$  fs) owing to coupling with intrachromophore high frequency modes and faster initial decay is due to sharpening of band shape caused by a decrease of amplitudes of hfm along the reaction coordinate of twisting. Interestingly a new transient absorption at ca. 700 nm was discovered and attributed to unreactive population of rhodopsin. These experimental evidences suggest that after the electronic transition on the absorbing excited states, the photo-excited population is not directly driven to only the photo-isomerization pathway, but stretching motion extended to the central double bond may be populated, leading the molecule to a different minimum located on a bound excited state, where population should live longer and may contribute to the emission or transient absorption. The following results provide new benchmarks for theoretical evaluation of the isomerization mechanism and dynamics in rhodopsin.

From the  $S_0$  equilibrium geometry we calculated the vertical excitations on the two low-lying single states  $S_1$  and  $S_2$ , by using single-state and state average computations for both set-up. Higher-lying single states ( $S_3$ – $S_9$ ) have also been investigated as possible contributors to the spectroscopic properties as the experimental  $S_2$  transient absorption data.



**Figure 8.3** This is a schematic outline of S<sub>2</sub> trap effect, due to the population splitting S<sub>2</sub>/S<sub>1</sub> with the transient absorption S<sub>2</sub>-S<sub>6</sub> and fluorescence S<sub>2</sub>-S<sub>1</sub> energies reported.

Wave function analysis confirms that, at FC point, if on one hand S<sub>1</sub> state (a Bu-like state) is dominated by a single  $\pi \rightarrow \pi^*$  excitation, with a ionic character, while S<sub>2</sub> state, corresponds to an optically forbidden state with a dot-dot covalent character (a 2A<sub>g</sub> state). Nevertheless S<sub>2</sub> is an optically active state and the oscillator strength value of 0.27 for the S<sub>0</sub>-S<sub>2</sub> transition, reflects its electronic transition. The calculated S<sub>0</sub>-S<sub>2</sub> vertical excitation of 82 kcal/mol (CASPT2 level) nicely reproduce the experimental absorption (80 kcal/mol). At this point, we should keep in mind that the computed vertical transition does not have an experimental counterpart. In other words, there are no experimental vertical transitions to be compared with. Thus, in order to perform a correct comparison to the observed band maximum, we have to compare the experimental data with the 0-0 transition. Therefore to ensure that S<sub>2</sub> can be populated by initial irradiation, vibrational relaxation on this state should be computed. No minimum energy path was computed because not yet available the IRC procedure at QMMM level: a SA (3 roots) geometry optimization starting from FC lead directly the system towards a planar minimum located 30 kcal/mol lower in energy. Herein pumping at 500nm makes possible the population of both S<sub>1</sub> and S<sub>2</sub> (the S<sub>0</sub>-S<sub>2</sub>(0-0)) transition and the vertical emission from S<sub>2</sub> are computed to be 560 and 591 nm respectively. The equilibrium structure found in protein is similar to the one in vacuo, characterized by a bond length alternation completely different from the S<sub>1</sub> stationary point.

This is due to the covalent nature of S<sub>2</sub>, that in contrast with the charge-transfer state S<sub>1</sub>, upon excitation to S<sub>2</sub>, the positive charge remains localized into the C-N tail (like S<sub>0</sub>), acquiring a biradical character. The planar minimum located on S<sub>2</sub>, and its bound character confirms the S<sub>2</sub> population to be unreactive and living longer it may contribute to emission and transient absorption. However the high photoisomerization quantum yield, implies a much S<sub>2</sub> smaller population compared to S<sub>1</sub>. Although internal conversion of S<sub>2</sub>-S<sub>1</sub> cannot be ruled out in principle (finally leading to the trans photoproduct), this process should occur in a much longer than the sub-picosecond timescale thus competing with radiative events. The S<sub>2</sub> emission calculated of 545 nm is in agreement with the experimental measurements. Concerning with this, a clear evidence of the S<sub>2</sub> participation to fluorescence, is provided from the  $\lambda_{\max}$  emission dependency from excitation wavelengths, used to pump the molecule, (wavelength values in a 530-780 nm range). The experimental  $\lambda_{\max}$  emission dependency clearly ensure S<sub>2</sub> being involved, since population of S<sub>2</sub> should depend on the excitation wavelength. Interestingly S<sub>2</sub> may contribute to both the fluorescence, as just discussed, and transient absorption. To individuate the higher excited candidate for absorbing from S<sub>2</sub>, additional number of states have been computed at the relaxed minimum. Remarkable the electronic state with the lowest intense transition ( $f_{3,7}=1.15$ ) is S<sub>6</sub> and the transient absorption value of 627 nm is not too close to the experimental data of 700 nm. This is probably due to the CASSCF wavefunction inadequate description, maybe affected from the surrounding effect of protein, leading the lower ionic state to mix with the higher S<sub>2</sub>.

