Calculation of Excited States: Molecular Photophysics and Photochemistry on Display

15

Luis Serrano-Andrés and Juan José Serrano-Pérez

Contents

Dedicated to the memory of Luis Serrano–Andrés, great scientist, teacher and friend, who passed away recently.

L. Serrano-Andrés
Instituto de Ciencia Molecular, Universitat de Valéncia, Valencia, Spain

J.J. Serrano-Pérez (✉)
Department of Chemistry, Imperial College London Computational and Structural Research Group, London, UK
e-mail: Juan.Serrano@uv.es

J. Leszczynski et al. (eds.), *Handbook of Computational Chemistry*,
DOI 10.1007/978-3-319-27282-5_14

Abstract

Excited states participate in photoinduced events as well as in thermally activated reactions, even in many cases in which only the ground state is believed to be involved. Life on Earth also depends, both directly and indirectly, on the influence that light has on chemistry. The energy of the Sun's visible and ultraviolet radiation promotes processes that not only permit the continued existence of life on the planet, but which are keys for evolution by means of mutations. To study a system in an excited state, far away from its optimum situation, is a challenge for chemists, both experimentalists and theoreticians. This chapter is focused on the practical aspects related to the calculation of excited states in molecular systems by using quantum-chemical methods, a type of study that escapes in many cases from the well-established computational strategies used for the molecular ground states, both because of the complexity of the problem itself and for the methodological requirements. A short review of the spectroscopic and photochemical panorama will be provided first in order to explain which are the main parameters and processes to be determined, followed by a compact description of the most relevant and employed quantum-chemical methods and computational strategies for excited states. A number of applied examples of actual calculations on paradigmatic excited state problems will be provided in the different subchapters, followed in each case by comments on practical issues occurring in the calculations. With these cases we will try to demonstrate that in the last years the quantum-chemical studies on excited states have reached the required maturity to interpret and predict, at a molecular level, different types of chemical situations.

Introduction

Computing electronic excited states with quantum chemical methods is much more complex than doing it for ground states, because it implies not only coping with higher solutions of the electronic Hamiltonian with diverse character and therefore requiring more complex methods, but also to solve a plethora of new situations, such as hypersurface crossings and coupling between the states, that usually requires to abandon convenient approximations like Born–Oppenheimer (see Szabo and Ostlund 1996), for instance.

Molecules, consisting of electrically charged nuclei and electrons, may interact with the oscillating electric and magnetic fields of light. Spectroscopic experiments demonstrate that energy can be absorbed or emitted by molecules (and atoms) in discrete amounts, corresponding to precise changes in energy of the molecule (or atom) concerned. As matter, light is a form of energy that exhibits both wave- and particle-like properties. Absorption of the relevant frequencies from incident radiation raises molecules from lower to higher levels. Electrons in molecules occupy molecular orbitals (MOs) with precise energy levels. Transitions from lower, filled orbitals, to upper (higher energy), empty orbitals usually involve absorption of radiation in the UV and visible parts of the spectrum. Much smaller quantities of

energy are linked to changes in the vibrational and rotational energy of the molecule. The understanding of the spectroscopic phenomena in the light of molecular orbital theory has opened new avenues in the comprehension of the photoinduced events.

The molecule in excited state is often prone to react in an easier way than in the ground state. The excess energy of an excited species can alter its reactivity and is particularly significant in the case of electronic excitation because of the energies involved are of similar order of magnitude as bond energies. Electronic excitations can then have a considerable effect on the structure of a species. Accordingly, the energies correspond roughly with typical activation energies for many reactions, which are too high to be reached from the ground but not from the excited state. The new electronic rearrangement may be also the key of the reactivity since the molecule in an excited state may exhibit nucleophilic or electrophilic properties different than those of its ground state.

Three modern developments have been produced in the last years that are the key for the comprehension of the photophysics and photochemistry of many chemical and biochemical phenomena: (1) rapid advances in quantum-chemical methods allow to study the excited states with high accuracy; (2) improved molecular beams techniques permit studies of isolated molecules, despite their sometimes low vapor pressure and propensity for thermal decomposition, and (3) the revolutionary impact that femtosecond laser and multiphoton techniques have had on the study of the electronic energy relaxation processes. Indeed, now it is possible to get information about reaction intermediates at very short times from femtochemical techniques, and, more than ever, the participation of quantum chemistry to interpret such findings has become crucial. A constructive interplay between theory and experiment can provide an insight into the chemistry of the electronic state that cannot be easily derived from the observed spectra alone.

From the theoretical viewpoint, the calculation of excited states is still a very complex task. Considering the many different electronic structure situations occurring in the potential energy hypersurfaces (PEHs) of the excited molecular systems, the only methods generally applicable to all of them are the multiconfigurational approaches. The application of these procedures requires a lot of skill and experience, and the limitations on the size of the problem are noticeable. Single-reference (black-box) methods only work in certain regions of the PEHs. In general, the excited state problem can be considered heavily multiconfigurational. New tools and strategies are required for excited states at the highest levels of calculation: optimization of minima, transition states, hypersurface crossings (conical intersections), and reaction paths, whereas states couplings (nonadiabatic, electronic, spin-orbit) need to be computed. This solves only the first part of the problem, that is, the solution of the time-independent Schrödinger equation. Once the potential represented by the PEHs is obtained, time-dependent equations have to be solved to finally determine reaction rates, states lifetimes, or populations. Coupling at the proper level those two types of calculations, static and dynamic approaches representing the electronic structure and reaction dynamics problems, respectively, is still a task under development.

Spectroscopy Overview

The concept of PEHs comes from the Born–Oppenheimer approximation, based on the separation of electronic and nuclear motion due to the large difference in mass between these particles, and assumes that the electrons follow the nuclei instantaneously during the motion of the latter. Therefore, an electronic and a nuclear Hamiltonian can be defined. Solving the electronic Schrödinger equation provides us a description of the movement of electrons, whereas we consider the rotation, vibration, and translation of the molecule solving the nuclear counterpart. The solution of the electronic Schrödinger equation is the energy of a particular nuclear configuration. The total energy for fixed nuclei must also include the constant (within this approximation) nuclear repulsion potential. The value of this total potential energy for every possible nuclear configuration is specifically the potential energy hypersurface. Photophysical and photochemical processes take place through interactions between PEHs. For the specific case of close degeneracies between the surfaces, where ultrafast energy transfers occur, the Born–Oppenheimer approximation breaks down and special methods in order to localize, optimize, and study the crossing structures are required.

The absorption of photons for a molecule is hardly a static problem. After the absorption (ABS) of one photon, a state of the same multiplicity as the ground state is mainly populated. Direct absorption to states of different multiplicity are only possible if the states heavily interact, for instance, by spin-coupling effects. Actually, in the general case, the energy goes to a vibrational excited state of an electronic excited state of the molecule. Straight afterward, a non-radiative decay occurs, with emission of heat (IVR, intramolecular vibrational relaxation), toward more stable structures of the state PEH, in many cases the state minimum. It might frequently happen that along the decay other states cross and, if appropriately coupling, the system can evolve toward other electronic states of the same multiplicity via a non-radiative internal conversion (IC). Finally, the molecule arrives to the lowest-lying singlet excited state, S_1, from which the molecule may emit (F, fluorescence) and return to the ground state. Alternatively, a non-radiative transition between two states of different multiplicity is also possible (ISC, intersystem-crossing). By successive internal conversions the system reaches the lowest-lying triplet excited state, T_1, from which the molecule may emit (P, phosphorescence).

Figure 1 contains a simplified Jablonski diagram summarizing the main photophysical and photochemical effects undergone by a molecular system. It is common to reserve the word photophysics to processes not involving the generation of new photospecies, that is, just to decays leading to emission or returns to the ground state. However, nonradiative internal conversions or intersystem crossing are also considered photochemical processes, therefore the use of both terms is somewhat loose. A proper nomenclature for excited states is not easy to establish. The less ambiguous (and less informative too) form is purely enumerative: S_0, S_1, T_1, T_2, where S represents a singlet state and T stands for the triplet states, and the states are ordered by energy. In symmetric systems, it is convenient to use the labels

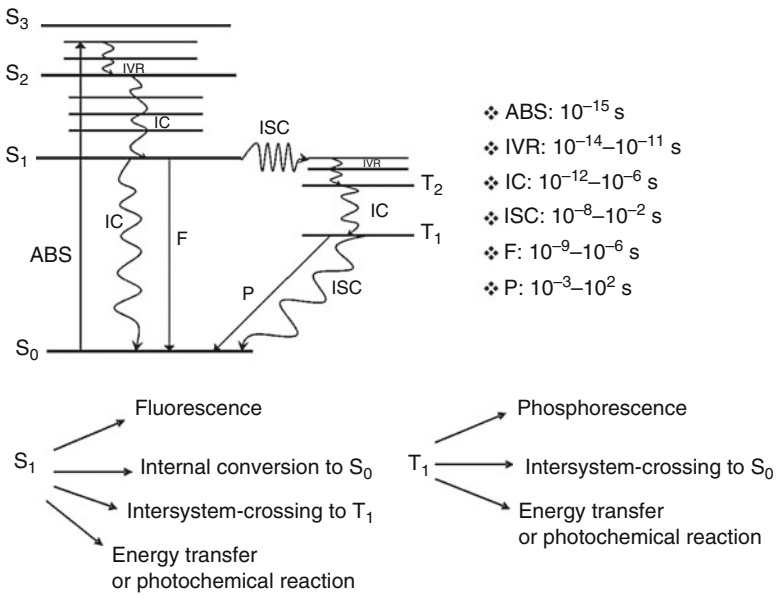

Fig. 1 Jablonski's diagram, lifetimes of the basic photophysical processes and deexcitation pathways from the lowest-lying excited states of a molecule

derived from group theory. A state is then described in terms of the behavior of the electronic wave function under the symmetry operations of the point group which the molecule belongs to $1^1A_g, 2^3B_{3u}$ …, including the energy ordering, multiplicity, and symmetry label. Classical nomenclatures, such as those developed by Mulliken (N, V, T), Kasha $\left({}^1\pi\pi^*\right), {}^3\sigma\pi^*\Big)$, and Platt ($^1L_a$, ^3B_b), highlighting the type of orbitals implied in the excitation, can be useful in different cases.

The basic information required to rationalize a photoinduced phenomenon is the energy levels of the excited solutions and the probability of energy (population) transfer from one state to the other. In the semiclassical treatment of the interaction radiation-matter, whereas we treat the molecule quantum-mechanically, the radiation field is seen as a classical wave obeying Maxwell's equations.

The electric and magnetic fields of the radiation will interact with the atomic or molecular electrons giving a time-dependent perturbation. Solving the time-dependent Schrödinger equation provides us the comprehension of absorption and stimulated emission, whereas to explain spontaneous emission we need the machinery of quantum electrodynamics. The resonance condition provided by the energy differences between the different PEHs of the corresponding states relates to the absorbed or emitted energy quanta. Regarding the transfer probability, it is related with the strength of the interaction between the time-dependent field and the multipolar (the dipole, d, approach is usually enough) charge distribution of the

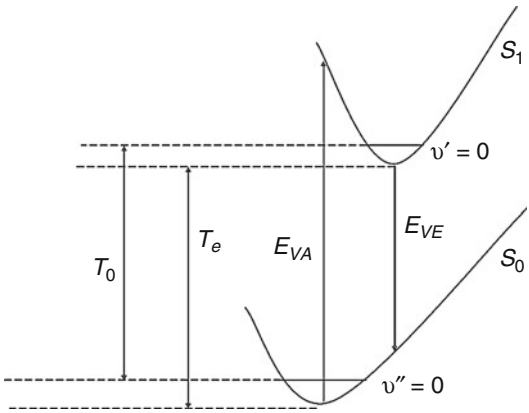

Fig. 2 Vertical energies and band origins. Vibrational states are labeled with the greek character υ

molecular system. Such strength is proportional to the transition dipole moment (TDM) (Fig. 2).

As an initial feature, the electronic states produce a superposition of bands which characterize the absorption spectrum. The range of absorbed energies fluctuates between the vertical absorption energy, E_{VA} (difference between the minimum of the ground state and the excited state at the same geometry, that is, the Franck–Condon transition), and the adiabatic transition or band origin, T_e (difference between the excited state and the ground state at their respective optimized equilibrium geometries): it is the minimal energy difference allowed in absorption if the assumption that all excitations begin from the relaxed ground state is considered, as well as the largest energy emitted from the relaxed excited state. In many cases, the determination of T_e provides enough information to assign band origins; however, the zero-point vibrational energy (ZVE) has to be included in both initial and final states to get the vibrational band origin, T_0, which can be directly compared to the experimental value, at least that obtained in the gas phase or in molecular beams. In addition, there is another magnitude named vertical emission energy, E_{VE}, which is the difference between the excited state and the ground state at the relaxed geometry of the former. The Franck–Condon principle stipulates that the vertical absorption can be related with the experimental band maximum. In fact, this is hardly the case, except when the ground and excited states have very similar geometries, and in this case the T_0 transition ($\mathrm{S}_0\upsilon'' = 0 \rightarrow \mathrm{S}_1\upsilon' = 0$) yields the most intense peak. The vertical excitation has, however, no experimental counterpart, whereas to get a true band maximum, the band vibrational profile must be computed. The only direct comparison relates the theoretical and experimental band origins, T_0 actually. Trying to asses the quality of a theoretical approach by comparing theoretical vertical excitations and experimental band maxima is one of the most frequent mistakes seen in the literature.

Regarding the transfer probability, far from the conical intersection regions the Fermi's Golden Rule is employed, in which the one-photon (optical) transition probability between two states $|n\rangle \rightarrow |m\rangle$ is proportional to the square of the

TDM between such states: TDM $= \langle m \,|d|\, n \rangle$. Two- and higher-order multiphoton probabilities can be also obtained. Based on symmetry considerations, selection rules for electronic transitions have been developed because only the totally symmetric matrix elements yield allowed nonzero probabilities (at first order). Using the so-computed electronic TDM, it is useful to estimate the electronic oscillator strength as $f = (2/3)\, E_{VA}\mathrm{TDM}^2$, with E_{VA} being the vertical energy difference. The oscillator strength is a classically derived magnitude that represents the relative area of the electronic transition band and that it can be compared with the experimental estimation based in shapes and bandwidths. On the other hand, the vibrational contributions to the band intensity (or, in general, the strength of the transfer) can be obtained by computing the TDM between vibrational states. If belonging to the same electronic states, infrared or Raman intensities can be produced, otherwise electronic band vibrational profiles can be obtained. The vibrational TDM is proportional to the vibrational overlap term between the electronic states, $\langle \upsilon_m | \upsilon_n \rangle$, which are called the Franck–Condon factors (the probability of transition is proportional to its square). The vibrational profiles are basically related to the differences in geometry existing from the initial to the final electronic state and, therefore, the most intense progressions proceed through the normal modes which trigger the aforesaid changes. Within the harmonic approach, the complete TDM with respect the nuclear coordinates Q_i is defined as the Herzberg–Teller expansion:

$$\mathrm{TDM} = \mathrm{TDM}\,(Q_0)\,\langle \upsilon_m(Q) | \upsilon_n(Q) \rangle + \sum_K \left(\frac{\partial \mathrm{TDM}(Q)}{\partial Q_K} \right)_{Q_0} \langle \upsilon_m(Q)\,|Q_K|\,\upsilon_n(Q) \rangle + \ldots.$$

Each one of the terms has an electronic and a vibrational component. The neglect of all terms except the first one is known as the Condon approximation, a usual way to proceed but only applicable for the one-photon dipole-allowed transitions. Otherwise, TDM(Q_0) is zero by symmetry and the first term vanishes. Nonetheless, this approximation is valid only when the Born–Oppenheimer approximation is also valid. Otherwise, the phenomenon of vibronic coupling arises, which gives rise to other approximations (the term vibronic should be reserved to solutions obtained in non-Born–Oppenheimer cases, although is frequently used improperly). From the calculation of transition dipole moments, radiative lifetimes can also be obtained, both in fluorescence and phosphorescence by using the Einstein coefficients (A_{21}) and the Strickler–Berg relationships (Strickler and Berg 1962):

$$A_{21} = \frac{1}{\tau_{\mathrm{rad}}} = 2.142005 \cdot 10^{10} E_{\mathrm{VE}}^3 \mathrm{TDM}_{1\to 2,}^2 \tag{1}$$

where τ_{rad} is the radiative lifetime measured in s^{-1}. The use of T_e instead of E_{VE} is more representative of the energy of the emission.

To study photophysical and photochemical processes on theoretical grounds, we need to determine the topography of the potential surfaces of the implied states (see Fig. 3). According to the different reaction paths through what a system

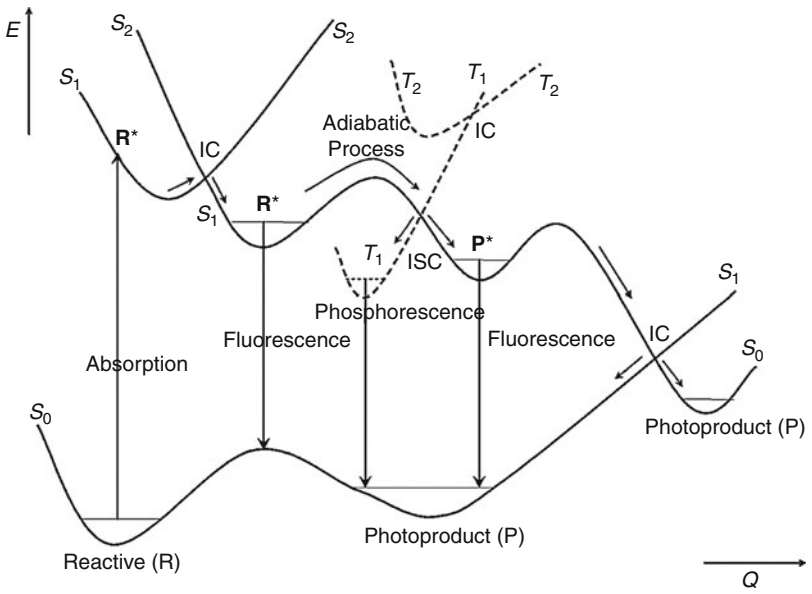

Fig. 3 Scheme of the main photophysical and photochemical molecular events. Notice that the order of the states changes along the nuclear coordinate, *Q*

might evolve, one can normally make the following classification that defines the photochemical panorama: adiabatic and non-adiabatic photochemistry. In an adiabatic reaction path, once vertical absorption takes place, the system proceeds along the hypersurface of the excited state to reach a local (or absolute) minimum leading in some cases to an emitting feature. On the contrary, in a non-adiabatic photochemical reaction, one part of the reaction takes place on the higher-state hypersurface and after a nonradiative jump at the surface crossing (or funnel) continues on the lower-state hypersurface. In a typical closed-shell ground-state molecule, the reaction usually begins on the potential energy surface of the excited state (S_1 or higher) at the Franck–Condon geometry (i.e., at the ground-state equilibrium geometry) and evolves either to the S_1 state minimum, from which it might emit, or to a crossing region with the ground state. Depending on the properties of such crossing, the process will end up on the reactant minimum or a new photoproduct minimum on the ground-state surface (S_0). The crossings between the excited state relevant from the photochemical view and the ground state are frequent, and they represent the basis of the spectroscopic phenomena. Therefore, a molecule evolving through the PEH of an excited state may well enter in a crossing region between two hypersurfaces during the lifetime of such an excited state. Hence, the lifetime of excited states is determined by the barriers that separate the excited states at the Franck–Condon geometry from the low-lying crossings. There are two types of crossings: conical intersections (CI), when the two interacting hypersurfaces have the same multiplicity, and (within the nonrelativistic approximation) singlet-triplet crossings (STC) or any other crossing between states

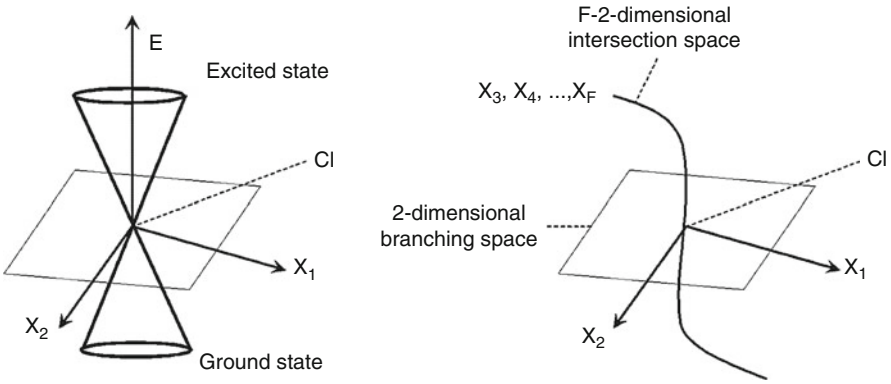

Fig. 4 Description of a conical intersection. The vectors x_1 and x_2 span the branching space. They are called the gradient difference (GD) vector and the derivative coupling (DC) vector. The DC vector measures the distortion of the system providing the maximum coupling between the electronic states involved in the crossing. The GD vector measures the distortion of the system leading to the largest variation of the energy difference between the two electronic states involved in the crossing

that have different multiplicity (Bearpark and Robb 2007; Domcke et al. 2004; Herzberg and Longuet-Higgins 1963; Robb et al. 1996; Teller 1937). Therefore, internal conversions take place through CIs, and intersystem crossings through STCs. The name of conical intersection (see Fig. 4) reflects the fact that a cone-shaped crossing is obtained when the energy of the states is plotted against the two privileged coordinates, the gradient differential vector, x_1, and the nonadiabatic coupling vector, x_2. Thus the total coordinate space F-dimensional is divided in two: the intersecting space (of dimension $F-2$), in which both states are degenerated, and the branching space (of two dimensions). The most pronounced difference between the slopes of both hypersurfaces occurs along x_1, whereas along x_2 the optimum nuclear displacement which mixes the two adiabatic wave functions in the CI point takes place. Actually, the intersection space is an hyperline that consists of an infinite number of CI points, i.e., a $(F-8)$-dimensional intersection space. To locate a CI point is equivalent to minimize the energy in the intersection space. In the case of STCs (if the nonrelativistic Hamiltonian is considered), we have only one privileged coordinates since the nonadiabatic coupling vector vanishes. Therefore we should refer this feature as an hyperplane, since we are moving along a $(F-1)$-dimension space.

As the Born–Oppenheimer approximation is not valid in regions where the electronic states become too close, it is logical that the nonadiabatic transfer is faster than the radiative relaxation. In essence, the former is based on the structure of the vibronic states. Certainly, the smaller is the gap between the states, the larger is the transfer probability.

Conical intersections may be peaked or sloped (see Atchity et al. 1991; Ben-Nun and Martínez 2000, and Fig. 5).

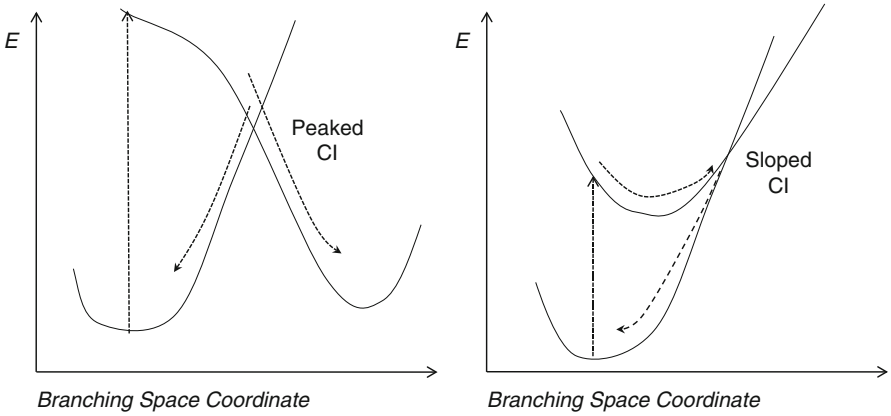

Fig. 5 Scheme of the limit characters displayed by the conical intersections. For a peaked CI, the two minima lie on the two different sides of the CI point and the gradients of the two intersecting PEHs are directed toward different directions. Conversely, for a sloped CI, the excited-state and ground-state minima lie on the same side (the gradients of the two PEHs point in the same direction), then the probability of re-crossing is higher and the efficiency of funneling is lower

If the connection between the two surfaces is sloped, the funnel may well be more efficient. In addition, the degeneracy at a crossing point can also be lifted at second order. As a consequence, we can choose a coordinate system in which to mix the branching and intersection space coordinates to remove this splitting and preserve the degeneracy to second order. These new coordinates give the curvature of the conical intersection hyperline and determine whether one has a minimum or a saddle point on it. These studies may also provide the vibrational modes that must be stimulated in order to enhance nonradiative decay because they decrease the energy gap and can lead to a CI (see Paterson et al. 2004; Sicilia et al. 2007).

A computational strategy can be designed, namely the Photochemical Reaction Pathapproach, in which the mechanism of the photoinduced process is accounted for by determining the fate of the energy on the populated state by computing the reaction profile. The whole process can be described by computing Minimum Energy Paths (MEPs), describing the lowest-energy, and therefore most favorable, although not unique, path for energy decay. The MEP is often built as steepest descent paths, guaranteeing the absence of barriers along the path. Each step requires the minimization of the PEH on a hyperspherical cross-section of the hypersurface centered on the initial geometry and characterized by a predefined radius. The optimized structure is taken as the center of a new hypersphere of the same radius, and the procedure is iterated until the bottom of the energy surface is reached. Mass-weighted coordinates are used, therefore the MEP coordinate corresponds to the so-called Intrinsic Reaction Coordinate (IRC), measured in a.u., that is, bohr $\cdot$ (amu)$^{1/2}$. The end of the path and the states crossed along the computed profile will inform about the fate of the energy, and, in particular, of the location of possible radiative minima and surface crossings, CIs and STCs. More crucial than

the presence of a crossing is its accessibility. The path of available energy should reach the crossing region to take place. Otherwise, if a too high energy barrier hinders the access to the crossing, the feature could be totally ineffective.

Electronic Structure Calculations

The development of Quantum Mechanics was spread over by Erwin Schrödinger, Werner Heisenberg, and Paul Dirac in the 1920s. The wave and particle aspects of matter are reconciled by the Schrödinger equation for stationary states, $H\Psi = E\Psi$. The Hamiltonian operator, $\mathbf{H}$, is associated to the total energy of a physical system and is the sum of the kinetic energy and the potential energy operators associated with electrons and nuclei ($\mathbf{H} = \mathbf{T}_e + \mathbf{T}_N + \mathbf{V}_{NN} + \mathbf{V}_{Ne} + \mathbf{V}_{ee}$). This is an eigenvalue problem, in which wave functions Ψ are the eigenfunctions of $\mathbf{H}$ and E stands for the corresponding eigenvalues (energies). The main challenge in Quantum Chemistry is that we cannot solve exactly the Schrödinger equation, except for one-electron systems, due to the electron repulsion term present in the Hamiltonian. Quantum-chemical methods look for approximate solutions of the equation, employing computational numerical methods typically based on the variational principle and perturbation theory. A point worth bearing in mind is that none of these models is applicable under all circumstances. Actually, we should get the best method in order to find what it has been wisely defined as "the right answer for the right reason."

The physics of electron correlation is hidden in the Hamiltonian itself. The Coulomb repulsion given by the term r^{-1} present in the $\mathbf{V}_{ee}$ energy, the inverse distance between two electrons, increases enormously in the regions close to $r_{ij} = 0$, preventing that two electrons may occupy the same space. Therefore, the motion of any two electrons is not independent but it is correlated. The phenomenon is known as electron correlation. Moreover, the statement that two electrons are correlated is equivalent to express that the probability of finding two electrons at the same point in space is zero. The instantaneous position of electron i forms the center of a region that electron j will avoid. For this reason, it is stated that each electron, as described by the exact wave function Φ, is surrounded by a Coulomb hole. However, electron correlation is not taken into account properly by many approximate methods. The effect of neglecting electron correlation partly in approximate quantum-chemical approaches has great impact in the computed molecular spectroscopic properties of interest.

General Overview

We can group computational–chemical methods in three basic categories (see for instance Atkins and Friedman 1997; Helgaker et al. 2004):

- Ab initio methods, in which the complete Hamiltonian is used, all the integrals are solved numerically, and no essential parametrization is employed

- Semiempirical methods, in which a simpler Hamiltonian is used or integrals are adjusted to experimental values or ab initio results
- Molecular mechanics, in which we solve Newton's equation of motion, only valid for situations where no bonds are broken or formed, i.e., conformational changes

Obviously, the larger is the system under study, the less accurate is the available method. Despite their inherent drawbacks, classical semiempirical methods are still employed in large systems, whereas modern semiempirical methods, based on the Density Functional Theory, have a widespread use. A combined approach, QM/MM (Quantum Mechanics/Molecular Mechanics) treats an internal part of the problem with QM methods, whereas the surroundings or a large part of a macromolecule (for instance, a protein) is introduced using classical mechanics.

According to the number of configurations used to build the reference wave function, the ab initio methods can be classified in the following two categories (see Fig. 6):

- *Single-configuration methods.* They are typically based in the single Hartree–Fock (HF) reference, which determines the optimal ground-state energy and MOs (molecular orbitals). Post-HF methods introduce the electron correlation usually at the Configuration Interaction (CI), Coupled-Cluster (CC) or perturbative (PT) Møller–Plesset (MP, or PT in general) levels. The coupled-cluster methods with singly and doubly excited configurations, including the effect of triple excitations by perturbation theory CCSD(T), as well as related approaches, yield accurate results in well-defined ground-state situations and are considered as benchmark

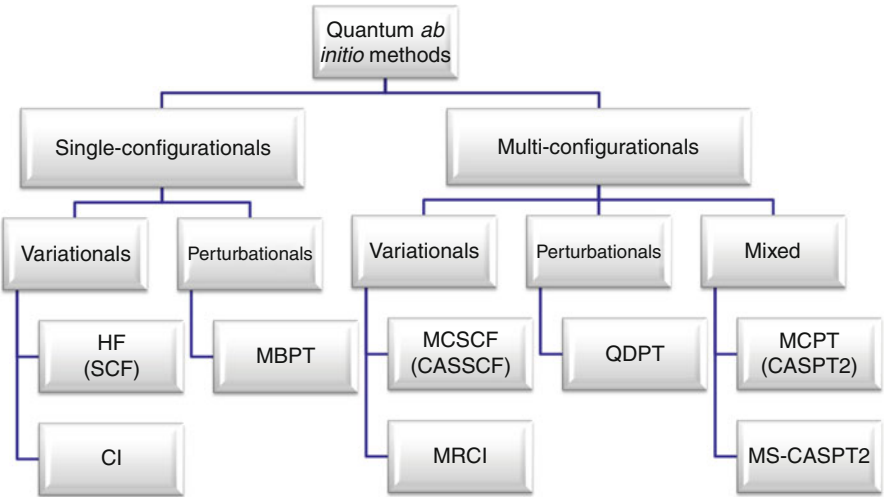

Fig. 6 Ab initio methods

results for small to medium molecules. In general, the applicability of the methods in this group is restricted to situations where a single reference wave function is adequate for the description of a chemical process, something not generally true for bond-breaking cases, degeneracies, and excited states

- *Multiconfigurational methods.* Part of the electronic correlation is already included in the reference wave function, normally by using a Multiconfigurational Self-Consistent-Field (MCSCF) wave function, which determines a set of MOs. The remaining electron correlation effects are accounted for by MRCI, MRCC, or MRPT techniques, where MR stands for multireference. They have a more ample range of applicability (ground state, excited states, transition states) than single-reference methods

The variation principle states that given a normalized wave function that satisfies the appropriate boundary conditions, then the expectation value of the Hamiltonian is an upper bound to the exact ground-state energy. In the linear variational problem, the trial function is a linear combination of basis functions, in general using the Linear Combination of Atomic Orbitals (LCAO) approach. On the other hand, in perturbation theory, the total Hamiltonian of the system is divided into two pieces: a zeroth-order part, which has known eigenfunctions and eigenvalues, and a perturbation part. The exact energy and wave function are then expressed as an infinite sum of contributions of increasing complexity. If we have chosen the zeroth-order Hamiltonian wisely, then the perturbation is small and the expansion (i.e., the sum of the 1st-, 2nd-, …, *n*th-order energies) converges quickly.

