

VNIVERSITAT^{IS} VALÈNCIA

INSTITUTO DE CIENCIA MOLECULAR

Programa de doctorado: Química Teórica y Computacional



Determinación teórica de la espectroscopia y fotorreactividad de biocromóforos



Tesis Doctoral presentada por:
María Teresa Climent Santamaría

Dirigida por:
Dra. Manuela Merchán Bonete
Dr. Luis Serrano Andrés

Valencia 2012

TESIS DOCTORAL



VNIVERSITAT
DE VALÈNCIA

INSTITUTO DE CIENCIA MOLECULAR

**Determinación teórica de la espectroscopia y
fotorreactividad de biocromóforos**

Tesis Doctoral presentada por
María Teresa Climent Santamaría
Valencia 2012

Programa de doctorado: Química Teórica y Computacional

Dña. Manuela Merchán y **D. Luis Serrano Andrés**, Catedrática y Profesor Titular, respectivamente, del Departamento de Química Física y del Instituto de Ciencia Molecular de la Universitat de València, certifican que la memoria presentada por la licenciada María Teresa Climent Santamaría con título **“Determinación teórica de la espectroscopia y fotorreactividad de biocromóforos”** corresponde a su Tesis Doctoral y ha sido realizada bajo su dirección, autorizando mediante este escrito la presentación de la misma.

En Valencia, a marzo de 2012

Dra. Manuela Merchán

Dr. Luis Serrano Andrés
(1966-2010)

Doctoranda: María Teresa Climent Santamaría

A mis padres, que me dieron la vida.
A mis hijos, que serán mi continuación en ella.

Agradecimientos

Siendo licenciada en Bioquímica y en Farmacia, muchas son las veces que me he preguntado quién me mandaría hacer una tesis en Química Cuántica, y más a mi edad. Aun a pesar de no entender nada, absolutamente nada, la primera vez que leí una publicación de Manuela Merchán, decidí intentarlo porque ella misma me dijo que “todo es cuestión de lenguaje” y que una vez aprendido todo sería más fácil. Han pasado unos cuantos años de trabajo desde aquel día y supongo que algo del lenguaje cuántico he aprendido, al menos para ser capaz de entender a Manuela. También recuerdo que aquel mismo día le dije que siempre había soñado con escribir una tesis que mis padres fueran capaces de entender. Uff, creo que eso será imposible, pero espero que al menos no piensen que el tema de mi tesis es otro con tantas “excitaciones”, “degeneraciones”, “tripletes” y “desexcitaciones posteriores”. En fin, sea como sea, al final hemos compartido muchos momentos de charla profesional y personal que nos han enriquecido a las dos y sobre todo se han creado unos vínculos de amistad que perdurarán más allá de lo puramente profesional. Gracias Manoli por haberme metido en este mundo y por haberme enseñado a caminar por él. Y sobre todo gracias por tu infinita paciencia conmigo.

Gracias a Luis Serrano[♦] por haber estado siempre ahí, y digo bien “siempre”, para resolver cualquier duda que se me planteara, por haberme orientado y corregido, por tus programas para hacernos la vida más fácil y sobre todo por entender el MOLCAS hasta el más recóndito de sus rincones. Gracias Luis por ser tan gran persona.

[♦] El 15 de septiembre de 2010 recibí la terrible noticia de que Luis había fallecido repentinamente. He dejado la dedicatoria y los agradecimientos tal como estaban escritos en ese momento, aun a pesar de que el cuerpo me pedía dedicársela a él, y de que los agradecimientos están escritos como si tuviera delante a la persona, porque siento que desde donde esté, lo estará recibiendo. Al final del presente manuscrito he incorporado un pequeño “in memoriam”.

Gracias a Remedios González por haberme iniciado en el JobIph y en el RasOrb, en los inputs y en los outputs. Siempre que recuerdo mis inicios, me vienen a la cabeza tus comentarios, las dos de cara a la pantalla de tu ordenador en tu despacho del edificio blanco, y sobre todo tu letra, ¡descifrarla fue todo un reto! Mil gracias por tus sabios consejos, por tus valiosísimas aportaciones en la corrección del presente manuscrito, y por tus palabras de ánimo cuando a veces me venía abajo y quería tirar la toalla. Reme, hemos compartido más que cálculos, ya que la adolescencia de nuestros hijos también ha sido motivo de largas charlas. Manuela, tú, y yo, podríamos, perfectamente, incluir en nuestro curriculum que somos Investigadoras Asociadas en el campo del Desarrollo de la Adolescencia y Relaciones Humanas.

Gracias a Israel González, porque al tenerte que contar cómo hacer los cálculos he tenido que poner las ideas en claro con antelación. Se aprende mucho cuando uno tiene que enseñar y, haciendo uso de tus palabras, he tenido el “favor de Dios” de poder trabajar contigo. Ánimo, que tú vas detrás y ya queda poco para que seas doctor.

Gracias a Gloria Olaso, Daniel Roca y Juan José Serrano por haberme acogido entre vosotros como a una más y contestar a mis dudas cada vez que os necesitaba, por vuestros ánimos y vuestro cariño. Sé que vais a desarrollar unas brillantes carreras como investigadores, y espero que algún día también como docentes, en esta maravillosa casa que es la Universitat de València.

Gracias a mi familia, a mis seres queridos, a mis amistades, a mis compañer@s, a mis maestr@s. Gracias por dejarme compartir con vosotr@s todas las experiencias, buenas y no tan buenas, de la convivencia. En definitiva, gracias a la vida, que me ha dado tanto.

Contenido

Prólogo.....	11
Capítulo 1. Metodología.....	15
Estructura de la materia. Breve resumen histórico.....	15
Metodología cuántica.....	16
Métodos <i>ab initio</i>	19
Conjuntos de base monoelectrónica.....	27
Hipersuperficies de energía potencial.....	32
Interacción de la luz con la materia.....	33
Cruces de igual simetría espacial y de espín.....	39
Cruces de igual simetría espacial y distinta multiplicidad.....	41
Determinación estructural: el problema de la correlación diferencial....	43
Error por superposición del conjunto de base	45
CASPT2 y sus modificaciones level-shift e IPEA-shift al	
Hamiltoniano de orden cero.....	46
Level-shift CASPT2.....	47
IPEA-shift CASPT2.....	49
Referencias.....	51

Parte I

Luz visible y fotocromismo

Capítulo 2. Introducción y objetivos.....	57
Receptores de luz y fotomorfogénesis.....	57
Fitocromos.....	59
Fototropinas.....	61
Criptocromos.....	64
Objetivos.....	65
Referencias.....	66
Capítulo 3. Fitocromos: Rutas de desactivación.....	69
Detalles computacionales.....	71
Selección del modelo molecular PΦB.....	71
Selección y validación de los métodos computacionales.....	71
Optimización de las geometrías del estado fundamental.....	73
Optimización de la geometría del estado excitado.....	75

Estimación de las energías del estado fundamental y estados excitados.....	77
Absorción y emisión.....	79
Rutas de desactivación.....	80
Conclusiones del capítulo 3.....	87
Referencias.....	89
Capítulo 4. Isoaloxacina, el anillo básico de las flavinas.....	93
Detalles computacionales.....	95
Estructuras de equilibrio.....	96
Espectros Verticales.....	99
Energías de emisión y de transición no vertical.....	102
Formación del estado triplete.....	104
Conclusiones del capítulo 4.....	109
Referencias.....	111

Parte II

Luz ultravioleta y mutaciones genéticas

Capítulo 5. Introducción y objetivos.....	117
Los ácidos nucleicos: Descripción y estructura.....	117
Estructura del ADN.....	120
Estructura del ARN.....	123
Implicación de la radiación UV en la formación de biodímeros.....	125
Excímeros, las moléculas responsables de la formación de fotoproductos: definición y caracterización.....	129
Objetivos.....	131
Referencias.....	133
Capítulo 6. Decaimiento ultrarrápido de las bases pirimidínicas....	135
Detalles computacionales.....	137
Modelo unificado sin barrera energética.....	138
Uracilo.....	140
Timina.....	141
Citosina.....	143
Caminos alternativos de relajación energética.....	147
Conclusiones del capítulo 6.....	149
Referencias.....	150
Capítulo 7. Población del estado triplete en las bases pirimidínicas	153
Detalles computacionales.....	155
Mecanismos de población del estado triplete en las bases pirimidínicas.....	156
Uracilo.....	157

Timina.....	164
Citosina.....	169
Conclusiones del capítulo 7.....	173
Referencias.....	175
Capítulo 8. Formación de ciclobutanos del ADN/ARN.....	179
Detalles computacionales.....	183
Excímeros de uracilo.....	184
Formación de ciclobutiluracilo a través del intermedio triplete.....	186
Fotodimerización de las bases nucleicas pirimidínicas.....	189
Conclusiones del capítulo 8.....	193
Referencias.....	194
Glosario de siglas y abreviaturas.....	197
A la memoria de Luis (In memóriam).....	201
Publicaciones.....	203

Prólogo

El sol constituye la mayor fuente de energía, concretamente energía electromagnética, del sistema en el que se encuentra nuestro planeta Tierra. El espectro electromagnético se extiende desde radiaciones de longitud de onda muy corta, como las proporcionadas por los rayos X y los gamma, pasando por la luz ultravioleta (UV), la luz visible y los rayos infrarrojos, hasta ondas electromagnéticas de menor frecuencia, como son las de radio.