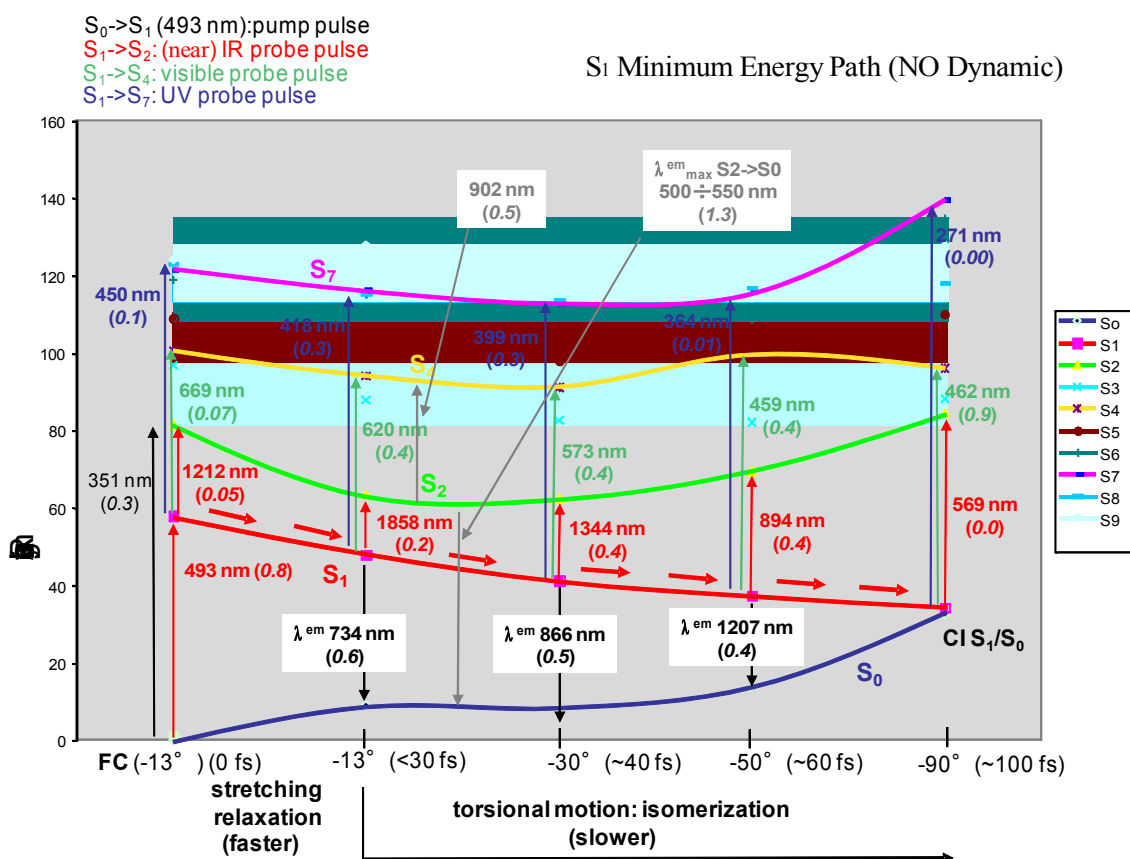
Concerning with this, it is, at this point, worth mentioning, how QMMM calculations reveal that S<sub>2</sub> and S<sub>1</sub> are almost degenerate at the optimized minimum: surprisingly the S<sub>2</sub>-S<sub>1</sub> energy gap depends deeply on the number of states included in the state average



calculations. Therefore through an  $S_2$ - $S_1$  conical intersection the  $S_2$  population may be relax towards the lower  $S_1$ , but this does not support the trap state theory. However as the  $\Delta E(S_2-S_1)$  strongly depends on the state average level, such artificial effect may be due to the protein environment and the eventual polarization which affects the chromophore wavefunction, providing energies quantitatively incorrect. For these reasons, we think our statement is more than a simple suggestion: it is a founded hypothesis where, for the first time, all these evidences can be accounted for in a simple model.

### 8.4 Transient absorption along the $S_1$ photoisomerization

In the previous section we have analyzed and discussed the transient absorption taking place from  $S_2$  minimum. Nevertheless a huge amount of recent experiments have began to consider quite interesting also the transient absorption in the near IR occurring from  $S_1$ . First, as already explained for  $S_2$ , it was necessary to individuate the electronic state associated with the lowest intense transition with  $S_1$ . At  $S_1$  stationary point the most intense transition is the  $S_1$ - $S_9$ . Therefore a series of CASSCF//CASPT2 (12,12) state average calculations including ten higher roots, have been carried out on some of the  $S_1$  optimized scan points along the torsional coordinate defining the  $S_1$  photoisomerization minimum energy path (discussed deeply in the chapter 7).



**Figure.8.4** There are reported all the  $S_1$ - $S_n$  transient absorption and  $S_1$ - $S_0$  the fluorescence energies computed on some of the optimized scans along the  $S_1$  reaction pathway. After  $S_0$ - $S_1$  excitation, starting from FC, the system first relaxes towards a stationary point moving faster along a stretching relaxation coordinate (dihedral angles frozen), following the isomerization proceeds on a barrierless reaction pathway populating the torsional normal mode.

The obtained results reflect the experimental data, in particular the transient absorption of the near IR is well reproduced comparing them with the measures or bovine rhodopsin. According with the experiments are also the  $S_1$ - $S_0$  fluorescence and the torsional degree of the dihedral angle associated with particular  $\lambda_{\max}$  of emission.