The simplest wave function to describe a many-electron system is a Slater determinant built by orthogonal one-electron wave functions. Electrons are fermions and accordingly they have to be described by an antisymmetric wave function. For an N-electron system, the Slater determinant has the form:

$$\Psi\left(x_1, x_2, \ldots, x_{\mathrm{N}}\right) = (\mathrm{N}!)^{-1/2} \begin{vmatrix} \chi_i\left(x_1\right) & \chi_j\left(x_1\right) & \ldots & \chi_k\left(x_1\right) \\ \chi_i\left(x_2\right) & \chi_j\left(x_2\right) & \ldots & \chi_k\left(x_2\right) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_i\left(x_{\mathrm{N}}\right) & \chi_j\left(x_{\mathrm{N}}\right) & \ldots & \chi_k\left(x_{\mathrm{N}}\right) \end{vmatrix} = \left|\chi_i \chi_j \cdots \chi_k\right\rangle \tag{2}$$

without specifying which electron is in which orbital. To simplify the notation, a normalized Slater determinant is represented by only showing the diagonal elements of the determinant. The constant $(\mathrm{N}!)^{-1/2}$ is a normalization factor. The wave function for an electron that describes both the spatial distribution and its spin is called spin orbital, $\chi_i(x_i)$. Since the Hamiltonian employed does not depend on the electronic spin, each spin orbital can be expressed by multiplying the spatial orbital, $\psi_j(r_i)$, by the spin function, ω (α – spin up and β – spin down): $\chi_i\left(x_i\right) = \psi_j\left(r_i\right) \cdot \omega$. A complete set for describing the spin of an electron consists of two orthogonal functions $\alpha(\omega_i)$ and $\beta(\omega_i)$.

A single-determinant wave function has several interesting properties. Firstly, it is worth noting that spin orbitals must be linearly independent, otherwise the value

of the determinant is zero. It is obvious that interchanging two rows of the Slater determinant, which is equivalent to interchanging the coordinates of two electrons, changes the sign of the determinant. The requirement of the antisymmetry principle is automatically fulfilled. Having two columns of the determinant identical, that is, two electrons occupying the same spin orbital, makes the determinant zero. Thus, no more than one electron can occupy a spin orbital (Pauli exclusion principle). When a linear transformation of the set $\{\chi_i\}$ is carried out, $\chi_i' = \sum_j \chi_j \cdot A_{ji}$, where A_{ji} is an element of the matrix $\mathbf{A}$ of dimension $\mathrm{N} \times \mathrm{N}$, with a value for its determinant, $\det(\mathbf{A})$, different from zero, then $\Psi' = \det(\mathrm{A})\,\Psi$. The wave functions Ψ' and Ψ differ just in a constant and, therefore, represent the same physical situation. Since the set of spin orbitals is linearly independent, we can always choose a transformation matrix $\mathbf{A}$ so that the resulting spin orbitals χ_i' become orthonormal. Therefore, no restriction at all is imposed when we choose from the beginning an orthonormal set of spin orbitals. It just makes the computation of the Hamiltonian matrix elements involving Slater determinants easier. A Slater determinant is completely specified by the spin orbitals used to build it, and any unitary transformation of them is equally valid. Two sets of spin orbitals related by a unitary transformation $\left(\mathbf{A}^{\dagger} = \mathbf{A}^{-1}\right)$, which keeps the orthonormality of the spin orbitals, yield the same Slater determinant. Slater determinants formed from orthonormal spin orbitals are normalized and N-electron Slater determinants that have different spin orbitals are orthogonal. In other words, a Slater determinant is completely specified by the spin orbitals used to build it and any unitary transformation of them is equally valid (each result of such a linear transformation represents the same physical situation).

There are two types of spin orbitals: restricted spin orbitals, which are constrained to have the same spatial function for α and β spin functions; and unrestricted spin orbitals, which have different spatial functions for α and β spins. A restricted set of spin orbitals has the form: $\chi_i(x) = \psi_j(r) \cdot \alpha(\omega)\,//\,\psi_j(r) \cdot \beta(\omega)$, whereas an unrestricted set has the form: $\chi_i(x) = \psi_j^{\alpha}(r) \cdot \alpha(\omega)\,//\,\psi_j^{\beta}(r) \cdot \beta(\omega)$.

Essentially, all practical calculations for generating solutions to the electronic Schrödinger equation are performed with molecular orbital methods: the zeroth-order wave function is constructed as one or two Slater determinants and the MOs are expanded in a set of atomic orbitals, the basis set. In a subsequent step, the wave function may be improved by adding electron correlation by either CI, MP, or CC methods. Nevertheless, there is another equivalent theory to get approximate solutions of this eigenvalue equation: the valence bond (VB) theory. The main drawback is that this theory leads to awkward calculations. On the positive side, conceptually it is much closer to the experimentalist's language since it may be regarded as a quantitative version of the resonance structure of electronic structure, with ionic structures and covalent structures of molecules. The VB method assumes that the wavefunction of a molecule may be written as a linear superposition of mathematical functions which represent canonical electronical structures: structures in which electrons are assigned to specific atoms and then paired (spinwise) leading to covalent and ionic structures, the combination of which represents the wave

function. The VB description of a bond as the result of two overlapping and localized orbitals is in contrast to the MO approach where a bond between two atoms arises as a sum over (small) contributions from many delocalized orbitals.

Classical valence bond theory is very successful in providing a qualitative explanation for many aspects. One of the great merits of VB theory is its pictorially institutive wave function that is expressed as a linear combination of chemically meaningful structures. It is this feature that has made VB theory so popular in the 1930–1950s. However, VB theory was "defeated" by MO theory for two main reasons: (1) mathematical simplicity and (2) several "failures" of the VB theory (actually due to misuse of very simplified versions of the theory). Currently, VB theory is coming of age in all the branches of theoretical chemistry (excited states, dynamics, environmental effects, and so on), with the development of faster, and more accurate ab initio VB methods, and with generation of new post-Pauling concepts. The Renaissance of VB theory is marked by surge in the following two-pronged activity: (1) creation of general qualitative models based on VB theory and (2) development of new methods and softwares that enable applications to moderate sized molecules.

Methods, Advantages, and Drawbacks

The Hartree–Fock method is the simplest ab initio approach. We can equate closed-shell HF theory to single determinant theory, and we are thus interested in finding a set of spin orbitals $\{\chi_a\}$ such that the single determinant formed from them, $|\Psi_0\rangle = |\chi_1\chi_2\chi_3\cdots\chi_a\chi_b\cdots\chi_N\rangle$ is the best possible approximation to the ground state of the N-electron system described by an electronic Hamiltonian $\mathbf{H}$. The expectation value of the energy, $E_0 = \langle\Psi_0|H|\Psi_0\rangle$, is a linear combination of one-electron integrals, $\langle i|h|j\rangle$ and two-electron integrals, $\langle ij|kl\rangle$. According to the properties of a Slater determinant, electrons are not uncorrelated in a HF wave function. At least the probability of finding two electrons with parallel spins at the same point in the space is zero, the so-called exchange correlation, which is incorporated by the antisymmetric condition of the wave function for fermions. The phenomenon is known as the Fermi hole. We are, therefore, facing a model of independent particles where the behavior of certain electrons is not fully independent because the Fermi hole simulates somehow the Coulomb hole. Consequently, only Fermi correlation is accounted for by the HF wave function.

By minimizing E_0 with respect to the choice of spin orbitals $\{\chi_a\}$, varying them with the only restriction that they remain orthonormal, $\langle\chi_a|\chi_b\rangle = \delta_{ab}$, one can arrive to the Hartree–Fock conditions. In doing so, one obtains an equation that defines the best spin orbitals, the ones that minimize E_0. This equation for the best (Hartree–Fock) spin orbitals is the Hartree–Fock integro-differential equation $f\,|\,\chi_a\rangle = \varepsilon_a|\,\chi_a\rangle$, for the N-occupied spin orbitals. Each of the solutions $\{\chi_j\}$ has a spin orbital energy ε_j. The N spin orbitals with the lowest orbital energies are just the spin orbitals occupied in $|\Psi_0\rangle$ for which we use the indices a, b, etc. The remaining

infinite number of spin orbitals with higher energies are the virtual spin orbitals, which we label with the indices *r*, *s*, etc.

The Fock operator f is the sum of a core-Hamiltonian operator $h(1)$, which is the kinetic energy and potential energy for attraction to the nuclei of a single electron, and an effective one-electron potential called the Hartree–Fock potential $v^{\mathrm{HF}}(1)$, which contains the Coulomb term (the total averaged potential acting on the electron in χ_a, arising from the $\mathrm{N}-1$ electrons in the other spin orbitals) and the exchange term (which arises because of the antisymmetric nature of the determinantal wave function). This is the essence of the HF approximation: to replace the complicated many-electron problem by a one-electron problem in which electron–electron repulsion is treated in an average way. What is more, since the Fock operator has a functional dependence, through the coulomb and exchange operators, on the solutions $\{\chi_a\}$ of the pseudo-eigenvalue equation, Hartree–Fock equations are actually nonlinear equations and will need to be solved by iterative procedures. The matrix representation of the Fock operatorin the basis of spin orbitals eigenfunctions is diagonal with diagonal elements equal to the orbital energies.

The exact solution to this integro-differential equation corresponds to the "exact" HF spin orbitals. In practice, it is only possible to solve this equation exactly (i.e., as an integro-differential equation) for atoms. In practice, for molecules, the spin orbitals are expanded as a combination of one-electron atomic base functions and the set of matrix equations are solved. Only as the basis set approaches completeness, i.e., as one approaches the HF limit, will the spin orbitals be the exact HF orbitals.

It can be shown that any single determinant wave function $|\Psi_0\rangle$ formed from a set of spin orbitals $\{\chi_a\}$ retains a certain degree of flexibility in the spin orbitals; the spin orbitals can be mixed among themselves without changing the expectation value $E_0 = \langle \Psi_0 | H | \Psi_0 \rangle$. Furthermore, for a single determinant wave function, any expectation value is invariant to an arbitrary unitary transformation of the spin orbitals, since a transformed single determinant $|\Psi_0'\rangle$ can at most differ from the original determinant $|\Psi_0\rangle$ by a phase factor and, obviously, any observable property depends on $|\Psi|^2$, as we stated before. Thus the spin orbitals that make the total energy stationary are not unique, and no particular physical significance can be given to a particular set of spin orbitals. Localized spin orbitals, for instance, are not more "physical" than delocalized spin orbitals. Indeed, there exist a set of spin orbitals for which the eigenvalue matrix is diagonal, and we obtain just the equation $f\,|\,\chi_a\rangle = \varepsilon_a|\,\chi_a\rangle$, and this set is called "canonical spin orbitals." The canonical spin orbitals, which are a solution to this equation, will generally be delocalized and form a basis for an irreproducible representation of the point group of the molecule.

It was in 1951 that Roothaan published his equations (see Roothaan 1951), considering molecular orbitals that were restricted to be linear combinations of a set of three-dimensional one-electron functions, ϕ_μ. Thus, ψ_i is a linear combination of these functions ϕ_μ (basis set): $\psi_i = \sum C_{\mu i} \cdot \phi_\mu$. If the set $\{\phi_\mu\}$ was complete, this would be an exact expansion, and any complete set $\{\phi_\mu\}$ could be used. Unfortunately, this one is always restricted, for practical computational reasons,

to a finite set of N basis functions. As the basis set becomes more and more complete, this expansion leads to more and more accurate representations of the "exact" molecular orbitals, which are eigenfunctions of the Fock operator. For any finite basis set, we will obtain molecular orbitals which are exact only in the space spanned by the aforesaid finite basis. In addition, since the electronic energy (the output of the quantum mechanical calculation) is variational, the better the basis set, the lower the total energy. As the basis set becomes more and more complete, the total energy approaches the Hartree–Fock limit. Of course, by the variational principle, the Hartree–Fock-limit energy is still above the "exact" energy, which here can be taken as the energy obtained form an exact solution to the nonrelativistic Schrödinger equation in the Born–Oppenheimer approximation.

Variation of the total energy was then carried out respect to the coefficients $C_{\mu i}$ of such a linear combination. This leads to a set of algebraic equations which can be written in matrix form, $\mathbf{FC} = \mathbf{SCE}$, where **F** is the matrix representation of the Fock operator (which can be divided into the core-Hamiltonian matrix, that is, integrals involving the one-electron operator $h(i)$ describing the kinetic energy and nuclear attraction of the electron i, plus the two-electron terms), **C** is the matrix of coefficients, **S** is the overlap matrix, and **E** is the diagonal energy matrix. The columns of **C** describe the molecular orbitals, i.e., the coefficients describing ψ_1 are in the first column, with the corresponding energy ε_1, those describing ψ_2 are in the second column, and so on. If **S** is the unit matrix (i.e., if we have an orthonormal basis set), then we would have $\mathbf{FC} = \mathbf{CE}$, and Roothaan's equations would just have the form of the usual matrix eigenvalue problem. Increasing the flexibility of the one-electron basis set $\{\phi_\mu\}$, the HF energy E_0 will progressively reach the Hartree–Fock limit.

The approximation has its limitations, in particular the lack of electronic correlation energy. For instance, Restricted Hartree–Fock (RHF) solution predicts incorrectly the dissociation of molecules into open-shell fragments (like $H_2 \rightarrow 2H$) given that both electrons are forced to occupy the same spatial molecular orbital (MO), when they should belong to infinitely separated atoms in the dissociated solution. This means that some electron–electron repulsion remains even at infinity, which is the source of a spurious term in the total energy (different from twice the energy of two hydrogen atoms). Alternatively, the products of dissociation are not just 2H·, but also include, incorrectly, H^- and H^+. However, the poor behavior of closed-shell RHF calculations upon dissociation to open-shell products does not detract from their utility as reference function in the region of equilibrium: the potential surface obtained from a closed-shell RHF calculation will not be parallel in regions of the surface characterizing dissociation limit, but it will be reasonably parallel near the region of the equilibrium geometry. It is worth mentioning that the VB wave function for H_2, $\Psi = C_{\text{cov}}\Psi_{\text{cov}} + C_{\text{ion}}\Psi_{\text{ion}}$, with the covalent and ionic counterparts, contains "left-right" correlation which is necessary for a correct description of dissociation of the H_2 molecule.

Within the HF approach, a slight improvement can be achieved in dealing with open-shell systems with the restricted open-shell (ROHF) and the unrestricted Hartree–Fock (UHF) procedures. In the former, all electrons, except those that

are explicitly required to occupy open-shell orbitals, occupy closed-shell MOs. In the latter, electrons of different spin, in general, are described by different sets of spatial orbitals, although the obtained wave function is typically contaminated with higher spin state solutions. The approximation leads to two integro-differential eigenvalue equations which are coupled and cannot be solved independently. The introduction of a basis leads to the Pople–Nesbet equations, $\mathbf{F}^{\alpha}\mathbf{C}^{\alpha} = \mathbf{S}^{\alpha}\mathbf{C}^{\alpha}\mathbf{E}^{\alpha}$ and $\mathbf{F}^{\beta}\mathbf{C}^{\beta} = \mathbf{S}^{\beta}\mathbf{C}^{\beta}\mathbf{E}^{\beta}$. $\mathrm{N}^{\alpha} = \mathrm{N}^{\beta}$, a restricted solution to the Roothaan equations is a solution to the unrestricted Pople–Nesbet equations. In general those approaches, which can serve as reference for post-HF methods in ground-state cases, are not accurate enough to compute excited states.

The lack of correlation is the actual source of all errors. In particular, a Slater determinant incorporates exchange correlation, i.e., the motion of two electrons with parallel spins is correlated (the so-called Fermi correlation). Unfortunately, the motion of electrons with opposite spins remains uncorrelated. It is common to define correlation energy, E_{corr}, as the difference between the exact nonrelativistic energy of the system, ε_0, and the Hartree–Fock energy, E_0, obtained in the limit that the basis set approaches completeness: $E_{\mathrm{corr}} = \varepsilon_0 - E_0$. The simplest manner to understand the inclusion of the correlation effects is through the method of configuration interaction (CI). The basic idea is to diagonalize the N-electron Hamiltonian in a basis of N-electron functions: we represent the exact wave function as a linear combination of N-electron trial functions and use the linear variational method.

$$|\Phi\rangle = C_0|\ \Psi_0\rangle + \sum_{ra} C_a^r|\ \Psi_a^r\rangle + \sum_{a<b;r<s} C_{ab}^{rs}|\ \Psi_{ab}^{rs}\rangle + \ldots. \tag{3}$$

If the basis were complete, we would obtain the exact energies of all the electronic states of the system. In spite of providing the exact solution of a many-electron problem, we can handle only a finite set of N-electron trial functions. As a result, the CI method provides only upper bounds to the exact energies. Specifically, the lowest eigenvalue, ε_0, will be an upper bound to the ground-state energy of the system. When all the N-electron wave functions are taken into account, the calculation is named full configuration interaction (FCI) and the corresponding eigenvalues and eigenvectors computed are exact within the space spanned by the finite basis set. Despite the great advances in FCI technology in the last few years, the size of the eigenvalue problem becomes rapidly too large to be handled by modern computers. As a result, FCI solutions are only available for very small molecular systems. In contrast to HF, the FCI energy of H_2 properly describes the dissociation. However, within a minimal basis set approach, the FCI potential curve does not agree very well with the exact one provided by Kolos and Wolniewicz due to the lack of flexibility of such a basis set. The truncation of both N-electron basis and one-electron basis is the main source of inaccuracies in quantum-chemical calculations (Fig. 7).

As mentioned earlier, for computational reasons, we have to truncate the full CI matrix or equivalently the CI expansion for the wave function, considering only those configurations which differ from the HF ground-state wave function

N-electron basis set

$$\langle \Phi \rangle = C_0 |\Psi_0\rangle + \sum_{ra} C_a^r |\Psi_a^r\rangle + \sum_{a<b;r<s} C_{ab}^{rs} |\Psi_{ab}^{rs}\rangle + \ldots$$

1-electron basis set

$$\langle \Phi_i \rangle = |\chi_i \chi_j \cdots \chi_N\rangle \qquad \chi_i = \phi_i \vec{\omega} \,/\, \vec{\omega} = \alpha, \beta$$

$$\phi_i(\vec{r}) = \sum_{\mu=1}^{K} C_{\mu i} \phi_\mu(\vec{r}) \qquad \phi_\mu(\vec{r}) = \sum_K d_{K\mu} g_K$$

Fig. 7 Many-electron expansion (CI) and one-electron expansion (basis set). The total wave function, Φ, is a linear combination of N-electron wave functions Ψ_0, Ψ_a^r, etc. Each one of these functions is an antisymmetrized and normalized product of spin orbitals, χ_l. Each of them is constituted by a one-electron wave function, ψ_l, and a spin function, ω. Each one-electron wave function is defined as a linear combination of a set of basis functions, ϕ_μ, which are used to be contracted gaussian functions, CGTFs (linear combinations of a set of primitive functions, g_K)

by no more than a predetermined number of spin orbitals. The simplest version of this scheme is to deal with single and double excitation out of $|\Psi_0\rangle$: a SDCI calculation. SDCI and, in fact, all forms of truncated CI deteriorate as the number of electrons increases. The truncation of the CI expansion leads to problems of size-extensivity. For instance, in a DCI calculation, by definition the wave function of each monomer contains double excitations within the monomer. If we restrict the supermolecule trial function to double excitations, we exclude the possibility that both monomers are simultaneously doubly excited, since this represents a quadruple excitation in the supermolecule. It is now common to differentiate size-extensivity and size-consistency. The latter was originally employed by Pople as a criterion for a well-constructed quantum-chemical method, indicating that the energy computed for two noninteracting molecules should be exactly twice that calculated for only one isolated molecule, that is, a property that described the additive separability of the wave function. Recently, the concept has been extended to include not only the fragmentation limit but the entire process, that is, qualitatively all regions of the potential energy hypersurface must be qualitatively correct. For instance, both RHF and UHF wave functions are size consistent in the first sense, considering that they properly describe a separated H_2 dimer system, but for a closed-shell molecule dissociating intoopen-shell fragments the RHF descriptions fails, and therefore does not conform to the broader size-consistency concept. On the other hand, size extensivity, analogously to the extensivity concept in thermodynamics, refers to the correct, linear, scaling of a method with the number of electrons, and its fulfillment leads to methods in which calculations with differing number of electrons can be compared, like those related to ionization processes. Otherwise, the error of the method increases with the size of the system. At the noninteracting limit size extensivity implies size consistency, but this latter property has the requirement of correct fragmentation, which does not depend on the mathematical scaling. Therefore, RHF and UHF approaches are always size extensive, but RHF does not provide

proper dissociation in open-shell cases, and it is therefore non-size-consistent, as any other method such as single-reference coupled-cluster (CC) or perturbation approaches (PT) which use RHF as reference wave function. On the other hand any truncated configuration interaction approximation is not size extensive, whereas CC and PT are. In the case of the multiconfigurational methods, the size-consistency depends on the selected reference – generally they are size consistent – and the size extensivity varies with the approach. MRCI is formally a truncated CI approach and therefore is not size extensive, as CASPT2 – multiconfigurational complete active space perturbation theory up to second order – and similar methods, which are formally not size extensive either, but in practice the effects are irrelevant, in particular for the most common implementation (see Taylor 2007).

It is worth mentioning that using a better-variational method only ensures to get a better variational energy, but a priori we cannot say nothing regarding properties other than the energy. For instance, the dipole moment of CO is ca. zero, and to obtain the correct sign of the vector (actually, greater than zero: +0.044 a.u. experimentally) is a tricky problem. A DCI calculation with the 200 most important double excitations has a lower energy than a SDCI procedure with 138 doubles + 62 singles. Still, the DCI calculation predicts the wrong sign of the dipole moment (−0.072 a.u.), just the contrary that the SDCI calculation (+0.030 a.u.). To improve the result it is necessary to employ CCSD(T) (+0.047 a.u.) (Paschoal et al. 2009).

There is another clever way to introduce in the wave function enough correlation effects to get a good account of the nature of the chemical problem, for instance, in a dissociation path of a diatomic molecule. Such procedure implies to improve the reference wave function by including more than one determinant or electronic configuration in its description. Adding more than one configuration means to get their weights in a linear combination. Therefore, the central idea of a multiconfigurational self-consistent field (MCSCF) calculation is to build the wave function as a truncated CI expansion

$$|\Psi_{\mathrm{MCSCF}}\rangle = \sum_I c_I |\ \Psi_I\rangle, \tag{4}$$

in which both the expansion coefficients and the orthonormal orbitals contained in $|\Psi_I\rangle$ are optimized simultaneously. For a closed-shell system, if only one determinant is included in such an expansion, the MCSCF and HF methods become identical. In the minimal basis set description of H_2, Ψ_{MCSCF} is identical to the FCI wave function. However, if an extended basis set is used, the MCSFC energy will be above and nearly parallel to the FCI energy but below the energy obtained from any two-configuration CI expansion based on canonical HF MOs.

The correlation energy introduced in a MCSCF wave function is usually named nondynamic, static, or large-range correlation energy, and the corresponding wave function must be build in such way that includes most of near-degenerate electronic configurations, that is, all those which basically contribute to define the reference function along the chemical process under study. An example is a bond breaking. A proper MCSCF description should include both the bonding and antibonding dis-

tribution of the electrons, and the corresponding MOs. The optimization procedure at each region of the PEH will determine the relative weight of each configuration at that point and the proper MOs. Improving the reference in such manner helps to the post-HF approaches to account for the remaining correlation effects, named dynamic or short-range. It is like starting the race from a more advantageous position.

One of the most successful and systematic procedures to account for the correlation energy, which is not variational but is size extensive at each order, is the perturbation theory. Using different zeroth-order Hamiltonians, Møller–Plesset perturbation theory (MPPT), also named many-body perturbation theory (MBPT), yields the energy and the wave function corrected up to n order following a Taylor series expansion. Although, in general, higher perturbation orders improve the result, it has been proved that it leads to divergences at infinite order. Second- and fourth-order perturbation (MP2, MP4) results are much better than the SCF and SDCI ones. The most recent family of methods based on a single reference is that based on the coupled-cluster (CC) approach, in which the correlation energy is introduced size-extensively by increasing the excitation level. CC methods for the ground state, especially when including up to triple excitations, have become, in practice, the most accurate quantum-chemical approaches for many systems.

No matter how good is the quality of the post-HF method for the inclusion of the correlation energy, when based on a single HF reference, the accuracy of the results largely decreases in many situations, like those related with dissociations, degeneracies, and excited states. For instance, regarding the dissociation problem, the potential curve calculated at any level based on a RHF reference is not satisfactory at long range. The problem is the starting point: the exact wave function at large R cannot be described by a single determinant. The results largely improve its accuracy when a multiconfigurational or multireference (MR) wave function is used as a reference, for instance, in the MRCI, MRCC, or MRMP approaches. Recovering correlation energy does not require so large effort in this case, for instance, lower perturbation (MP) or excitation (CC) levels. As a drawback, the complexity of the methods increases.

On the other hand, the semiempirical DFT approach has become the most widespread quantum-chemical method in the last years. It allows us to treat large systems in a relatively short time taking into account electron correlation. The basic idea is that the energy of an electronic system can be written in terms of the electron density, ρ. For a system of N electrons, $\rho(\mathbf{r})$ denotes the total electron density at a particular point in space $\mathbf{r}$. The electronic energy E is said to be a functional of the electron density, denoted $E[\rho]$, in the sense that for a given function $\rho(\mathbf{r})$, there is a single corresponding energy. It is not necessary to compute a wave function, but, unfortunately, the Hohenberg–Kohn theorem does not inform about the form of the functional dependence of energy on the density: it confirms only that such a functional exists. So the procedures should resort to approximate derivations with adjustable parameters. This is the reason why functionals proliferate to hundreds in the literature. Some of them, like the Becke3-Lee-Yang-Parr (B3LYP) functional, is the most popular in the literature, not because its overall accuracy but of its balance

and its controlled flaws. There is no systematic way to improve a functional, just to derive another one with different parameters. It is typical to use one functional per each situation, something that makes the predictability of the results somewhat questionable. On the contrary, ab initio calculations, if allowed by the size of the problem, can be hierarchically improved in a known direction enlarging the N-electron and/or the one-electron basis.

Basis Sets

The choice of the one-particle space is a most important decision when setting up any calculation, and there is no point in trying to improve the result if the selection of the one-electron basis set is not adequate. This is especially true for the calculation of excited states, in which states of very different nature (for instance, compact and diffuse) have particular requirements that must be fulfilled simultaneously when selecting the basis set. In a strict mathematical sense, many different types of basis set functions $\{\phi_\mu\}$ can be used. Gaussian-type functions (GTFs, $g \propto e^{-\alpha r^2}$) are the more widespread in spite of being ill-behaved in both close and far away of the nuclei. Nevertheless, four-center integrals (that is, two-electron repulsion integrals) are very easy to evaluate with GTFs, and, in addition, the basis set can be improved by employing contracted Gaussian functions, CGTFs, which are linear combinations of primitive GTFs. To build a CGTFs, the primitive functions have to be optimized, both exponents and contraction coefficients. According to the number of basis functions, it is common to distinguish among minimal basis set (a single basis function per occupied atomic orbital), double-ζ (two basis functions per atomic orbital), triple-ζ, split-valence (duplication only affects to valence layer whereas minimal basis set is used for core orbitals), etc.

Ground-state calculations have made extensive use of the so-called Pople basis sets: STO-G, 3-21G, 6-31G. These basis sets are built using segmented contraction, which means that each primitive functions contributes to one or a limited number of contracted function only. Exponents and coefficients were optimized using the HF approach. Valence basis sets are too limited to get accurate results, and therefore additional functions have to be added, named polarization functions (6-31G* or 6-31G(d), which means that polarization functions are added to second-period atoms; and 6-31G** or 6-31G(d,p), which means that, additionally, polarization functions are added to hydrogen atoms). These are functions of higher angular momentum than those of valence type and are necessary in order to describe the changes of the electronic density of an atom in the molecule. Diffuse functions (6-31+G, 6-31++G) can be also added which are more extended in the space, that is, with smaller exponents, to describe better situations such as anions or Rydberg states. A 6-31G basis set for the C atom ($1s^2\ 2s^2\ 2p^2$) is described as (10s4p)/[3s2p], i.e., 10 primitives s and 4 primitives p (22 primitives: $1s \times 6 + 2s \times 3 + 2s \times 1 + 2p \times 3 + 2p \times 1$) are contracted to three combined functions s and two combined functions p (nine contracted functions: 1s, 2s, 2s′,

2p, 2p′). The basic flaw of Pople's basis sets is that they have been built with a method lacking correlation energy. The consequence is that they are very poor in recovering correlation effects in post-HF methods. In order to get better values, more polarization of diffuse functions have to be added to compensate the lack of flexibility of these basis set.

If the ultimate goal is to perform correlated calculations, it would seem preferable to include correlation in the construction of the basis set. For correlated calculations, the basis set requirements are different and more demanding since we must also provide a virtual orbital space capable of recovering a large part of the correlation energy. It is often sufficient to correlate only the valence electrons. This is the case for correlation consistent (cc) or Atomic Natural Orbitals (ANO) basis sets. Correlation consistent basis sets are designed so that functions which contribute in similar amounts of correlation energy are included at the same stage, independently of the function type. We can add polarization functions (cc-pVDZ \(\ldots\)) or diffuse functions (augmented, aug). They need fewer primitives than ANOs, and each contracted basis comes from a different primitive set and the exponents of polarization functions are optimized by correlated calculations within a segmented contraction scheme.

On the other hand, ANO-type basis sets arise as eigenfunctions of the first-order reduced density matrix of the atom. They give the most rapidly convergent CI expansion, and to obtain a given accuracy one requires fewer configurations formed from NOs than configurations formed from any other orthonormal basis. The occupation number is a reference of its significance in the wave function and the truncation of the orbital space by elimination of the ANOs whose occupation numbers are small produces the least possible error on the wave function. The coefficients of ANOs obtained in a correlated calculation of the ground state of the atom represent the coefficients of the basis functions within a general contraction scheme (all primitives on a given atom and of a given angular momentum enter all the contracted functions having that angular momentum, but with different contraction coefficients, which improves the flexibility). This yields basis of any size with a unique calculation with the primitive set, and since the same primitive set for all contractions is used, a smaller basis is always a subset of a larger basis. In other words, we can enlarge our basis set without changing of vector space (Almlöf and Taylor 1986, 1991). An improvement of the original scheme takes into account an average density matrix to build the basis set (including ground state, excited state, cation, anion, atom into an electric field, see Widmark et al. (1990)). In general, ANO basis set will require less contracted functions than the other basis to get the same results. Recently, all electron ANO basis sets, including correlation effects for all periodic system elements, have been generated (Roos et al. 2005). It has also been observed that for extremely accurate calculations, only ANO basis sets can provide the best answer (see, e.g., Martin et al. 1997).

A final type of approach regarding the one-electron basis sets is the combination of inner-shell pseudopotentials (including typically relativistic effects) with valence basis sets and the use of embedded potentials to represent atomic environments.

Methods for Excited States

Based on the methods described in the previous section, specific algorithms to deal with electronic excited states have been developed. In this section, the most commonly employed approaches will be briefly summarized (see also Merchán and Serrano-Andrés 2005; Serrano-Andrés and Merchán 2005).

Starting by the single-configurational methods that use the HF solution as reference wave function, an approach which has lost popularity because of its poor performance is the method named Configuration Interaction-Singles (CIS). The essence of the method is to consider that a excited state can be described by a singly excited determinant formed by replacing, with respect to the reference wave function, an occupied spin orbital with a virtual spin orbital. The drawbacks of such a description may be partially compensated if a linear combination of all possible single excited determinants is used to build the excited state wave function. In general, the CIS excitation energies are largely overestimated due to the absence of correlation energy effects. It is more common to find in the literature cases in which CIS yields the wrong order and state nature than the proper ones, simply because the differential correlation energy affects the excited states unevenly and because the intrinsic character of the states is multiconfigurational (Foresman et al. 1992).