La luz visible es tan solo uno de los muchos tipos de radiación, que abarca un pequeño intervalo dentro de la totalidad del espectro electromagnético, aunque, el hecho de que podamos percibirla directamente con nuestros ojos haya hecho que resaltemos su importancia. Pero en realidad, la verdadera importancia de esta pequeña franja del espectro, reside en que ha desempeñado, y desempeña, un papel primordial en la evolución y desarrollo de la vida en la Tierra.

Las plantas, y algunas bacterias y hongos, contienen dentro de sus células sistemas moleculares cuya función es absorber selectivamente luz de una determinada longitud de onda y aprovechar su energía para convertirla en una fuente de reacción fotosintética. Podría decirse, que estos organismos explotan la flexibilidad conformacional de dichos sistemas moleculares, llamados fotorreceptores, para hacer un ajuste fino en el dial de longitudes de onda, absorbiendo la luz que les interesa y actuando como “interruptores fotocromicos”, que permiten transformar a la molécula en una forma fisiológicamente activa, desde una forma fisiológicamente inactiva. La luz, absorbida por el fotorreceptor, inicia los cambios conformacionales del

sistema molecular, que a su vez desencadenarán una cascada de señales reguladoras de funciones fundamentales en dichos organismos; funciones que abarcan desde el potencial fotosintético y pigmentación en bacterias, hasta el crecimiento y desarrollo de las plantas, controlando por ejemplo la germinación de sus semillas, el desarrollo de sus cloroplastos y la fotomorfogénesis, o crecimiento alejándose de la sombra.

Por otro lado, una radiación electromagnética no percibida por el ojo humano, pero de vital importancia para la evolución de las especies tal como las conocemos, es la luz ultravioleta, directamente relacionada con la aparición de un gran número de mutaciones en los organismos vivos, bien por interacción directa de ésta con su material genético, o bien indirectamente, a través de la reacción de dicho material con otras especies químicas generadas por la interacción de la radiación con moléculas del entorno celular.

El objetivo de la presente Tesis Doctoral es profundizar, desde un punto de vista químico-cuántico, en los mecanismos moleculares que expliquen los primeros pasos que se suceden tras la absorción de luz visible y UV en distintos fotorreceptores biológicos. Es decir, se pretende estudiar de forma teórica la fotorreactividad de ciertos biocromóforos. Para ello el presente libro se ha dividido en dos partes: parte I, **Luz visible y fotocromismo**, y parte II, **Luz ultravioleta y mutaciones genéticas**. A su vez, ambas partes se han subdividido en varios capítulos que contienen los objetivos, metodología, y análisis realizados, así como las correspondientes conclusiones y referencias bibliográficas.

En la parte I se aborda el estudio de las rutas de desactivación-isomerización de los fitocromos, investigación realizada con un modelo de cuatro anillos del cromóforo y el estudio de la fotoquímica de la isoaloxacina, anillo formador del cromóforo de las flavinas.

En la parte II se realiza una descripción de la composición y estructura de los ácidos nucleicos y se analiza la importancia de la formación de excímeros, agregados precursores de la formación de fotoproductos, al interactuar con la radiación UV y su implicación en las mutaciones genéticas.

Se inicia el Libro con un capítulo, que bajo el título de **Metodología**, describe los métodos y técnicas teóricas empleadas en los cálculos y se completa con un **Glosario** con la relación de siglas y abreviaturas.

Se adjunta un CD-ROM en el que se incluye el presente documento de Tesis y la copia, en formato pdf, de los artículos publicados en los que se fundamenta.

1 Metodología

Estructura de la materia. Breve resumen histórico

Desde que en el siglo V a. C. Demócrito postulara que la materia estaba formada por unidades indivisibles, a las que denominó átomos,^{*} muchas han sido las teorías para intentar describir la estructura de la materia. Su teoría no obtuvo buena aceptación entre los filósofos de la época y no fue hasta el siglo XIX, cuando el científico inglés John Dalton, enunció la teoría atomista. En el transcurso de ese siglo, diferentes descubrimientos, con experiencias realizadas por el físico y químico británico Michael Faraday sobre electrolisis y por el físico francés Antoine Henri Becquerel sobre radioactividad, llevaron al descubrimiento de los electrones, protones y neutrones, dando paso a otras teorías sobre la estructura de la materia, con modelos como el de Joseph John Thomson y el de Ernest Rutherford.

Conforme se ampliaban los conocimientos sobre el átomo, se fueron sucediendo los distintos modelos atómicos, cada uno reemplazando al anterior y proporcionando una visión cada vez más próxima de la realidad.

Las observaciones experimentales de principios de siglo XX demostraron que el movimiento de los electrones y otras partículas submicroscópicas no podían tratarse mediante la mecánica clásica con las leyes de Newton, y se desarrolló una nueva herramienta teórica denominada mecánica cuántica.

Con una interesante mezcla de teoría clásica y cuántica, Niels Bohr postuló que para el átomo de hidrógeno:

^{*} **Átomo** (del latín *atōmum* y éste del griego *ἄτομον*, *sin partes*; también, se deriva de "a" no, y "tomo" divisible; no divisible. Fuente: Diccionario la Lengua. Real Academia Española; <http://www.rae.es>)

1. El electrón se mueve en órbitas circulares alrededor del núcleo con un movimiento descrito por la física clásica.
2. El electrón sólo tiene un conjunto de órbitas permitidas y en ellas su momento angular, J , ($J = nh/2\pi$) tiene valores determinados, es decir, n debe ser un número entero.
3. Un electrón sólo puede pasar de una órbita permitida a otra. En estas transiciones están implicadas cantidades discretas y fijas, o cuantos, de energía absorbida o emitida.

Normalmente, el electrón de un átomo de hidrógeno se encuentra en la órbita más próxima al núcleo ($n = 1$). La energía, en este caso, es la más baja permitida y el átomo se encuentra en su estado fundamental. Cuando un electrón adquiere un cuanto de energía pasa a un nivel más alto ($n = 2, 3, \dots$) o estado excitado, mientras que si pasa de una órbita de número alto a otra menor emite una determinada cantidad de energía. La energía del fotón emitido o absorbido se calcula de acuerdo con la ecuación de Planck ($\Delta E = h\nu$).

La teoría de Bohr permitió comprender el espectro de líneas del átomo de hidrógeno, pero no puede explicar el espectro de emisión de los átomos o iones con más de un electrón, ni el efecto de los campos magnéticos sobre los espectros de emisión.

Metodología cuántica

La mecánica cuántica moderna substituyó a la teoría atómica de Bohr y a partir de 1926 dos ideas fueron las que propiciaron esta nueva aproximación. En primer lugar, Louis de Broglie, considerando la naturaleza de la luz y de la materia, formuló que las partículas microscópicas, como los electrones, a veces pueden mostrar propiedades de ondas (dualidad onda-partícula). En segundo lugar, el principio de indeterminación de Heisenberg indica que, en los términos de la física clásica, no se puede medir, simultáneamente y con precisión, la posición y la cantidad de movimiento lineal y, por tanto la velocidad de un objeto dado. Schrödinger sugirió que los electrones, o cualquier otra partícula que posea propiedades de onda, podían ser descritos por una función matemática, que

denominó función de onda, y que designó con la letra Ψ . Esta función describe los diferentes estados disponibles para el electrón del átomo de hidrógeno, que quedan caracterizados por cuatro números, llamados números cuánticos. Los números cuánticos tienen un importante papel en la descripción de los niveles de energía y la forma de los orbitales. Así pues, el principal, n , describe el nivel de energía del electrón; el secundario, l , describe la forma de la región que ocupa el electrón; el magnético, m_l , designa la orientación del orbital en el espacio y el de espín, m_s , se refiere al giro del electrón y a la orientación del campo magnético producido por ese espín.