State	Franck Condon (PSB11 <sub>dist</sub> )	PSB11 <sub>relax</sub> 13°	30°	50°	90°	$S_2$ Minimum
$S_0$	----	38.94(734) f <sub>1-0</sub> (0.62)	33.01(866) f <sub>1-0</sub> (0.52)	23.69(1206) f <sub>1-0</sub> (0.363)	1.61(17720) f <sub>1-0</sub> (6E-06)	56.87(502) f <sub>2-0</sub> (1.28)
$S_1$	57.96(493) f <sub>0-1</sub> (0.75)	----	----	----	----	9.91(2885) f <sub>2-1</sub> (0.016)
$S_2$	<u>81.54(350)</u> f <sub>0-2</sub> (0.33)	15.38(1858) f <sub>1-2</sub> (0.22)	21.28(1343) f <sub>1-2</sub> (0.36)	31.98(894) f <sub>1-2</sub> (0.45)	50.29(568) f <sub>1-2</sub> (0.003)	----
$S_3$	96.93(294) f <sub>0-3</sub> (0.0303)	40.04(714) f <sub>1-3</sub> (0.069)	41.34(691) f <sub>1-3</sub> (0.132)	44.71(639) f <sub>1-3</sub> (0.24)	<u>53.90(530)</u> f <sub>1-3</sub> (6E-04)	25.76(1109) f <sub>2-1</sub> (0.003)
$S_4$	100.66(184) f <sub>0-4</sub> (0.082)	46.15(619) f <sub>1-4</sub> (0.35)	49.86(573) f <sub>1-4</sub> (0.41)	62.33(458) f <sub>1-4</sub> (0.362)	<u>61.92(461)</u> f <sub>1-0</sub> (0.888)	31.71(901) f <sub>2-3</sub> (0.546)
$S_5$	109.04(262) f <sub>0-5</sub> (0.007)	53.97(529) f <sub>1-5</sub> (0.019)	56.85(502) f <sub>1-5</sub> (0.009)	63.22(452) f <sub>1-5</sub> (0.009)	<u>75.86(376)</u> f <sub>1-0</sub> (0.002)	40.25(710) f <sub>2-5</sub> (0.003)
$S_6$	118.89(240) f <sub>0-6</sub> (0.048)	67.19(425) f <sub>1-6</sub> (0.0045)	68.76(415) f <sub>1-0</sub> (0.098)	71.08(402) f <sub>1-6</sub> (0.064)	<u>83.93(340)</u> f <sub>1-0</sub> (1E-04)	45.57(627) f <sub>2-6</sub> (0.046)
$S_7$	121.55(235) f <sub>0-7</sub> (0.301)	67.36(424) f <sub>1-7</sub> (0.113)	71.72(398) f <sub>1-0</sub> (0.267)	76.79(372) f <sub>1-7</sub> (0.259)	<u>93.64(305)</u> f <sub>1-0</sub> (0.001)	48.49(589) f <sub>2-7</sub> (0.088)
$S_8$	123.04(232) f <sub>0-8</sub> (0.006)	68.44(417) f <sub>1-8</sub> (0.325)	72.49(394) f <sub>1-0</sub> (0.005)	78.47(364) f <sub>1-8</sub> (0.014)	<u>100.52(284)</u> f <sub>1-0</sub> (0.010)	56.20(508) f <sub>2-8</sub> (0.007)
$S_9$	124.12(230) f <sub>0-9</sub> (0.005)	79.86(358) f <sub>1-9</sub> (0.046)	82.17(347) f <sub>1-0</sub> (0.061)	79.34(360) f <sub>1-9</sub> (0.0923)	<u>105.45(271)</u> f <sub>1-0</sub> (3E-04)	75.70(377) f <sub>2-9</sub> (0.012)

**Table8.1** There are collected all the transient absorption energy SA-10 roots, for twisted points optimized along the  $S_1$  MEP, and the minimum  $S_2$ .

## 8.6 Conclusions.

To conclude, we have provided computational evidences to two relevant effect controlling the photoisomerization efficiency of retinal in protein. First trough QMMM potential we were able to describe the of interaction the  $\beta$ -ionone ring with the surrounding residues (namely Trp-265), that may origin (upon planarization) the spring effect necessary to trigger the efficient ultrafast isomerization seen in Rh (or at least it

may contribute to that). This was accomplished by comparing full retinal with a dihydroretinal, hydrogenated in the ring-double bond. Furthermore, we have detected an accessible low lying state ( $S_2$ ) which behaves as a trap (i.e. molecules pumped there do not isomerize) and which may be possibly involved in the formation of the unreactive population recently observed. Finally, these results suggest novel scenarios for the understanding of the spectroscopic observations recorded in Rh (including time resolved experiments) and the catalytic mechanism responsible for the highly efficient and ultrafast (i.e. sub-picosecond) photoisomerization.

### References:

- 1) Kandori, H.; Shichida, Y.; Yoshizawa, T. *Biochemistry-Moscow* 2001, 66, 1197.
- 2) Needleman, R. *Bacteriorhodopsin and Rhodopsin*. In *CRC Handbook of Organic Photochemistry and Photobiology*; ed.; Horspool, W. M., Song, P.-S., Eds.; CRC Press: Boca Raton, FL, 1995; Vol. pp 1508-1515.
- 3) Ottolenghi, M.; Sheves, M. *Isr. J. Chem.* 1995, 35, U3.
- 4) Wald, G. *Science* 1968, 162, 230.
- 5) Mathies, R.; Lugtenburg, J. *The Primary Photoreaction of Rhodopsin*. In *Molecular Mechanism of Vision*; Stavenga, D. G., DeGrip, W. J., Pugh, E. N. J., Eds.; Elsevier Science Press: New York, 2000; Vol. 3, pp 55-90.
- 6) Yoshizawa, T.; Kuwata, O. *Vision: Photochemistry*. In *CRC Handbook of Organic Photochemistry and Photobiology*; ed.; Horspool, W. M., Song, P.-S., Eds.; Boca Raton, FL, 1995; Vol. pp 1493-1499.
- 7) a) Kandori, Ichioka, Sasaki, *Chem.Phys.Lett.* 354 2002 251-255b) Kandori, Maeda A. *Biochemistry*, 1995 34 14220-14229 c) Bartl, Fritze, Hofmann, and Ernst. *Joun of Biol.Chem.* 280,40, 2005,34259
- 8) Garavelli M, Bernardi F, Olivucci M, Vreven T, Klein S, Celani P, Robb MA (1998) *Faraday Discuss* 110:51-70
- 9) González-Luque R, Garavelli M, Bernardi F, Merchán M, Robb MA, Olivucci M (2000) *Proc Natl Acad Sci USA* 97:9379-9384
- 10) Cembran A, Bernardi F, Olivucci M, Garavelli M (2003) *J Am Chem Soc* 125:12509
- 11) De Vico L, Page CS, Garavelli M, Bernardi F, Basosi R, Olivucci M (2002) *J Am Chem Soc* 124:4124-4134
- 12) Garavelli M, Vreven T, Celani P, Bernardi F, Robb MA, Olivucci M (1998) *J Am Chem Soc* 120:1285-1288
- 13) Garavelli M, Celani P, Bernardi F, Robb MA, Olivucci M *J Am Chem Soc* 119:6891-6901
- 14) Cembran A, González-Luque R, Altoè P, Merchán M, Bernardi F, Olivucci M, Garavelli M (2005) *J Phys Chem A* 109:6597-6605
- 15) a) Teller, D. C., Okada, T., Behnke, C. A., Palczewski, K., and Stenkamp, R. E. (2001) *Biochemistry* 40, 7761-7772. b) Okada, T., Fujiyoshi, Y., Silow, M., Navarro, J., Landau, E. M., and Shichida, Y. (2002) *Proc. Natl. Acad. Sci. U.S.A.* 99, 5982-5987
- 16) . T. Okada, M. Sugihara, A.N Bondar, M. Elstner, P. Elstnel, V. Buss. *J. Mol. Biol.* 342, 571-583 (2004)
- 17) Cembran, Andres, Merchán, Garavelli. *Theor.Chem.Acc.* 2007, 118, 173-183.