In the light of the previous discussion, the logical method to deal with excited states should be multireference CI (MRCI) (Buenker et al. 2000). Consider a wave function of CI type expanded in a many-electron basis set of determinants. As in the H_2 molecule, it is possible to select a number of determinants to describe the correct dissociation limit. When the energy is minimized with respect to the coefficients of the expansion, the configuration interaction (CI) method is employed. It should be kept in mind that actual calculations are performed using either spin-adapted Configuration State Functions (CSFs: appropriate linear combination of Slater determinants) or determinants as a N-electronic basis set (Slater determinants lead to more efficient CI algorithms), whereas CSFs lead to shorter CI expansions and, obviously, spin eigenfunctions, i.e., singlet, doublet, triplet states . . .). In case that the expansion contains more than one configuration, the process is denoted as MRCI. The wave function $|\Psi\rangle$ is a multireference function and, at least, the singly and doubly excited determinants generated from each reference determinant $|m\rangle$ are taken into account. Unfortunately, MRCI-based approaches are restricted to molecular systems of small molecular size (containing up to 6–7 atoms), where dynamic correlation can be fully retrieved. Otherwise, truncation of the many-electron functions involved in the MRCI expansion might easily lead to large errors in the computed excitation energies.

Another type of methods applied to the calculation of excited states are the propagator approaches (Oddershede 1987). The underlying technique, also called Green's function approach, equation-of-motion or linear response theory in its different forms, can be applied to various types of methodologies: single- or multi-reference configuration interaction, coupled-cluster, or density functional.

The essentials of the technique consider that once a molecule is subjected to a linear time-dependent electric field fluctuating with frequency ω, a second-order property as the frequency-dependent ground-state polarizability of the system is well approximated by

$$\alpha_\omega = \sum_{i\neq 0}^{\text{states}} \frac{|\langle \Psi_0 | r | \Psi_i \rangle|^2}{\omega - \Delta E_i}, \tag{5}$$

where the denominator of the expression involves the frequency of the field and the excitation energies (ΔE_i) characterizing the excited states (Eq. 1), while the numerator of each term is the square of the transition dipole moment between the ground and the corresponding excited state. Using complex function analysis, it is possible to obtain the poles of the expression, that is, the values for which the frequency corresponds to the excitation energies and the denominator goes to zero, while the residues provide the numerators, in this case, the one-photon absorption matrix elements. The peculiarity of the propagator approaches is that the wave functions of the individual states are not necessarily computed to obtain excitation energies and transition probabilities, while its quality relies on the type of reference wave function. Except for very elaborated implementations, usual propagator approaches have also the same problems as all single-reference procedures.

The family of methods CCS, CC2, and CC3 is based on response theory (Koch and Jørgensen 1990). The CCS approach is equivalent to the single excited configuration interaction or Tamm–Dancoff approach. The iterative hybrid CC2 and CC3 procedures introduce approximations to account for the level of excitation. They have been defined for systems with closed-shell ground-states, although some Equation of Motion EOM-CC procedures also deal with open-shell ground state cases. In order to get accurate excitation energies and properties, the single-configuration coupled-cluster methods should include high excitation levels to compensate both the poor reference wave function and the multiconfigurational character of the excited states. CC-based methods are, up-to-date and in practice, the most accurate methods to compute excited states in small to medium size molecules with closed-shell ground states, but only for those states which are well described by singly excited configurations, in systems where the ground state has a clear single-configuration character, and close to the equilibrium geometry. Triple excitations have to be anyway included in the cluster expansion if accuracy is intended. The precision of the single-reference CC methods decreases in systems with open-shell ground states and vanishes (up to several eVs of error) for multiconfigurational cases, like the $2^1\ A_1$ state of ozone. The inclusion of quadruple excitations, unpractical so far, would improve some of those results, although the only solution in prospect to beat in accuracy the lower level and less expensive multireference perturbation approaches such as CASPT2, is to use multireference coupled-cluster (MRCC) methods, in which the required excitation level will be expected lower. However, the development of such methods is still in their infancy, and the initial results are not very promising.

Before continuing with the ab initio procedures, in particular the MRMP theory most employed in this chapter, the use of DFT approaches for excited states will be discussed. The implementation, known with the unfortunate name of Time-Dependent DFT (TD-DFT) approach (no time-dependency is accounted for) should be expected to be able to deal with large systems were ab initio methods become too expensive. Unfortunately, these methods fail dramatically in too many situations: charge transfer states, multiconfigurational states, doubly or highly-excited states, and even introduce large and systematic errors in valence states of large π extended systems. In some cases, the deviations can be as large as 5–6 eV, especially in cases in which the ground state is poorly defined by the HF configuration. Even when many different parameterizations have been tried and different functional developed in order to correct the flaws of the method, so far no single functional is able to solve most of them simultaneously. Much worst than that (and this is common for all single-configurational approaches), they cannot describe at all degenerate situations like conical intersections, which are the core of the quantum-chemical description of the excited states.

An hybrid derivation, known as DFT/MRCI is unarguably the best DFT-based procedure for excited states. It consists of an MRCI expansion with multiconfigurational wave functions replacing the HF orbitals with Kohn–Sham orbitals in the building of CSFs. The Hamiltonian, however, is heavily modified with empirical parameters, and only a few situations can be handled. The objective is to recover dynamic correlation by means of DFT and static correlation by means of MRCI. In this way, severe size-extensivity problems can be avoided even for systems with many valence electrons. Currently, optimized parameter sets for the effective DFT/MRCI Hamiltonian are available in combination with the BHLYP functional. The accuracy in energy is about 0.2–0.3 eV, and other properties are well balanced as well. Additionally, it is very appropriate to compute SOC, including a spin-dependent Hamiltonian (Grimme and Waletzke 1999; Kleinschmidt et al. 2001).

As already mentioned, a multiconfigurational description of the reference wave function helps enormously to recover the overall correlation effects. Starting from a good reference, the method used for such a recovery does not need to be extremely elaborated. That is why multiconfigurational second-order Møller–Plesset theory, the CASPT2 method, based on a Complete Active Space SCF (CASSCF) zeroth-order wave function, has become the most successful method for excited states. As this is the procedure in which most of the examples of the chapter will be based on, a more detailed description of the procedure will be given. Nowadays, the CASPT2//CASSCF methodology has proved the best ratio quality of the results/computational cost. At the CASSCF level (a particular case of MCSCF), both the many-electron-function coefficients of the MCSCF expansion and the coefficients included in the expansion of each molecular orbital are optimized simultaneously (see Eq. 3). Their variations are considered as rotations in an orthonormalized vector space. In the CASSCF method, the orbitals are classified in three categories, depending on the role they play in building the N-electron wave function: inactive, active, and secondary orbitals. Inactive and active orbitals are occupied in the wave functions, whereas the remaining of the orbital space, given by

the size of the one-electron basis set employed, is constituted by secondary orbitals, also called external or virtual. Inactive orbitals are doubly occupied in all the CASSCF configurations. The rest of the electrons (called active electrons) occupy active orbitals. The CASSCF wave function is formed by a linear combination of all the possible configurations that can be built by distributing the active electrons among the active orbitals and are consistent with a given spatial and spin symmetry. That is, in the configuration space spanned by the active orbitals, the CASSCF function is complete (CAS-CI space, equivalent to FCI). Inactive orbitals are also optimized in the variational process, but they are treated as in the restricted HF function. The CASSCF energy is invariant to rotations among the active orbitals (Roos et al. 1979).

Essentially, we construct for a given state a multiconfigurational wave function which includes all configurations generated by a set of active orbitals and active electrons that fulfils spin and spatial symmetry requirements. This provides the nondynamic correlation effects due to configurations which are very close in energy. Several states that belong to a same symmetry are usually computed by means of a State-Average (SA) CASSCF calculation, where a functional of energy is defined as average of a number of states, that, if required, although it is not recommended, can be weighted. From a SA-CASSCF calculation, a single set of average orbitals and a number of orthogonal wave functions equal to the number of roots used in the average process are obtained. In this manner, it is sometimes possible to overcome the problem of "root flipping," that is, the interchange of roots along the CASSCF optimization procedure. For a given spatial and spin symmetry, the treatment of excited states is preferably performed by using SA-CASSCF calculations. In principle, it is also possible to make a single CASSCF calculation for higher roots, optimizing just one state. Nevertheless, experience shows that in most cases it can only be achieved for the second root of a given irreducible representation. Wave functions obtained in a SA-CASSCF calculation are orthogonal among them, whereas those obtained from different CASSCF calculations are not. In the latter cases, the wave function is rather poor and cannot be used for a further correlated calculations, although sometimes it may be useful to perform CASSCF optimizations.

On the other hand, the Restricted Active Space Self-Consistent Field (RASSCF) method is a more general extension of the CASSCF method. Now, there are three subspaces within the active orbitals: RAS1 (orbitals that are doubly occupied except for a maximum number of holes allowed in this orbital subspace), RAS2 (in these orbitals, all possible occupations are allowed), and RAS3 (orbitals that are unoccupied except for a maximum number of electrons allowed in this subspace). CASSCF calculations can be performed by allowing orbitals only in the RAS2 space. A single reference SDCI wave function is obtained by allowing a maximum of two holes in RAS1 and a maximum of two electrons in RAS3, while RAS2 is empty (Malmqvist et al. 1990; Olsen et al. 1988; Fig. 8).

Either by using CASSCF or RASSCF, the active space provided by the user of a CASSCF calculation represents a key point to obtain accurate theoretical predictions once dynamic correlation has subsequently been taken into account,

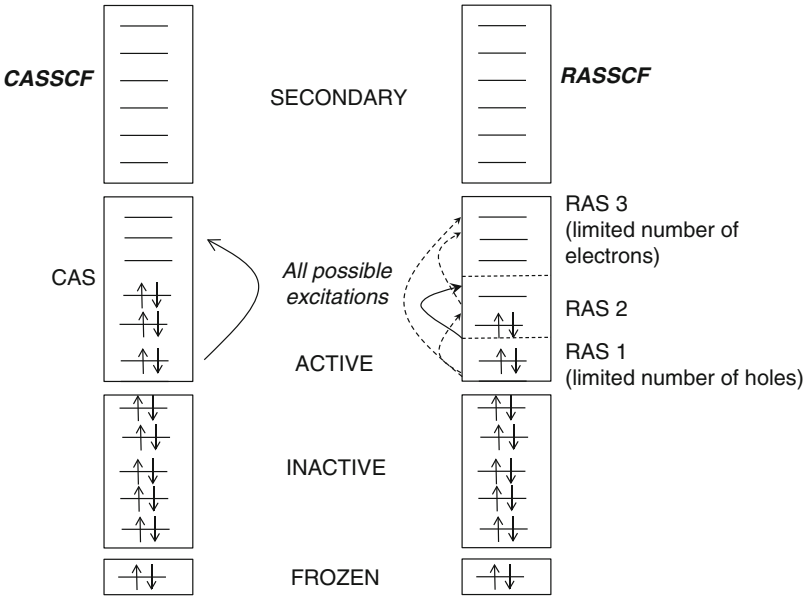

Fig. 8 CASSCF and RASSCF methods

for instance, at the CASPT2 or RASPT2 levels. The properties of a CASSCF wave function depend on the active space. Thus, a valence CASSCF is size-extensive (the computed energy within a given methods scales linearly with the number of particles) and the corresponding CASPT2 results become also nearly size-extensive (formally the CASPT2 method is not size-extensive, but it is in practice, in particular the MOLCAS implementation). As in any quantum-chemical approach, one has to make sure that the method has enough flexibility, i.e., the active space is the appropriate one, to describe the chemical process under consideration. It is important to mention here the power of the CAS State Interaction (CASSI) approach, which provides orthogonal wave functions and transition densities from CASSCF or RASSCF wave functions optimized independently for a number of excited states of whatever symmetry (Malmqvist and Roos 1989).

Valence-bond methods have increased its applicability recently. One example is the CASVB (complete active space valence bond) method. A CASVB wave function can be obtained simply by transforming a canonical CASSCF function and readily interpreted in terms of the well-known classical VB resonance structures. The total CASVB wave function is identical to the canonical CASSCF wave function. In other words, the MO description and the VB description are equivalent, at least at the level of CASSCF. The CASVB method provides an alternative tool for describing the correlated wave functions.

The CASPT2 (complete active space perturbation theory to second order) method (Andersson et al. 1992) includes the remaining dynamic correlation due to short-range electronic interactions. This method can be seen as a conventional non-

degenerate perturbation theory, that is, a single state is independently considered, with the particularity that this zeroth-order wavefunction is multiconfigurational (CASSCF). The wave function is corrected up to first order and, consequently, the energy is corrected up to second order. The set of functions required to compute the first-order correction of the wave function is formed by those that interact with the zeroth-order wave function through the Hamiltonian in the Rayleigh–Schrödinger perturbation theory, and it is known as the first-order interacting space. Taking into account the one and two particle nature of the Hamiltonian, the first-order interaction space, called hereafter V_{SD}, comprises the functions generated by singly and doubly excited configurations from the zeroth-order (CASSCF) wave function. As a matter of fact, the configurational space is divided in four subspaces: V_0 (one-dimension space expanded by the reference function $|0\rangle$ of the studied states), V_K (orthogonal to $|0\rangle$ in the restricted FCI subspace used to generate the CAS function), the aforesaid V_{SD} (space expanded by the singles and doubles replacements from $|0\rangle$), and $V_{\mathrm{TQ}\ldots}$ (space that contains the excitations of higher order). Regarding orbitals, they are classified into frozen (doubly occupied, not correlated), inactives (doubly occupied in the reference function), actives (with any occupation, between 0 and 2, in the reference function) and secondary (empty in the reference function). Recently, a new shifted zeroth-order Hamiltonian named IPEA has been designed and set up with a value 0.25 a.u. into the default for the MOLCAS implementation (Ghigo et al. 2004). The new formulation solves previous effects for open-shell cases in which the correlation effects were overshoot. The net effect of this correction is to slightly increase the excitation energies. In any case care has to be taken when comparing present with previous CASPT2 results.

The normalized wave function is corrected up to first order. The weight of the reference function, C_0^2, can be used as a simple and rapid criterion of quality for the perturbation treatment carried out. Ideally, in order to get a fast convergence in the perturbation series, the weight should be close to unity. Nevertheless, its value depends on the number of correlated electrons. Thus, upon enlarging the molecular system the reference weight decreases. The electronic excited states considered should have a similar magnitude for the weight as compared to the ground state, employing the same active space. Sometimes intruder states appear in the second-order calculation, which are normally related to the occurrence of large coefficients in the first-order expansion, leading to a low value for the reference weight. Analysis of the states with large coefficients (intruder sates) may give a hint about the type of reformulation in the perturbation partition necessary to overcome the problem. Thus, a new CASSCF calculation might be designed comprising in the active space the MOs implied in the description of the previous intruder states. It is the proper action to be taken when intruder states are strongly interacting with the CASSCF reference wave function, because it points out to obvious deficiencies in the choice of the active space. Intruder states are often present in the treatment of excited states of small organic compounds when the active space does not include the full π valence system. Thus, the low weight for the zeroth-order wave function in such a case just tells us that the active space has to be enlarged in a way that previous intruder states would be treated variationally, that is, they should be moved to the

active space. It is also frequent to find calculations where the reference weight of the excited state is "somewhat low" compared to that of the ground state, but a particular state cannot be identified as intruder in the first-order wave function, which is instead characterized by a large number of low-energy minor contributions. It occurs often in the simultaneous computation of valence and Rydberg states, where the one-electron valence basis set has been augmented with Rydberg-type functions. This is a typical case when using large and diffuse basis sets like ANOs. We have to face then accidental near-degeneracy effects, implying weakly interacting intruder states, and the level-shift (LS) technique is especially useful in order to check the validity of the perturbation treatment performed. The level-shift CASPT2 (LS-CASPT2) method removes efficiently weak intruder states by the addition of a shift parameter ε, to the zeroth-order Hamiltonian and a subsequent back correction of its effect to the second-order energy.

Many times, one has to apply both strategies: enlargement of the active space to overcome the problem of severe intruder states, and, with the enlarged active space, the LS technique is applied in order to minimize the effect of weakly interacting intruder states. Since a constant added to a linear and Hermitian operator (like $\mathbf{H}$) does not affect its eigenfunctions but it is added to its eigenvalues (let k be a constant; as long as $\mathbf{H}\Psi = E\Psi$, then $(\mathbf{H} + k)\,\Psi = \mathbf{H}\Psi + k\Psi = E\Psi + k\Psi = (E + k)\,\Psi\Big)$, the energies of the states may be altered, but this is a price that must be paid. However, it is desirable that the shift parameter is as small as possible. For instance, results at $\varepsilon = 0.0$ (standard CASPT2), 0.1, 0.2, 0.3, and 0.4 a.u. are sufficient to establish the proper behavior of the LS-CASPT2 results. It is extremely dangerous to rely on just one result because the appearance of an accidental near degeneracy might lead to large errors in the excitation energies. In order to demonstrate the proper performance of the LS-CASPT2 technique, calibration calculations of that type always have to be carried out. The best choice for ε is the lowest possible value capable of removing intruder states. In order to avoid singularities, currently the technique so-called imaginary level-shift should be employed. The dependence of the energy on the imaginary level-shift parameter is of minor relevance (see Forsberg and Malmqvist 1997; Roos et al. 1996).

Finally, the multistate CASPT2 (MS-CASPT2) procedure has to be mentioned. It represents an extension of the CASPT2 method for the perturbation treatment of chemical situations that require two or more reference states. The procedure implies the use of an effective Hamiltonian in which the diagonal terms are the single-root, and non-orthogonal, CASPT2 solutions, whereas the interacting terms form the off-diagonal components. The diagonalization of the Hamiltonian, after symmetrization, provides a set of MS-CASPT2 energies and orthogonal states ready to describe situations such as avoided crossings and near-degeneracy of valence and Rydberg states, which cannot be fully accounted for by just using a single-reference perturbation treatment. A new wave function, named Perturbatively Modified CAS-CI (PMCAS-CI), is obtained built as a combination of the previous SA-CASSCF states, which has been in many cases improved from the previous set. This approach is not free of problems, as it will be discussed along this chapter (Finley et al. 1999; Serrano-Andrés et al. 2005).

The CASPT2 approach, especially when combined with ANO-type basis set functions, has proved to give a balanced and accurate description of all types of excited states and electronic structure cases, independently of their nature. The energy and the wave function are treated in general in an unbiased way. As all other quantum-chemical methods, the results heavily rely on the proper determination of the structural parameters, that is, the geometry optimizations. Because of their computational cost, geometries are obtained at lower levels of calculations than energies. In many cases, analytical gradients, more convenient for optimization processes, are not available at the highest level of calculation (they are, for instance, at the MRCI level, see Shepard et al. (1992) or for CASPT2 applicable to small molecules, see Celani and Werner (2003)). This is why most of the calculations use the mixed strategy, for instance, CASPT2 energies and CASSCF geometries (CASPT2//CASSCF), but also MRCI//CASSCF or CC3//CC2. Some problems related with this inconsistency will be discussed during the chapter.

It can be finally commented that, recently, the RASPT2 approach, in which a RASSCF wave function is used as reference for multiconfigurational perturbation theory has been made available, using all tools related with CASPT2, for instance allowing MS-RASPT2 calculations (Malmqvist et al. 2008). Benchmark studies are currently under way in order to determine the best partition schemes for the RASSCF active spaces. Prospectively, the approach will largely extend the applicability of the MRMP calculations.

How to Compute Excited States

How to Start: Selection of Goals, Methods, Geometries

Quantum-chemical methods provide information for excited states directly applicable to explain and predict the spectroscopy, photophysics, and photochemistry of molecular systems. A balanced description of the different electronic states is required in order to obtain the initial, basic data, that is, energy differences and transition probabilities, in an accurate way. This goal is a much more difficult task for excited states as compared to the ground state. First, one has to deal with many classes of excited states, each one showing different sensitivity to the amount of electronic correlation energy and, also, flexible one-electron basis functions able to describe all effects simultaneously are required, in general larger than that used in ground-state quantum chemistry. Then, it is necessary to compute extremely complicated potential energy hypersurfaces where the number of minima, transition states, and surface crossings like conical intersections, is multiplied. Because of the inherent complexity of the problems, the methods and algorithms to compute excited states are not as widespread as for ground states or are still under development.

In this section, we are going to illustrate different examples for the calculation of excited states using presently available quantum-chemical techniques. In each subsection, we will proceed in the same manner, first, by describing an actual

example taken from the literature, and, after, practical aspects which explain the choices made. The results obtained will be then discussed, including new cases if necessary.

Before introducing the first example, let's make some comments on the initial way to tackle the problem. Our goal, in principle for a typical ground-state closed-shell case, will be focused in the lowest-lying singlet and triplet excited states of the molecule. Most of the interesting spectroscopy and photochemistry will take place in the low-lying states, although higher states are relevant in other contexts. Properties of interest provided by static electronic structure quantum-chemical calculations that help to rationalize the photophysical and photochemical processes in a molecule are: molecular structures, charge and spin distributions, electronic and vibrational energies, oscillator strengths, dipole and transition moments and their directions, radiative lifetimes, nonadiabatic, vibronic, spin-orbit, and electronic couplings. In a second step, in reaction dynamics calculations, reaction rates, lifetimes, and population distributions can be provided.

Which method is the most appropriate for the purpose of that research? The initial questions to answer is what is the goal of the study, how far can it go, how many questions can be solved, and, especially, which is the accuracy required in the study. All this depends in most cases on the complexity and the size of the system. Of course, the best ratio quality/CPU time should be looked for, but it is useless to carry out calculations, no matter how cheap they are, with methods that cannot provide accuracy to make conclusive predictions. Apart from checking carefully the literature, calibrations will be typically necessary. Both methods and basis sets should be tested, and several steps to approach to the problem are maybe necessary. We can start by using less expensive semiempirical classical or TD-DFT methods in order to understand the problem and its requirements, and then move to more sophisticated coupled-cluster or multiconfigurational approaches, depending on the goals of the research. The same with basis sets. 6-31G-type basis sets are practical because of their reduced time requirements, but they are less accurate and in many cases using them may degrade too much our results. Calibrating the obtained values with fully correlated methods or higher-quality basis sets like Dunning's correlation consistent or ANO-type basis sets is always a requirement. Obviously the program package we should use has to balance all aspects we have mentioned. Normally a combination of them will be necessary: versatile programs, such as GAUSSIAN, GAMESS, or QCHEM; packages specialized in DFT implementations, like ADF, TURBOMOLE, or NWCHEM; suites focused on CC algorithms such as DALTON or ACES II; or programs especially designed for multiconfigurational methods like MOLCAS or MOLPRO, these more generally applicable to the calculation of excited states.

A point to bear in mind is that there are different types of excited states, and each of them has different theoretical requirements. A valence state can be viewed as a promotion from one occupied molecular orbital to a virtual one. On the other hand, a Rydberg state, in a simple MO theory, is the result of exciting one electron from one occupied molecular orbital to an atomic-like orbital of higher quantum number. Valence excited states are more compact than the diffuse Rydberg states. A valence

state can be denoted as covalent or zwitterionic according to the type of valence bond (VB) structures. Covalent and zwitterionic states are described by hole-hole and hole-pair VB structures, respectively. What is more, the label ionic is reserved to states with an actual charge separation. On the other hand, if a pair of nearly degenerated nonbonding orbitals is occupied with a total of two electrons in the ground state, the molecule is called biradical. In addition, there are charge transfer states, in which an electron is transferred from one occupied orbital of a molecule to a virtual orbital of another molecule (or the same molecule, if it is intramolecular).

Excited states can be also classified according to the main types of configurations involved in their description, in relation to the ground state: singly excited states, doubly excited states, and so on. Rydberg states are usually well described by only one singly excited configuration. Multiconfigurational singlet excited states with a large contribution of doubly excited configurations are normally covalent. Singlet excited states of zwitterionic character are described by one or several singly excited configurations. In anionic systems, new types of excited states are present, such as the resonance states or the multipole-bound (dipole, basically) states, where the bonding energy is the result of the interaction between the additional electron and the multipole moments of the molecule.

Molecular Photophysics: Computing Absorption and Emission Spectra

An Application Example: Psoralen

The simplest approach to the electronic state problem is the determination of the photophysics of an isolated molecule, that is, to reproduce its absorption and emission spectra. As mentioned, let's first describe in its full extent a case, the psoralen molecule, and later the choices made can be analyzed (see Serrano-Pérez et al. 2006, 2008a, b, c, d).

Furocoumarins (also named psoralens, see Fig. 9) are a class of heterocyclic compounds with a known phototherapeutic activity. These systems have been found to possess mutagenic properties when applied in conjunction with near UV-A light (320–400 nm) exposure. The technique so-called PUVA therapy (psoralen + UV-A) has been specifically designed to treat different skin disorders such as psoriasis and vitiligo. It is generally assumed that there is an oxygen-independent mechanism, which implies a [2 + 2]-photocycloaddition of psoralen and a pyrimidine DNA base monomer, and there is also an oxygen-dependent mechanism, in which energy transfer between the furocoumarin and molecular oxygen present in the cellular environment is produced generating cytotoxic singlet molecular dioxygen. It is believed that the state protagonist of the photosensitizing action, in both cases, is the lowest-lying triplet excited state of psoralen. Initially, our main goal is to describe the absorption spectrum of the molecule, and we do it by computing the lowest-lying singlet and triplet states at the optimized structure of the ground state (i.e., those distances, angles, and dihedral angles which make the energy minimum), what is

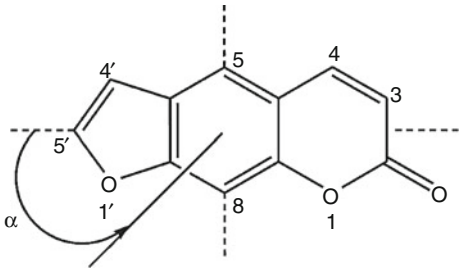

Fig. 9 Psoralen labeling. The *arrow* defines the positive angles of both the dipole moment and electronic transition moment directions with respect to the pseudosymmetry long axis

typically known as the Franck–Condon geometry. Later, the mechanisms for triplet state population can be determined.

Using CASSCF multiconfigurational wave functions as reference, the second-order perturbation theory through the CASPT2 method was employed to include dynamic correlation energy in the calculation of the electronic excited states. The imaginary level-shift technique was employed in order to prevent the effect of intruder states. A shift parameter of 0.3 a.u. was selected by checking the stability in the excitation energies. The molecular symmetry was constrained to the C_S point group. An atomic natural orbital (ANO-L)-type basis set, contracted to C,O [4s3p1d]/H [2s1p] was used throughout. The carbon and oxygen 1 s core electrons were kept frozen in the second-order perturbation step. Geometries were obtained by computing analytical gradients at the RASSCF level of calculation for the ground and the lowest singlet and triplet excited states. In the optimization of the A' states, an active space of 14 a'' active orbitals and 16 electrons, i.e., the full π system of the molecule, has been employed, and up to quadruple excitations were considered (eight orbitals in RAS1 space and six orbitals in RAS3 space). Within the irreducible representations (a', a'') of the C_S group, this active space can be labeled as (0, 14). An additional oxygen lone-pair orbital was included in the active space (1, 14) in order to optimize the lowest A'' excited state. In all the remaining calculations, CASSCF wave functions were generated as state-average (SA) CASSCF roots of a given symmetry. Based on preliminary RASSCF calculations and using the criterion of the largest natural orbitals occupations for the states of interest, the CASSCF active space was reduced to include 12 active electrons and 12 active orbitals (0, 12) for A' roots and 14 active electrons and 13 active orbitals (1, 12) for A'' roots. The CAS state interaction method (CASSI) was used to compute transition properties, including the spin-orbit coupling (SOC) elements between selected states.

At the Franck–Condon (FC) geometry, the lowest singlet excited states $2^1A'(\pi\pi^*)$, $1^1A''(n\pi^*)$, and $3^1A'(\pi\pi^*)$ lie at 3.98, 5.01, and 5.03 eV, respectively (see Table 1). Whereas the transition to the $n\pi^*$ state is predicted with negligible intensity, the $\pi\pi^*$ states have related oscillator strengths of 0.027 and 0.107. Unlike other states, the $3^1A'$ state has a high dipole moment, 8.70 D, differing by more than 2.5 D from that of the ground state, and therefore the associated transition is expected to undergo a red shift (bathochromic effect) in polar environments. The recorded absorption spectra in different solvents, from cyclohexane to water,

Table 1 Computed excitation energies ΔE (eV) at CASPT2 level, oscillator strengths f, dipole moments $\mu(D)$, dipole moment directions μ_{dir} (deg), and transition dipole moment directions TDM_{dir}(deg) for the low-lying electronic transitions of psoralen

State	ΔE	f	μ	μ_{dir}	TDM_{dir}
$1^1A'$	–	–	6.25	128	–
$2^1A'(\pi\pi^*)$	3.98	0.027	6.50	132	4
$1^1A''(n_o\pi^*)$	5.01	0.000	2.07	130	–
$3^1A'(\pi\pi^*)$	5.03	0.107	8.72	132	5
$4^1A'(\pi\pi^*)$	5.22	0.064	7.07	−23	−10
$5^1A'(\pi\pi^*)$	5.30	0.331	7.07	−37	102
$6^1A'(\pi\pi^*)$	5.70	0.091	7.37	−37	30
$1^3A'(\pi\pi^*)$	3.27	–	5.40	124	–
$2^3A'(\pi\pi^*)$	3.55	–	5.35	125	–
$3^3A'(\pi\pi^*)$	4.08	–	5.48	127	–
$1^3A''(n\pi^*)$	4.85	–	2.15	111	–

display a weak and structured band ranging from 360 to 270 nm (3.44–4.77 eV). Depending on the band resolution and the environment, one or two maxima near 330 and 280 nm (3.76 and 4.43 eV, respectively) have been described. The present computed results suggest that this set of features can be better assigned just to the $2^1A'(\pi\pi^*)$ transition, with the weak $n\pi^*$ band lying beneath. In that case, the observed band profile should be attributed to vibrational structure. This explanation is not unlikely, considering that a noticeable rearrangement of the molecular bond distances occurs at the $2^1A'$ minimum.

Although less relevant regarding the phototherapeutic properties, we will describe the higher-energy region of the absorption spectrum. Transitions to the $4^1A'$ (5.22 eV) and $5^1A'$ (5.30 eV) ($\pi\pi^*$) excited states have oscillator strengths of 0.064 and 0.331, respectively. The recorded spectra show a single and sharp band peaking near 248 nm (5.00 eV) in cyclohexane and 240 nm (5.16 eV) in ethanol and water. In principle, the observed feature can be assigned to transition to the $3^1A'$ state at 5.03 eV. Moreover, taking into account our computed results, an additional and more intense band can be expected at higher energies (5.30 eV). The measured band is probably a combination of both transitions.