La probabilidad de encontrar al electrón en una determinada zona alrededor del núcleo la proporciona Ψ^2 . Se denomina orbital atómico (*atomic orbital*, AO) a la región del espacio en el que la probabilidad de encontrar al electrón es alta.

La correcta descripción de un sistema molecular depende, evidentemente, de lo que interesa saber de dicho sistema y es, por tanto, que desde un punto de vista químico-cuántico se pretende encontrar soluciones aceptables para los estados estacionarios del sistema. El sistema, en este caso, está formado por electrones (i, j, \dots) moviéndose en el campo creado por un conjunto de núcleos (A, B, \dots) y cuyas interacciones más importantes son las de carácter electrostático. Los estados estacionarios de dicho sistema vienen descritos por las soluciones encontradas cuando se resuelve la ecuación de Schrödinger independiente del tiempo:

$$\hat{H}\Phi = \varepsilon\Phi \quad (1.1)$$

En el contexto de una formulación no relativista y dentro de la aproximación de Born-Oppenheimer,^{1,2,3,4} el operador Hamiltoniano, \hat{H} , se define a través de la siguiente expresión:

$$\hat{H} = -\sum_{i=1}^N \frac{1}{2} \nabla_i^2 - \sum_{i=1}^N \sum_{A=1}^M \frac{Z_A}{r_{iA}} + \sum_{i=1}^N \sum_{j>i}^N \frac{1}{r_{ij}} + \sum_{A=1}^M \sum_{B>A}^M \frac{Z_A Z_B}{R_{AB}}, \quad (1.2)$$

que consta de los términos correspondientes a: la energía cinética de los electrones, la atracción de Coulomb entre electrones y núcleos, la repulsión

entre electrones y una constante de repulsión nuclear para una geometría dada, respectivamente. La energía total, \mathcal{E} , para los núcleos fijos proporciona un potencial para el movimiento nuclear. La ecuación 1.1 constituye el problema electrónico, la cuestión inicial básica a resolver. Mediante los métodos de la Química Cuántica se pueden obtener soluciones aproximadas de diferente nivel de calidad.

El Hamiltoniano esconde la física de lo que se conoce como “correlación electrónica”. La repulsión de Coulomb dada por el término r_{ij}^{-1} , la inversa de la distancia entre dos electrones, aumenta en las regiones cercanas a $r_{ij} = 0$, previniendo la ocupación de dos electrones en un mismo lugar del espacio. Esto supone que el movimiento de cualquier par de electrones no es independiente, sino correlacionado. Decir que dos electrones están correlacionados es equivalente a expresar que la probabilidad de encontrarlos en el mismo punto del espacio es nula. La posición instantánea del electrón i forma el centro de una región que el electrón j evitará ocupar. Por esta razón, se suele decir que cada electrón, descrito mediante la función de onda exacta Φ , está rodeado por un “agujero de Coulomb” o “*Coulomb hole*”. La consideración de la correlación electrónica no es sencilla, por lo que no es tratada adecuadamente por muchos métodos aproximados.

La función más simple empleada para describir un sistema polielectrónico es el “determinante de Slater”, construido por funciones monoelectrónicas ortogonales. Los electrones son fermiones y, por tanto, deben ser descritos mediante funciones antisimétricas. La función para un electrón que describe tanto la distribución espacial como su espín recibe el nombre de “espín-orbital”. Debido a que el Hamiltoniano empleado no depende del espín electrónico, cada espín-orbital puede expresarse mediante el producto del orbital espacial y la función de espín. El determinante de Slater cumple además la ley básica de simetría derivada del “principio de indistinguibilidad”, ya que permite la descripción de N electrones ocupando N espín-orbitales sin especificar qué electrón ocupa cada orbital. Desde un punto de vista físico, el uso de un determinante de Slater para describir un sistema de múltiples electrones implica que nos encontramos en un modelo de electrones independientes, no correlacionados, descartando por tanto el hueco de Coulomb. No obstante, la incorporación de la “condición de antisimetría” de la función de onda para fermiones hace que, en el

determinante de Slater, el movimiento de dos electrones con la misma función de espín esté correlacionada, es decir, que la probabilidad de encontrar dos electrones con espines paralelos en el mismo punto del espacio sea cero, la llamada “correlación de intercambio”. Este fenómeno se conoce como el “hueco de Fermi”. Nos encontramos, por tanto, ante un modelo de partículas independientes, donde el comportamiento de ciertos electrones no es totalmente independiente, ya que el hueco de Fermi simula hasta cierto punto el hueco de Coulomb. Como el movimiento de los electrones con función de espín diferente permanece sin correlacionar (existe una probabilidad finita de encontrar dos electrones con espines opuestos en el mismo punto del espacio), la función de onda correspondiente a un solo determinante de Slater es conocida comúnmente como “función de onda no correlacionada”.

Métodos *ab initio*

Los métodos cuánticos *ab initio*^{2,5} usualmente consideran como punto de partida la aproximación **Hartree-Fock** (HF),^{6,7} la cual constituye, en la mayoría de los casos, el primer paso hacia procedimientos más precisos, jugando un papel crucial en la Química Cuántica moderna. De hecho, muchos de los métodos químico-cuánticos pueden considerarse como simplificaciones del método HF o mejoras del mismo. El método HF proporciona las herramientas matemáticas necesarias para obtener los espín-orbitales desconocidos para la construcción del mejor determinante de Slater dentro del marco del “principio variacional”, aplicando la condición de estacionario. Mediante dicho principio se establece que la mejor función de onda (en este caso de tipo monodeterminantal) es aquella que proporciona la energía más baja. Por tanto, mediante la minimización de la energía con respecto a la elección de los espín-orbitales se puede llegar a las condiciones Hartree-Fock. La ecuación HF no es lineal y debe resolverse de forma iterativa a través del Procedimiento de Campo Autoconsistente (*Self-Consistent Field*, SCF). En la práctica, la ecuación HF se resuelve introduciendo un conjunto finito de funciones de base espaciales que dan lugar a distintas ecuaciones matriciales: las ecuaciones de Roothaan⁸ en el

caso de determinantes restringidos de capa cerrada, las ecuaciones de Pople-Nesbet⁹ para determinantes no restringidos y las ecuaciones Roothaan-Hartree-Fock¹⁰ para determinantes restringidos de capa abierta.

Obviamente, ni todas las moléculas, ni todos los estados electrónicos de moléculas con un número par de electrones pueden describirse mediante pares de electrones situados en orbitales de capa cerrada. Es por ello que el formalismo de capa cerrada tiene que ser generalizado para situaciones en las cuales el sistema de estudio tiene uno o más electrones desapareados, generando una capa abierta. Mediante esta generalización surgen los Procedimientos de Capa Abierta Restringidos (*Restricted Open-shell Hartree-Fock*, ROHF) y No Restringidos (*Unrestricted Hartree-Fock*, UHF), mencionados en el párrafo anterior. En el primero, todos los electrones, excepto aquellos explícitamente requeridos para ocupar los orbitales de capa abierta, son situados en orbitales capa cerrada. La ventaja de este procedimiento es que las funciones de onda obtenidas son funciones propias del operador de espín, \hat{S}^2 . Sin embargo, la solución de las ecuaciones ROHF es menos inmediata que en el caso UHF y por ello este último suele preferirse. En esta variante, se introducen formas específicas no restringidas de los espín-orbitales y se derivan las ecuaciones espaciales que determinan los orbitales espaciales no restringidos. El método UHF se suele elegir, no sólo por su simplicidad y generalidad, sino por la correcta descripción que proporciona en situaciones tales como la disociación homolítica, inadecuadamente tratada mediante el método restringido de capa cerrada. En este caso, la solución no restringida permite acomodar el desapareamiento de los electrones inherente a la rotura del enlace. No obstante, el cálculo no restringido presenta como deficiencia el hecho de que no produce estados de espín puros. Las funciones de onda obtenidas como resultado no son funciones propias de espín del operador \hat{S}^2 y normalmente están contaminadas con contribuciones de funciones de onda de espín más elevado. En respuesta a este inconveniente, se han desarrollado métodos de proyección (o aniquilación) de la “contaminación de espín”, de tal forma que se eliminan los estados de espín superiores. A pesar de eso, la contaminación puede ser bastante intensa, incluyendo estados de muy elevado espín, y entonces el proceso, que ya resulta de por sí computacionalmente caro y tedioso, puede resultar inútil. Las técnicas de

proyección se han aplicado principalmente a nivel UHF (Hartree-Fock Proyectado o *Projected Hartree-Fock*, PHF) y a los métodos de la Teoría de Perturbaciones de Møller-Plesset hasta Segundo Orden, MP2, (MP2 Proyectado o *Projected MP2*, PMP2),^{11,12} los cuales se describen más adelante.