- 18) Kandori, Furutami, Nishimura, Chem. Phys. Lett. 334 2000 271-276.
- 19) Altoè, P.; Stenta, M.; Bottoni, A.; Garavelli, M. Theor. Chem. Acc. 2007, 118, 219.
- 20) González-Luque, R.; Garavelli, M.; Bernardi, F.; Merchán, M.; Robb, M. A.; Olivucci, M. Proc. Natl. Acad. Sci. U. S. A. 2000, 97, 9379.
- 21) Malmqvist, P.-Å.; Roos, B. O. Chem. Phys. Lett. 1989, 155, 189.
- 22) Andersson, K. et al. MOLCAS, 6.0; Department of Theoretical Chemistry, Chemical Centre, University of Lund: Lund, 2004.
- 23) Andersson, K.; Malmqvist, P. A.; Roos, B. O. J. Chem. Phys. 1992, 96, 1218
- 24) G. Haran et al, J. Phys. Chem A 1999, 103, 2202-2207
- 25) H. Kandori et al, Chem. Phys. Letters 2001, 334, 271-276

## - CHAPTER 9 -

### Conclusions

Concluding, during this Phd, computational investigations on different organic compounds have been carried in order to model and rationalize the origin of the photochemical properties. Interestingly we have investigated different compounds in the field of photochemistry. On one hand, small isolated systems were the subject of our research: the potential energy surfaces of the photochemically relevant states have been mapped to locate and characterize the reaction channels unveiling their photoreactivity.

On the other hand QMMM calculations were employed to study larger systems as protein, with the aim of increase our knowledge of primary events, responsible for twilight vision. Therefore we have summarized the following concluding remarks:

- The open to close-ring electrocyclization of 2-5, dimethyl-furyl-fulgide (FF-3) has been investigated using QM accurate *ab initio* calculations. The design of efficient, thermally stable fatigue-resistant photochromic systems<sup>1,2</sup> was pursued by the mechanistic picture presented in this thesis. If covalent state usually are the state driving the polyene<sup>3</sup> photochemistry, in this

molecule a ionic “charge transfer state” is the state triggering efficiently the  $S_1$  photo-excited population towards the  $S_1/S_0$  degenerate region<sup>4</sup>. On one hand, covalent state leading to the formation of a biradical minimum, and with a necessary barrier to access the  $S_1/S_0$  CI complicates the photocyclization. On the other hand, for FF-3 the “charge-transfer” state does lead to a direct barrierless pathway for an accessible photocyclization. Only one reaction coordinate is involved in this process giving reason of the simple photoreactivity. Finally the origin for the intrinsic high stability of the ionic state in furyl-fulgides stands in the high electron withdrawing (EW) and electron releasing (ER) ability of the anhydride and heteroaromatic groups, respectively, that stabilize this state and its underlying charge transfer. Very remarkably, this understanding opens unprecedented scenarios for the de novo design of efficient related photochromic systems.

- We also, have examined the complex stilbene photoreactivity, in order to gain more insight into the radiationless event origin from the phantom state  $S_1$ . Despite mono-determinantal, the particular influence of the  $\sigma$ - $\pi$  excitation on the  $S_1$  stilbene wave function, dominated by a  $\pi \rightarrow \pi^*$  character, reveals how the non-dynamical correlation energy was important to correctly describe the excited state behavior and the topological aspect of its potential energy surface<sup>5</sup>. The CASSCF//CASPT2, correlating only the active electrons<sup>6</sup>, fails to describe accurately the ionic/covalent conical intersection geometries. This approach does not provide the proper topology of the excited  $S_1$  potential energy surface for stilbene.

The inclusion of  $\sigma\pi$  correlation was shown to be necessary, and RASSCF has proved to be a successful approach to achieve this. The valence–core correlation energy is recovered by adding single  $\sigma$  excitations to all the  $\sigma$  configurations. This gives a balanced description of the differential electron correlation and greatly improved results for the  $S_1$  topology in stilbene. On the basis of the preliminary results obtained, (that seems to well agree with experiments<sup>7,8</sup>) we are confident that this strategy is the most appropriate for supply the CASSCF inadequacy for ionic state treatment.

- Finally, hybrid QM(CASPT2//CASSCF/6-31G\*)/MM<sup>9</sup>(Amber) computations have been used to map the photoisomerization path of the retinal chromophore in Rhodopsin<sup>10-15</sup> and explore the reasons behind the photoreactivity efficiency and spectral control in the visual pigments. It is shown that while the electrostatic environment plays a central role in properly tuning the optical properties of the chromophore, it is also critical in biasing the ultrafast photochemical event: it controls the slope of the photoisomerization channel as well as the accessibility of the  $S_1/S_0$  crossing space triggering the ultrafast decay. The role of the E113 counterion, the E181 residue and the other amino acids of the protein pocket are explicitly analyzed: it appears that counterion quenching by the protein environment plays a key role in setting up the chromophore’s optical properties and its

photochemical efficiency. A unified scenario is presented that discloses the relationship between spectroscopic and mechanistic properties in rhodopsins, and allows to draw a solid mechanism for spectral tuning in color vision pigments: a tunable counterion shielding appears as the elective mechanism for L $\leftrightarrow$ M spectral modulation, while a retinal conformational control must dictate S absorption. It is, also, shown that this model may contribute to shed new light into mutations-related vision deficiencies that opens innovative perspectives for experimental biomolecular investigations in this field. Furthermore, we have accurately estimated the interaction the  $\beta$ -ionone ring with the surrounding residues of the binding pocket, origins the denominated "lever effect"<sup>16</sup> catalyzing the S<sub>1</sub> photoisomerization. Finally we have detected an accessible low lying state (S<sub>2</sub>) which behaves as a trap (i.e. molecules pumped there do not isomerize) and which may be possibly involved in the formation of the unreactive population recently observed.