Three other $^3A'\pi\pi^*$ states are next in energy at 3.55, 4.08, and 4.66 eV. The nature of the low-lying transitions of each symmetry, which are those basically responsible for the photophysical properties of psoralen, can be graphically described by computing the differential electron density plots as displayed in Fig. 10. Transition to the $S_1\pi\pi^*$ state is mainly benzene-like, with the charge migration concentrated in the central benzenoid ring. On the contrary, that related to the $T_1\pi\pi^*$ state has its major contributions in the pyrone ring, with high participation of the carbonyl oxygen and a shift in the density away from the pyrone ring C_3–C_4 bond, which will be later discussed as an essential feature of the photophysics of the system. Also in Fig. 10, we find the expected differential density plots of the $n\pi^*$ states centered on the carbonyl group. Transition dipole moment directions (TDM_{dir}) indicate that the three lowest $\pi\pi^*$ features have nearly parallel polarizations, that is, they are aligned with the long axis of the molecule, while

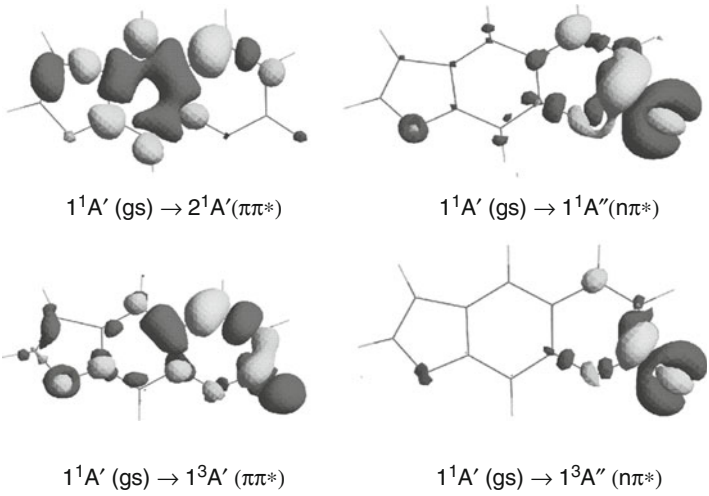

Fig. 10 Differential electron density for the main valence transitions in psoralen computed at the ground-state optimized geometry. The electron density is shifted upon light-induced excitation from *darker* to *lighter* regions

transition to the $5^1A'$ state has essentially perpendicular polarization. Regarding the vertical excitations to the triplet states, the $1^3A'\pi\pi^*$ (T_1) state lies at 3.27 eV, near 0.7 eV below the $2^1A'(S_1)$ state.

Fluorescence has been reported for psoralen in polar solvents starting (T_0) at 350 nm (3.54 eV) with a maximum at 409 nm (3.03 eV). Phosphorescence has also been recorded in solution with band origin at 456 nm (2.72 eV) and a maximum between 460 and 490 nm (2.7–2.5 eV).

The fluorescence and phosphorescence quantum yields were measured in ethanol as $\Phi_F = 0.019 - 0.02$ and $\Phi_P = 0.13$, respectively. The obtained ratio Φ_F/Φ_P is approximately 7.1. The total phosphorescence decay time (τ_P) has been reported 1.1 s in glycerol-water and 0.66 s in ethanol. With those data, the phosphorescence radiative lifetime τ_{rad} ($= \tau_P \cdot \Phi_P$) can be therefore expected between 8 and 5 s.

The low-lying singlet excited state $2^1A'(\pi\pi^*)$ is responsible for the lowest-energy absorption and emission fluorescence bands (see Table 2). Vertically, at the ground-state geometry, the transition energy is computed to be 3.98 eV and, upon relaxation of the geometry, the band origin (T_e) decreases to 3.59 eV. This means that the range of absorption goes from 3.59 to 3.98 eV, well within the PUVA action. A similar relaxation is observed experimentally between the lowest-energy absorption band maximum and the band origin. The structural changes of the computed equilibrium geometries for the ground (S_0) and the $2^1A'(\pi\pi^*)$ states affect the bond alternation of the system, mainly in the central ring (cf. Figs. 10 and 11), as expected from the differential charge density plots. By using the Strickler–Berg relationship, a fluorescence radiative lifetime of 74 ns is calculated for the S_1 state. The low-lying $1^1A''(n\pi^*)$ state (vertically S_2) becomes relaxed by more than

Table 2 Computed and experimental excitation energies (eV) and emission radiative lifetimes (τ_{rad}) relevant for the photophysics of psoralen

Theoretical (CASPT2)				
State	E_{VA}	T_e	E_{VE}	τ_{rad}
$2^1A'(\pi\pi^*)$	3.98	3.59	3.45	74 ns
$1^1A''(n_o\pi^*)$	5.01	3.91	2.78	3 μs
$1^3A'(\pi\pi^*)$	3.27	2.76	2.29	28 s
$1^3A''(n_O\pi^*)$	4.85	3.84	2.79	9 ms
Experimental (data in ethanol)				
State	Abs_{max}	T_0	E_{max}	τ_{rad}
$2^1A'(\pi\pi^*)$	3.7-4.3	3.54	3.03	–
$1^1A''(n_O\pi^*)$	–	–	–	–
$1^3A'(\pi\pi^*)$	–	2.7	2.7	5–8 s
$1^3A''(n_O\pi^*)$	–		–	

E_{VA} vertical absorption, T_e adiabatic electronic bandorigin, E_{VE} vertical emission, Abs_{max} experimental absorptionmaximum, T_0 experimentalband origin, and E_{max} emission maximum

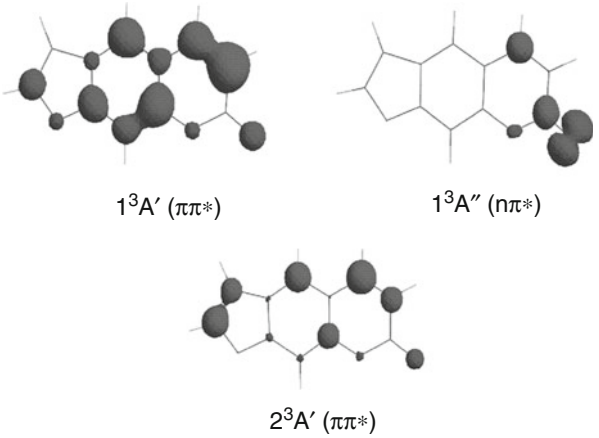

Fig. 11 Spin density for the low-lying triplet states in psoralen computed at the ground-state optimized geometry

1 eV upon geometry optimization. Although the $1^1A''(n\pi^*)$ minimum belongs to the S_1 hypersurface, the final T_e value is about 0.3 eV higher in energy than the computed and measured band origin for $2^1A'(\pi\pi^*)$. Therefore the $n\pi^*$ state is not a plausible candidate for the fluorescence, which is better attributed to the $\pi\pi^*$ state.

The $1^3A'(\pi\pi^*)$ state is the clear protagonist of the phosphorescence. The computed band origin at 2.76 eV perfectly relates to the observed value in solution at 2.72 eV. The relaxation energy is about 0.5 eV. The largest structural change is produced in the C_3–C_4 bond of the pyrone ring, which enlarges by near 0.13 Å from the ground-state value. The computed spin population, displayed in Fig. 11,

is mainly placed on each of the carbon atoms forming the bond. In that way, psoralen becomes highly reactive in its lowest triplet state through its pyrone C_3–C_4 bond. This finding is the cornerstone of the photophysics of psoralen, which has been repeatedly proposed to take place through a reactive triplet state. The computed phosphorescence radiative lifetime is 28 s, somewhat higher than those estimated experimentally from the quantum yield and the total relaxation time, 5 and 8 s. For the $2^3A'(\pi\pi^*)$ state, the spin population is placed mainly on the carbon atoms forming both C=C bonds, that is, C_3–C_4 (pyrone) and $C_4, -C_5$, (furan).

Practical Aspects

Selection of Geometries

Any quantum-chemical calculation starts with the definition of the molecular geometry. As mentioned in section "Methods, Advantages, and Drawbacks," we can safely assume that the absorption spectrum is initiated at the ground state equilibrated geometry. An optimization is therefore required, in principle at the highest reasonable level of theory. DFT (with a proper functional) works well for most closed-shell cases, and it is a fully correlated method. Other ab initio methods that also include all correlation energy will be more accurate, but also more computationally expensive: MP2, CCSD, CASPT2 (with numerical gradients for small molecules). Although the best choice is not to use any symmetry restriction and let the system to find the lowest-energy solution, it is possible to make approximations and, for instance, let the system to be planar (C_s point group), like in the case of psoralen and save computational time, especially if experiment give some indication along these lines. In order to confirm that the obtained point is a true minimum, it would be desirable to perform a frequency analysis, although computing Hessians (second derivatives) is an expensive task. If too demanding, maybe is feasible to compute the Hessian at lower levels of theory and check the outcome. Of course, if possible, one should compare the result with experimental (X-ray or electron diffraction, for instance) data. To compute emission, we have to optimize excited states too, a more difficult task than for ground states only at reach of few methods: CASSCF, RASSCF, CASPT2, and CCSD, to mention the most reliable ones (the two latter only for small systems). If the ground state is optimized with a better method than the ground state, some unbalance in the results may occur. It can be a good idea to get the different states optimized at the same level, like CASSCF or RASSCF. The choice of RASSCF in the psoralen example is motivated by the impossibility of including all π space into a CASSCF calculation. Although safe in many cases, the exclusion of some MOs, like, for instance, the lowest-energy nodeless π orbital, may lead to problems if the excluded orbital, instead of becoming delocalized and having then no biased effect, tends to localize unbalancing the calculations. It is worth noticing that CASSCF (RASSCF) gives very good geometries in π-conjugated systems, probably because of some compensation between the obtained single and double bonds lengths.

Accuracy of the Excitation Energies and How to Compare with Experiment

It is time to decide if the method and strategy employed is accurate enough to solve the problem. In principle, the accuracy should be established by previous experience if the electronic structure problem is under control. We cannot expect a TD-DFT calculation to be more accurate than 0.5–0.4 eV, much less if the problem has no closed-shell ground state or a charge transfer character (except for specific functionals like CAMB3LYP, see Peach et al. (2008)). Choosing one functional for each molecule makes theoretical chemistry useless as a predictive tool. It is much better to stick into one procedure and know their limitations. Coupled-cluster results, EOM-CC(T) or CC3, are much more accurate, sometimes near 0.2–0.1 eV if pushed to the limit, but this also relies on being far from degenerate situations or multiconfigurational ground states (see Grein (2009), 1^1B_2 state of ozone: CCSD(T) 1.81 eV vs. best estimation 4.11 eV, or Kowalski (2005), 2^1A_1 state of ozone: EOMCCSD 9.82 eV vs. best estimation 4.33 eV). Multiconfigurational calculations, CASPT2 in particular, can reach also 0.2–0.1 eV in most cases, although the result may degrade if the reference is not well described. More important that all we have said: accuracy with respect to what? Many studies discussing the precision of a method for excited states compare their vertical excitation energies with what experiment initially offers, absorption maxima. This is wrong and misleading. A vertical excitation energy computed at the electronic ground-state equilibrium geometry is a purely theoretical concept without experimental counterpart. The Franck–Condon principle relates such value with the absorption (not emission) maximum, but this approximation will be true only in the case that the ground and excited minima take place at the same geometry. Apart from that, an experimentally determined absorption maximum depends on the concentration of the sample, the apparatus, the environment, the temperature, the pressure, etc. In the most favorable case, vertical absorption and emission maximum will differ by near 0.1 eV from an experimental gas-phase or molecular beam maximum. Typically, those cases display as maximum feature the band origin or T_0 transition. Indeed, the true experimental data to compare with this band origin, that theory computes as the energy difference between the states minima, also named the adiabatic energy difference, T_e if it is just electronic energies or T_0 – the true comparable value – if the zero-point energy correction has been included. The situation is even worse for the emission, where the FC principle is hardly fulfilled. In the psoralen example, a good agreement has been obtained for the band origins: computed T_e 3.57 and 2.76 eV and measured T_0 3.54 and 2.7 eV for fluorescence and phosphorescence, respectively. In most cases, T_e and T_0 differ by close to 0.1 eV. On the other hand, the vertical emission is typically much lower than the emission maxima.

How to Deal with Symmetry

Unfortunately, the topic is too extensive to be treated here. For practical cases, the reader is referred to the advances examples in the MOLCAS program manual (see http://www.teokem.lu.se/molcas). Group theory may be very useful in a calculation on excited states, especially because it allows dividing the states in each one

of the irreducible representations (since a symmetry operation leaves a molecule unchanged) and consequently reduce the computational effort. Unfortunately, the molecular symmetry is typically only found for ground states, and most of the photochemistry takes place at distorted asymmetric structures, especially for organic systems. Linear molecules or many inorganic complexes can however be studied making an extensive use of the symmetry, as well as other systems if absorption, and sometimes emission, spectra are studied. Symmetry is particularly useful in multiconfigurational calculations to reduce the active space requirements. Also, symmetry permits prediction of intensities when following selection rules or analysis of the polarization or dipole moment directions. Normally, one should have at hand the character tables and also the Wigner–Witmer rules, as those found in Herzberg books. Finally, it is necessary to warn about symmetry breaking problems, in which a bad selection of the method or the active space may lead to find spurious lower-energy solutions.

How to Obtain Intensities and Band Shapes: Vibrational Contributions

Apart from the energies, computing electronic oscillator strengths, f, provides information about the relative intensity of the different transitions, initially for those states allowed in one-photon (optical) spectroscopy. Group theory indicates that one-photon allowed transitions of the molecule are those in which the direct product of the symmetries of the initial state, the corresponding dipole moment component (x, y or z), and the final state belong to the totally symmetric irreducible representation of the point group of the molecule:

$$\int \psi_m^* \widehat{d}_{\mathrm{el}} \psi_n d\tau_{\mathrm{el}} \Rightarrow \Gamma(\psi_m) \otimes \Gamma\left(\widehat{d}_{\mathrm{el}}\right) \otimes \Gamma(\psi_n). \tag{6}$$

For instance, in the case of a C_S molecule like psoralen (in the XY plane), the allowed transitions are those in which the aforesaid direct product is A'. Otherwise, the transition is forbidden. In other words, from the ground state A', the accessible states are A' if the light is x-polarized or y-polarized, and A'' if the light is z-polarized. If the light is x-polarized, for instance, the transition $A' \rightarrow A'$ is allowed, whereas the transition $A' \rightarrow A''$ is forbidden. The magnitude of the adimensional f has to be compared to the area beneath the band representing the transition, not to the height. Frequently, especially in gas-phase spectra, weak bands such as the Rydberg transitions are the highest because they sit on top of the more intense, although broader, valence bands. Notice that the range of energies of a band depends on the length of the vibrational progressions, which, in turn, relies on the difference in geometries with the ground state. The larger is the difference, the broader and extended will be the band. Flexible systems like NH_2 or C=O-based spectra like that of formaldehyde extends their valence bands and spread their intensity by several eVs. In many occasions it may be required to compute the vibrational profile of different electronic bands and plot them together. This is an expensive task because, requires the calculation of Hessians, FC factors, and sometimes the derivatives of such factors and of TDMs in order to obtain vibrational TDMs. This is, however,

the only procedure to elucidate complex cases, for instance, that of the lowest single valence 5.5–6.5 eV band of pyrrole, which contains two valence and four Rydberg bands which have all to be resolved (see Roos et al. 2002). Ideally, and in order to get all actual excitation bands, calculations on vibrational TDMs should be complemented with the inclusion of vibronic couplings, in which more than one electronic state is considered at the same time (breaking the Born–Oppenheimer approach), an example of what can be found in the pioneering studies of Domcke and coworkers on the pyrazine molecule (Domcke et al. 1993).

How to Add Environmental Effects

Adding the effects of the environment for excited states accurately is, if possible, even more complex than for the ground state. Usual procedures use cavity models such as Onsager's or the Polarized Continuum Model (PCM), with the additional consideration of the non-equilibration of the electronic response for the excited states that leads to divide the reaction field in slow, inertial, and fast, optical, parts. Results obtained with cavity models cannot be expected to be as accurate as those for the isolated system when compared with gas-phase results, among other things, because using large basis sets as those required for excited states will force the charge to leave the cavity and provide non-physical results. In many cases, the information yielded by the dipole moment of the states will be informative enough for qualitative purposes. In the psoralen example, it was discussed how states with dipole moments larger than that of the ground state were expected to stabilize in polar solvents (and undergo a spectral red-shift) than those with smaller dipole moments (blue-shift), typical case of the $n\pi^*$ states, which, additionally, tend to directly interact with protic solvents forming hydrogen bonds and pushing the excitation energy up in energy, sometimes even 0.5 eV. These interactions cannot be included by the cavity models and specific molecules have to be applied, even several solvation shells. Careful microhydration experiments allow nowadays comparison with such type of calculations.

Solvation is a very dynamical phenomenon which requires also the inclusion of statistical effects. More sophisticated studies require the employment of dynamical approaches making use of statistical mechanics, such as Monte Carlo type of calculations. Solvent molecules can be then simulated by point charges (like in QM/MM approaches as it will be discussed later) and dynamical time shots with their positions taken for a subsequent quantum chemical calculation. The required property will be obtained as an average of the different conformations, as some studies on 2-aminopurine reported recently have shown.

Active Spaces for Multiconfigurational Methods

As a large part of studies on excited states, and especially in photochemistry, employ multiconfigurational approaches, it is necessary to understand the process of selection of an appropriate active space (AS) for such computations. An active space should contain all orbitals and electrons relevant for the chemical process under study. The size and nature of the AS define the type of states and processes the multiconfigurational method will compute. Obviously, there is no single choice,

and the size is limited to 14–15 or 30–30 orbitals/electrons in CASSCF and RASSCF calculations, respectively, depending on the type of partitions made. It is important to emphasize that the selection depends on the problem. If one is only interested in low-lying excited states, maybe a small AS is enough for that purpose. On the other hand, more states imply larger ASs. In photochemistry the situation is worse, because the AS must be flexible enough to include all MOs participating in the process in all different regions of the PEHs which are chemically relevant. For instance, each bond that breaks means that the AS should contain both the corresponding bonding and antibonding MOs. In the psoralen example above the highest-lying $\pi\pi^*$, MOs plus one lone-pair MO from the carbonyl oxygen were selected because the purpose was to compute the low-lying spectra. The selection was a combination of experience and use of tools like performing a RASSCF calculation on several roots with enlarged active spaces to analyze the MOs participation in the different states and eliminate those orbitals with occupation numbers <0.05 and >1.95 in all roots of interest. If Rydberg states were to be computed, they should be added into the active space too. Making a TD-DFT calculation also helps to select the AS. As a general rule, it is wise to have in the AS one "virtual correlating" MO for each heavily occupied MO, for instance, each π MO with its corresponding "correlating" π^* MO. This is not necessary for lone-pair orbitals. Other well-known requirements relates to the need to include a second "correlating" d shell (4d) when computing the excited states of first-row transition metals, or more specific rules for very heavy elements. One big advantage of using symmetry in the calculations is the possibility of splitting the AS following the symmetry requirements, always when the energy differences are computed between states computed with the same AS size. The reader is referred to the specialized literature and the MOLCAS manual.

The recently developed RASPT2 method has opened the field to the use of larger RASSCF ASs. Special care has to be taken in the proper distribution of the MOs into the three RAS subspaces, and in the inclusion of an excitation level high enough (triples typically) from and to RAS1/RAS3 to get accurate results. The method is still under calibration, but it is a big step forward for the application of the methodology to larger and more complex systems. Typical improvements include the placement of Rydberg MOs or the second d-shell into the RAS3 space.

Computing Rydberg States

An Application Example: Water

How do we understand a Rydberg state? If we extract one electron from each of the MOs of a neutral molecule, the system becomes positively charged. An electrostatic interaction is therefore established between the molecular cation and the negative electron up to the moment that this ends up as a free electron, that is, at the energy named ionization potential (IP). While the electron leaves the molecule many metastable situations, the Rydberg states, take place, one series from each of the MOs converging to the respective IPs. Obviously, the electrons will be located

relatively far from the molecule, therefore the Rydberg states will differ in extension from the compact valence ones and will require basis sets that generate large and diffuseorbitals able to represent them. Typically, the Rydberg states are labeled, and even represented, in the united-atom approach, that is, by using atomic-type orbitals, because they "see" the molecule as a single atom. The Rydberg states are of interest in gas-phase photochemistry; because of their diffuseness, they are strongly affected by external fields and solvation. Solution chemistry does not contain Rydberg transitions.

Treating simultaneously valence and Rydberg states is not that simple. It requires methods able to deal with the mixing of configurations and orbitals in the wave function. In the past, the literature has been plagued with discussions about how physical is the effect of the valence-Rydberg mixing. In most cases, such mixing was just a consequence of the lack of electron correlation in the calculation of the wave function. Then, both valence and Rydberg orbitals and configurations mix, and valence states, become more diffuse and Rydberg states more compact, with specially dangerous consequences for the valence states, which are extremely sensitive to the mixing. Therefore, if a calculation does not provide clear and compact valence states the result is always suspicious.

Not always the mixing is spurious. One intriguing case is the water molecule. A comprehensive ab initio study performed in 1974 by Goddard and Hunt characterized all the computed states below 11.7 eV as having Rydberg nature, a result supported at the configuration interaction (CI) level. The electronic spectrum of water in the gas phase is currently interpreted as composed either of different Rydberg series or implying excited states with a significant Rydberg character at the ground-state equilibrium geometry. The lowest-energy band of the gas-phase electronic spectrum of water is broad (6.8–8.2 eV), with poorly defined vibrational progressions, and has its maximum absorption around 7.4 eV. There is a unanimous agreement in assigning the lowest-energy band to the 1^1B_1 state, but the valence, Rydberg or mixed valence-Rydberg, nature of the state is still under debate.

In C_{2v} symmetry, the ground state of the water molecule is mainly described by the electronic configuration $(1a_1)^2(2a_1)^2(1b_2)^2(3a_1)^2(1b_1)^2$. The latter orbital is nonbonding. The valence orbitals comprise the four highest occupied MOs listed above and two unoccupied orbitals of a_1 and b_2 symmetries ($4a_1$ and $2b_2$ orbitals). The valence excited states can interact with the corresponding Rydberg states of the same symmetry and close in energy. For instance, the valence singly excited $1b_1 \rightarrow 4a_1$ configuration could be mixed with the $1b_1 \rightarrow 3s$, $1b_1 \rightarrow 3p_z$, $1b_1 \rightarrow 3d_{x^2-y^2}$, and $1b_1 \rightarrow 3d_{z^2}$ Rydberg states.

The properties of the lowest-lying electronic states were studied in the light of CASPT2//CASSCF procedure (see Rubio et al. 2008). The coupling of the CASSCF wave functions via dynamic correlation was dealt by using the MS-CASPT2 method. In this way, all the states of a given symmetry are allowed to interact under the influence of dynamic correlation, and the possible erratic valence-Rydberg mixing can be removed. Properties of the states were determined from the PMCAS-CI wave functions.

Table 3 Computed MS-CASPT2 vertical transition energies (ΔE, eV), oscillator strengths (f), and orbital extensions ($\langle r^2 \rangle$, a. u.) from the PMACAS-CI wave functions for the lowest-lying singlet states of the water molecule at the ground-state equilibrium geometry

State	ΔE	f	$\langle r^2 \rangle$
1^1A_1			13
1^1B_1 ($1b_1 \rightarrow 3s/4a1$)	7.50	0.033	30
1^1A_2 ($1b_1 \rightarrow 3p_y/2b_2$)	9.27	Forbidden	43
2^1A_1 ($3a_1 \rightarrow 3s/4a_1$)	9.86	0.032	45

The next table lists some properties of the low-lying electronic states of the water molecule, employing (after calibration) a one-electron basis set O[5s4p2d1f]/H[3s2p1d] + (2s2p2d), being the diffuse 2s2p2d set, located at the oxygen atom only, necessary to properly describe Rydberg states. Indeed, if no Rydberg orbital is included into the active space, the CASSCF calculation will only yield roots corresponding to valence states. Rydberg and valence orbitals must be treated simultaneously, and this is not possible if there is no Rydberg orbital in the active space. The full valence active space comprises six orbitals with eight electrons and corresponds to (3120), labels that represents each one of the irreducible representations of the point group ($a_1\, b_1\, b_1\, a_2$).

Regarding the first excited state, 1^1B_1, it has a diffuse nature according to the value of $\langle r^2 \rangle$, more than twice that of the ground state, but appreciably lower than that obtained for a pure Rydberg state as the 2^1B_1 ($1b_1 \rightarrow 3pz$) state ($\langle r^2 \rangle = 64$ a.u.). As the state is mainly described by the singly excited $1b_1 \rightarrow 3s/4a_1$ configuration (75 % in the PMCAS-CI wave function), the valence-Rydberg character of the state arises mainly from the nature of the excited orbital, labeled here as $3s/4a_1$ to highlight its intermediate character valence ($4a_1$)-Rydberg (3s), and not from configurational mixing (Table 3).

We can get further insight into the nature of such a state analyzing its evolution with respect to the O–H internuclear distance. For the sake of simplicity, we have considered the symmetrical stretching of both O–H bonds, preserving thus the C_{2v} symmetry of the system. MS-CASPT2 calculations were performed for the three lowest roots of $^1B_1$ symmetry using the O[5s4p1d]/H[2s1p] + (2s2p2d) basis set (since now we will use the same basis sets for each state, one for vertical transitions, another for potential curves analysis) and the (5120) active space (full valence active space plus two extra Rydberg orbitals of a_1 symmetry). The bond angle was kept fixed at 104.5°.

The second and third $^1B_1$ states have Rydberg character at all distances examined since they show high values of $\langle r^2 \rangle$ in comparison with the ground state. However, the situation is clearly different for the 1^1B_1 state. The evolution of the computed $\langle r^2 \rangle$ with the O–H distance shows that the state is getting more and more valence character when stretching the O–H bond as a result mainly of the changes undergone by the virtual orbital: the energy of the antibonding $4a_1$ orbital decrease when lengthening the O–H distance separating this orbital from the related 3 s-Rydberg orbital. In other words, the nature of the 1^1B_1 state evolves form a valence-Rydberg

mixing type ($1b_1 \rightarrow 3s/4a1$) at the ground-state equilibrium geometry to a valence character ($1b_1 \rightarrow 4a_1$) at O–H bond lengths around 1.25 Å and longer, which corresponds only to a symmetrical bond stretching of 0.3 Å from the ground-state equilibrium value. Accordingly, the second $^1B_1$ root becomes the Rydberg state ($1b_1 \rightarrow$ 3s) at O–H distances longer than 1.32 Å. This state constitutes an example of MO Rydbergization postulated by Mulliken. The reason for such behavior is probably that both valence and 2 s Rydberg MOs share the same energy and position in space, yielding therefore a common mixed state (Fig. 12).

A similar analysis of the 1^1A_2 state shows another example of Rydbergization, as its nature evolves from $\left(1b_1 \rightarrow 3p_y/2b_2\right)$, with a 73 % weight at PMCAS-CI wave function at geometries close to the ground-state equilibrium geometry to

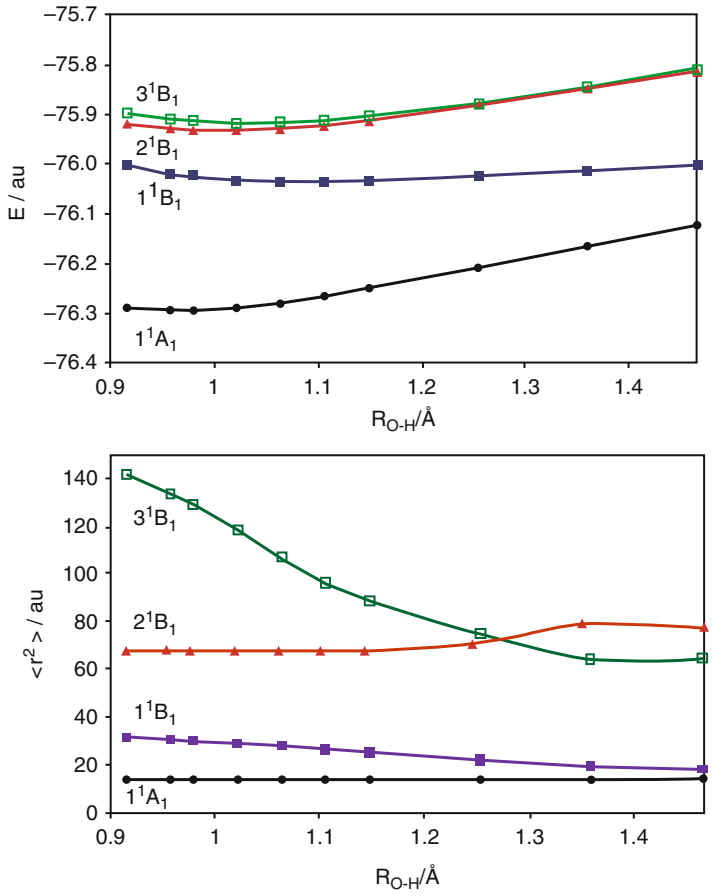

Fig. 12 Symmetrical stretching of the O–H bonds in the ground and low-lying $^1B_1$ states of the water molecule. Potential energy curves referred to the corresponding ground-state minimum (*left*) and computed $\langle r^2 \rangle$ values

($1b_1 \rightarrow 2b_2$) at O–H distances longer than ~1.217 Å, where the second $^1A_2$ state becomes the $1b_1 \rightarrow 3p_y$ Rydberg state. Indeed, the diffuse nature of the 1^1A_2 state decreases when increasing the O–H distance. Calculations of the stretching potential energy curves were performed using the active space (3141) with eight active electrons.

On the other hand, the 2^1A_1 state is characterized mainly by means of the $3a_1 \rightarrow 3s/4a_1$ configuration, with a weight of 70 % at the PMACAS-CI wave function, and by $1b_1 \rightarrow 3p_x$, with a weight of 28 %. This state is closely related to the 3^1A_1 state, which is characterized by the same configuration but in different proportion (69 % and 28 %, respectively). Since the $1b_1 \rightarrow 3p_x$ one-electron promotion is a pure Rydberg excitation, its contribution increases the configurational mixing of the state. Calculations carried out using the active space (4220) show that the valence character of the state increases with the stretching the O–H bonds due to the increment of the weight of the singly excited $3a_1 \rightarrow 3s/4a_1$ configuration in the PMACAS-CI wave function of the state together with the transformation of the excited orbital toward a valence 4 a_1 orbital. To sum up, the 2^1A_1 state is a valence-Rydberg state at the ground-state equilibrium geometry and the mixing is not only due to the nature of the excited orbital ($3s/4a_1$), but also to configurational mixing involving the Rydberg $1b_1 \rightarrow 3p_y$ excitation.

In the light of these results, it is concluded that the electronic spectrum of gas-phase water has a predominant Rydberg character at the Franck–Condon geometry, in line with the general view, although a certain degree of valence character has been shown to exist in the three lowest singlet states, a property that increases far from the FC region.

Practical Aspects

How to Solve Valence-Rydberg Mixing

Many methods are unable to solve properly the valence-Rydberg mixing because they do not include enough correlation effects in the calculation of the wave function, for instance, TD-DFT, CASSCF, RASSCF, or even CC2. In many cases, the problem is aggravated by the use of excessively diffuse basis set functions. For instance, aug-cc- or 6-31++G-type basis sets are already too diffuse, becoming even more when atom-centered and combined in the LCAO approach. The excess of uncontracted diffuse functions has been found necessary to compensate the poor recovery of the correlation energy of the underlying valence basis sets. A procedure to reduce the mixing problems and even to make the obtained Rydberg solutions identifiable is to employ molecule-centered uncontracted functions, as explained in the previous example. Typically, orbitals, excitations, and even population analysis are more clearly represented in such cases.

If the mixing persist, and it can be observed by comparing the orbital extension (second Cartesian moment $\langle r^2 \rangle$) of the valence state with that of the ground state, which should be similar and different from that of the Rydberg state, the only solution is to increase the level of correlation used to compute the wave

function. In the CC treatments, it might mean to include the triple excitations. For multiconfigurational studies, it means to make the CASPT2 states to interact and display orthogonal solutions via the MS-CASPT2 method. Ethene is the most paradigmatic case. A CASPT2(2e,9MOs)/ANO-L C[5s4p2d]/H[3s2p] + 1s1p1d calculation yielded an expected $^1B_{1u}$ valence and Rydberg states at 8.45 and 8.93 eV with orbital extensions $\langle x^2 \rangle$ 52 and 50 a.u., and with related oscillator strengths 0.261 and 0.166, respectively (see Finley et al. 1998; Müller et al. 1999). Several errors plague these results. First, the ground state extension was 12 a.u., therefore none of the two states seemed to have proper valence character. Also, Rydberg states cannot have oscillator strength values close to those of valence states. Second, the correct energetic values are known to be ∼8.0 and 9.33 eV, respectively, leading then to deviations of ±0.4 eV. A clear mixing of states was obtained at the CASSCF level (in which our wave function was obtained) that CASPT2 could not solve properly for the energies. Using the multistate MS-CASPT2 method makes both states to interact. By diagonalization of the effective interaction Hamiltonian, a new set of states is obtained. The new wave functions, PM-CASCI, are linear combination of the previous SA-CASSCF references built with the eigenvectors of the MS treatment. The new results provide two new solutions for the $^1B_{1u}$ valence and Rydberg states at 7.98 and 9.33 eV with orbital extensions $\langle x^2 \rangle$ of 20 and 82 a.u., and with related oscillator strengths 0.360 and 0.051, respectively, leading to a proper interpretation of the spectrum.