La aproximación Hartree-Fock no proporciona una descripción precisa de las propiedades moleculares, debido principalmente a que no tiene en cuenta la energía de correlación electrónica. La “energía de correlación” (E_{corr}) se define como la diferencia entre la energía exacta no relativista del sistema dentro de la aproximación Born-Oppenheimer (ε_0) y la energía HF (E_0) en el límite en el cual el conjunto de funciones de base empleado cumpla la condición de que sea un conjunto de base completo.²

$$E_{corr} = \varepsilon_0 - E_0. \quad (1.3)$$

Como ε_0 es menor que E_0 , la energía de correlación es negativa. El uso de un conjunto de base completo resulta prohibitivo o simplemente imposible. Es por ello que la energía de correlación exacta de un sistema no puede calcularse, salvo en el caso en que su tamaño molecular sea pequeño. En la práctica, la definición de E_{corr} corresponde entonces a la diferencia entre la energía calculada a un nivel dado del tratamiento de la correlación electrónica y la energía HF correspondiente, ambas obtenidas con el mismo conjunto de base monoeléctronica, suficientemente flexible. La función de onda HF, siendo un determinante de Slater, no está correlacionada, cosa que supone ciertas limitaciones en su aplicabilidad. De cara a introducir la correlación electrónica se han desarrollado todo un conjunto de procedimientos, conocidos como métodos “*post-Hartree-Fock*”, cuya calidad se encuentra valorada en función de la correlación electrónica que incluyen (además de por el conjunto de base monoeléctronica).

El método HF determina el mejor determinante de Slater, Φ_{HF} , dentro de un conjunto de base dado. Una posible mejora del procedimiento HF, que aumenta la flexibilidad variacional del problema, supone la construcción de una función de onda de partida que contenga más de un determinante de Slater. Esa función de onda multideterminantal puede escribirse de forma genérica como

$$\Psi = a_0 \Phi_{\text{HF}} + \sum_{i=1} a_i \Phi_i, \quad (1.4)$$

donde el cuadrado de los coeficientes a_i es el peso de cada uno de los determinantes en la expansión. La base polielectrónica Φ_i usada para expandir la función multideterminantal se construye mediante el reemplazo de Orbitales Moleculares (*Molecular Orbitals*, MOs), o espín-orbitales, que están ocupados en el determinante de Slater Φ_{HF} por MOs que se encuentren desocupados (con N electrones y Q funciones de base, la solución HF da lugar a N MOs ocupados y $Q - N$ desocupados o virtuales). Estos Φ_i se describen de acuerdo con la cantidad de MOs ocupados, en el determinante HF, que han sido reemplazados por MOs sin ocupar, es decir, Simples (*Single*, S), Dobles (*Double*, D), Triples (*Triple*, T), etc. El truncamiento del conjunto de determinantes de Slater limita la descripción de la energía de correlación. La diferencia entre los métodos *post*-Hartree-Fock se encuentra en la manera en la cual esa energía de correlación es incluida. A este respecto, los procedimientos *ab initio* más frecuentes son: Interacción de Configuraciones (*Configuration Interaction*, CI), Teoría de Perturbaciones Multipartícula (*Many-Body Perturbation Theory*, MBPT) y *Coupled Cluster* (CC).

El método de **Interacción de Configuraciones** (CI)¹³ se basa en el principio variacional (análogamente al procedimiento HF) y usa como función de partida la función de onda, Ψ_{CI} , formada por la combinación lineal de determinantes de Slater. Estos determinantes se construyen a partir de MOs tomados, por ejemplo, de un cálculo previo Hartree-Fock, y se mantienen fijos en la expansión CI. En este procedimiento, los coeficientes a_i de la ecuación 1.4 se determinan mediante el requerimiento de que se cumpla la condición de estacionario. En el caso de que todos los posibles determinantes excitados (simples, dobles, triples, etc.) se incluyan en la función CI de prueba, el método se conoce como CI Completo (*Full CI*, FCI) y resulta en la solución exacta dentro del conjunto de base monoeléctronica empleado. En la práctica, debido a las restricciones computacionales, el número de determinantes excitados que pueden considerarse en la expansión CI debe reducirse, dando lugar a los niveles de

cálculo CID (incluyendo las dobles excitaciones), CISD (simples y dobles), CISDT (simples, dobles y triples), etc. Uno de los problemas más importantes que presentan los métodos CI truncados es que no escalan apropiadamente con el número de partículas o, como suele decirse, no son “*size-extensive*”.

Como una extensión y refinamiento del procedimiento *ab initio* CI, podemos encontrar la **Metodología Multiconfiguracional de Campo Autoconsistente** (*Multi-Configurational Self-Consistent Field*, MCSCF)^{3,4}

En este caso, la función de onda Ψ_{MCSCF} tiene la misma forma que la Ψ_{CI} y es sometida también al principio variacional, pero tanto los coeficientes de la expansión como los MOs usados para construir los determinantes son optimizados en el proceso. La función de onda MCSCF no sufre el problema de la contaminación de espín ya que las funciones usadas en la expansión MCSCF, denominadas Funciones de Estado Configuracional (*Configuration State Functions*, CSFs) o Configuraciones Adaptadas de Espín (*Spin-Adapted Configurations*, SACs), son estados puros de espín. Una variante del procedimiento MCSCF que se ha hecho particularmente popular, debido a su simplicidad conceptual y técnica, es el método SCF de Espacio Activo Completo (*Complete Active Space SCF*, CASSCF)^{14,15,16}. En este caso, la selección de las configuraciones que se incluirán en la función de prueba (Ψ_{CAS}) se lleva a cabo clasificando los MOs en tres conjuntos: orbitales inactivos, activos y externos. Los MOs activos serán típicamente algunos de los espín-orbitales topológicamente equivalentes a los MOs ocupados y desocupados más altos y más bajos en energía, respectivamente, procedentes de un cálculo HF. Sin embargo, hay que tener en cuenta que la elección del espacio activo dependerá del problema en cuestión. Los MOs inactivos tienen siempre dos electrones, mientras que los MOs externos (también denominados secundarios o virtuales) no presentan ningún electrón. Entre los MOs activos se lleva a cabo un proceso CI completo o *full* CI (CAS-CI), incluyendo en la optimización MCSCF todas las CSFs de simetría espacial y de espín apropiadas. La energía de correlación introducida por el procedimiento CASSCF usualmente corresponde a la “correlación no-dinámica” o “correlación electrónica estática” e incluye las CSFs cuasidegeneradas. La restante es la “correlación dinámica”. El objetivo de los métodos CASSCF (y también de otros procedimientos

MCSCF) no es recuperar una gran parte de la energía de correlación total, sino más bien describir cualitativamente todos los cambios que ocurren en la energía de correlación para un proceso dado, cosa que se consigue mediante una apropiada elección de los MOs a correlacionar.

Una función de onda MCSCF puede ser usada también como referencia en procedimientos de cálculo de la correlación dinámica, por ejemplo, en un cálculo CI. En este caso, un CISD implicaría excitaciones de uno o dos electrones a partir de todos los determinantes que forman la función MCSCF y estaríamos hablando del método **Interacción de Configuraciones Multirreferencial** (*Multi-Reference Configuration Interaction*, MRCI).^{17,18} La ventaja de estos métodos es que pueden generar funciones de onda muy precisas si la referencia MCSCF incluye suficientes configuraciones. Los inconvenientes son, por una parte, que no escalan linealmente con el número de partículas (no son *size-extensive*), debido a que siguen siendo procedimientos truncados y, por otra, su elevado coste computacional, demasiado alto incluso truncando la expansión MRCI al nivel de las simples y dobles excitaciones.