#### References:

- 74) a). Irie, M. *Chem. Rev.* 2000, *100*, 1685. b) Bouas-Laurent, H. Bouas-Laurent, H. ; Durr, H. *Pure Appl. Chem.* 2001, *73*, 639
- 75) Alfimov M., Fedorova O, Gromov S, *Jour. Of Photochemistry and Photobiology A*, 2003, *158*, 183-198
- 76) G. Tomasello, Francois Ogliaro, Michael J. Bearpark, Michael A. Robb, and Marco Garavelli *J. Phys. Chem. A*; 2008; *112*(41) pp 10096 – 10107.
- 77) .M. Handschuh, M. Seibold, H. Port and H.C. Wolf. *J. Phys. Chem. A* 1997, *101*, 502-506.
- 78) Martial Boggio-Pasqua, Michael J. Bearpark, Michael Klene, and Michael A. Robb. *Jour. of Chem. Phys.* 2004, *120*, 17, 7849.
- 79) J. Quenneville, T. J. MartRnez, *J. Phys. Chem. A* 2003, *107*, 829 – 837.
- 80) W. Fuß, C. Kosmidis, W. E. Schmid, S. A. Trushin, *Chem. Phys. Lett.* 2004, *385*, 423 – 430.
- 81) W. Fuß, S. Lochbrunner, A. M. Miller, T. Schikarski, W. E. Schmid, S. A. Trushin, *Chem. Phys.* 1998, *232*, 161 – 174.
- 82) Altoè, P.; Stenta, M.; Bottoni, A.; Garavelli, M. *Theor. Chem. Acc.* 2007, *118*, 219.
- 83) Kandori, H.; Shichida, Y.; Yoshizawa, T. *Biochemistry-Moscow* 2001, *66*, 1197.

- 84) Needleman, R. Bacteriorhodopsin and Rhodopsin. In CRC Handbook of Organic Photochemistry and Photobiology; ed.; Horspool, W. M., Song, P.-S., Eds.; CRC Press: Boca Raton, FL, 1995; Vol. pp 1508-1515.
- 85) Ottolenghi, M.; Sheves, M. *Isr. J. Chem.* 1995, 35, U3.
- 86) Wald, G. *Science* 1968, 162, 230.
- 87) Mathies, R.; Lugtenburg, J. The Primary Photoreaction of Rhodopsin. In *Molecular Mechanism of Vision*; Stavenga, D. G., DeGrip, W. J., Pugh, E. N. J., Eds.; Elsevier Science Press: New York, 2000; Vol. 3, pp 55-90.
- 88) Yoshizawa, T.; Kuwata, O. Vision: Photochemistry. In *CRC Handbook of Organic Photochemistry and Photobiology*; ed.; Horspool, W. M., Song, P.-S., Eds.; Boca Raton, FL, 1995; Vol. pp 1493-1499.
- 89) Cembram, Andres, Merchà, Garavelli. *Theor.Chem.Acc.*2007,118,173-183.

# Appendix A

## A.1 The QM/MM method

We have developed a new code based on an hybrid Quantum Mechanical/Molecular Mechanical potential (QM/MM)<sup>34</sup> to study the chemical and photo-chemical reactivity of large systems. We have presented in details our method elsewhere<sup>49</sup> and it is briefly described here.

Following our approach we divide the system into three layers (*high*, *medium* and *low*) as depicted in Figure S1 by adopting a hydrogen *atom-link* scheme<sup>50</sup> (i.e. free valences of the *high* (QM) layer are saturated with hydrogen atoms to give *model-H*, see Figure S1). Energies and forces of the *high* layer are computed at the QM level with an electrostatic embedding scheme to account for the electrostatic influence of the surrounding MM region on the QM (*high*) layer. That is, QM computation are performed on *model-H* surrounded by the atomic point charges of the MM layers (i.e. the *emb* charges). In our notation *pod* and *emb* are two subsets of atomic point charges of the MM region, see Figure S1. The *medium* and *low* layers are treated at the MM level and, analogously, the effect of charge changes occurring during the (photo)chemical process in the *high* (QM) layer is accounted for by using in the MM calculations the QM atomic point charges coming from QM computations. The energy values for the QM and MM region are added together to compute the correct QM/MM energy, and a subtractive scheme similar to ONIOM<sup>51</sup> is used for this purpose:

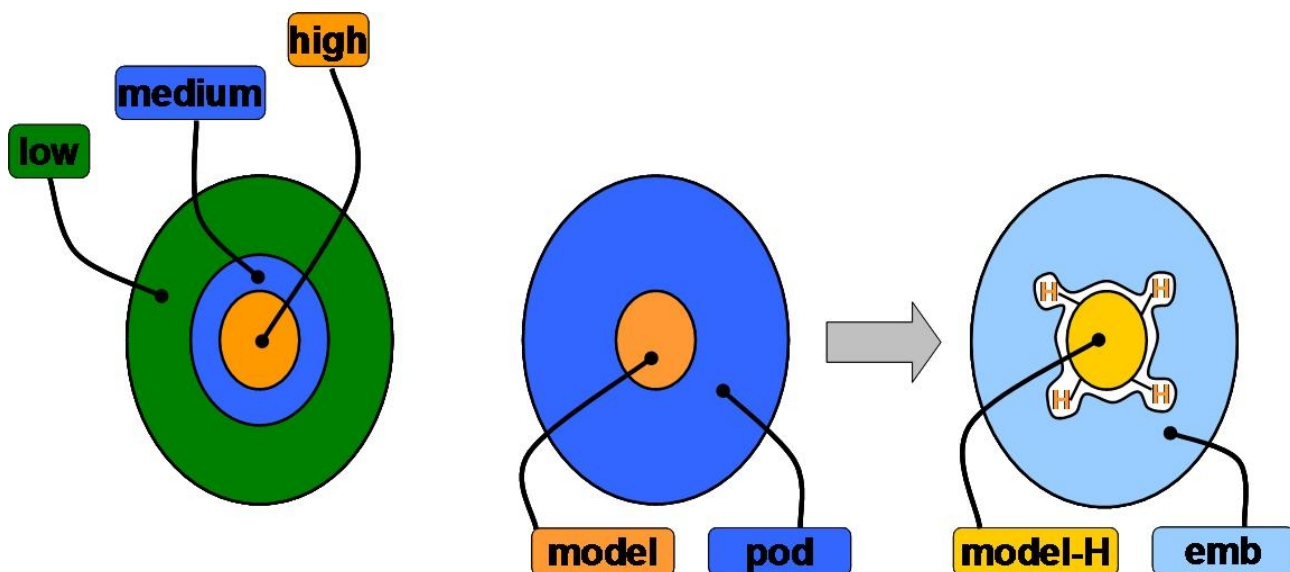
$$E^{tot} = E_{QM}^{model-H} + E_{MM}^{real} + E_{QM}^{el,model-H+emb} - E_{MM}^{el,model/pod} - E_{MM}^{model-H}$$