A proof that the valence-Rydberg mixing problem is just a problem of lack of correlation in the wave function is obtained when performing RASSCF/RASPT2 calculations on the ethene molecule. When a calculation including all valence σ, π, π^*, σ^* and Rydberg orbitals in the RAS active spaces is carried out, the mixing vanishes even at the RASSCF level and compact valence and diffuse Rydberg states are obtained separated. The RASPT2 step, without the need of a multistate treatment, produces already the correct results.

Focusing on the Valence States

Is it possible to ignore the Rydberg states and focus on the valence spectrum? Logically, this depends on the system and also on our goals. A system with a solvated or proteinic environment perturbs so much the diffuse transitions that it is safe to ignore the Rydberg solutions. For isolated systems, the simultaneous treatment of valence and Rydberg states seems unavoidable, for instance, in pyrrole, where the lowest-energy excited single state is of Rydberg character. Using poor basis sets like 6-31G* type or excluding in general the diffuse functions from the basis set or the Rydberg orbitals from an active space may seem a solution (the Rydberg states cannot be represented in such case), but the obtained "valence" solutions will most probably be of mixed type.

For larger systems the, exclusion of the Rydberg orbitals from the basis or the active space may be less dangerous. Valence states decrease much more in energy with the enlargement of the molecule or the complexity of the chromophore than Rydberg states. It is unusual to find Rydberg series below 4.5–5.0 eV. Therefore, if one is interested in computing low-lying valence states in a large molecule, it

might be safe to first estimate where the lowest-energy Rydberg state is placed and consider that valence states below such energy are treated accurately with compact basis sets. From one to other molecule, this estimation can be done using the respective IPs, which are a good measure of where the Rydberg series are being located energetically. On the other hand, the geometry of the Rydberg states tends to be similar to that of the molecular cation, typically different from that of the valence states. A way to simplify the calculations in many systems may be to use a compact valence basis set to get the geometry of valence states that, far from the FC region, might not suffer of mixing with the Rydberg states. A recent example can be found in the literature for 1,1′ -bicyclohexyliden (see Pérez-Hernández et al. 2008).

Electronic States of Anionic Systems

An Application Example: p-benzosemiquinone Radical Anion

When computing excited states in anionic systems, several cautions should be taken into account. In some molecules, certain states in the anion may lie below the ground state of the neutral system, for instance, those in which the actual electron affinity is positive. Most of the electronic states in anions, including in many cases the ground state, are, however, higher in energy than the ground state of the neutral system and represent temporary anion states, which means that they are unstable with respect to electron detachment. They are typically named Temporary Negative Ion (TNI) states, resonances, or anionic valence-bound states. Conventional quantum-chemical techniques cannot be applied in general to the study of these temporary states since they lie in the continuum of the neutral species plus the free electron. It has been shown that it is, however, possible to obtain well-localized solutions with ab initio methods such as CASSCF and others. Those eigenvalues can be regarded as discrete representations of the TNI or anion resonance states. The metastable states of anions can be classified as either shape or core-excited resonances. From the electronic structure standpoint, shape resonances originate in the attachment of the electron to a virtual orbital of the neutral ground state. Alternatively, they can be viewed as the result of the promotion of the unpaired electron of the LUMO to higher-lying virtual orbitals. Core-excited resonances can be seen, on the other hand, as arising from the attachment of an extra electron to an excited state of the neutral molecule. They can be divided into Feshbach and core-excited shape resonances, depending on their energetic position with respect to the excited state of the neutral molecule involved. The former lie below the parent state of the neutral, whereas the latter are located above.

The theoretical treatment of temporary negative ions can proceed via the use of scattering theory or the employment of modified bound-state quantum chemical techniques. The latter requires the use of stabilization methods for the obtained solutions, for instance, by decreasing the exponents of the diffuse functions. It is of general knowledge that any quantum chemical calculation in anions requires a basis set including extra diffuse functions. Besides the discrete valence TNI states, also named valence-bound (VB) anion states, diffuse states described by a singly

excited configuration involving a diffuse orbital can be also obtained, and they have been sometimes erroneously interpreted as Rydberg states. Singly excited Rydberg states are not present on monoanions because there is no specific direct electrostatic interaction between the neutral molecule and the electron; doubly excited Rydberg states would be instead present at high energies. Then, if not Rydberg states, what is the nature of the diffuse states that a quantum-chemical calculation finds interleaved between the valence excited states of the anion? In the cases where the neutral system has a dipole moment larger than a critical value established between 1.26 and 2.50 D, metastable solutions known as dipole-bound (DB) anion states should be found consequence of the interaction between the additional electron and the dipole moment of the neutral system. These states will have small binding energies (i.e., they will be more stable than the neutral ground state), but in some cases, they may appear higher than the neutral state because of deficiencies of the quantum-chemical treatment. Upon improvement of the level of the calculation, they should end up with a positive binding energy. In systems with smaller dipole moments, however, spurious diffuse solutions are also present together with the VB and DB states, but they are an artificial consequence of the calculation, in particular of the structure of the one-electron basis set, which forces a confinement effect (known as the basis set cage effects) leading to erroneous results. In the low dipole moment molecules, these diffuse solutions for the excited states of the anion try to simulate the neutral molecule plus a free electron by placing the extra electron into the most diffuse orbital available. The computed energy for these states strongly depends on the diffuseness of the basis set employed. In practice, the only way to check if the obtained result is correct is to analyze the stability of the energy, for instance, with the increase of the diffuseness of the basis set, as it will be discussed below.

We have a hint of how to analyze ionic systems in the monodeterminant HF approach. The Koopman's theorem stated that given an N-electron HF single determinant $\left|^{\mathrm{N}}\Psi_0\right\rangle = |\chi_1\chi_2\ldots\chi_c\ldots\chi_N\rangle$, the ionization potential (IP) to produce and $(\mathrm{N}-1)$-electron single determinant by removing an electron from spin orbital χ_c, and the electron affinity (EA) to produce and $(\mathrm{N}+1)$-electron single determinant by adding an electron to spin orbital χ_r, are just $\mathrm{IP}={}^{\mathrm{N}-1}E_c-{}^{\mathrm{N}}E_0 = -\varepsilon_c$ and $\mathrm{EA}={}^{\mathrm{N}}E_0-{}^{\mathrm{N}+1}E_r = -\varepsilon_r$, respectively; that is, the corresponding orbital energies of the protagonist spin orbitals. This "frozen orbital" approximation assumes that the spin orbitals in the $(\mathrm{N}\pm 1)$-electron states are identical with those of the N-electron state, neglecting the relaxation of the spin orbitals in the ionized states. As a result, Koopman's theorem calculations tend to produce too positive IPs and too negative EAs. In addition, we should take into account the correlation effects, which one obtains in going beyond the HF approximation, which will produce further corrections. In general, Koopman's ionization potentials are reasonable first approximations to the experimental ones. On the contrary, Koopman's electron affinities are often inaccurate. Obviously, correlated methods are needed to determine quantitatively such properties.

The ground state of the anion can be then lower (positive EA) or higher (negative EA) in energy than the ground state of the neutral system. How about the excited states of the anion? They will lie in general higher than the neutral ground state,

in the continuum region. As mentioned, it is possible to obtain discrete solutions of the Hamiltonian corresponding to the valence-bound states of the anion, but, because of the confinement effect of the basis set functions, other solutions in which the electrons are placed in the diffuse orbitals will be also present. In a series of calculations on the *p*-benzosemiquinone radical anion ($p\text{BQ}^-$, nonpolar in the neutral form), a number of experiments were performed (see Pou-Amérigo et al. 2000). Apart from the clearly localized valence-bound anionic states, a number of supposedly diffuse states were obtained interleaved with the valence states. What are those solutions? They are spurious solutions caused by the cage effect of the basis set. By decreasing the orbital exponents in the C and O basis sets, it was observed that, while the valence states remained in energy, the "diffuse" solutions started to decrease their absolute energy. The instability remained until their energy converged to the energy of the neutral system, while the electron becomes free and detached from the molecule, an effect that is observed because it always ends up on the most diffuse orbital available. Once identified, the stable solutions above the neutral ground state energy can be TNI states or anion resonances, whereas the instable solutions can be neglected as spurious. The apolar character of $p\text{BQ}^-$ prevents the molecule to have dipole bound anion states (although it has quadrupole-bound states).

The CASPT2 study on $p\text{BQ}^-$ used a basis set of ANO-L type C[4s3p1d]/H[2s1p] plus a set of 1s1p1d diffuse functions centered in the molecule and an active space of 9 MOs of valence $\pi\pi^*$ type, (03010301), distributed into the D_{2h} irreducible representations ($a_g b_{3u} b_{2u} b_{1g} b_{1u} b_{2g} b_{3g} a_u$). The adiabatic electroaffinity (AEA), obtained as the difference between the ground states of the neutral and anion *p*BQ at their respective geometries was computed as −0.33 eV (CASSCF) and 2.01 eV (CASPT2), as compared to the 1.9 eV measured in experiment. The result clearly highlights the fundamental role of the correlation effects. Two series of calculations on the excited states were performed to analyze the spectrum of the system at the geometries of neutral and anion *p*BQ. The former results can be compared with those obtained from photodetachment measurements, whereas the latter are better related to the electron attachment or electron transmission spectroscopy (ETS). Table 4 summarizes the results.

Whereas most of the assignments, made to $\pi\pi^*$ valence-bound states, seem to be explained, the observed band near 2.7–2.9 eV in the $p\text{BQ}^-$ absorption spectrum could not be initially resolved because the degeneracy of the 1^2B_{3u} and 1^2A_u states, both leading to one-photon allowed transitions. This broad band had two main peaks at 2.7 and 3.1 eV, which could be attributed to one or other states. In principle, the transition to the 1^2A_u was computed with larger oscillator strength, but that does not guarantee that the highest peaks correspond to such a state. The vibrational profile of the transitions to both states was therefore computed at the CASPT2//CASSSCF level of calculation by obtaining vibrational energies and TDMs, which required geometries optimizations and Hessians for both states and resolution of the vibrational Hamiltonian. As a result, it was confirmed that the two highest bands at 2.7 and 3.1 eV corresponded to the initial quanta in the vibrational progression of the breathing mode in the 1^2B_{3u} state, whereas the 1^2A_u transition,

Table 4 Computed and experimental excitation energies (eV) and oscillator strengths of the p-benzosemiquinone radical anion

	Anion absorption					Attachment energies[a]			
	Theoretical[b]		Experimental[c]			Theoretical	Experimental[d]		
State	E_{VA}	f	A_{max}	f	State	AE[e]	SF_6	ETS	EI
1^2B_{2g}	–	–	–	–	1^2B_{2g}	−1.64			
1^2B_{3g}	2.25	Forbidden	2.41	Forbidden	1^2A_u	0.91	0.70	0.72	0.77
1^2B_{3u}	2.80	0.05	2.7/3.1	0.15/0.06	1^2B_{2u}	0.96			
1^2A_u	2.82	0.17	2.7/3.1	0.15/0.06	1^2B_{3u}	1.31	1.35	1.46	1.6
1^2B_{1g}	3.25	Forbidden	–	Forbidden	2^2B_{3u}	1.87	1.90	2.15	2.0
2^2B_{3u}	3.56	0.32	3.8/3.9	0.50/0.35	1^2B_{1g}	1.99			

[a]Electronattachment energies at neutral ground state geometry
[b]E_{VA}:vertical excitation energy at the anion ground state geometry
[c]Absorptionmaxima in acetonitrile/water
[d]SF_6 scavenger spectra (SF_6),electron transmission spectroscopy (ETS), and vibrational excitation by electronimpact (EI)
[e]Adiabaticenergy difference: energy difference between the anion state and the neutralground state geometry at such geometry

even when broader, extended to higher energies displaying a larger band area but not so high peaks. Such result proves the danger implicit in the use of vertical excitation energies matching band maxima to assign molecular spectra.

Practical Aspects

Basis Sets and Spurious Solutions in Anions

Closely controlling the quality of our basis set and its effects on our states will confirm the validity of our results. Let us illustrate this aspect with another example. Neutral nitromethane (CH_3NO_2) has a dipole moment of 3.46 D. A standard ANO-L C,N [4s3p2d]/H[3s1p1d] basis set is used supplemented with a set of 1 s diffuse functions (eight primitives) with exponents for the diffuse functions explicitly optimized to deal with Rydberg states. At the CASSCF optimized geometry for the neutral molecule, the lowest state of the anion is 0.530 eV above the neutral ground state at the MS-CASPT2 level. Just by scaling the exponents of the diffuse functions by a factor 0.5 in both the neutral ground state and the lowest anion state, the energy difference decreases to 0.250 eV. Scaling instead the exponents by 0.1 leads to an energy difference of −0.022 eV, that is, the lowest state of the anion has converged below the energy of the neutral system. Increasing further the diffuseness of the basis set does not vary the result, which corresponds to a dipole-bound state of the anion with a small binding energy of 0.022 eV (experimental value 0.012 eV). The lowest valence-bound anion state is computed, adiabatically, 0.23 eV below the neutral ground state of the molecule at the CCSD(T) level. This state remains basically stable with the increase of the diffuseness of the basis set and it is, in general, better represented by using diffuse functions localized in the positive sites of the molecule. To test the stability of the solutions obtained, other techniques have

been developed, such as increasing the effective positive nuclear charge, adding specific counter ions to fix the negative charge, or using penalty functions or electric fields, such the dielectric continuum cavities, which will highly perturb the diffuse solutions.

Negative Electron Affinities

Ionization potentials and electron affinities are intrinsic properties of the DNA and RNA nucleic acid bases (NABs) whose determination enables a deep understanding of all phenomena related to the electron donor and acceptor abilities of the NABs, such as those involving charge transfer and transport along the DNA strand. Determination of the EAs of NABs is difficult both experimentally and theoretically, and the uncertainties range up to several eV, including also changes in the sign of the energies, because both VB and DB anions may be located within a small range of energies, and therefore determination of accurate EAs is uncertain, especially because the type of anion formed may vary with the different experimental conditions. Apart from that, new difficulties interfere the experimental determination of EAs of nucleobases as the presence of different tautomers of the nucleobases which are close in energy in the gas phase. In particular, the canonical (keto) form of guanine, which is the biologically relevant tautomer, has a very low concentration in the vapor, and there is no direct experimental value reported for the corresponding EA.

The electron affinity of a neutral molecule is the energy required for detaching an electron from a singly charged negative ion, or equivalently, the energy released when an electron is attached to the neutral system. Thus, the electron affinity of a neutral molecule can be defined as the energy difference between the ground state of the neutral system and that of the anion. A positive EA implies that the anion is more stable than the neutral species. As in the case of the ionization potential, three theoretical magnitudes (see Fig. 13) are used for describing this transition: the vertical electronic energy difference (VEA_e or VEA) between the ground states of the neutral system and the anion at the equilibrium geometry of the neutral molecule, the adiabatic energy gap (AEA_e) between the minima of the neutral and anion molecule ground states, and the corrected adiabatic property (AEA_0 or AEA) with the addition of the zero-point vibrational energy correction (ZPE). Thus, positive VEAs indicate that the molecule acts as a trap for an excess electron, with an attachment energetically favored and, therefore, the anion can be created spontaneously. In this case, positive AEAs follows, and the system becomes stable, that is, it does not undergo autodettachment and can take part in chemical reactions. On the other hand, negative values for VEAs and AEAs represent the TNI states, existing in short periods of time and becoming prone to photodetachment.

Analysis of the experimental literature on nucleobase EAs shows an extremely confuse situation, ranging from clearly negative values (−0.56 eV) up to largely positive energies (1.51 eV), and including EAs close to 0 eV. In general, determination of EAs represents a technical challenge, especially when they have negative values, and in many cases it is based on indirect measurements. Negative electron affinities can be experimentally measured by electron transmission spectroscopy.

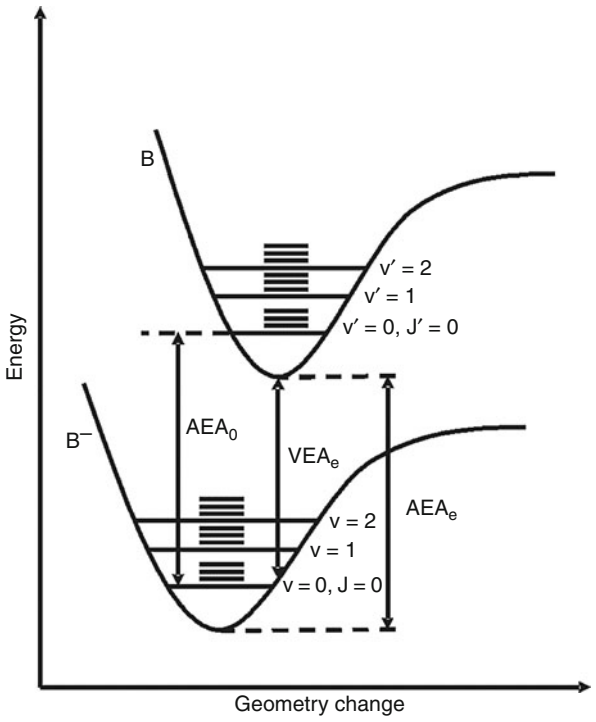

Fig. 13 EA diagram ($B + 1e^- \rightarrow B^-$). Definitions of the theoretical magnitudes related to EA are graphically shown through the electronic, vibrational, and rotational potential energy levels. Those magnitudes are VEA_e (vertical electronic electron affinity, from the neutral ground-state minimum), AEA_e (adiabatic electronic electron affinity, from minimum to minimum), and AEA_0 (adiabatic electron affinity including the zero-point vibrational corrections of the minima)

The technique is able to detect negative ion resonance states, which are energetically unstable with respect to electron autodetachment. It is unclear when the experiment is measuring vertical (VEA) or adiabatic (EAs) attachments, or if the indirectly obtained data truly represent the molecular EAs. In general, however, ETS is the only direct experimental technique which is expected to provide actual VB anions in the region of the resonance states. In particular for gas-phase NABs (except G, which cannot be isolated), ETS measurements report EAs values clearly in the negative region (from −0.22 to −0.54 eV).

In order to compute accurate theoretical results for the VEA and AEA of NABs different levels of theory were taken into account, employing the MP2, CCSD, and CCSD(T), and the CASSCF and CASPT2 methods, in conjunction with the 6-31G*, cc-pVDZ, aug-cc-pVDZ, ANO-L C,N,O[4s3p1d]/H[2s1p] (hereafter ANO-L 431/21), and ANO-L C,N,O[4s3p2d1f]/H[3s2p1d] (hereafter ANO-L 4321/321) basis sets. Methods and basis sets were selected to obtain the most accurate values from preliminary calculations on atomic systems, in which the required levels of highly flexible enough basis sets and strongly correlated methods to obtain predictive EAs were determined. Geometry optimizations of both neutral and anionic NABs were carried out at the MP2/6-31G(d), MP2/aug-cc-pVDZ, CASSCF/cc-pVDZ, CASSCF/ANO-L 431/21, and CCSD/aug-cc-pVDZ levels of theory. No symmetry restrictions (C_1 symmetry) were imposed, whereas all minima

were characterized by computing second derivatives at the same level, except in the case of CCSD/aug-cc-pVDZ where the geometries were tested comparing with the optimized parameters at the other levels of theory. At the respective equilibrium structures, additional CASPT2 and CCSD(T) calculations were performed to account for the most accurate energy values. ZPE corrections were included at different levels using the harmonic approach. The active space for the CASSCF calculations in geometry optimizations comprises the full π-valence system, except the molecular orbital localized mainly on the nitrogen atom of the NH_2 group in the case of cytosine, adenine, and guanine, whose occupation number is very close to two. This MO is further included in conjunction with the lone pair electrons and orbitals of the heteroatoms in the final CASSCF and CASPT2 calculations of VEAs and AEAs, except when the large ANO-L 4321/321 basis set was employed (see Roca-Sanjuán et al. 2006, 2008a).

The analysis of the results shows that neither DFT procedures nor the MP2 method have the required accuracy, either by the known problems of DFT to deal with negative centers or by the spin contamination problem that affects MP2. CCSD(T)//CCSD/aug-cc-pVDZ and CASPT2(IPEA)/ANO-L 4321/321//CASSCF/ANO-L 431/21 will be established as the most accurate procedures, both for vertical and adiabatic EAs. It must be emphasized here that the main factor to achieve accurate results for the VEA of NABs by using ab initio methods is the employment of atomic one-electron basis sets flexible enough to describe both the spatial distributions of electrons and their correlation effects and including functions decaying slowly with the radial distance. The CASPT2//CASSCF strategy has the advantage to compute several states of the system, and it is possible to easily distinguish between the different solutions. Single-reference methods are not free of problems, because they can only obtain the lowest solution. For instance, at the corresponding geometry of the neutral species, adenine and guanine, in which the VB anion state lies much higher than other diffuse states, the CCSD and CCSD(T) computations lead initially to a diffuse and low-energy spurious solution in which a delocalized electron is located far from the molecule in a diffuse orbital. Finally, the sequence of stable solutions obtained for the VEAs of NABs using the reference CCSD(T) and CASPT2(IPEA) levels of theory is established as U $\approx$ T $>$ C $>$ A $>$ G, ranging from -0.61 eV (U) to -1.14 eV (G). Therefore, for guanine is less favorable to accept an electron at the neutral molecule geometry.

Regarding gas-phase adiabatic EAs, those for uracil and thymine were determined very close to zero, whereas cytosine has a small negative AEA. The sequence of AEAs for isolated NABs can be established as 0 eV $\sim$ U $\sim$ T $>$ C $>$ G $>$ A. Purines are much less favorable than pyrimidines to retain the electron attached to the neutral nucleobase, and after geometry relaxation adenine becomes the poorest electron acceptor of all NABs, in contrast to what occurred for the vertical EAs, in which guanine had the more negative value. To understand the differences among these compounds, it can be also analyzed how the inductive effect (and the number of stable resonance structures) makes more or less stable the new center of negative charge created by the addition of the new electron. It should not be forgotten

that the accepting properties of the systems will largely change in solvated or biological environments. For instance, a recent calculation performed for the EAs of the cytosine molecule within a chain of oligomers, $dC_{18} \times dG_{18}$, employing a CASPT2/MM approach led the AEA of cytosine to change from −0.26 eV in vacuo to 0.69 eV in the biological surrounding, becoming therefore the molecule a strong electron acceptor.

Photochemistry: On the Trail of the Energy

An Application Example: Thymine

The energy absorbed by a molecule can be released radiatively, that is, slowly emitted via fluorescence or phosphorescence, or nonradiatively. In that case, it can give rise to productive photochemistry, yielding photoproducts different from the initial species, or it can become unproductive, meaning that it is dissipated to the environment through the vibrational degrees of freedom, ending in some cases in the initial ground state of the system. Photophysics and photochemistry is typically a combination of all such processes. From the theoretical viewpoint, the best initial strategy to understand the photochemical processes is trying to follow closely the path of the energy from the initially populated states at the FC region toward favorable regions of the PEHs. That means to trace the lowest-energy possible pathway until reaching an energy barrier, that is, a minimum, or a transition state, or a hypersurface crossing, in particular conical intersections (CIs), the protagonists of the ultrafast radiationless nonadiabatic energy transfers between PEHs. The only strategy that guarantees finding PEH points along the lowest-energy path is the computation of a minimum energy path (MEP), that is, a steepest-descendent pathway. The procedure is performed generally in mass-weighted coordinates, and it is equivalent to the Intrinsic Reaction Path approach. Once localized, the crossings and CIs, the estimation of their accessibility, and the calculation of the interstate couplings (like the SOC between singlet and triplet states) and transfer probabilities will help understand and predict favorable IC and ISC processes. The described strategy is named as the Photochemical Reaction Path approach, and it can be applied to solve a large number of photochemical problems.

An interesting example of nonproductive photochemistry can be found in important compounds such as the DNA/RNA natural nucleobases, which were determined a long time ago as basically nonfluorescent. Modern femtochemical techniques have determined fluorescence decay times in the DNA/RNA nucleobase monomers, nucleosides, and nucleotides in different media to be ultrafast. In particular, molecular beam measurements reported two main decay lifetimes near 100 fs and few ps in all five natural NABS: thymine, uracil, cytosine, adenine, and guanine (Crespo-Hernández et al. 2004). Finding ultrafast decays in a molecule suggests the presence of extremely efficient internal conversion channels, and therefore the presence of accessible conical intersections, in particular one connecting the initially populated bright spectroscopic $\pi\pi^*$ singlet and the ground state, yielding therefore nonfluorescent species. The ultrashort lifetime of nucleobases

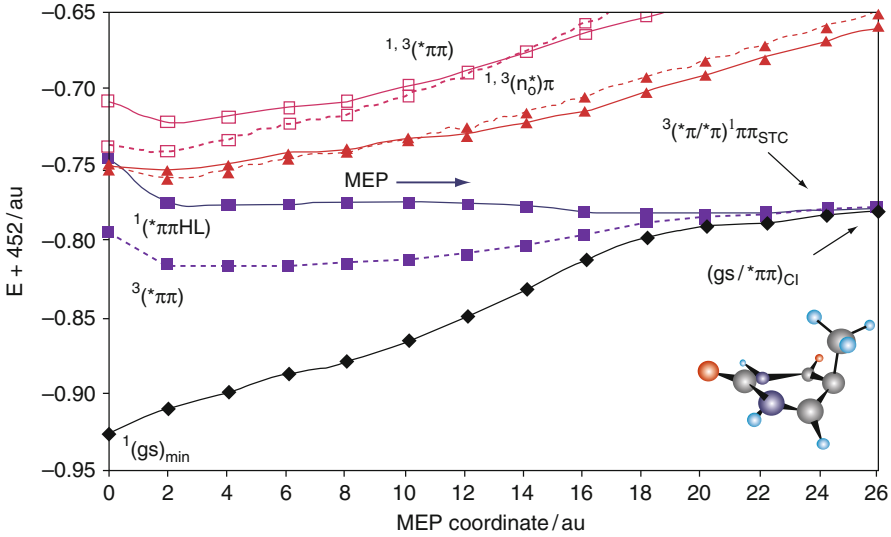

Fig. 14 Evolution of the ground and lowest singlet excited states for thymine from the FC geometry along the $^1(\pi\pi^* HL)$

is an intrinsic molecular property as it has been proved in recent years both from theory and experiment. Indeed, such photoprotective properties may well be very important at the beginning of life in our planet given that there are evidences that point out that life began on Earth millions of years before the development of the ozone layer. The ultrafast decay channels of the nucleobases may have favored photostable natural nucleobases against less stable derivatives (Serrano-Andrés and Merchán 2009). This type of photochemical mechanisms may have been operative as natural selection processes reaching in that way the genomic stability until the development of the ozone layer and complexes mechanisms of reparation of DNA.

Thymine photophysics is here selected as an illustration of the computational strategies mentioned above (see Merchán et al. 2006; Serrano-Pérez et al. 2007b). Figure 14 displays the CASPT2 energies of the lowest-lying singlet and triplet states of thymine along the MEP calculated for thymine on the bright $^1\left(\pi\pi^*\right.$ HL, HOMO $\rightarrow$ LUMO) state from the FC geometry. The level of the calculation was CASPT2//CASSCF(14e$^-$,10MOs)/ANO-S C,N,O[3s2p1d]/H[2s1p]. Apart from the radius of the hypersphere controlling the distance from the initial geometry, not any other restriction was imposed to the calculations.

From thymine ground state and upon near-UV absorption at the FC region, most of the population reaches initially singlet excited states, and in particular, it is the transition to the $^1(\pi\pi^* HL)$ excited state at 4.89 eV which has the largest oscillator strength up to 6 eV, that is, 0.167. The ultrafast nonradiative decay undergone by thymine in the femtosecond range can be rationalized by the barrierless character of the path leading from the FC region toward a CI seam with

the ground state, $\left(\mathrm{gs}/^1\pi\pi^*\right)_{\mathrm{CI}}$. Unlike simple geometry optimizations, the use of the MEP technique guarantees the absence of energy barriers along the lowest-energy path. The structure of the CI at the end of the MEP can be characterized as ethene-like, a diradical species (as many CI are) involving combined stretching and twisting of the ethylenic bond and leading to a screw-boat (S) conformation $^5S_6$ for the six-membered ring. The presence of an accessible CI explains also the low fluorescence quantum yield $\left(\sim \phi_F = 10^{-4}\right)$ detected for thymine with band origin at 4.5 eV in water. This weak emission can be related to the presence of a high-lying planar $^1(\pi\pi^* HL)$ minimum computed adiabatically from the ground state at 4.49 eV, whereas a nonfluorescent $^1(n_O\pi^*)$ minimum is found at 4.05 eV with a minor contribution to the emissive properties. As recently estimated by means of reaction dynamics, the near 100 fs decay detected in natural nucleobases can be related to the barrierless path from the FC to the CI region. Although other slower singlet decay pathways could beestimated proceeding through the low-lying $n_O\pi^*$ state, that crosses the MEP at high energies and has its own CI with the ground state, still the main relaxation path for the energy runs along the computed $^1(\pi\pi^* HL)$ barrierless MEP, which will transfer most of the energy toward the ground state and back to the original species. Such behavior means that the systems are largely photostable, as it has been proved for natural and methylated nucleobases, and that the mechanism can be considered an intrinsic property of the systems. Figure 15 illustrates such mechanism for the pyrimidine nucleobases, thymine, uracil, and cytosine, although the basics also hold true for the purine systems, adenine and guanine (Merchán et al. 2006; Serrano-Andrés et al. 2006, 2008; ?). Further insight can be only obtained when reaction dynamic calculations are performed on larger regions of the PEHs (see, for instance, Szymczak et al. (2009)).

The decay along the singlet manifold is not the only procedure for energy relaxation. Efficient population of the triplet manifold can also take place, essentially through intersystem crossing (ISC) processes taking place along the main decay process on $^1(\pi\pi^* HL)$ or via photosensitization from endogenous or exogenous species. Triplet states are frequent intermediates in different types of photoinduced reactions. Both their usual diradical character and long lifetimes make them reactive species prone to interact with other systems. Among the most important reactions involving triplet states, those related to DNA/RNA purine and pyrimidine nucleobases have undoubtedly attracted more attention, in particular the photodimerization of pyrimidine nucleobases, considered to be the most frequent DNA lesion taking place after UV light irradiation. Through their triplet states, DNA/RNA nucleobases may not be as photostable as expected. As seen in Fig. 14, along the $^1(\pi\pi^* HL)$ state, MEP is clear that two singlet-triplet crossings (STC) are accessible, and therefore two ISC processes may take place: at 4.8 eV with the $^3(n_O\pi^*)$ triplet state, $\left(^3n_O\pi^*/^1\pi\pi^*\right)_{\mathrm{STC}}$, and at 4.0 eV, further along the relaxation path and near the ethene-like CI with the ground state, directly with the lowest $\left(^3\pi\pi^*\right) T1$ triplet state, $\left(^3\pi\pi^*/^1\pi\pi^*\right)_{\mathrm{STC}}$, a structure displaying the same type of screw-boat puckered geometry with a stretched and twisted double bond $C_5{=}C_6$ as at the $\left(\mathrm{gs}/^1\pi\pi^*\right)_{\mathrm{CI}}$ CI. Efficient ISC requires both small singlet-triplet energy gaps and large spin-orbit coupling elements at the regions of degeneracy.