Los métodos de la **Teoría de Perturbaciones Multipartícula** (MBPT)^{2,3} se basan en la aplicación de técnicas de perturbaciones a las funciones de onda obtenidas como soluciones de otros métodos *ab initio*, con la finalidad de añadir correcciones a esas funciones de onda y a la energía. La energía y función de onda resultantes pueden escribirse en base a un desarrollo de Taylor en potencias de un parámetro de perturbación, surgiendo correcciones a distintos órdenes (primero, segundo, etc.) en función de los términos considerados de la expansión de Taylor. Entre estos métodos, los más comunes son los correspondientes a la teoría de perturbaciones de Møller-Plesset (MP)¹⁹ en sus diferentes órdenes de perturbación incluidos (MP2, MP3, etc.). El método MP2 (que usa como referencia la función de onda HF) típicamente cubre alrededor de 80–90% de la energía de correlación, aún siendo el método más barato (computacionalmente hablando) para tratar la correlación electrónica.⁵ Como muchas de las aproximaciones relativas a la teoría de perturbaciones, MP funciona bien cuando la perturbación es suficientemente pequeña. Una de las grandes ventajas que presenta este método respecto a los CI es que MP escala adecuadamente con el número de partículas del sistema molecular (es *size-*

extensive). Por el contrario, la desventaja respecto a CI se encuentra en el hecho de que mientras estos últimos métodos dan lugar a una aproximación por encima o *upper bound* a la energía total exacta del sistema (consecuencia del procedimiento variacional empleado), no se tiene esa garantía en los métodos de perturbaciones. Otro inconveniente de los métodos MP es la contaminación de espín cuando se usa una función de onda UHF como referencia para describir una molécula de tipo capa abierta, como en el caso de los sistemas radicalarios. Como se mencionó anteriormente al describir el método UHF, se puede recurrir a técnicas de proyección que eliminen esa contaminación de espín, dando lugar a los métodos PMP.

De la misma forma que el método monorreferencial CI puede extenderse a MRCI, es posible usar métodos de perturbaciones sobre una función de onda de referencia multideterminantal para obtener el nivel de teoría MR-MBPT. Uno de los métodos más populares dentro de esta categoría es el CASPT2,^{20,21} basado en una función de referencia CASSCF. Este método, concretamente la estrategia de cálculo CASPT2//CASSCF, se suele emplear en el tratamiento de estados excitados debido a su generalidad y a sus exitosas predicciones (0.1–0.2 eV) a un nivel de coste computacional moderado.¹

El tercer grupo de procedimientos *ab initio post-Hartree-Fock* se basan en la teoría **Coupled Cluster** (CC), desarrollada por Čížek y Paldus^{22,23,24} en los años 60. Mientras que los métodos de perturbaciones añaden todo tipo de correcciones a la función de onda de referencia hasta un orden de perturbación dado, la idea en los métodos CC consiste en incluir todas las correcciones de un tipo dado a orden infinito.²⁵ Ejemplos de ellos son CCSD, CCSDT o CCSDTQ. La demanda computacional de esta metodología resulta ser alta. Así, mientras que CCSD aún puede aplicarse a sistemas de tamaño medio, los métodos CCSDT o CCSD(T), donde, en este último caso, las contribuciones triples se evalúan mediante la teoría de perturbaciones y se añaden a los resultados CCSD, no pueden llevarse a cabo normalmente, a pesar de su gran precisión en situaciones claramente de una sola referencia. Al igual que los métodos MP, los procedimientos CC tienen la ventaja de escalar correctamente con el número de partículas (método *size-extensive*) y la desventaja de sufrir el problema de la

contaminación de espín en situaciones de capa abierta. Sin embargo, la naturaleza de orden infinito de los métodos CC hace que éstos sean mucho mejores a la hora de reducir dicha contaminación. El inconveniente es que son difícilmente aplicables en la práctica por la forma en la que escalan al aumentar el nivel de excitación. La teoría estándar CC se basa en una función de onda de referencia monodeterminantal y, por tanto, del mismo modo que MP, funciona apropiadamente si la función de onda de orden cero es suficientemente buena. Pero de nuevo, debido a la suma de contribuciones a orden infinito, CC es más tolerante que MP a una función de referencia pobre. En este sentido, ya que los determinantes correspondientes a las excitaciones simples permiten relajar los MOs y con ello simular en cierta medida el carácter multirreferencial en la función de onda, la magnitud de los parámetros de amplitud de las simples excitaciones resulta ser un indicador adecuado de lo bueno que es el determinante único HF como referencia. El “diagnóstico T_1 ”²⁶ es un parámetro basado en la amplitud de las simples y, aunque en la mayoría de textos se diga que tiene en cuenta el carácter multiconfiguracional de la función de onda de referencia, en realidad es un indicativo de la falta de correlación en el procedimiento.²⁷ Esa falta de correlación se puede solucionar de dos formas, bien aumentando el nivel de excitación (triples, cuádruples, etc.) o mejorando la referencia, a través de un método multiconfiguracional. Con una mejor referencia se recupera más correlación. Concretamente, si T_1 es mayor que 0.02, se debería tomar una de estas dos opciones.

El uso de los métodos CI ha declinado, en los últimos años, hacia la utilización de los procedimientos MP y, especialmente, CC. Una de las razones que ha impulsado este hecho es el reconocimiento del correcto escalado con el número de partículas (*size-extensivity*) como característica metodológica importante a la hora de obtener resultados precisos. Algunos procedimientos capaces de corregir este defecto en los cálculos CI, como el método de Interacción de Configuraciones Autoconsistente y Consistente con el Tamaño incluyendo las Dobles excitaciones (*Self-Consistent Size-Consistent Double Configuration Interaction*, (SC)²DCI)²⁸ no se han generalizado. Los estados excitados, sin embargo, son difíciles de tratar con métodos de perturbaciones o de tipo *Coupled Cluster* y, como se ha comentado previamente, el procedimiento MCSCF es el preferido como

punto de partida en estos casos. Finalmente, en términos de precisión, con un conjunto de base de tamaño medio se observa a menudo el siguiente orden:⁵

$$\text{HF} \ll \text{MP2} < \text{CISD} < \text{CCSD} < \text{MP4} < \text{CCSD(T)}$$

Todos estos métodos son monodeterminantes. Los niveles de teoría que incluyen múltiples referencias son de difícil clasificación ya que la calidad de los resultados depende fuertemente de la naturaleza y tamaño de la referencia.

Conjuntos de funciones de base monoeléctronica

Los métodos *ab initio*, normalmente, emplean MOs para la descripción de los electrones del sistema de estudio. Estos orbitales, o funciones monoeléctronicas, se suelen construir, en general, como expansión lineal a partir de un conjunto inicial de funciones de base atómicas (centradas en los átomos), lo que se conoce como aproximación “Orbitales Moleculares como Combinación Lineal de Orbitales Atómicos” (Molecular Orbitals as Linear Combination of Atomic Orbitals, MO-LCAO). Sin embargo, una base completa en este espacio vectorial de Hilbert[♦] supone un número infinito de funciones. Cuando se usa una base finita, sólo pueden representarse las componentes de los MOs a lo largo de los ejes de coordenadas correspondientes a la base seleccionada. En principio, cuanto más pequeña sea la base, más pobre será la representación y también la precisión del cálculo, aunque la cancelación de errores puede dar lugar a un resultado acorde con los datos experimentales.

Hay dos tipos de funciones de base monoeléctronica comúnmente usadas en los cálculos de estructura electrónica:

- Orbitales de Tipo Slater (*Slater Type Orbitals*, STOs).²⁹ Presentan una dependencia exponencial con la distancia entre los

[♦] Se recomienda en este punto la lectura del capítulo 0 “Antecedentes” de la Tesis doctoral del profesor Rosendo Pou Américo, en el que se amplía la visión de los progresos habidos durante la historia de la matemática y las personas que lo hicieron realidad.

núcleos y electrones. Este tipo de funciones son más eficientes y más precisas en lo que se refiere a la representación de los orbitales moleculares.

- Orbitales de Tipo Gaussiano (*Gaussian Type Orbitals*, GTOs).³⁰ Presentan una dependencia exponencial con el cuadrado de la distancia entre los núcleos y electrones. Estas funciones suponen un menor tiempo computacional en la evaluación de las integrales dielectrónicas y son ampliamente empleadas en los cálculos de estructura electrónica.