The cartesian forces of the two regions are merged, then the total QM/MM energy and forces are transferred to an accurate optimization algorithm. In our code the *high-medium* layers are simultaneously optimized using the BFGS<sup>S3</sup> algorithm, while the *low* region may be kept frozen or be fully optimized at each optimization step with a fast and rough algorithm, e.g. steepest descent (this feature nearly resembles the so called “micro-iteration” technique of ONIOM)<sup>51</sup>. The capability of handling coupled together the *high-medium* region allows the explicit treatment of large molecular motions around the reactive region without increasing the computational cost, because energy and forces of non-reacting (MM) atoms included in the *medium* region are computed at the MM level. The hydrogen atom-link approach is adopted to handle the boundary region between the QM and MM sub-systems.

In summary, all the energy components of the *high* layer (QM region) are obtained from QM calculations, including the electrostatic term (that is accounted for by an Electrostatic Embedding scheme). The bonded (stretching, bending, torsions) and non-bonded (electrostatic and van der Waals) interactions within the *medium* and *low* layers (MM region) are handled at the MM level. The bonding and van der Waals cross terms are treated at the MM level, while the electrostatic interaction term between the QM and the MM region is computed classically for MM atoms.

**Figure A.1.** Partitioning scheme of the whole system.





## [A.2 CASPT2 vs. MS-CASPT2: validation of MS-CASPT2 results for the ion pair](#)

Both the standard CASPT2 and the Multi State (MS)-CASPT2 methods have been used with the 6-31G\* basis set to evaluate the energies of all the structures discussed in the text. Interestingly, MS-CASPT2 energies for the ground and ionic excited states almost perfectly replicate (within 3 kcal in general) CASPT2 results. This occurs for all the structures investigated in the paper but for the ion-pair system (i.e. PSB11<sub>dist</sub>-E113 and Rh<sub>113</sub>), see Table A.1. That is, CASPT2 and MS-CASPT2 gives the same energies except for those situations where we do still have a full counterion effect (PSB11<sub>dist</sub>-E113) or where the counterion is not yet quenched (Rh<sub>113</sub>). Here, S<sub>1</sub> and S<sub>2</sub> are almost degenerate at the CASPT2 level and the S<sub>0</sub>-S<sub>1</sub> energy gap is quite large, while it appears that S<sub>1</sub>-S<sub>2</sub> degeneration is removed and the S<sub>0</sub>-S<sub>1</sub> energy gap is significantly reduced after MS-CASPT2 correction. In other words, this appears as a multi state problem.

**Table A.1.** CASPT2 vs. MS-CASPT2 relative energies (kcal/mol) of the first singlet states at the FC structure in different environments. 6-31G\* basis set.

FC chromophore structure in: sphere	VACUO	Rh	ION-PAIR	r=3.0	r=3.5	Buss-
CASPT2 / MS-CASPT2 (2-roots state average)						
S0 not available	0/0	0/0	not available	not available	not available	not available
S1 not available	50.4/53.8	57.8/60.2	not available	not available	not available	not available
CASPT2 / MS-CASPT2 (3-roots state average)						
S0 0/0	0/0	0/0	0/0	0/0	0/0	0/0
S1 59.2/60.0	50.3/54.5	57.6/61.0	75.7/65.7	62.2/59.6	58.7/58.8	
S2 81.6/89.1	79.3/81.9	82.0/84.9	74.7/95.2	79.6/93.9	80.9/93.2	

On the other hand, very similar CASPT2 and MS-CASPT2 energies are computed at the twisted structures (corresponding to the  $S_1/S_0$  CI in Rh), regardless the charge status (i.e. vacuo, ion-pair, Rh<sub>113</sub> or Rh), see Table S2. That is, we never have a multi state problem here (including the ion-pair case). This is reasonable and expected as the ground  $S_0$  and excited  $S_1$  states have a clean well-defined and very different electronic character (covalent/diradical vs. ionic/charge-transfer, respectively) that prevents any interaction even at the MS level.

**Table A.2.** CASPT2 vs. MS-CASPT2 relative energies (kcal/mol) of the first singlet states at the twisted (CI) structure in different environments. 6-31G\* basis set.