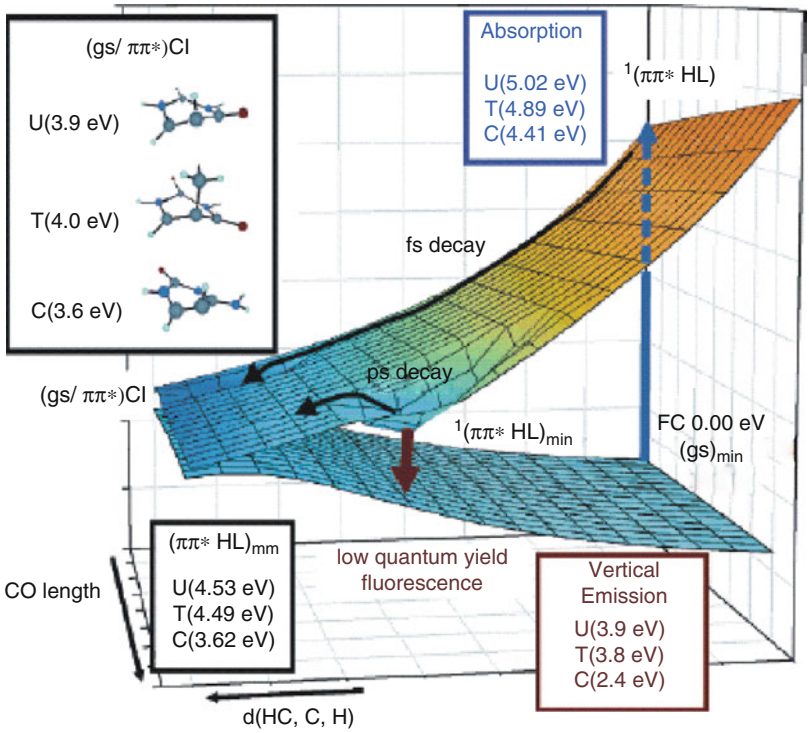

Fig. 15 Global scheme of the photochemistry of uracil (U), thymine (T) and cytosine (C) as suggested by the CASPT2 calculations

As compared to IC, taken place essentially in small zones where the seam of CIs becomes accessible, the regions of the potential energy hypersurfaces for effective ISC are more extensive.

Figure 16 includes a scheme describing the population of T_1 based on our CASPT2 calculations, and three are the suggested competitive mechanisms from the initially populated singlet state by means of STC processes:

$$\begin{array}{llll} \surd & {}^1\pi\pi^* & \rightarrow & \left({}^3n_O\pi^*/{}^1\pi\pi^*\right)_{\mathrm{STC}} \rightarrow \left({}^3n_O\pi^*/{}^3\pi\pi^*\right)_{\mathrm{CI}} \\ \surd & {}^1n_O\pi^* & \rightarrow & \left({}^1n_O\pi^*/{}^3\pi\pi^*\right)_{\mathrm{STC}} \\ \surd & {}^1\pi\pi^* & \rightarrow & \left({}^3\pi\pi^*/{}^1\pi\pi^*\right)_{\mathrm{STC}} \end{array}$$

Regarding the higher-energy ISC process, the ${}^3(n_O\pi^*)$ triplet state can be populated from ${}^1(\pi\pi^*)$ in the STC crossing region, in which a high SOC of 8 cm^{-1} has been computed favoring the process. From such region, a MEP on the ${}^3(n_O\pi^*)$ state leads directly to its energy minimum, ${}^3(n_O\pi^*)_{\mathrm{min}}$, placed at 3.93 eV adiabatically from the ground state. Basically degenerated, we have located a conical intersection connecting the triplet states ${}^3(n_O\pi^*)$ and ${}^3(\pi\pi^*)$, facilitating the population switch

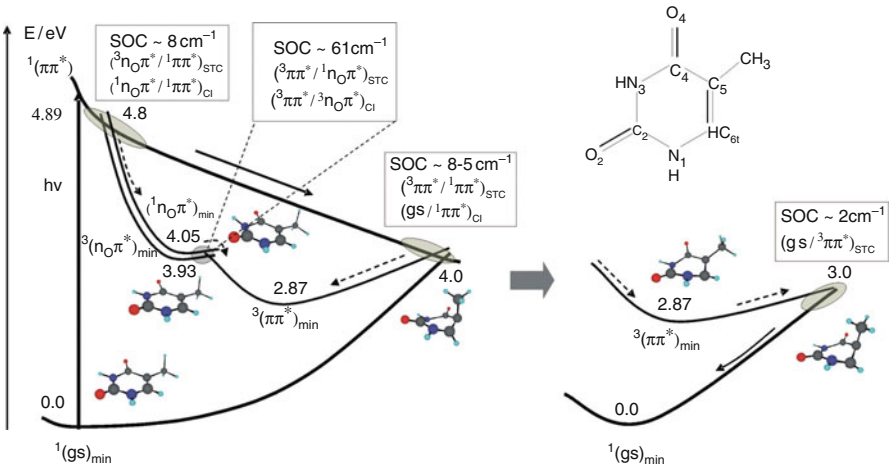

Fig. 16 Scheme, based on CASPT2 results, of the photochemistry of thymine focused on the population of the lowest-energy triplet state

toward the lowest $^3(\pi\pi^*)$ state. Another computed MEP on the T_1 hypersurface leads from $\left(^3n_O\pi^*/^3\pi\pi^*\right)_{CI}$ to the $^3(\pi\pi^*)$ state minimum. A second ISC channel also relates with an $n\pi^*$ state. In our results, as well as in previous theoretical studies, the presence of the low-lying $^1(n_O\pi^*)$ state is confirmed, lying almost isoenergetic with the lowest ethene-like $\left(gs/^1\pi\pi^*\right)_{CI}$ CI near 4.0 eV. The $^1(n_O\pi^*)$ state may then be related to the observed dark state in pyrimidine nucleobases in solution. Additionally, we have found that, as both singlet and triplet $^{1,3}(n_O\pi^*)$ states have almost parallel hypersurfaces because of the typical small singlet-triplet splitting of $n\pi^*$ -type states, near degeneracy structures with the lowest $^3(\pi\pi^*)$ state are equal to both states. Therefore, close to the $\left(^3\pi\pi^*/^3n_O\pi^*\right)_{CI}$ CI, we have found a singlet-triplet crossing $\left(^3\pi\pi^*/^1(n_O\pi^*)_{STC}\right.$. The extremely large computed SOC terms, ~ -61 cm^{-1}, guarantee also an efficient ISC process in the region, confirming this mechanism as relevant for the overall process. In principle, in different environments such as in polar solvents, it is expected that the $n\pi^*$ -type excited state will destabilize with respect to $\pi\pi^*$ -type excited states. Despite those effects, both singlet and triplet $n_O\pi^*$ -type states are estimated to lie in the solvent below the $^1(\pi\pi^*\text{HL})$ state at the FC geometry, confirming the existence of the STC crossing upon decay along the $^1(\pi\pi^*\text{HL})$ state. Finally, a third ISC channel directly connecting the lowest $\pi\pi^*$ states is found at low energies with SOC values ranging 5–8 cm^{-1}. The profile of the computed MEP in Fig. 16 suggests that the lowest-energy ISC mechanism may enhance its efficiency with respect to the other nucleobases resulting in larger quantum yields because the region for the energy transfer near to the end of the MEP seems to be much more extended. The presence of three basic ISC funnels in thymine (occurring also in uracil and adenine but not in guanine and cytosine where the higher-energy $n\pi^*$ -mediated channels are

absent) successfully explains for the first time the wavelength dependence on the measured ISC quantum yield (ϕ_{ISC}) reported in thymine (uracil and adenine too) on the basis of the location and accessibility of the two STC crossing regions upon the initial excitation conditions. In the case of thymine, the ϕ_{ISC} value increases from 3.9×10^{-3} at 280 nm (4.43 eV) to 5.2×10^{-2} at 240 nm (5.17 eV). At low excitation energies, only the lowest-lying ISC computed to take place close to 4.0 eV will be activated. At higher-energies, however, the channels near 4.8 eV will be additionally activated, increasing the overall triplet formation yield.

From any of the STC or CI regions, the lowest triplet state is populated by any of the previous ISC processes and reaches its minimum, as proved by the computed MEPs. The high reactivity attributed to this triplet state originates from its partial diradical character on C_5 and C_6. The minimum is placed at 2.87 eV, adiabatically from the ground state optimized minimum. As a final aspect of the evolution along the triplet manifold in thymine, we have located the singlet-triplet crossing connecting the $^3(\pi\pi^*)$ and the ground state, and mapped the MEP leading from such STC toward $^3(\pi\pi^*)_{min}$. The crossing is placed at near 3.0 eV from the ground state minimum, what means that there is a relatively small barrier of 0.13 eV (3.0 kcal mol^{-1}) to reach $\left(gs/^3\pi\pi^*\right)_{STC}$ from $^3(\pi\pi^*)_{min}$. The computed electronic SOC is, however, somewhat low, $\sim$2 cm^{-1}, predicting for the triplet state a long lifetime and a slow relaxation, becoming therefore prone to undergone reactivity (Merchán et al. 2005; Serrano-Pérez et al. 2007b).

Practical Aspects

Reaction Paths: MEP Verses Other Approaches

Computing minimum energy paths (MEP) is an expensive procedure in which each point of the MEP is obtained from individual optimizations. Special care is required in the selection of active spaces, which should be flexible enough and appropriate along the whole path, and the radius of the hypersphere, which should be not too small or too large. When the steepest-descendent MEP ends finding an energy barrier, for instance, as frequently in the minimum of the state, other procedures should be employed. It must occur also that the calculation of the MEP is too expensive. It is possible to continue using the MEP algorithm upward, that is, by increasing the hypersphere radius increasing energy paths can be computed. Other procedure is to search for a transition state (TS) along the path that represents the energy barrier. More general strategies are normally required to map the full profile of the reaction linking two predefined regions of the PEHs. One possibility is performing a relaxed scan, that is, varying by specific amounts one internal coordinate which is expected to be the leading structural modification along the path and, at each selected value, to perform a restricted geometry optimization in which only such coordinate is frozen. Relaxed scans are extremely dangerous and should be used with caution, because they often produce unconnected and useless paths. There is no guarantee that the lowest-energy value for each of the selected frozen-coordinate structures belongs to a connected path. In many cases, the results will

jump from one region of the PEH to another along a profile that should represent a reaction path but does not at all. In general, the only relatively safe case in which a relaxed path can be used is that leading to a dissociation of an atom, because in this case the path has a descendent character while the bond length is increased.

Instead of a relaxed scan is preferable to make a linear interpolation in internal coordinates (LIIC). This is a predefined path in which each structure is generated from the previous one by adding linearly to each internal coordinate the geometrical change of the initial and final structures divided by the number of steps. The advantage of this profile is that the points belong to a connected path. The disadvantage is that it will most probably give rise to too high energy barriers, although they will always be upper bounds of the actual barrier. For instance, in a study of the photophysics of psoralen, it was determined that the MEP from the FC region led the system to the minimum of the initially populated ${S_1}^1(\pi\pi^*)$ state. Psoralen is a strongly emitting system, especially by phosphorescence (in a ratio 7/1 with respect to fluorescence). Emission from the ${S_1}^1(\pi\pi^*)$ minimum explains the fluorescence of the system, but, what about phosphorescence? The population of the triplet manifold seems not to take place close to the FC region, and no relevant STC crossing takes place soon along the MEP. Therefore, an effective triplet formation has to take place by depopulating the singlet state minimum by an accessible STC. The $T_1(T_\pi)$ state lies much lower in energy than the $S_1\left(S_\pi\right.$ state, both vertically ($\sim$0.6 eV) and adiabatically ($\sim$1.2 eV). Direct interaction between the singlet and triplet vertical excited states is unlikely, and therefore another mechanism has to be found, involving most probably population of higher-energy triplet states and subsequently internal conversion toward T_1. The qualitative El-Sayed rules indicate that the spin-orbit coupling is large between states of $\pi\pi^*$ and $n\pi^*$ types and small between states of the same character. In addition, it is known that the ISC process is enhanced in molecules with heavy atoms (the heavy-atom effect) because the SOC terms increase. There are two necessary conditions to be fulfilled for an effective ISC to take place: low energy gap and high spin-orbit coupling between a singlet and a triplet state. Thus, the nonradiative decay to a triplet state should occur along the relaxation pathway of S_π, starting from the Franck–Condon region, and in close vicinities of a singlet-triplet crossing. A LIIC profile was computed (see Fig. 17) from $(S_\pi)_{\mathrm{MIN}}$ to the minimum of the $1^1A''$, $(S_n)_{\mathrm{MIN}}$, a path considered to be appropriate to find the singlet-triplet crossing, since both S_n and T_n states share the same basic structure and energetics (see Serrano-Pérez et al. 2007a). At half of the path the S_π and ${T_n}^3$ $(n\pi^*)$ states cross, and the SOC value is high enough (9 cm^{-1}) to guarantee an efficient ISC. A barrier of 0.4 eV, surely an upper bound, has to be overcome to reach the T_n state.

The geometry closer to the crossing point was used as initial geometry to initiate a search at the CASSCF level for a singlet-triplet crossing (STC) between $S\pi$ and T_n, that is, $(S_\pi/T_n)_X$. As expected, the main change in geometry from the minimum of S_π to the optimized structures of S_n or T_n is related to the length of the C=O bond, much larger ($\sim$0.15 Å) than the ground state value for the $n\pi^*$ states. It is not surprising that the STC structure displays an intermediate value of 1.220 Å. The

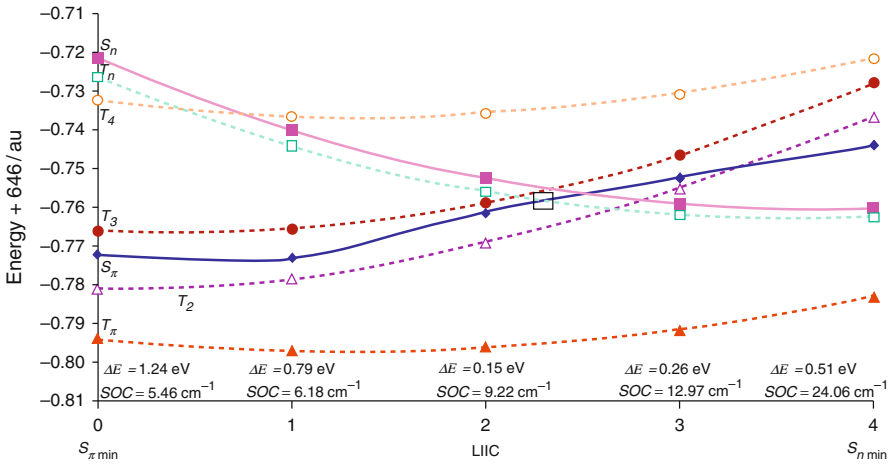

Fig. 17 LIIC path for psoralen. The crossing region between S_1 and T_n is highlighted with a square

SOC term in STC is computed as 6.4 cm^{-1}. As displayed in Fig. 18, the barrier from $(S_\pi)_{MIN}$ to $(S_\pi/T_n)_X$ was computed 0.36 eV at the CASPT2 level. This is still a large energy barrier; however, the crossing is now below the energy of the S_π state at the FC geometry, and therefore the system has enough excess energy to access the STC region and make the ISC process to take place.

Once the T_n state is efficiently populated through a favorable ISC mechanism, it will quickly evolve toward the nearby energy minimum, $(T_n)_{MIN}$, placed 0.11 eV below $(S_\pi/T_n)_X$. At those geometries, T_n is the second-energy excited triplet state and can be expected that the energy follows a pathway for favorable internal conversion (IC) toward the low-lying T_π state. Actually, a conical intersection has been found, $(T_n/T_\pi)_{CI}$, placed isoenergetic with $(T_n)_{MIN}$. An efficient IC will therefore take place transferring the population to T_π, which will subsequently evolve to its own minimum, $(T_\pi)_{MIN}$, from where phosphorescence (P) will take place. Additionally, the molecule may react with thymine or transfer its energy to the molecular oxygen exerting the photosensitizing action. The family of furocoumarins (8-MOP, 5-MOP, khellin, TMP, and 3-CPS) was studied at the same grounds showing that all the compounds share the same mechanism of population of T_π. The differences in the values of the barriers (to reach the crossing regions on the hypersurfaces), besides the values of SOC, are actually the key points in considering the photosensitizing ability of each compound.

A final remark should be introduced. Considering that $n\pi^*$ states tend to destabilize in polar solvents, the (T_n)-mediated mechanism here proposed will surely decrease its importance in polar media, remaining a plausible candidate for the efficient population of the low-lying T_π triplet state in the gas phase. Indeed, the relevance of out-of-plane displacements for the relaxation of the S_π state has been emphasized recently at DFT/MRCI level of theory, together with the enhancement

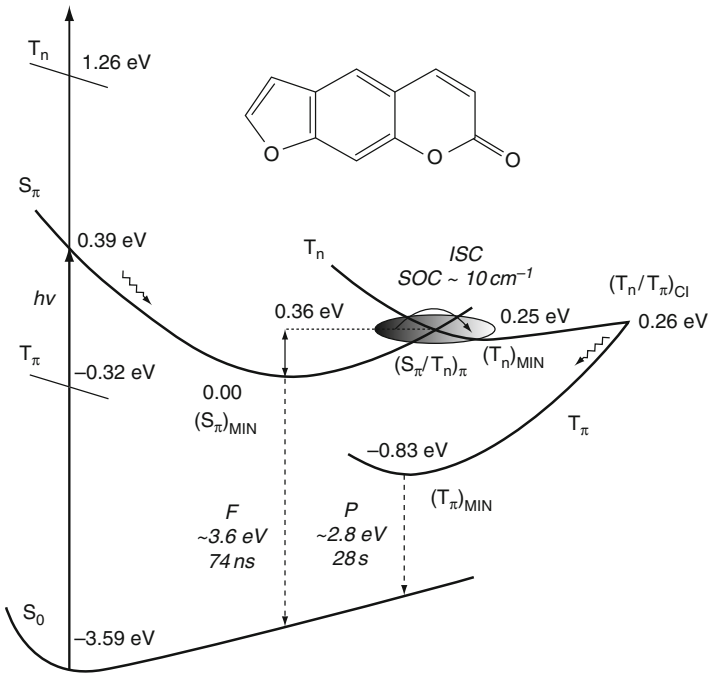

Fig. 18 Scheme of the gas-phase photochemistry of psoralen based on quantum-chemical CASPT2 calculations. Energies are referred to $(S_\pi)_{MIN}$

of the SOC terms between low-lying S_π and T_π states by vibronic coupling effects involving out-of-plane modes, as an alternative mechanism for efficient ISC.

How to Compute Conical Intersections

There is hardly any good recipe to make the calculation of CIs more efficient, in particular those relating the ground and excited states that are usually the most relevant features for the photochemistry of the systems. Experience orientates toward certain types of CIs common in organic compounds, like the so-named ethane- or methanamine-like CIs, taking place by the stretching and twisting of C=C and N=C bonds, respectively, or situations of C=O stretching, in both cases creating an almost diradical species. In many other cases, one has to let the computational algorithm to find the right path. One initial approach is trying to find the minimum energy CI (MECI), considered to be the most relevant for the photochemical process. Technically, it is easier to look for MECIs, because one can simply search for the lowest-energy crossing point fulfilling the nonadiabaticity rules. Calculation of nonadiabatic terms, the so-called branching space, is complex. Time-consuming projected-gradient or Barry's phase algorithms have to be used to guarantee the true CI nature. In practical cases, there are some implementations

available at a high level of theory, like that at the MRCI level in the COLUMBUS code (see Lischka et al. 2004).

In many cases, however, it is safe enough just to compute the minimum energy crossing point (MECP), discarding the calculation of nonadiabatic coupling, because the obtained solution already defines the crossing region accurately enough. Even when the MECI (or MECP) is found, nothing guarantees that this is going to be the relevant structure from the photochemical viewpoint. It might happen that the MECI is placed in a region that it cannot be easily accessed from the main energy decay paths of the molecule. Therefore, the accessibility of the region is the key point making a CI (or a seam of CIs) relevant for the photochemistry. The next example can properly illustrate this idea. In the same manner as described for thymine, the MEP from the FC region on the bright $^1(\pi\pi^*$HL) state was computed for guanine at the CASPT2//CASSCF (14e$^-$,12MOs)/ANO-S C,N,O[3s2p1d]/H[2s1p] level of calculation (see Fig. 1, top), leading, through a barrierless path, to a crossing with the ground state at the end of the MEP. Independently, a MECI search was carried out to find the lowest-energy CI (Fig. 19, bottom, right), which provided a geometry somewhat different from that obtained at the end of the MEP. To solve the problem, a CASSCF path was computed along the seam of CIs (Fig. 19, bottom). The points of this profile have been computed to belong to the seam, that is, both the ground and the $^1(\pi\pi^*)$ states are degenerated, and by successive optimizations in which the radius of the hypersphere has been increased from the initial MECI structure. It is clear that the MECI energy is lower than any other point of the seam, but also that a barrier separates that structure from a CI which is higher in energy but accessible from the end of the MEP (the actual source of the energy) and whose structure is quite similar to that of the end of the MEP (Fig. 19, bottom, left). That CI is the relevant funnel for the photochemistry of guanine.

Other serious problems can be found when computing PEHs crossings, although it is as general issue in quantum chemistry. Determination of molecular structures is one of the most time-consuming steps in computational chemistry. It is therefore a typical strategy to use lower-level methodologies to locate singular-point geometries (minima, transition states, conical intersections . . .) and correct then the energetics by employing higher-level methods. Implicit in this procedure is the problem of the differential correlation effects. We shall check this issue in the case of the location of conical intersections (or any other hypersurface crossing). A usual protocol employed for that purpose is named CASPT2//CASSCF, meaning that geometries have been determined at the CASSCF level and, at those structures, the energies have been computed at the CASPT2 level. CASSCF includes the static or long-range correlation effects and provides wave functions which describe qualitatively the electronic states and form the reference for subsequent calculations. The remaining short-range correlation effects, dynamical electron correlation, are included by means of a second-order perturbative treatment, CASPT2. Depending on the nature of the state, the effects of the static and dynamical correlation energies differ, affecting directly to the description of the hypersurface. Figure 20 summarizes what we have called the differential correlation effects on the location of surface crossing within the CASPT2//CASSCF protocol. It is important to recall that this problem

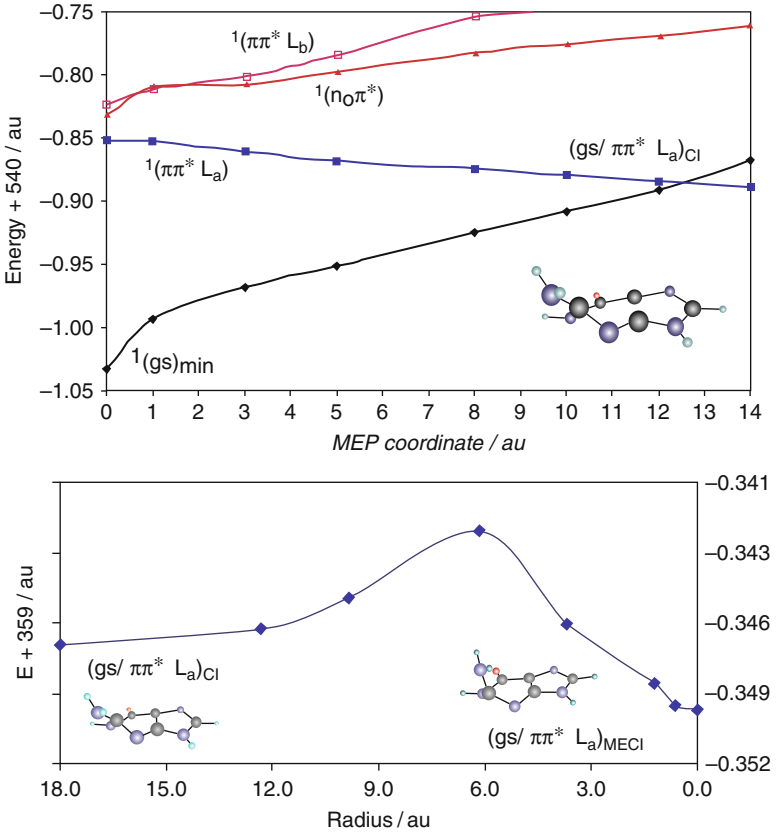

Fig. 19 CASPT2 low-lying singlet states along the MEP from the FC region on the $^1(\pi\pi^*\text{HL})$ state in guanine (*top*) and seam of CIs (*bottom*) computed from the MECI structure (*bottom, right*) to the CI close to the end of the MEP (*bottom, left*)

is general in quantum chemistry if two levels of theory are being used to compute geometries and energies, e.g., CCSD(T)//CCSD or MRCI//CASSCF.

Turning to our example, two main situations do actually occur. In cases where the PEHs computed at the CASSCF and CASPT2//CASSCF levels behave approximately parallel (Case A), the CASSCF optimized geometries will be in general correct, despite they have been computed at a lower level of theory. It means that the dynamical correlation contributions are quite regular and similar for the two states in ample regions of the PEH. In these situations, two-step computational strategies like CASPT2//CASSCF can be confidently applied. When dynamical correlation is markedly different for the states considered and varies significantly along the PEH of interest, geometry optimization has to be carried out at the highest correlated level (Case B). Otherwise, the uneven contributions of dynamical correlation may lead to unphysical crossings and interactions between the two electronic states.

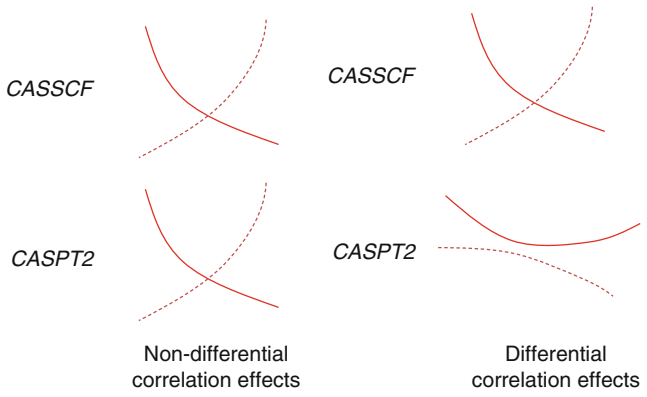

Fig. 20 Different effects of correlation within the CASPT2//CASSCF protocol

The crossing can be then placed at a different geometry or it may never occur! Normally, states of ionic (in the valence-bound sense) nature have larger amounts of dynamical correlation in contrast to covalent states, in which the static correlation is able to recover large fractions of the total correlation energy. When both types of states are involved in a crossing, it is clear that a CASSCF description will give an unbalanced picture of the hypersurfaces and, obviously, of the crossing, which can be recovered only by inclusion of the dynamical correlation, with CASPT2 for instance, even in the searching of states crossing. In difficult cases, it can be convenient to map a grid of CASPT2 points in the region of the CASSCF CI to find the degeneracy at the highest level, or, if possible, to compute the CI with CASPT2, more specifically, with MS-CASPT2, considering that CASPT2 provides non-orthogonal and therefore mixed solutions. This is a tricky issue that was studied recently and requires careful consideration.

Figure 21 displays the result of CASPT2/MS-CASPT2/ANO-L Li [4s2p]/F [4s3p1d] calculations on the lowest-lying $^1\Sigma^+$ states of the LiF molecule along its dissociation path (see Serrano-Andrés and Merchán 2005). The lowest state has ionic and covalent character at short and long distances, respectively, because the lowest dissociation channel leads to the ground states of the neutral atoms, whereas the opposite occurs for the second state. At intermediate distances, the lowest ionic and next covalent states interact and change their character smoothly thanks to the multiconfigurational treatment. According to the non-crossing rule for states of the same spin and spatial symmetry that holds true for diatomic molecules, the two curves cannot become degenerate and instead an avoided crossing should take place. In the sample calculations, the results of three different active spaces and two levels of calculation, CASPT2 and MS-CASPT2, have been displayed. In cases A and B, the CASPT2 curves (dotted lines) incorrectly cross or touch. The reason is that in a single-state perturbation theory each root is not orthogonal to the others.

Neglecting the state interaction leads to unphysical crossing situations. On the contrary, the multistate treatment of MS-CASPT2 (solid lines) makes the two states

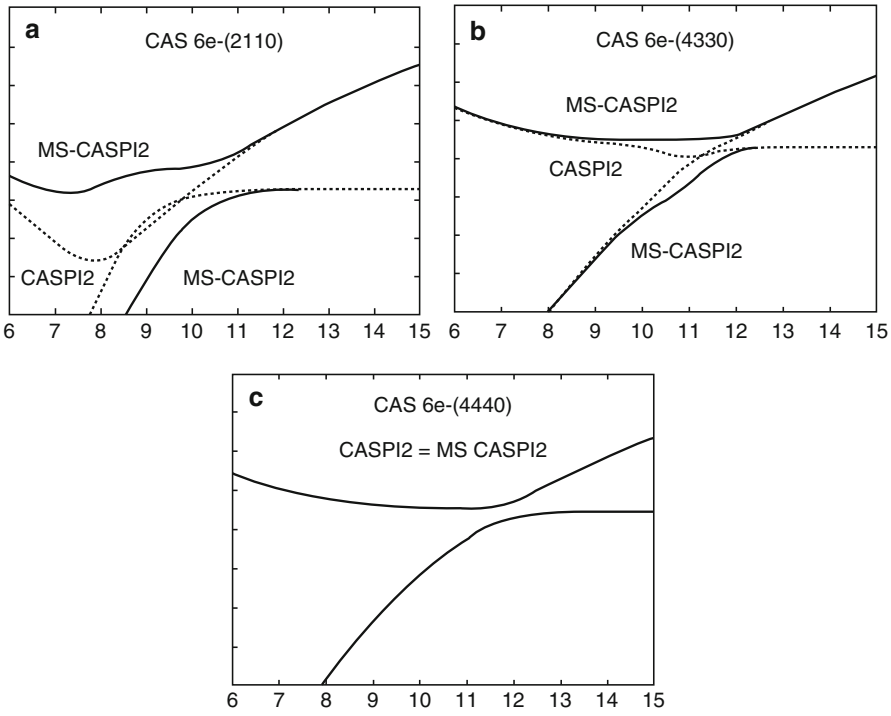

Fig. 21 CASPT2 (*dotted lines*) and MS-CASPT2 (*solid lines*) calculations on the lowest-lying $^1\Sigma^+$ states of the LiF molecule

to interact. The corresponding orthogonal solutions display therefore a "correct" avoided crossing. In case A the employed AS is the minimal one (one 2 s orbital from Li and three 2p orbitals from F). Although the MS-CASPT2 solution yields a formally correct avoided crossing, the state interaction is overestimated, giving rise to a large energy splitting at too short distances (as compared to higher-level MS-CASPT2 or MRCI calculations) because of the too large value of the off-diagonal elements in the multistate effective Hamiltonian. A better treatment is obtained for enlarged active spaces in B and, especially, C, in which the increased space allows for angular correlation in the AS and the CASPT2 and MS-CASPT2 solutions match, undoubtedly because the initial CASPT2 states were including most of the interaction effects. Unfortunately in most cases the enlargement of the active space will not be possible. A good strategy is testing different spaces to check the stability of the results. The right solution seems to be between the CASPT2 and the MS-CASPT2 ones. In the worst case, the CASPT2result is generally not so bad after all.