El concepto de base mínima implica que sólo se emplea el mínimo número de funciones necesarias, por átomo, en la descripción de sus orbitales atómicos ocupados. Para el hidrógeno, esto significa una única función s , mientras que se requieren dos funciones s ($1s$ y $2s$) y un conjunto de funciones p ($2p_x$, $2p_y$ y $2p_z$) para el siguiente periodo en la tabla periódica. El siguiente paso para mejorar el conjunto de base supone doblar todas las funciones, dando lugar al conjunto de base de tipo Doble Zeta (*Double Zeta*, DZ). En este caso se emplean dos funciones s para el hidrógeno ($1s$ y $1s'$) y cuatro funciones s y dos conjuntos de funciones p para los elementos de la primera fila. Una variante de ese tipo de bases es aquella que sólo dobla las funciones de valencia (más importantes químicamente), produciendo el conjunto de base Doble Zeta de Valencia (*Valence Double Zeta*, VDZ). El siguiente paso, en cuanto al aumento del número de funciones, es el que da lugar a la Triple Zeta (*Triple Zeta*, TZ) que, si sólo se mejora la capa de valencia, resulta en el conjunto de base Triple Zeta de Valencia (*Valence Triple Zeta*, VTZ).

En algunos casos, aumentar sólo el número de funciones s y p para cada átomo no es suficiente para describir el fenómeno bajo estudio, y se requieren funciones de momento angular superior. Estas funciones se denominan funciones de polarización. La adición de un solo conjunto de ellas (funciones p para el hidrógeno y funciones d para los átomos pesados) a las bases DZ y TZ forma los conjuntos de base Doble Zeta de Polarización (*Double Zeta Polarization*, DZP) y Triple Zeta de Polarización (*Triple Zeta Polarization*, TZP), respectivamente. En ciertas situaciones, en las cuales se

pretende describir con mayor precisión la parte externa de la función de onda multielectrónica, se añaden funciones difusas (funciones de base con exponentes pequeños) al conjunto de base o incluso funciones no centradas en los átomos, como las funciones de enlace (*bond functions*). Las funciones difusas son requeridas especialmente allí donde se encuentran presentes electrones débilmente enlazados, como, por ejemplo, en aniones o estados excitados muy externos, puesto que deben representarse situaciones con electrones lejanos al núcleo.

Con la finalidad de reducir el coste computacional significativamente, sin una pérdida importante de precisión, las funciones de base GTOs Primitivas (*Primitive Gaussian Type Orbitals*, PGTOs), se agrupan mediante la formación de combinaciones lineales fijas (proceso denominado contracción del conjunto de base), para dar lugar a un conjunto más pequeño de funciones llamadas GTOs Contraídas (*Contracted Gaussian Type Orbitals*, CGTOs):

$$\phi_{CGTO} = \sum_i a_i \phi_{i,PGTO} . \quad (1.5)$$

La contracción de un conjunto de base supone entonces una reducción del número de parámetros variacionales y, por tanto, una ganancia de la eficiencia computacional. La notación en el proceso de contracción del conjunto de base es la siguiente:

(“funciones primitivas para los elementos de la primera fila”/“funciones primitivas para el hidrógeno”)→[“funciones contraídas para los elementos de la primera fila”/“funciones contraídas para el hidrógeno”].

Un ejemplo de contracción es $C,H(10s4p1d/2s1p) \rightarrow C,H[3s2p1d/2s1p]$ para el carbono y el hidrógeno, respectivamente. Hay dos formas diferentes de contraer un conjunto de PGTOs, la contracción general y la segmentada. En el primer caso todas las primitivas (de un átomo dado) de un momento angular determinado participan describiendo todas las funciones contraídas de ese momento angular, pero con coeficientes de contracción distintos. En la contracción segmentada cada orbital o función primitiva participa en una o unas pocas funciones contraídas.

El tipo de funciones de base más extendido en cuanto a su uso, por su presencia en la mayoría de programas computacionales químico-cuánticos y su extenso calibrado disponible, son las llamadas funciones de base de Pople.³¹ Este conjunto de base usa contracción segmentada y permite llevar a cabo cálculos de bajo coste computacional. Ejemplos de funciones Pople son:

- STO- n G.³² Orbitales de tipo Slater formados por n PGTOs. Corresponde a una base mínima. El conjunto de base más común de este tipo es STO-3G.
- 6-31G³³ Este tipo de base corresponde a las bases de valencia doblada o *double split valence basis*, donde la parte interna de los orbitales de valencia viene representada por una función contraída de tres PGTOs y la parte externa está representada por una PGTO. La contracción es C,H(10s4p/4s)→C,H[3s2p/2s].
- 6-311G.³⁴ Este conjunto de base presenta un *core* con una contracción de seis PGTOs y una valencia dividida en tres funciones (*triple split valence*), representadas por tres, una y una función PGTO, respectivamente.

Las funciones de polarización se indican tras la letra “G” con una nomenclatura separada para los átomos pesados y el hidrógeno. Así, por ejemplo, 6-311++G(2df,2pd) es un conjunto de base de valencia triplicada con funciones difusas *sp* añadidas y dos funciones *d* y una *f* para los átomos pesados y difusas *s* y dos funciones *p* y una *d* en los hidrógenos. En el caso de sólo usar un conjunto de funciones de polarización, se suele usar otra notación alternativa basada en el símbolo “*”. La base 6-31G* es idéntica entonces a 6-31G(*d*) con sólo una única función de polarización de tipo *d* para los átomos del segundo periodo. La adición de funciones difusas se denota mediante el símbolo “+” (6-31+G) o “++” (6-311++G) antes de la letra “G”, con el primer “+” indicando un conjunto de funciones difusas *s* y *p* para los átomos pesados y el segundo (“++”) corresponde a la adición extra de una función *s* en los hidrógenos.

Al contrario que los conjuntos de base de tipo Pople, las bases más modernas presentan la ventaja de que sus coeficientes de contracción han sido optimizadas mediante el uso de métodos químico-cuánticos que tienen en cuenta la energía de correlación y presentan la habilidad de generar una secuencia de conjuntos de base que converja hacia el límite del conjunto de base (los errores HF y de correlación se van reduciendo en cada paso de aumento de calidad). Los conjuntos de base Consistentes con la Correlación (*Correlation Consistent*, cc) de Dunning,^{35,36} también de contracción segmentada, fueron desarrolladas con la finalidad de recuperar la energía de correlación de los electrones de valencia. Su diseño es tal que aquellas funciones que contribuyen con cantidades similares a la energía de correlación son incluidas conjuntamente, independientemente del tipo de función. Ejemplos de ellas son cc-pVDZ, cuya contracción es $C,H(9s4p1d/4s1p) \rightarrow C,H[3s2p1d/2s1p]$ o cc-pVTZ con $C,H(10s5p2d1f/5s2p1d) \rightarrow C,H[4s3p2d1f/3s2p1d]$. Estos conjuntos de base pueden crecer en tamaño añadiendo funciones difusas, en cuyo caso se añade el prefijo “aug-“ al acrónimo.³⁷ Este aumento consiste en añadir una función extra con un exponente más pequeño por cada momento angular, es decir, aug-cc-pVDZ tiene funciones 1s, 1p y 1d añadidas y aug-cc-pVTZ presenta funciones 1s, 1p, 1d y 1f extra para los átomos que no sean de hidrógeno.