CI chromophore structure in:	VACUO	Rh	ION-PAIR
	CASPT2 / MS-CASPT2 (2-roots state average)		
S0	24.2/25.7	31.1/31.4	0/0
S1	33.9/35.4	35.2/37.7	26.4/26.4
	CASPT2 / MS-CASPT2 (3-roots state average)		
S0	23.1/25.3	32.5/33.3	32.6/35.8
S1	35.1/37.3	34.5/35.8	57.1/59.9
S2	86.9/89.1	84.7/86.0	84.7/88.4
	CASPT2 / MS-CASPT2 (4-roots state average)		
S0	not available	not available	30.1/35.9
S1	not available	not available	54.6 /57.0
S2	not available	not available	80.3/92.3

The above results suggest that MS-CASPT2 computations are trustable in general for evaluating the ground and the excited ionic state energies. In particular, when the  $S_1$  and  $S_2$  CASPT2 energies are far apart (which turns to be the case for all structures but the ion pair) they reproduce CASPT2 results, while when  $S_1$  and  $S_2$  are degenerate (i.e. the untwisted ion-pair), MS-CASPT2 removes that degeneracy giving energies that deviate significantly from CASPT2 ones, still are more reasonable (thus suggesting that a MS problem really emerges in this case and that MS-CASPT2 should be the choice for the ion pair): i) they better fit literature data from other correlated methods,<sup>46,47</sup> as well as the results on the ion pair reported by Buss<sup>45</sup> (that is the reference paper on this topic up to now); ii) MS-CASPT2  $S_1$  Abs energy in the ion-pair (430 nm) matches nicely experimental energies for PSB11 in apolar solvent (440-460 nm) where analogous ion-pair systems should exist (on the other hand, standard CASPT2 results give an unrealistic much more blue shifted value (380 nm)).

As pointed out above, when the ionic and covalent excited  $S_1$  and  $S_2$  states are close in energy, a MS problem emerges. More specifically,  $S_1$  and  $S_2$  appear quite mixed in the FC region (note in fact that also  $S_2$  absorbs and has a significant oscillator strength, see Table 1). Thus, it is not unreasonable that they can interact and their energy be affected by MS corrections. Actually, it is  $S_1$  and  $S_2$  that are mostly affected by multi state, as can be seen in Table S3 below for a 3-roots and 4-roots state average approach. It is apparent that most of the mixing is on between  $S_1$  and  $S_2$ , that go down/up by roughly the same amount of energy (see bold numbers in Table S3). On the other hand,  $S_0$  stays at the same energy at both (CASPT2 and MS-CASPT2) levels. This conclusion is confirmed by a 4-roots state average approach, revealing that the above reasoning is independent on the number of roots included in the state averaged procedure.

**Table A.3.** CASPT2 vs. MS-CASPT2 absolute energies and corresponding energy variations (hartree) for the first singlet states at the ion-pair FC structure. 6-31G\* basis set.

(3-roots state average)		
CASPT2	MS-CASPT2	CASPT2->MS-CASPT2 total energy variation
-871,03860	-871,044115	-0,005515
-870,91793	-870,939471	<b>-0,021541</b>
-870,91949	-870,892433	<b>+0,027057</b>

(4-roots state average)		
CASPT2	MS-CASPT2	
-871,0345689	-871,0467815	
-870,9161644	-870,9436755	
-870,9094712	-870,8901305	
-870,8875126	-870,8655421	

On the other hand, as shown above, MS and PT2 behave the same at both the CI and FC for Rh and vacuo (as well as at the twisted structure of the ion pair). That is, PT2 and MS energy differences are similar and we can trust both results. Table S4 shows the absolute energy values for the FC point in Rh and Vacuo (corresponding relative energies have been already reported in Table S1). As already pointed out, MS-CASPT2 replicates CASPT2 total energies within 3 kcal/mol.

**Table A.4.** CASPT2 vs. MS-CASPT2 absolute energies (hartree) for the first three singlet states at the FC structure in Rh and vacuo. 6-31G\* basis set.

Rh		
(3-roots state average)		
CASPT2	MS-CASPT2	
-871,0624737	-871,0643268	
-870,9707168	-870,9671055	
-870,9318768	-870,928956	

VACUO		
(3-roots state average)		
CASPT2	MS-CASPT2	
-870,88778	-870,891344	
-870,80761	-870,804523	
-870,76135	-870,760864	

For seek of completeness, we have also recalculated the MEP in Rh by MS-CASPT2 and compared it to the standard CASPT2 results (see Table S5 and the corresponding Figure A.1). Remarkably, the trend and whole behavior of the  $S_1$  MEP is the same at both MS and PT2 levels, further supporting the reliability of a MS approach in those systems. That is, MS behaves well and is trustable in general, which implies we can take MS results (and quantitatively trust them) when the CASPT2 approach alone seems to be unreliable, as it occurs in the ion-pair untwisted system (where a MS problem emerges indeed). Here,  $S_1$  and  $S_2$  are degenerate and heavily mixed, i.e. much more than in other systems. MS corrects for that mixing by returning very clean wavefunctions with almost pure ionic ( $S_1$ ) and covalent ( $S_2$ ) character. This calls for MS to be the choice in those situations.

**Table A.5.** CASPT2 and MS-CASPT2 QM/MM absolute (hartree) and relative (kcal/mol) energies for the first three singlet states along the  $S_1$  photoisomerization MEP computed in Rh. 6-31G\* basis set.

CASPT2 absolute QM/MM energies (3-roots state average)						
state	FC	MinS1	20°	40°	60°	90°
S0	-876,071	-876,060	-876,060	-876,058	-876,044	-876,020
S1	-875,980	-875,997	-876,002	-876,009	-876,014	-876,016
S2	-875,941	-875,969	-875,970	-875,966	-875,951	-875,936

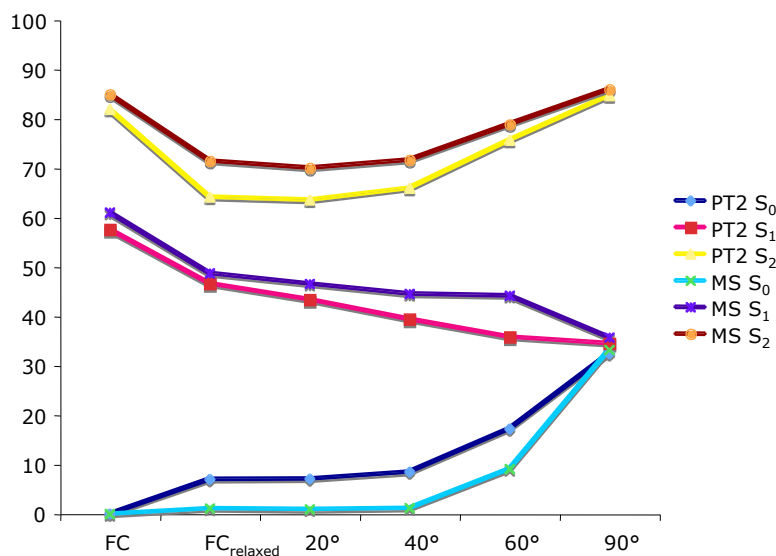
  

MS-CASPT2 absolute QM/MM energies (3-roots state average)						
state	FC	MinS1	20°	40°	60°	90°