Table 5 compiles the results of CASPT2 and MS-CASPT2 calculations on the two lowest singlet states of the C_{2h} ethylene dimer using a cc-pVDZ basis set computed at the MECP point upon the increase of the AS (see Serrano-Andrés and Merchán 2005). In order to provide a proper multistate treatment, the off-diagonal

Table 5 Computed MS-CASPT2 energies and off-diagonal Hamiltonian matrix elements for the MECP point of the $1^1A_g/2^1A_g$ states in the C_{2h} ethylene dimer[a]

	ΔE_{CASPT2}	$-H_{12}$	$-H_{21}$	$-(H_{21}+H_{21})$	$\Delta E_{\text{MS-CASPT2}}$
Active spaces	(kcal mol^{-1})	(kcal mol^{-1})	(kcal mol^{-1})	(kcal mol^{-1})	(kcal mol^{-1})
2,020, 4e$^-$	0.41	1.96	2.44	2.20	4.42
3,030, 4e$^-$	2.08	2.09	2.89	2.49	5.40
4,040, 4e$^-$	3.35	2.60	5.12	3.86	8.41
4,040, 8e$^-$	2.76	2.51	5.48	4.00	8.46
4,242, 12e$^-$	2.13	1.27	0.03	0.65	2.49

[a]The MS-CASPT2 calculations comprise two roots at the CASPT2 (2,020,4e$^-$) optimized $1^1A_g/2^1A_g$ MECP geometry

elements of the effective Hamiltonian have to be small and similar, otherwise the matrix symmetrization will be erroneous, whereas the energy split undergone by the original states once diagonalized the matrix will be excessive. Indeed, although the CASPT2 energy differences between the states remained always small (<3.35 kcal mol^{-1}), the splitting largely increases at the MS-CASPT2 level, especially when the off-diagonal elements H_{12} and H_{21} most differ. It is only with the largest active space (angular $\sigma - \pi$ correlation has been included in the active space in this case, 4,242) that the elements become smaller and the MS-CASPT2 reduces. Values close to 2 kcal mol^{-1} can be considered small enough to identify such feature as a CI point within the accuracy of the method. What happens when a CI search is performed blindly using MS-CASPT2 solutions? Typically, it occurs that the obtained solutions are too separated and, apparently, the closest-energy solutions correspond to an avoided crossing. This is probably a wrong answer derived from that fact that the active space is too small and the MS-CASPT2 method is overshooting the interaction of the CASPT2 solutions. If it is not possible to increase the AS (may be MS-RASPT2 is a solution), it is better trusting in the CASPT2 energies.

Setting the Path for Dynamics

Mapping the PEHs of the molecular systems is not but the initial stage of the calculations. Once the topography of the hypersurfaces is obtained by solving the time-independent Schrödinger equation, the evolution of the nuclei on that potential must be described. Studying the reaction dynamics (RD) of the system implies considering time dependence and the statistical distribution on a sample of molecules. Several procedures can be followed, from the most expensive approach, solving the time-dependent Schrödinger equation, typically along specific reaction coordinates, to the semiclassical approaches, in which quantum-chemical information is used to solve the classical Newton equations. When properly averaged, the obtained data yield, in principle, predictive data for reaction rates, state lifetimes, and population distribution. We will not discuss dynamical calculations here, but will simply try to emphasize that the quality of a RD calculation, no matter how sophisticated, strongly relies on the quality of the underlying quantum-chemical description of the PEHs.

In other words, the dynamics is always as good as the quantum chemistry employed to describe the PEH. Expensive RD calculations tend to use analytical potentials fitted with quantum-chemical parameters, a technique filled with problems. The other approach is known as direct or "on the fly" reaction dynamics, meaning that the quantum-chemical information is obtained at each of the points of a trajectory and passed to the RD step. The basic difficulty stays: how good is the description of the PEHs and how this affects to the RD results?

In the study of thymine described above, it was found that when using lower levels of calculation, for instance, a 6-31G* basis set or a small active space, the CASSCF MEPs run in wrong regions of the PEHs, something that only occurred in thymine and uracil, not in purine bases. The MEP FC^1 ($\pi\pi^*$) leads in this case to the high-energy planar minimum and not to the low-lying CI with the ground state, an outcome obtained only when increasing the quality of the calculation by using ANO-type basis sets, larger active spaces or CASPT2 gradients. It is obvious that the RD description will be strongly affected by the quality of the underlying quantum-chemical description. As an illustration, Fig. 22 includes two sample trajectories computed in the guanine molecule on the lowest-energy $^1(\pi\pi^*)$ state starting near the FC region. The calculations used a Verlet-type algorithm for the classical part (which gives the solution of the equations of motion, i.e., how position, velocity, and acceleration change over time) and CASSCF(12,10) or CASPT2(12,10)/6-31G(d,p) quantum-chemical energies and gradients computed "on the fly" at intervals of 5 fs. In this preliminary study, most of the computed trajectories starting close to the FC region were shown to follow a similar path than the MEP, that is, leading to the region of CI with the ground state, and this is done in times close to 100 fs, explaining the ultrafast lifetime measured for the molecule in molecular beams (see Serrano-Andrés et al. 2008). From that point, the complications continue with the selection of the surface hopping algorithm for the population transfer, the diabatic description of the crossing regions, the re-evaluation of the reaction step, etc. (see, just as one of the many examples the use of the NEWTON-X reaction dynamics approach (Barbatti et al. 2007)).

Photoinduced Reactions in Bimolecular Systems

The theoretical description of photoinduced bimolecular reactions requires specific strategies and procedures that will be illustrated here with two examples. First, the cytosine dimer, which forms an excimer in the ground state and it is also able to react yielding a photoadduct in the ground state. Second, the energy transfer reaction between psoralen and molecular oxygen, in which the lowest-energy triplet states are the protagonists.

An Application Example of Reactivity: Cytosine Dimer

One intriguing aspect of UV-irradiated DNA is the appearance of red-shifted long-lived emissive states not found in base monomers, whereas the DNA absorption spectra closely resemble that of the monomers. This phenomenon is called excimer

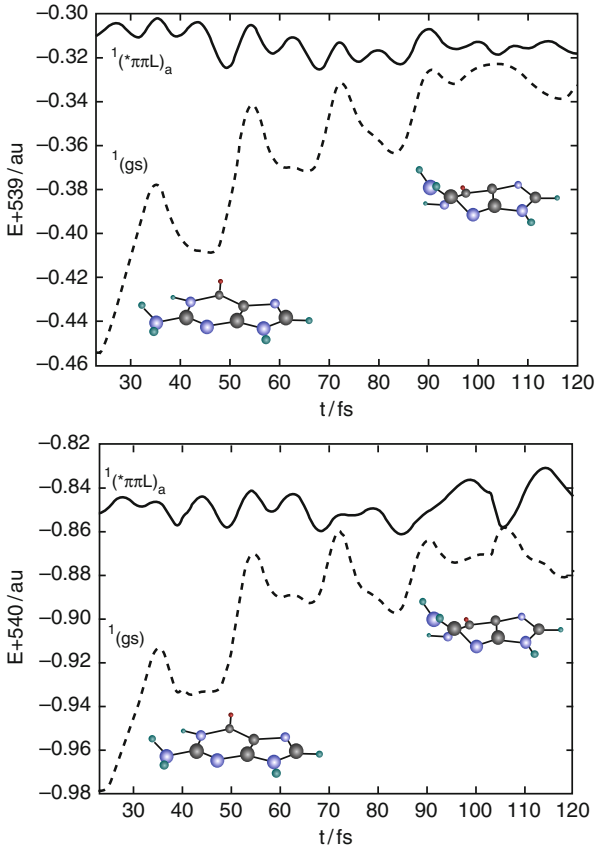

Fig. 22 Sample CASSCF (*top*) and CASPT2 (*bottom*) semiclassical trajectories run on the lowest-energy $^1(\pi\pi^*)$ state of guanine from the FC region

fluorescence, reflecting the relevant role assumed to be played by the corresponding excited dimer (excimer) of the biopolymer, given the similarity between the emission from dinucleotides and polynucleotides (Eisinger and Shulman 1968). Recent experimental results in the light of time- and wavelength-resolved fluorescent techniques using 80 picoseconds (ps) excitation pulses make readily apparent the longer-decay components and red-shifted emission that it was assumed to arise from excimer formation. Because of the slow rate of energy relaxation, these long-lived states associated to excimer-like states have been suggested as the precursors of the DNA photolesions, including photodimers (Schreier et al. 2007).

Classically, an excimer (homodimer) or exciplex (heterodimer) is defined as a dimer system which is bound in the excited state and dissociative in the ground state (see Fig. 23). The existence of excimers between DNA nucleobases was proven theoretically by means of CASPT2(12,12)/ANO-S C,N,O [3s2p1d]/H [2s1p] calculations (see Olaso-González et al. 2006). Potential energy curves (PECs)

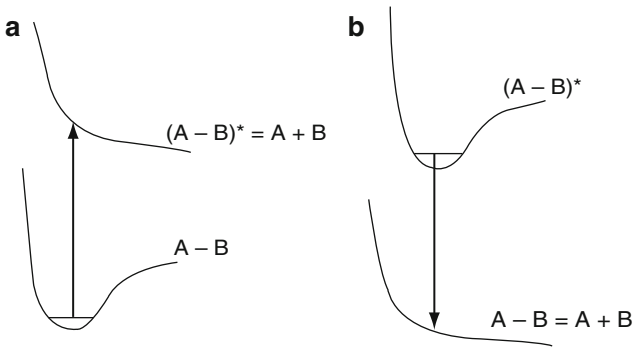

Fig. 23 (**a**) The nonexcited molecule (or a set of molecules; notice that A and B can be both atoms or molecules) is in a bound state; once excited, the molecules evolve to a dissociative state. (**b**) An excimer is a combination of atoms (or molecules) such that it is bound only in the excited state, and then it dissociates as soon as the system relaxes

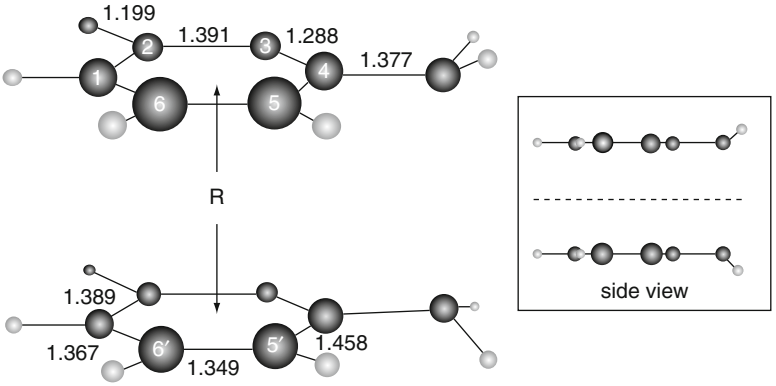

Fig. 24 Labeling for the cytosine dimer. Bond distances (in Å) correspond to the ground-state equilibrium geometry of the monomer. The homodimer system displays C_S symmetry, being the mirror symmetry plane represented by a *dashed line* in the side-view *inset*

with respect to the intermolecular separation (R) of two cytosine molecule kept at the ground-state geometries were built (see Fig. 24). The structure allows for an effective and natural interaction of two cytosine molecules in the biologically relevant *cis-syn* stereoisomer. The ground-state PEC is repulsive at the CASSCF level, whereas the lowest singlet excited state is weakly bound. In contrast, at the CASPT2 level (see Fig. 25), when dynamic correlation is taken into account, the ground-state S_0 and the three lowest singlet excited states (S_1, S_2, and S_3) have well defined minima with binding energies of a few tenths of an eV.

The Basis Set Superposition Error (BSSE) is corrected by means of the counterpoise procedure, an important aspect that will be discussed later. In the figure energies are referred to two ground-state cytosine molecules separated about 20 a.u. In the asymptotic limit, S_1 and S_2 become degenerate, which is consistent with the

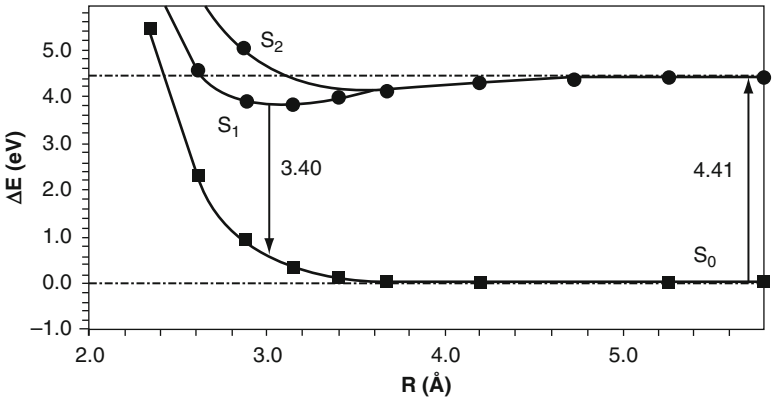

Fig. 25 BSSE-corrected CASPT2 (12,12) potential energy curves built with respect to the distance cytosine-cytosine arranged in a face-to-face disposition

fact that they are related to the equivalent situations $C + C^*$ and $C^* + C$, where C and C* stand for the ground-state cytosine and its lowest singlet excited state, respectively. The absorption $S_0 \rightarrow S_1$ calculated at 20 a.u. (4.41 eV) corresponds to the monomer absorption. On the other hand, the vertical emission from the minimum of S_1 is calculated at 3.40 eV, and it can be considered the source of emission observed in DNA or oligomer samples. In particular, the fluorescence maximum was reported in aqueous solution for the dimer $d(C)_2$ and the 15-mer $d(C)_{15}$ at 3.22 eV, considerably red-shifted as compared to that of the monomer (3.96 eV).

Since that theoretical determination, the existence of excimers in DNA and their role as precursors for phenomena like the red-shifted DNA emission, charge transports in the DNA strand, the formation of nucleobase photoadducts, and, in general, the control of the decay dynamics of excited DNA has been determined both theoretically and experimentally. Crespo-Hernández et al. (2004), for instance, have shown by using femtosecond transient absorption spectroscopy that excimers are formed in high yields in a variety of synthetic DNA oligonucleotides and conclude that excited-state dynamics of $A \cdot T$ DNA is controlled by base stacking. CASPT2 calculations explained also the dynamics of the adenine dimer by the presence of the excimers.

As an illustration of bimolecular reactions, the formation of pyrimidine nucleobase adducts in the excited state will be presented here (see Roca-Sanjuán et al. 2008b; Serrano-Pérez et al. 2008d). Among the possible photoreactions that pyrimidine (Pyr) bases of nucleic acids may undergo upon ultraviolet (UV) irradiation, cyclobutane thymine dimers (T < > T or CBT) formed by intrastrand adjacent thymine bases (see Fig. 26) constitute one of the major photoinduced lesions, particularly in cellular DNA, in spite of the direct reparation by a light-activated mechanism (DNA photolyase) in which abnormal bonds are cleaved. Unrepaired or misrepaired thymine or cytosine dimers and the resultant mutations

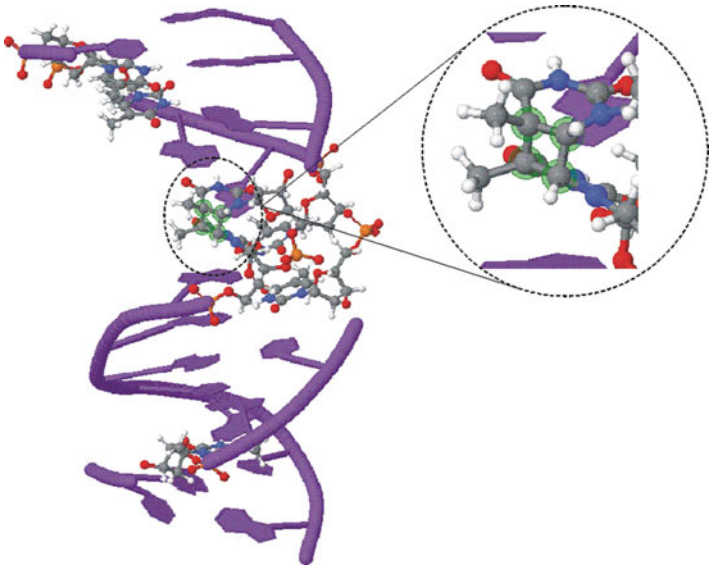

Fig. 26 Photocycloaddition between two adjacent molecules of thymine is a particular class of pericyclic [2 + 2] reaction allowed photochemically and not thermally. It is one of the most usual lesion on DNA. In this picture, taken from a structure obtained in the Protein Data Bank and visualized by Jmol program, the cyclobutane ring is highlighted

may well lead to the development of skin cancer. The major of Pyr < > Pyr photoproducts. The relatively smaller degree of flexibility of A-DNA compared to B-DNA to achieve the right orientations that become prone to react has been related to the greater resistance of A-DNA to Pyr < > Pyr formation.

Dimerization seems to occur only for thymine residues that are already in a reactive arrangement at the instant of excitation because the rate of formation by favorably oriented thymine pairs is much faster that the rate of orientation change. The dimerization reaction occurs both in the singlet and triplet manifolds (see Bosca et al. 2006; Cadet et al. 1992). Focusing on the singlet states, a CI must drive the ultrafast formation of the cycloadduct. Figure 27 displays an scheme of the energy levels of the cytosine dimer in its lowest-lying singlet states based on CASPT2(12,12)/ANO-S C,N,O[3s2p1d]/H[2s1p] results (BSSE corrected). The lines connecting the different solutions represent the different evolution paths followed by the system.

Initially, upon absorption of UV-light radiation in the 4.4 eV region, the system can undergo two types of processes. Nearly unstacked pairs will localize its excitation in one of the monomers and evolve in a ultrafast way toward the ground state via the CI of the monomer, located at 3.6 eV. Other pairs will form an excimer through the stacking interaction of the π clouds of the nucleobases, displaying much larger lifetimes, as it has been already determined experimentally. The binding interaction in the excimer will depend on the conformational arrangement of the

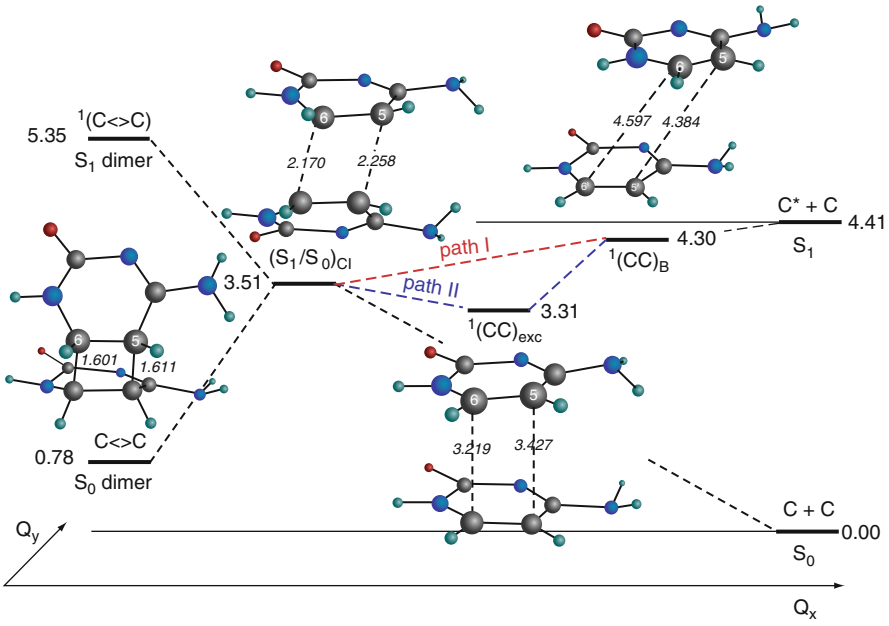

Fig. 27 Scheme based on CASPT2 calculations of the dimerization photochemistry of the cytosine dimer after UV-light absorption

nucleobase pair. Whereas the most common DNA confirmation, B-DNA, displays a very small interaction, strongly overlap situations like the face-to-face parallel arrangement yields the largest binding energy. The cytosine dimer has an excimer structure even lower than the CI leading the system to the formation of the ground state photoadduct CBT. This peculiarity and the presence of both S_0/S_1 CIs, that of the monomer and the one related to CBT, almost isoenergetic and therefore competitive, makes the yield of formation of CBT adducts much smaller than in the case of the thymine dimer, where the CI of the adduct is the lowest and easily accessible feature.

An Application Example of Energy Transfer: Psoralen + O_2

Photomedicine applies the principles of photobiology, photochemistry, and photophysics to the diagnosis and therapy of diseases. One of the most active research areas in this field is photodynamic therapy (PDT), in which the affected living tissue is treated with a combination of a photosensitizer, activated by UV light, and molecular oxygen in its ground triplet state, $^3\Sigma_g^-$. Oxygen is present in the cellular environment ready to transform into singlet oxygen $^1O_2$ ($^1\Delta_g$), which is a strong electrophilic species that reacts with different compounds including some components of the cellular membrane causing cell death by apoptosis. An energy transfer (ET) process triggered by electronic coupling between a molecule in an excited state, the donor (D*), and a molecule, the acceptor (A) or quencher within a

collision complex, is the mechanism through what the reaction takes place, a process that strongly depends on the inter-fragment distance. In general, ET processes, at large separation between the moieties (20–30 Å or even larger), the electronic coupling arises from the Coulomb interaction between electronic transitions that, under the dipole approximation, reduces to the known Förster's dipole–dipole coupling. The process is actually a nonradiative transfer of energy occurring whenever the emission spectrum of D overlaps with the absorption spectrum of A (although no intermediate photon takes part on it). It is the electric field around D*, behaving like a field generated by a classical oscillating dipole, the cause of the excitation of A. At larger separations than Förster's, fluorescence resonance ET (with photon emission by D* and subsequent absorption by A) becomes more efficient than excitation ET. At shorter interfragment distances, however, the so-called Dexter exchange coupling predominates, arising from the exchange integrals that account for the indistinguishability of the electrons in many-electron wave functions. This factor decreases steeply with separation. If the interaction is assumed weak and overlap between D* and A wave functions is produced, Fermi's Golden Rule for coupled transitions can be applied. Such processes have been studied theoretically in depth in recent years, in particular for singlet–singlet ET processes implying an exchange of electrons of the same spin but different energies, that is, the spin state of each fragment is conserved. In PDT, the actual mechanism is, on the other hand, an intermolecular triplet-triplet energy transfer (TET), that is, a process of exchanging both spin and energy between a pair or molecules or molecular fragments. These reactions are commonly used to efficiently populate the triplet states of many organic molecules.

TET processes can be therefore understood as two simultaneous ETs with spin exchange between the interacting fragments and it is similar to the Dexter coupling for singlet-singlet ET, in particular because, as it depends on an electron exchange mechanism, it only takes place at short donor-acceptor distances (<10 Å). In TET the Förster's mechanism will not contribute, because at short distances the dipole approximation breaks down and because the transitions are dipole forbidden. The electronic coupling is not the only key factor that determines the efficiency of the ET process, but also the resonance condition, that is, the energy available in the donor must be at least equal or higher than that required to populate the excited state of the acceptor. If this is the case, the process is usually controlled by diffusion and described as exothermic. In the opposite situation, that is, if the energy of the acceptor is lower than that of the donor, the process becomes thermally activated and lies in the endothermic region. That means that there is an energy barrier whose height will depend on the nature of the acceptor, either classical (for rigid systems) or nonclassical (flexible systems which might find conformations for efficient, non-vertical TET), with a corresponding larger or smaller, respectively, decay in the process rate.

Besides the reaction with DNA nucleobases, psoralen can also interact with molecular oxygen to exert its photosensitizing action. Singlet oxygen and other reactive species of oxygen induce photooxidation of lipids and are considered responsible for cell membrane damaging effects, causing also the appearance of

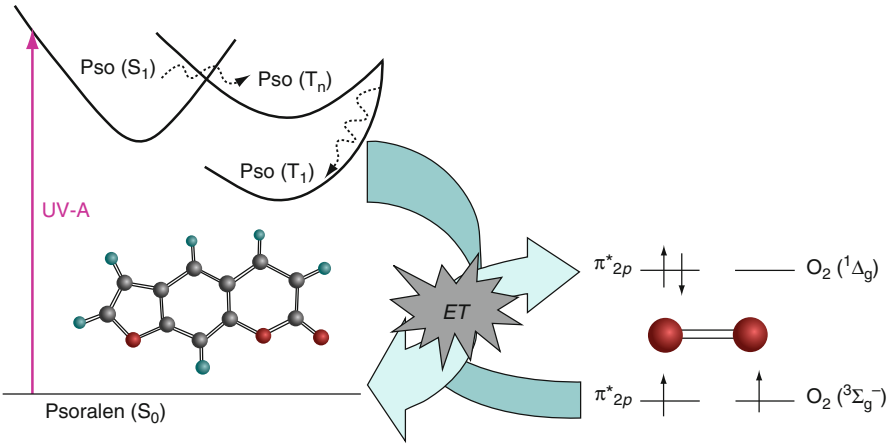

Fig. 28 Scheme of the oxygen-dependent PUVA mechanism

erythema and pigmentation activity in the human skin. In particular, the TET process taking place between psoralen and molecular oxygen is (see also Fig. 28):

$$^3\text{Pso}^*\,(T_1) + {}^3O_2\left({}^3\Sigma_g^-\right) \rightarrow {}^1\text{Pso}\,(S_0) + {}^1O_2^*\left({}^1\Delta_g\right), \tag{7}$$

where activated psoralen behaves as a donor in its lowest triplet state, and triplet ground-state oxygen is the acceptor. The lowest excited singlet state of molecular oxygen ($^1\Delta_g$) is located at 0.97 eV. Furocoumarins (psoralen, 8-MOP, 5-MOP, khellin, TMP, and 3-CPS) have their lowest-lying triplet T_1 state energy at least 1.4 eV higher than the oxygen singlet state, what makes the TET exothermic and diffusion-controlled, with molecular oxygen behaving as a rigid, classical acceptor.

A recent study (see Serrano-Pérez et al. 2009) performed employing the CASPT2/ANO-LC,O[4s3p1d]/H[2s1p] methodology estimated the electronic coupling at some specific disposition of the moieties (psoralen and O_2). Looking for an appropriate arrangement yielding the most effective TET process is nontrivial and, in general, not even relevant, in particular in diffusion-controlled systems which may form a collision complex at short distances. It is important, however, to estimate reaction rates and lifetimes at different intermolecular distances. Furthermore, $M-O_2$ interaction potentials are very weak, and the potential surfaces are generally characterized by multiple shallow minima. Hence, it is necessary to consider different orientations when approaching M and O_2 through a basic interfragment coordinate, here the distance R. Previous studies on systems composed by two ethylene molecules (Et–Et), by the methaniminium cation and ethylene ($MetN^+$–Et), and by ethylene and molecular oxygen (Et–O_2), show that the face-to-face (FF) arrangement is the most appropriate orientation (see Fig. 29).

The geometries of both psoralen and molecular oxygen were kept fixed at the CASSCF optimized triplet excited (T_1) state structure and the triplet ground

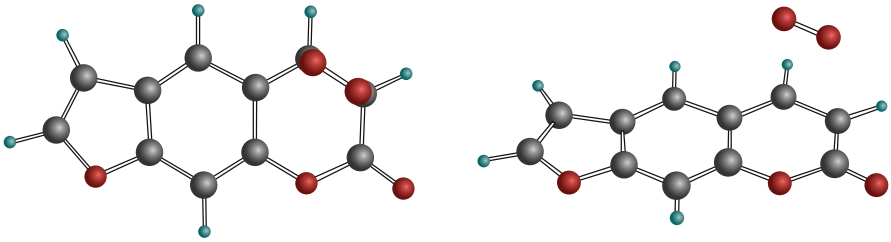

Fig. 29 Psoralen-O_2 supermolecule. The molecular oxygen was placed parallel to the reactive pyrone double bond of psoralen

$\left({}^3\Sigma_g^-\right)$ state experimental geometry, respectively. The active space employed was 14 electrons/11 orbitals (8/7 located in the furocoumarin and 6/4 located in O_2). The active space was validated after comparing the results with previous findings in the isolated psoralen and control calculations on the oxygen molecule with larger active spaces and basis sets. The four lowest singlet states and the three lowest triplet states of the supermolecule were computed. The molecule behaves as a donor in its triplet state, and it is capable to transfer its energy to molecular oxygen in its triplet ground state to generate the singlet ground state furocoumarin and excited singlet oxygen (${}^1\Delta_g$). The energy of the triplet state of the furocoumarin is much higher than the energy of the oxygen ${}^1\Delta_g$ state (computed 1.09 eV, experimental 0.97 eV), and the process falls clearly into the exothermic regime, expected to be controlled by diffusion. Figure 30 displays the potential energy curves of the lowest-lying singlet and triplet states of the supermolecule psoralen-O_2 in a FF arrangement with respect to the C_3=C_4 bond of the pyrone ring (the reactive bond in the triplet state, where the spin density is basically located at different intermolecular distances). The states of the supermolecule protagonist of the TET are 4^1A (T_1 of psoralen+ ${}^3\Sigma_g^-$ of O_2), as initial energy level, and 1^1A (S_0 of psoralen + ${}^1\Delta_g$ of O_2), as the final outcome of the process in which both moieties have changed spin and energy. Within the present approach, the electronic coupling (H') is obtained as half the difference $|\epsilon_\infty - \epsilon_i|$, where ϵ_∞ and ε_i are the energy gaps between the states 4^1A and 1^1A at infinite distance (at 10 Å in the current computation) and at the different interfragment distances, respectively. In this way, the coupling represents the perturbation introduced in each state due to the interaction within the dimer. It is important to notice how advantageous that definition is, because the results are in practice independent of the number of states considered. Notice that the BSSE problem does not affect the results of the coupling, which uses energy differences between states computed at the same geometry. As a result, the BSSE correction vanishes when the energy gap is computed. In other cases, known as non-vertical TET systems, the geometry of the acceptor can be distorted to better accept the photosensitization from the donor. In such situations, the search for regions of state degeneracy with enhanced effectiveness for the ET transfer may be convenient. This is not the case for the present example, in which the acceptor is a rigid oxygen molecule and the process is basically diffusion controlled.

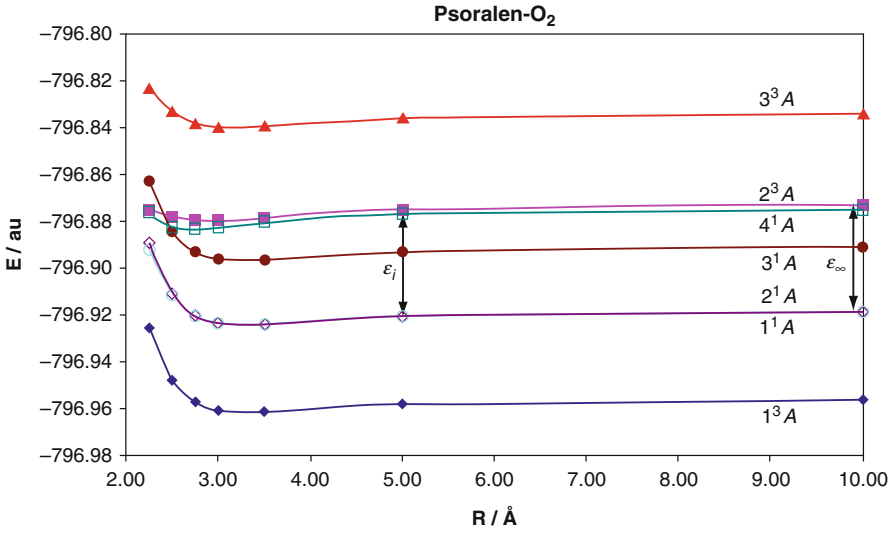

Fig. 30 Potential energy curves of the low-lying excited states of the supermolecule psoralen-molecular oxygen along the interfragment distance (R). The energy coupling H′ is obtained as half of the energy difference $|\epsilon_\infty - \epsilon_i|$ between the initial 4^1A (T_1 of psoralen and $^3\Sigma_g^-$ of O_2) and final 1^1A (S_0 of psoralen and $^1\Delta_g$ of O_2) states of the supermolecule at infinite distance (ϵ_∞, zero coupling situation) and at each of the distances (ε_i)

Table 6 ET analysis of Psoralen-O_2 system along the distance between the two moieties

R/Å	H'/eV	k_{TET}/s^{-1}	τ_{TET}/s
10.00	0.0000	–	–
5.00	0.0012	1.43×10^{9}	6.98×10^{-10}
3.50	0.0038	1.41×10^{10}	7.10×10^{-11}
3.00	0.0357	1.21×10^{12}	8.23×10^{-13}
2.75	0.0870	7.23×10^{12}	1.38×10^{-13}
2.50	0.1966	3.69×10^{13}	2.71×10^{-14}
2.25	0.3833	1.40×10^{14}	7.13×10^{-15}

Table 6 compiles the values of the electronic coupling computed at the different internuclear distances, related with the ET rates and lifetimes. In the weak coupling regime in which the electronic interaction is smaller than the vibrational reorganization energy, the rate for triplet–triplet energy transfer (k_{TET}), and the corresponding lifetime (τ_{TET}), between the donor and the acceptor can be estimated using Fermi's Golden Rule:

$$k_{TET} = \frac{1}{\tau_{TET}} = \frac{4\pi^2}{h} \left| \left\langle \psi_i \left| \widehat{H} \right| \psi_j \right\rangle \right|^2 \rho_E = \frac{4\pi^2}{h} (H')^2 \rho_E, \tag{8}$$

where the matrix element of the Hamiltonian, H', is the electronic part of the energy transfer (i.e., the electronic coupling) and ρ_E is the density of vibrational states in the

initial and final states and their spectral overlap. The inverse of the rate is the lifetime of energy transfer. To obtain the TET rates, we have taken values of $\rho_E = 200\ \mathrm{eV}^{-1}$ and $(4\pi^2/h) = 9.554 \times 10^{15}\ \mathrm{eV}^{-1}\ \mathrm{s}^{-1}$. This order of magnitude for the value of the density of states was used previously as a good estimation in systems of this size.