Otro tipo de conjuntos de base muy poderoso, optimizado con métodos correlacionados y que usa el esquema de contracción general, es el que engloba las bases de Orbitales Atómicos Naturales (*Atomic Natural Orbital*, ANOs).³⁸ De hecho, las bases segmentadas CC antedichas se inspiraron en la extraordinaria eficacia de los ANOs de contracción general. La idea que fundamenta este tipo de bases es la contracción de un conjunto grande de PGTOs en un número relativamente pequeño de funciones CGTO, mediante el uso de Orbitales Naturales (*Natural Orbitals*, NOs) obtenidos en cálculos correlacionados sobre átomos libres, típicamente a nivel CISD. Para aumentar la flexibilidad de la base se promedia entre NOs obtenidos de diferentes situaciones en el átomo: neutro, catión, anión, diversos estados excitados, con adición de un campo eléctrico, etc. Los NOs son aquellos orbitales que diagonalizan la matriz de densidad monoeléctronica y sus valores propios son los números de ocupación orbital. La contracción ANO

selecciona las combinaciones importantes de las funciones PGTO en función de la magnitud de los números de ocupación. Debido a su carácter difuso y flexible, las bases ANO, al igual que los conjuntos de base Extendida Consistentes con la Correlación (*Augmented Correlation Consistent*, aug-cc), son especialmente indicadas para el tratamiento de estados de tipo aniónico, donde se requieren conjuntos de base grandes, difusos y suficientemente flexibles.^{39,40,41} Usualmente, las bases ANO dan lugar a mejores resultados con un menor número de funciones que otras bases. Se han contraído distintos conjuntos de base de tipo ANO con diferente número de primitivas: ANO-L^{42,43,44} y ANO-S.⁴⁵

Hipersuperficies de energía potencial

La resolución de la ecuación de Schrödinger, para distintos valores fijos de las coordenadas atómicas del sistema molecular bajo estudio, va a dar lugar a todo un conjunto de resultados, que relacionan la energía calculada de los estados electrónicos con las coordenadas nucleares. La representación de este tipo de relaciones constituye la llamada Hipersuperficie de Energía Potencial (*Potential Energy Hypersurface*, PEH) donde el término hipersuperficie se emplea porque la representación gráfica de la energía frente a las posiciones atómicas del sistema molecular generalmente requiere de más de tres dimensiones. Sin embargo, para visualizar gráficamente los resultados se suelen seleccionar las coordenadas de dos en dos y la representación energética corresponde entonces a la Superficie de Energía Potencial (*Potential Energy Surface*, PES). Si sólo se escoge una coordenada, el resultado corresponderá a la Curva de Energía Potencial (*Potential Energy Curve*, PEC). Existen dos tipos de puntos estacionarios esenciales en las PEHs. Por un lado, se encuentran los mínimos relativos, correspondientes a estructuras (reactivos, productos) que van a poseer una cierta estabilidad, ya que se encuentran atrapadas energéticamente. En este caso, la matriz de las derivadas segundas de la energía (Hessiano) presenta todos sus valores propios positivos. Por otro lado, si este Hessiano tiene n valores propios negativos, se tiene un Punto Silla (*Saddle Point*, SP) de orden n , es decir, un punto que es máximo en n de los desplazamientos

nucleares y mínimo en el resto. El SP se relaciona químicamente con el Estado de Transición (*Transition State*, TS) de las reacciones adiabáticas o fotoadiabáticas, las cuales conectan, a través de este TS, diferentes regiones de una PEH en el estado fundamental o excitado, respectivamente.

Interacción de la luz con la materia

El objetivo de gran parte de las cuestiones a las que se enfrenta la química teórica aplicada, es el de poder explicar y predecir las propiedades espectroscópicas y fotoquímicas de los sistemas moleculares mediante los métodos teóricos.⁴⁶ En el proceso de interacción de la luz con la materia, se puede producir una absorción de energía, proporcional a la diferencia energética entre los estados electrónicos de la molécula implicados en la transición. Es por tanto que, el estudio de la energía de los puntos singulares, junto con las probabilidades de transición, o cruces entre las distintas PEHs, van a ser las dos características químico-cuánticas más importantes a la hora de describir los **Procesos Fotofísicos** (espectros de absorción y emisión, etc.) y **Fotoquímicos** (fotoestabilidad, aislamiento e identificación de fotoproductos, racionalización de mecanismos, etc.).

La excitación electrónica tras la absorción de luz es sólo el primer paso de un camino que puede llegar a desembocar en la formación de un fotoproducto. En la figura 1.1 se presenta un esquema general de los principales procesos fotofísicos y fotoquímicos que pueden desencadenarse tras la absorción de radiación electromagnética UV-visible. En tales procesos se ven envueltos los estados excitados que han sido poblados y que disponen de un tiempo de vida limitado, pues la molécula es inestable en esta situación energética y buscará los caminos de relajación adecuados para volver a su estado fundamental.

Los fenómenos descritos se corresponden básicamente con los encontrados frecuentemente en los diagramas de Jablonski⁴⁷ (véase figura 1.2), es decir, Absorción (*Absorption*, Abs) representada en la figura 1.2 con flechas ascendentes desde el estado fundamental, S_0 , hasta el primer estado excitado, S_1 , y segundo estado excitado, S_2 , Fluorescencia (*Fluorescence*, F), Fosforescencia (*Phosphorescence*, P), Conversión Interna (*Internal*

Conversion, IC) y cruce entre sistemas (*Inter-System Crossing*, ISC). La descripción detallada de los procesos que se producen tras la interacción de la radiación con la materia puede encontrarse en la bibliografía.^{1,48,49,50,51,52}

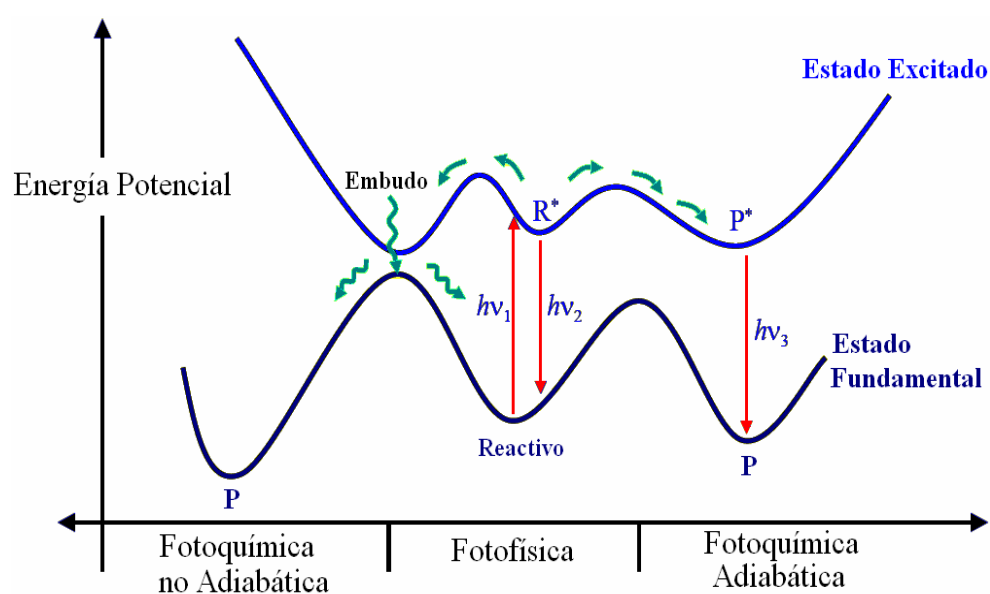


Figura 1.1. Descripción de los procesos fotofísicos y fotoquímicos utilizando perfiles de superficies de energía potencial.

Tras la absorción de la energía, que es el proceso más rápido de todo el mecanismo (del orden de los femtosegundos, 10^{-15} s) las moléculas pasan a desexcitarse a través de movimientos vibracionales, fenómeno que dura entre 10^{-14} y 10^{-11} segundos dependiendo del tamaño molecular. Las moléculas recorren un camino, desde su disposición inicial hacia el mínimo de su estado excitado, en el que pueden cruzarse con PHEs de la misma multiplicidad de espín, produciéndose una relajación no radiativa mediante un proceso de IC, o de distinta multiplicidad de espín por medio de un ISC. Por último, las moléculas que han quedado en la región del mínimo excitado acaban de relajarse por mecanismos radiativos, más lentos, que pueden producirse entre estados de igual multiplicidad de espín (fluorescencia, 10^{-9} – 10^{-6} s) o distinta (fosforescencia, 10^{-3} – 10^{-2} s).

Como resultado del proceso global puede presentar una energía radiante de mayor longitud de onda que la absorbida por las moléculas y en el que el rendimiento cuántico, Φ_F , de fluorescencia y/o fosforescencia, mide la relación entre la radiación absorbida y la emitida. Los sistemas que presentan una emisión intensa tienen un rendimiento cuántico cercano a la unidad y toda reducción de dicho rendimiento cuántico implica que la molécula se ha relajado por mecanismos no radiativos, es decir, que ha sufrido una desactivación o *quenching*.