S0	-876.073	-876.072	-876.0717569	-876.071	-876.059	-876.0199
S1	-875.976	-875.996	-875.9990961	-876.002	-876.003	-876.0164
S2	-875.938	-875.959	-875.9616963	-875.959	-875.948	-875.9363

CASPT2 relative QM/MM energies (3-roots state average)						
state	FC	MinS1	20°	40°	60°	90°
S0	0,000	7,030	7,115	8,480	17,232	32,453
S1	57,578	46,671	43,439	39,467	35,852	34,545
S2	81,951	64,184	63,573	65,979	75,659	84,713

MS-CASPT2 relative QM/MM energies (3-roots state average)						
state	FC	MinS1	20°	40°	60°	90°
S0	0,000	1,137	0,985	1,181	9,124	33,258
S1	61,007	48,756	46,581	44,560	44,181	35,768
S2	84,947	71,484	70,049	71,674	78,928	85,968



**Figure A.2.** CASPT2 and MS-CASPT2 corrected  $S_0$ ,  $S_1$  and  $S_2$  energy profiles along the relaxed QM/MM scan on  $S_1$  tracing the photoisomerization process in Rh, from the Franck-Condon (FC) point to the twisted conical intersection funnel (CI), through a point ( $FC_{\text{relaxed}}$ ) where only skeletal bond stretchings have been relaxed (i.e. dihedral angles have been frozen here). Energies are in kcal/mol.

As a final validation of the assumption above, we have recomputed CASPT2 energies for the ion-pair (FC) via using an ANO type basis set, which is a better (although more expensive) choice than 6-31g\* for (MS)CASPT2 computations, thus allowing a more accurate estimate of correlation energy. The purpose is to show that also standard CASPT2 calculations can return correct energies (in agreement with MS-CASPT2/6-31g\* results) provided a proper (enough flexible and extended) basis set is used, such as ANO. Very remarkably, the new CASPT2/ANO-s(3s2p1d/2s) computations on the FC ion-pair system returns energies in agreement with the previous MS results (see Table A.6):  $S_1$  and  $S_2$  are not degenerate any more, and the CASPT2/ANO energy of the ionic  $S_1$  state (69.4 kcal/mol) matches satisfactorily the old MS-CASPT2/6-31g\* energy (65.7 kcal/mol, see Table S1 and S3). Additionally, the electronic nature of  $S_1$  (ionic, single H $\rightarrow$ L excitation) and  $S_2$  (covalent, double H $^2\rightarrow$ L $^2$  excitation) appears now very well defined.

**Table A.6.** CASPT2/ANO-s(3s2p1d/2s) absolute (hartree) and relative (kcal/mol) energies for the first four singlet states at the FC structure of the ion-pair.

(4-roots state average)				
CASPT2	E	$\Delta E$ (kcal/mol)	Config	Weight
S0	-871.09713189	0.00	222222000000	0.687

S1	-870.98653377	69.40	2222u2d00000	0.583
S2	-870.97640303	75.76	222202200000	0.283
S3	-870.93528249	101.56	22u222d00000	0.224

The covalent excited state (i.e.  $S_2$  in vacuo and Rh) deserves a different discussion: MS-CASPT2 appears to overestimate it systematically vs. CASPT2 (while, as shown above,  $S_0$  and  $S_1$  energies appear good), see Tables A.1 and A.2. This is why we decided to use standard CASPT2 energies in general, but for those systems (i.e. the ion pair) where  $S_1$  and  $S_2$  are degenerate at this level (and a multi state problem emerges).

In conclusion, the results presented above support the validity of MS-CASPT2/6-31g\* results for the ion-pair that is the only system where standard CASPT2/6-31g\* results appear doubtful. Thus, while standard CASPT2/6-31g\* results (with a 3-roots state average approach) have been considered in general in this work, MS-CASPT2/6-31g\* computations are used for the problematic case of the ion-pair as they appear to return quantitatively accurate results.

### A.3. Complete references 51 and 55.

(51) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J., J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03*, Revision C.02; Gaussian, Inc.: Wallingford CT, 2004.

(55) Andersson, K.; Barysz, M.; Bernhardsson, A.; Blomberg, M. R. A.; Carissan, Y.; Cooper, D. L.; Cossi, M.; Fülcher, M. P.; Gagliardi, L.; de Graaf, C.; Hess, B.; Hagberg, G.; Karlström, G.; Lindh, R.; Malmqvist, P.-Å.; Nakajima, T.; Neogrady, P.; Olsen, J.; Raab, J.; Roos, B. O.; Ryde, U.; Schimmelpfennig, B.; Schütz, M.; Seijo, L.; Serrano-Andrés, L.; Siegbahn, P. E. M.; Stålring, J.; Thorsteinsson, T.; Veryazov, V.; Widmark, P. O. *MOLCAS*, 6.0; Department of Theoretical Chemistry, Chemical Centre, University of Lund: Lund, 2004.

## Publications:

1)G.Tomasello, Francois Ogliaro, Michael J. Bearpark, Michael A. Robb, and Marco Garavelli *J. Phys. Chem. A*; 2008; 112(41) pp 10096 – 10107.

2)G.Tomasello, G. Olaso-González, P. Altoe, M. Stenta, L. Serrano-Andrés, ‡ M. Merchán, \*, G. Orlandi, Andrea Bottoni and Marco Garavelli\*, † “Electrostatic Control of the Photoisomerization Efficiency and Optical Properties in Visual Pigments: On the Role of Counterion Quenching” Accepted on J.A.C.S