The whole process of generation of singlet oxygen from psoralen does not only depend on the efficiency of the TET from the triplet state of the photosensitizer, but also on the rate of formation of the triplet state itself. As shown previously, in psoralen the crucial step to populate the triplet manifold in the gas phase is the intersystem crossing (ISC) process between the initially populated singlet $S_\pi\ (\pi\pi^*)$ state and the lowest-lying triplet $T_n\ (n\pi^*)$ state. The latter state evolves subsequently toward the lowest triplet $T_1\ (\pi\pi^*)$ state via a corresponding (and essentially barrierless and ultrafast) internal conversion (IC).

In a similar manner as for Eq. 8, the estimation of the rate constant, here for nonradiative ISC (k_{ISC}), can be obtained as:

$$k_{\mathrm{ISC}} = \frac{4\pi^2}{h} |H_{\mathrm{SO}}|^2 \rho_E, \tag{9}$$

where H_{SO} stands for the spin-orbit coupling terms for the nonradiative transition $S_\pi\ (\pi\pi^*) \rightarrow T_n\ (n\pi^*)$. An estimated value of $200\ \mathrm{eV}^{-1}$ will be employed for ρ_E as used for psoralen in studies explicitly computing vibronic factors.

The ISC nonradiative process corresponds to the transfer $S_\pi\ (\pi\pi^*) \rightarrow T_n\ (n\pi^*)$ in each furocoumarin donor, a process which will take place more efficiently in the region of the $(\mathrm{S}_\pi/\mathrm{T}_n)_{\mathrm{STC}}$ singlet-triplet crossing, and that subsequently will give rise to an ultrafast population of the lowest $T_\pi\ (\pi\pi^*)$ by internal conversion in the triplet manifold. The spin-orbit coupling factors near the crossing point between S_π and T_n, along the LIIC performed in a previous study, is $9.2\ \mathrm{cm}^{-1}$ (see Fig. 17). As a result, the k_{ISC} rate computed was $3.58 \times 10^9\ \mathrm{s}^{-1}$. Additionally, to the H_{SO} strength, the presence of energy barriers in the potential energy hypersurfaces may strongly affect the value of the rate constants, which can be corrected using the Arrhenius exponential term in the framework of the transition state theory. As a qualitative estimation of these effects, a corrected ISC rate (k'_{ISC}) can be obtained from:

$$k'_{\mathrm{ISC}} = k_{\mathrm{ISC}} e^{-\Delta E/RT}, \tag{10}$$

where k_{ISC} is that computed from Eq. 9, ΔE is the energy of the barrier from the initial electronic singlet to the triplet state, R is the ideal gas constant, and T the temperature (298 K). In this particular context, ΔE will be estimated as the energy difference between the singlet $S_\pi\ (\pi\pi^*)$ state, populated at the Franck–Condon geometry, and the triplet $\mathrm{T}_n\ (n\pi^*)$ state in the computed crossing point with the singlet state $(\mathrm{S}_\pi/\mathrm{T}_n)_X$, considering that the energy obtained initially to populate S_π will be available, at least, to surmount the barrier and populate T_n.

In a previous section, we considered as a barrier the energy difference between the minimum of the S_π state and the crossing point $(\mathrm{S}_\pi/\mathrm{T}_n)_X$, which is 0.36 eV, but the present approach accounts for the excess vibrational energy available in the

system from the initially populated S_π state (at the Franck–Condon geometry) to the crossing point (0.03 eV), which seems to be more relevant in this context, where the interest focuses on rate constants. The results were 1.11×10^9 s^{-1} and 0.90 ns for k'_{ISC} and τ'_{ISC}, respectively.

Within this methodology, we can compute the photosensitizing effectiveness, regarding the generation of singlet oxygen, of other furocoumarins and other photosensitizers in general, analyzing the effectiveness of populating T_π as well. Since molecules (donors) that share the same basic structure are known to possess almost the same electronic coupling (against a common acceptor), the first part of this study may well be more interesting if different families of compounds are compared. On the contrary, we should rely on the different way to populate the lowest-lying triplet excited state. In studies performed on the family of furocoumarins (psoralen, 8-MOP, 5-MOP, khellin, TMP, and 3-CPS), the latter, 3-carbethoxypsoralen, was found with the largest ISC rate.

Practical Aspects

The Basis Set Superposition Error (BSSE) for Excited States

The use of finite basis sets derives in a specific defect of the quantum-chemical calculation known as the Basis Set Superposition Error (BSSE). The majority of the contribution to the energy of a system comes from the internal electrons. If the basis set of an atom is deficient in the core region, a molecular method recovers a large amount of energy correcting this deficient area with the basis set of the other atoms. The BSSE is therefore related with the improper inclusion of the correlation energy in a quantum-chemical calculation. Although present in all practical cases, as BSSE strongly depends on the separation between the different centers, its effects only become crucial in dimers and aggregates, that is, when energies at large internuclear distances are compared to those at short distances. In general, the result of ignoring BSSE is both a shortening of bond lengths and an increasing of bond energies because the net effect is an increase of the energy in absolute value. This error is a purely mathematical artifact owing to the fact that the supermolecule possesses a larger basis set than the isolated monomers and as a result the potential energy surface is altered. There are some methods to correct this error like the counterpoise (CP) or the Chemical Hamiltonian Approach (CHA) procedures.

In the previous study on the formation of the excimer of the cytosinedimer, the correction of the BSSE effect was necessary. Figure 31 displays the same PECs as in Fig. 25, but excluding the BSSE correction. All states seem now to be bound, including the ground state, and compared to the BSSE-corrected results, the vertical emission takes place at 3.19 eV instead of 3.40 eV. Table 6 collects the numerical results on the system.

The CASPT2 ground-state binding energy in the uncorrected result is substantial, 0.62 eV, but the system becomes unbound by −0.15 eV when the BSSE is included, that is, the ground state dimer at 3.416 Å is 0.15 eV above the sum of two ground-state monomers. The CP-BSSE corrections seem to be large. With inclusion of the BSSE, both S_1 and S_2 are bound ($CP - E_b$ is positive). Because of the cancelation

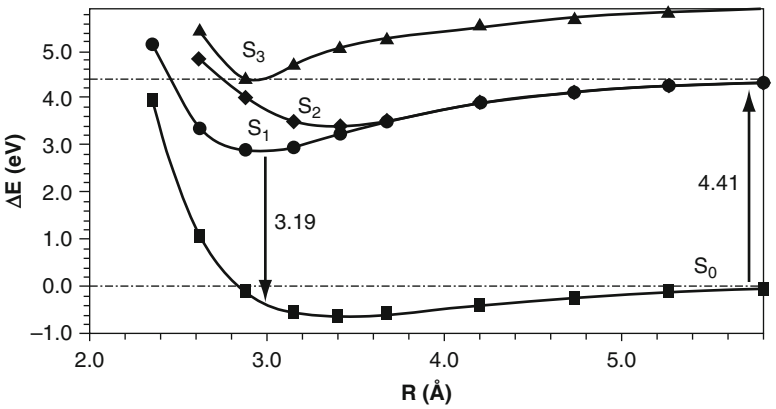

Fig. 31 CASPT2 (12,12) potential energy curves built with respect the distance cytosine–cytosine arranged in a face-to-face way without the inclusion of the CP correction for BSSE

Table 7 Binding energies (E_b), basis set superposition error (BSSE) obtained through the counterpoise method (CP-BSSE), and the corrected binding energy (CP-E_b), computed at the CASPT2 level

State	Geometry	R	E_b	CP-BSSE	CP-E_b
S_0	S_0	3.416	0.62	0.77	−0.15
S_1[a]	S_1	2.954	1.51	0.97	0.54
S_0	S_1	2.954	0.29	0.97	−0.68
S_2	S_2	3.376	0.99	0.74	0.25

[a]TheCASPT2 vertical emission (fluorescence), including the CP-BSSE correction leadsto 3.19 eV, as the result of (4.41 eV – 0.54 eV – 0.68 eV)

of BSSE corrections, the vertical emission remains like the direct CASPT2 result once that the BSSE has been taken into account.

The corrected minimum for S_1 is obtained at $R = 3.\ 076$ Å, with a vertical emission of 3.40 eV and a binding energy CP-E_b of 0.58 eV. Thus, our best estimate has a difference of 0.18 eV with respect to the emission maximum datum obtained experimentally. It is worth recalling that the computed vertical transition does not have experimental counterpart, and for a truly correct comparison with experiment, vibrational resolution of the band should be computed in order to determine the band maximum (Table 7).

The CP procedure has been followed here to take into account the BSSE, astrategy that is known to slightly overestimate the effect. The method corrects the energy at each geometry by subtracting from the full energy of the supermolecule twice the difference between the energy of the fragment using its own basis set and the energy of the fragment using the full basis, but placing the basis functions of the opposite fragment in ghost centers, that is, at the same position as the atoms but with no atoms on them. This subtracted energy is what is known as the BSSE in the CP correction procedure. Obviously the error will decrease when the distance between the fragments decreases, therefore at large internuclear distances the BSSE vanishes. The uncorrected binding energies will then be larger than they should

because they are computed as differences between the energy of the minimum of the supermolecule (with large and stabilizing BSSE effects) and the summed energy of the independent fragments (with no BSSE). The CP method was developed for ground states, and it can be directly extended to excited states only in those cases in which the excitation is clearly localized in one of the fragments. For delocalized situations, it is not possible to define a consistent procedure. As observed in Table 6, however, BSSE values are the same for ground and excited states at a common geometry, for instance here, 0.97 eV both for the S_0 and S_1 states at the S_1 geometry. This is a systematic behavior, in particular for methods like CASPT2, in which a previous common set of State-Average MOs has been used for both states. Therefore, it can be concluded that it is safe to assign the same BSSE correction for ground and excited states at the same geometry, considering also that the inter-fragment distance is the most important parameter to modulate the extent of the BSSE. Let's finally add that ANO basis sets, large and flexible, produce larger BSSE effects than lower quality basis sets.

Computation of Electronic Couplings

The calculation of the electronic coupling matrix element H' is the crucial part in the determination of ET rates and lifetimes. The extent of the coupling controls the energy transfer process, specifically the passage from one state to another, and it can be taken as a measure of the efficiency of the ET process. Different procedures to estimate the ET coupling have been developed based on diabatic localized dimer calculations, monomer transition densities or transition dipole moments, and a supermolecule ansatz of the dimer, whereas generalization of such approaches to determine TET couplings are also available. From all procedures, an energy gap based method such as the supermolecule dimer approach, in which the value of the coupling is obtained as half of the splitting or perturbation between the interacting states, has been shown to be convenient and accurate, and it will be valid in the weak regime where the interaction is not strong. It is clear that its accuracy strongly relies on the quality of the quantum-chemical method used to perform the electronic structure calculations, something guaranteed by the highly reliable and accurate CASPT2 method. It has also the additional advantage that their results are not affected by the BSSE because they are obtained as energy differences at the different internuclear distances.

Conclusions

Quantum Mechanics helps in explaining natural phenomena from theoretical physics (particles, strings) and theoretical chemistry (chemical reactions, intermolecular forces) to the most complex theoretical biology. The complexity level increases as the simplicity of models decrease due to the growing number of variables to be dealt with and the difficulty in simulating the environment. The challenge lies in the ascertainment that life takes place into hierarchically structured matter (macromolecules, cells, tissues, organs, and entities), and it requires the

action of several physical properties of its constituent elements on the whole: the interactions among them and with the environment.

Interaction of light with tissues or, in general, living or nonliving entities is a complex phenomenon which lies on the borders among physics, chemistry, and biology. Specifically, the quantum chemistry of the excited state has experienced an outstanding development in the last decades. Improved algorithms and computational strategies, and obviously faster computers, have contributed to increasing the accuracy of the theoretical description of the photochemical phenomena. Although quantitative determinations in realistic biological systems are still far into the future, medium-size systems and their processes are already at reach of quantum-chemical models. Even more, state-of-the-art experimental techniques, such those related to femtosecond lasers, cannot be properly interpreted without the support of accurate calculations, in a way that a constructive interplay between theory and experiment shows particularly rewarding.

We are witnesses of how the walls which blocked our understanding of photochemical and photophysical processes are being slowly demolished. Step by step larger systems and more complex processes can be studied with accuracy. In fact, it has been theory which has provided modern photochemistry with the proper framework formed by the potential energy hypersurfaces, their interaction through conical intersections and states coupling, and the molecular evolution along such potentials. Unfortunately there are not many methods able to compute properly excited states, and the available ones are complex and sometimes of limited applicability. Multiconfigurational approaches (MRCI, MRPT, and MRCC) give the most general and unbiased description of all types of excitations and situations, either valence, Rydberg or anionic states, bound and dissociated situations, involving open and closed-shell ground states. Among the available approaches, the CASPT2 method has proved to better balance applicability and accuracy at a relatively low computational cost, although new developments and refinements are soon to come, like RASPT2. Single configurational methods (CC, TDDFT, CIS . . .) must be used with caution because they do not give a balanced treatment of all situations in excited-state chemistry, for instance, for degenerate cases like conical intersections. Apart from having proper methods, they must be provided with the required tools, like geometry optimizers, algorithms to search for conical intersections, to compute state couplings or trajectories for the nuclei. One-electron basis sets must be also selected with caution. The simultaneous calculation of valence and Rydberg states is necessary in those regions where both types of states are present, and properly solving valence-Rydberg mixing requires specific strategies. The calculation of anionic excited states requires their own cautions, such as checking the convergence of the resonance solutions by means of different stabilization methods. On the other hand, when dealing with finite one-electron functions and comparing fragment energies with bonded situations, the problem of BSSE arises. In such a situation, the BSSE corrections may well be very important and the phenomenon should be taken into account.

In summary, the excited state quantum chemistry cannot be undertaken routinely because of its complexity. Some of the tools available for the ground state are not

even developed yet for the excited state. In order to get accurate and predictive results, one has to calibrate carefully the computational procedure and determine the requirements in each situation. In this manner, an accuracy of 0.2–0.1 eV is currently expected for the excitation energies, almost an order of magnitude larger than previously. Undoubtedly, the development and refinement of quantum-chemical methods (the software) and computers (the hardware) are making easier the study of the interaction radiation-matter, a still unfinished task with fascinating challenges in the foreseeable future.

Acknowledgments The authors thank to all the members of the Quantum Chemistry for the Excited State (QCEXVal) group from the University of Valencia for their valuable contributions and to all their coworkers along the years. Financial support is acknowledged from projects CTQ2007-61260, CTQ2010-14892, and CSD2007-0010 Consolider-Ingenio in Molecular Nanoscience of the Spanish MICINN/FEDER.

Bibliography

Almlöf, J., & Taylor, P. R. (1986). General contraction of Gaussian basis sets. I. Atomic natural orbitals for first- and second-row atoms. *The Journal of Chemical Physics, 86*, 4070–4077.

Almlöf, J., & Taylor, P. R. (1991). Atomic Natural Orbital (ANO) basis sets for quantum chemical calculations. *Advances in Quantum Chemistry, 22*, 301–373.

Andersson, K., Malmqvist, P.-Å., & Roos, B. O. (1992). Second-order perturbation theory with a complete active space self-consistent field reference function. *The Journal of Chemical Physics, 96*, 1218–1226.

Atchity, G. J., Xantheas, S. S., & Ruedenberg, K. (1991). Potential-energy surfaces near intersections. *The Journal of Chemical Physics, 95*, 1862–1876.

Atkins, P. W., & Friedman, R. S. (1997). *Molecular quantum mechanics* (3rd ed.). Oxford: Oxford University Press.

Barbatti, M., Granucci, G., Persico, M., Ruckenbauer, M., Vazdar, M., Eckert-Maksić, M., & Lischka, H. (2007). The on-the-fly surface-hopping program system Newton-X: Application to ab initio simulation of the nonadiabatic photodynamics of benchmark systems. *Journal of Photochemistry andPhotobiology A: Chemistry, 190*, 228–240.

Bearpark, M. J., & Robb, M. A. (2007). *Conical intersection species as reactive intermediates*. Hoboken: Wiley.

Ben-Nun, M., & Martínez, T. J. (2000). Photodynamics of ethylene: Ab initio studies of conical intersections. *Chemical Physics, 259*, 237–248.

Bosca, F., Lhiaubet-Vallet, V., Cuquerella, M. C., Castell, J. V., & Miranda, M. A. (2006). The triplet energy of thymine in DNA. *Journal of the American Chemical Society, 128*, 6318–6319.

Buenker, R. J., Hirsch, G., Li, Y., Gu, J., Alekseyev, A. B., Liebermann, H., & Kimura, M. (2000). *Ab initio calculations of excited state potential functions* (pp. 135–168). Chichester: Wiley.

Cadet, J., Anselmino, C., Douki, T., & Voituriez, L. (1992). New trends in photobiology: Photochemistry of nucleic acids in cells. *Journal of Photochemistry and Photobiology, B: Biology, 15*, 277–298.

Celani, P., & Werner, H.-J. (2003). Analytical energy gradients for internally contracted second-order multireference perturbation theory. *The Journal of Chemical Physics, 119*, 5044–5057.

Crespo-Hernández, C., Cohen, B., Hare, P., & Kohler, B. (2004). Ultrafast excited-state dynamics in nucleic acids. *Chemical Reviews, 104*, 1977–2019.

Domcke, W., Sobolewski, A. L., & Woywood, C. (1993). Internal conversion funnel in benzene and pyrazine: Adiabatic and diabatic representation. *Chemical Physics Letters, 203*, 220–226.

Domcke, W., Yarkony, D. R., & Köppel, H. (Eds.). (2004). *Conical intersections. Electronic structure, dynamics and spectroscopy*. River Edge: World Scientific.

Eisinger, J., & Shulman, R. G. (1968). Excited electronic states of DNA. *Science, 161*, 1311–1319.

Finley, J., Malmqvist, P.-Å., Roos, B. O., & Serrano-Andrés, L. (1998). The multi-state CASPT2 method. *Chemical Physics Letters, 288*, 299–306.

Finley, J., Malmqvist, P.-Å., Roos, B. O., & Serrano-Andrés, L. (1999). The multi-state CASPT2 method. *Chemical Physics Letters, 288*, 299–306.

Foresman, J. B., Head-Gordon, M., Pople, J. A., & Frisch, M. J. (1992). Toward a systematic molecular orbital theory for excited states. *The Journal of Physical Chemistry, 96*, 135–149.

Forsberg, N., & Malmqvist, P.-Å. (1997). Multiconfiguration perturbation theory with imaginary level shift. *Chemical Physics Letters, 274*, 196–204.

Ghigo, G., Roos, B. O., & Malmqvist, P.-Å. (2004). A modified definition of the zeroth-order Hamiltonian in multiconfigurational perturbation theory (CASPT2). *Chemical Physics Letters, 396*, 142–149.

Goddard, W. A., & Hunt, W. J. (1974). The Rydberg nature and assignments of excited states of the water molecule. *Chemical Physics Letters, 24*, 464–471.

Grein, F. (2009). Coupled cluster and density functional studies on geometries and energies of excited C_{2v} states of ozone. *The Journal of Chemical Physics, 130*, 124118.

Grimme, S., & Waletzke, M. (1999). A combination of Kohn–Sham density functional theory and multi-reference configuration interaction methods. *The Journal of Chemical Physics, 111*, 5645.

Helgaker, T., Jørgensen, J., & Olsen, J. (2004). *Molecular electronic-structure theory*. Chichester: Wiley.

Herzberg, G., & Longuet-Higgins, L. C. (1963). Intersection of potential energy surfaces in polyatomic molecules. *Discussions of the Faraday Society, 35*, 77–82.

Kleinschmidt, M., Tatchen, J., & Marian, C. M. (2001). Spin-orbit coupling of DFT/MRCI wavefunctions: Method, test calculations, and application to thiophene. *Journal of Computational Chemistry, 23*, 824–833.

Koch, H., & Jørgensen, P. (1990). Coupled cluster response functions. *The Journal of Chemical Physics, 93*, 3333–3344.

Kowalski, K. (2005). Completely renormalized EOMCCSD(T) method employing independent optimization of the cluster product terms. *Chemical Physics Letters, 411*, 306–310.

Lischka, H., Dallos, M., Szalay, P. G., Yarkoni, D. R., & Shepard, R. (2004). Analytic evaluation of nonadiabatic coupling terms at the MR-CI level. I. Formalism. *The Journal of Chemical Physics, 120*, 7322–7329.

Malmqvist, P.-Å., & Roos, B. O. (1989). The CASSCF state interaction method. *Chemical Physics Letters, 155*, 189–194.

Malmqvist, P.-Å., Rendell, A., & Roos, B. O. (1990). The restricted active space self-consistent-field method, implemented with a split graph unitary group approach. *The Journal of Physical Chemistry, 94*, 5477–5482.

Malmqvist, P.-Å., Pierloot, K., Shahi, A. R. M., Cramer, C. J., & Gagliardi, L. (2008). The restricted active space followed by second-order perturbation theory method: Theory and application to the study of CuO_2 and Cu_2O_2 systems. *The Journal of Chemical Physics, 128*, 204109.

Martin, J. M. L., Taylor, P. R., & Lee, T. J. (1997). The harmonic frequencies of benzene. A case for atomic natural orbital basis sets. *Chemical Physics Letters, 275*, 414–422.

Merchán, M., & Serrano-Andrés, L. (2005). *Ab initio methods for excited states*. Amsterdam: Elsevier.

Merchán, M., Serrano-Andrés, L., Robb, M., & Blancafort, L. (2005). Triplet-state formation along the ultrafast decay of excited singlet cytosine. *Journal of the American Chemical Society, 127*, 1820–1825.

Merchán, M., González-Luque, R., Climent, T., Serrano-Andrés, L., Rodríguez, E., Reguero, M., & Peláez, D. (2006). Unified model for the ultrafast decay of pyrimidine nucleobases. *The Journal of Physical Chemistry B, 110*, 26471–26476.

Müller, T. H., Dallos, M., & Lischka, M. (1999). The ethylene 1 $^1B_{1u}$ V state revisited. *The Journal of Chemical Physics, 110*, 7176–7184.

Oddershede, J. (1987). Propagator methods. *Accounts of Chemical Physics, 69*, 201–239.

Olaso-González, G., Roca-Sanjuán, D., Serrano-Andrés, L., & Merchán, M. (2006). Toward the understanding of DNA fluorescence: The singlet excimer of cytosine. *The Journal of Chemical Physics, 125*, 231102.

Olsen, J., Roos, B. O., Jørgensen, P., & Jensen, H. J. A. (1988). Determinant based configuration interaction algorithms for complete and restricted configuration interaction spaces. *The Journal of Chemical Physics, 89*, 2185–2192.

Paschoal, D., Costa, M. F., Junqueira, G. M. A., & Santos, H. F. D. (2009). Ab initio calculation of electric properties for the BH, CO, CS and N_2 molecules. *Journal of Molecular Structure: THEOCHEM, 913*, 200–206.

Paterson, M. J., Bearpark, M. J., Robb, M. A., & Blancafort, L. (2004). The curvature of the conical intersection seam: An approximate second-order analysis. *The Journal of Chemical Physics, 121*, 11562–11571.

Peach, M. J. G., Benfield, P., Helgaker, T., & Tozer, D. J. (2008). Excitation energies in density functional theory: An evaluation and a diagnostic test. *The Journal of Chemical Physics, 128*, 044118.

Pou-Amérigo, R., Serrano-Andrés, L., Merchán, M., Ortí, E., & Forsberg, N. (2000). A theoretical determination of the low-lying electronic states of the *p*-benzosemiquinone radical anion. *Journal of the American Chemical Society, 122*, 6067–6077.

Pérez-Hernández, G., González, L., & Serrano-Andrés, L. (2008). Rydberg or valence? The long-standing question in the UV absorption spectrum of 1,1′ -bicyclohexylidene. *ChemPhysChem, 9*, 2544–2549.

Robb, M. A., Bernardi, F., & Olivucci, M. (1996). Conical intersections as a mechanistic feature of organic-photochemistry. *Pure and Applied Chemistry, 67*, 783–789.

Roca-Sanjuán, D., Rubio, M., Merchán, M., & Serrano-Andrés, L. (2006). Ab initio determination of the ionization potentials of DNA and RNA nucleobases. *The Journal of Chemical Physics, 125*, 084302.

Roca-Sanjuán, D., Merchán, M., Serrano-Andrés, L., & Rubio, M. (2008a). Ab initio determination of the electron affinities of DNA and RNA nucleobases. *The Journal of Chemical Physics, 129*, 095104.

Roca-Sanjuán, D., Olaso-González, G., González-Ramírez, I., Serrano-Andrés, L., & Merchán, M. (2008b). Molecular basis of DNA photodimerization: Intrinsic production of cyclobutane cytosine dimers. *Journal of the American Chemical Society, 130*, 10768–10779.

Roos, B. O., Taylor, P. R., & Siegbahn, P. E. M. (1979). A complete active space SCF method (CASSCF) using a density matrix formulated super-CI approach. *Chemical Physics, 48*, 157–173.

Roos, B. O., Andersson, K., Fülscher, M. P., Serrano-Andrés, L., Pierloot, K., Merchán, M., & Molina, V. (1996). Applications of level shift corrected perturbation theory in electronic spectroscopy. *Journal of Molecular Structure: THEOCHEM, 388*, 257–276.

Roos, B. O., Malmqvist, P.-Å., Molina, V., Serrano-Andrés, L., & Merchán, M. (2002). Theoretical characterization of the lowest-energy absorption band of pyrrole. *The Journal of Chemical Physics, 116*, 7526–7537.

Roos, B. O., Lindh, R., Malmqvist, P.-Å., & Widmark, P.-O. (2005). New relativistic ANO basis sets for actinide atoms. *Chemical Physics Letters, 409*, 295–299.

Roothaan, C. C. J. (1951). New developments in molecular orbital theory. *Reviews of Modern Physics, 23*, 69–89.

Rubio, M., Serrano-Andrés, L., & Merchán, M. (2008). Excited states of the water molecule: Analysis of the valence and Rydberg character. *The Journal of Chemical Physics, 128*, 104305.

Schreier, W. J., Schrader, T. E., Soller, F. O., Gilch, P., Crespo-Hernández, C. E., Swaminathan, V. N., Carell, T., Zinth, W., & Kohler, B. (2007). Thymine dimerization in DNA is an ultrafast photoreaction. *Science, 315*, 625–629.

Serrano-Andrés, L., & Merchán, M. (2005). Quantum chemistry of the excited state: 2005 overview. *Journal of Molecular Structure: THEOCHEM, 729*, 99–108.

Serrano-Andrés, L., & Merchán, M. (2009). Are the five natural DNA/RNA base monomers a good choice from natural selection? A photochemical perspective. *Journal of Photochemistry and Photobiology, C: Photochemistry Reviews, 10*, 21–32.

Serrano-Andrés, L., Merchán, M., & Lindh, R. (2005). Computation of conical intersections by using perturbation techniques. *The Journal of Chemical Physics, 122*, 104107.

Serrano-Andrés, L., Merchán, M., & Borin, A. C. (2006). Adenine and 2-aminopurine: Paradigms of modern theoretical photochemistry. *Proceedings of the National Academy of Sciences of the United States of America, 103*, 8691–8696.

Serrano-Andrés, L., Merchán, M., & Borin, A. C. (2008). A three-state model for the photophysics of guanine. *Journal of the American Chemical Society, 130*, 2473–2484.

Serrano-Pérez, J., Serrano-Andrés, L., & Merchán, M. (2006). A theoretical insight into the photophysics of psoralen. *The Journal of Chemical Physics, 124*, 124502.

Serrano-Pérez, J., Serrano-Andrés, L., & Merchán, M. (2007a). Quantum chemical study on the population of the lowest triplet state of psoralen. *Chemical Physics Letters, 434*, 107–110.

Serrano-Pérez, J., González-Luque, R., Serrano-Andrés, L., & Merchán, M. (2007b). On the intrinsic population of the lowest triplet state of thymine. *The Journal of Physical Chemistry, 111*, 11880–11883.

Serrano-Pérez, J., Merchán, M., & Serrano-Andrés, L. (2008a). Photosensitization and phototherapy with furocoumarins: A quantum-chemical study. *Chemical Physics, 347*, 422–435.

Serrano-Pérez, J., González-Luque, R., Merchán, M., & Serrano-Andrés, L. (2008b). The family of furocoumarins: Looking for the best photosensitizer for phototherapy. *Journal of Photochemistry and Photobiology, A: Chemistry, 199*, 34–41.

Serrano-Pérez, J., Merchán, M., & Serrano-Andrés, L. (2008c). Photoreactivity of furocoumarins and DNA in PUVA therapy: Formation of psoralen-thymine adducts. *The Journal of Physical Chemistry B, 112*, 14002–14010.

Serrano-Pérez, J., González-Ramírez, I., Coto, P., Serrano-Andrés, L., & Merchán, M. (2008d). Theoretical insight into the intrinsic ultrafast formation of cyclobutane pyrimidine dimers in UV-irradiated DNA: Thymine versus cytosine. *The Journal of Physical Chemistry B, 112*, 14096–14098.

Serrano-Pérez, J., Olaso-González, G., Merchán, M., & Serrano-Andrés, L. (2009). Singlet oxygen generation in PUVA therapy studied using electronic structure calculations. *Chemical Physics, 360*, 85–96.

Shepard, R., Lischka, H., Szalay, P. G., Kovar, T., & Ernzerhof, M. (1992). A general multireference configuration interaction gradient program. *The Journal of Chemical Physics, 96*, 2085–2098.

Sicilia, F., Bearpark, M. J., Blancafort, L., & Robb, M. A. (2007). An analytical second-order description of the S0/S1 intersection seam: Fulvene revisited. *Theoretical Chemistry Accounts, 118*, 241–251.

Strickler, S. J., & Berg, R. A. (1962). Relationship between absorption intensity and fluorescence lifetime of molecules. *The Journal of Chemical Physics, 37*, 814–822.

Szabo, A., & Ostlund, N. S. (1996). *Modern quantum chemistry: Introduction to advanced electronic structure theory*. Mineola: Dover Publications.

Szymczak, J. J., Barbatti, M., Soo Hoo, J. T., Adkins, J. A., Windus, T.-L., Nachtigallová, D., & Lischka, H. (2009). Photodynamics simulations of thymine: Relaxation into the first excited singlet state. *The Journal of Physical Chemistry, A, 113*, 12686–12693.

Taylor, P. R. (2007). *Lecture notes in quantum chemistry: European summer school*. Berlin: Springer.

Teller, E. (1937). The crossing of potential surfaces. *The Journal of Physical Chemistry, 41*, 109–116.

Widmark, P.-O., Malmqvist, P.-Å., & Roos, B. O. (1990). Density matrix averaged atomic natural orbital (ANO) basis sets for correlated molecular wave functions. *Theoretica Chimica Acta, 77*, 291–306.