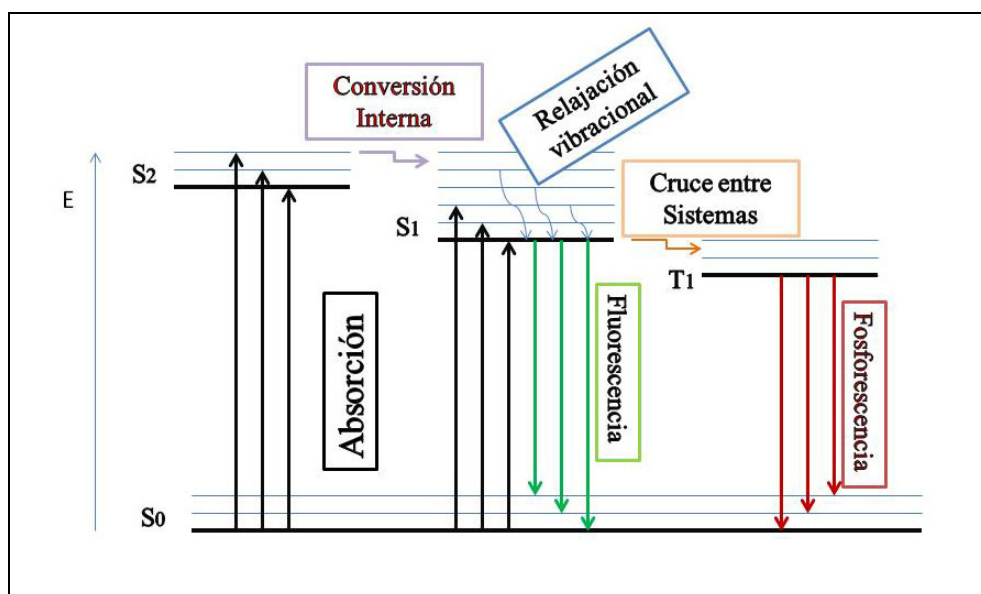


Figura 1.2. Diagrama de Jablonski: descripción esquemática de los diferentes procesos que pueden tener lugar en una molécula tras ser fotoexcitada desde su estado fundamental.

Desde un punto de vista teórico, el análisis adecuado de los espectros de absorción y de emisión implica calcular las energías relativas entre las distintas PEHs, en una serie de puntos significativos localizados previamente mediante algoritmos de optimización apropiados. Así pues, el intervalo de energía de una banda espectral vendrá determinado por la energía de absorción (*vertical absorption energy*, E_{VA}) calculada a la

geometría optimizada del estado fundamental, la energía de emisión (*vertical emission energy*, E_{VE}) calculada a la geometría optimizada del estado excitado y las diferencias de energía de los dos estados electrónicos implicados, calculadas a la geometría optimizada de sus mínimos (T_e).

Con la finalidad de proporcionar una descripción más precisa de los espectros hay que tener en cuenta que toda molécula tiene un umbral de energía vibracional, denominado punto cero, que es el estado de mínima energía en el que se puede encontrar dicha molécula. Esta Energía Vibracional de Punto Cero (*Zero-Point Vibrational Energy*, ZPVE) se suma a la energía electrónica en los distintos estados. Por razones prácticas, su evaluación se lleva a cabo dentro de la aproximación vibracional armónica,⁴⁸ que en la mayoría de los casos resulta razonable para evaluar las contribuciones energéticas de punto cero de la vibración fundamental ($v = 0$) del sistema. La adición de la energía ZPVE al origen de banda electrónico, T_e , da lugar a la magnitud adiabática con corrección vibracional, T_0 , conocida también como origen de banda 0-0. Estrictamente, sólo esta magnitud tiene correspondencia experimental con la que compararse, mientras que la vertical se relaciona típicamente con el máximo de banda, dentro del marco de la aproximación Franck-Condon (FC). Esta aproximación es válida únicamente cuando las geometrías de los estados electrónicos inicial y final son muy similares,^{1,48} pues en ella se supone que la transición más probable es aquella que tiene lugar sin cambios relativos en las posiciones de los núcleos de la molécula. Esto sucede porque la transición electrónica, debida a la absorción de energía, es muy rápida en comparación con el movimiento nuclear y, por tanto, mientras ésta se produce, los núcleos permanecen prácticamente congelados en la posición de equilibrio de su estado fundamental.

La interpretación teórica del espectro molecular requiere que, además del cálculo de las energías, se determine la probabilidad con la que ocurre la transferencia de energía entre los estados implicados.⁴⁸ En los procesos de absorción y emisión no sólo entran en juego los estados inicial y final, sino que también participa el campo electromagnético. El momento de transición, M_{if} , magnitud que depende del momento dipolar eléctrico, μ , ofrece información sobre la polarización de la transición causada por la radiación. La fuerza de oscilador, f , nos informa sobre la intensidad de las

bandas observadas en los espectros experimentales. La fuerza de oscilador de las transiciones electrónicas verticales, y suponiendo que la transición obedece el principio de FC, se puede estimar por la diferencia de energía entre el estado final e inicial y el cuadrado del momento de transición dipolar entre estos estados:

$$f = \frac{2}{3} |M_{if}|^2 \Delta E_{fi}$$

donde ΔE_{fi} corresponde a la diferencia de energía entre el estado final e inicial y M_{if} es el momento de transición dipolar entre estos estados, definido como

$$M_{if} = \langle \Psi_i | \hat{\mu} | \Psi_f \rangle \quad (1.6)$$

siendo $\hat{\mu}$ el momento dipolar eléctrico. El cálculo[♥] de M_{if} se realiza de forma eficiente mediante el método de Interacción de Estados CAS (*CAS States Interaction*, CASSI).^{53,54}

La intensidad de la banda espectral correspondiente a la transición espectroscópica entre los estados i y f es proporcional a la fuerza de oscilador entre estos estados y, por tanto, al cuadrado del elemento de la matriz de transición $\langle \Psi_i | \hat{\mu} | \Psi_f \rangle^2$ (Regla de Oro de Fermi).⁵⁵ Las reglas de selección para las transiciones se basan en encontrar las condiciones en las cuales estos elementos matriciales son distintos de cero. Las transiciones estarán permitidas o prohibidas en función de las características de simetría y espín de los estados i y f . Sólo las integrales M_{if} que pertenezcan a la representación irreducible totalmente simétrica podrán ser distintas de cero. En el caso del espín, el operador momento dipolar eléctrico, $\hat{\mu}$, no opera sobre éste y, por tanto, el resultado de la integración sobre las variables de espín dará siempre cero si las funciones de espín de los dos estados i y f son diferentes, siendo sólo permitidas aquellas transiciones entre estados de la misma multiplicidad ($\Delta S = 0$). El que una transición sea prohibida no

[♥] Tanto ΔE_{fi} como M_{if} se calculan en unidades atómicas (*atomic units*, au)

supone que no pueda observarse experimentalmente en ciertas condiciones. Así, las absorciones singlete-triplete se registran en los espectros, aunque con mucha menor intensidad que las absorciones singlete-singlete, y son bien conocidos los fenómenos de emisión fosforescente y el cruce entre sistemas que tienen lugar entre estados de distinta multiplicidad. Esto se debe a que las reglas de selección se basan en una serie de aproximaciones, bien sea en la resolución de la ecuación de Schrödinger o en el tratamiento de la interacción radiación-materia. Por ejemplo, el Acoplamiento Espín-Órbita (*Spin-Orbit Coupling*, SOC)⁵⁶ no se tiene en cuenta en el Hamiltoniano de la ecuación de Schrödinger no relativista usado en la definición de las reglas de selección. El SOC corresponde a la interacción entre el momento angular de espín (\vec{S}) y el momento angular orbital (\vec{L}), y posibilita la mezcla de estados de distinta multiplicidad de espín permitiendo la interacción entre ellos. La longitud de su vector para dos estados i y f viene dada por

$$SOC_{if} = \sqrt{\sum_u \left| \langle T_{i,u} | \hat{H}_{SO} | S_f \rangle \right|^2} / u = x, y, z \quad (1.7)$$

donde \hat{H}_{SO} es el Hamiltoniano espín-órbita, donde SOC_{if} indica la probabilidad de transición entre los estados de multiplicidad distinta.

Por otro lado, cabe señalar que el origen de banda puede relacionarse con el tiempo de radiación, τ_{rad} , a través de la aproximación de Strickler-Berg,^{57,58} que define el coeficiente de Einstein de emisión espontánea, A_{21} , como:

$$A_{21} = \frac{1}{\tau_{rad}} = 2.142005 \cdot 10^{10} T_e^3 M_{if}^2 \quad (1.8)$$

siendo T_e el origen de banda y M_{if} el momento de transición dipolar entre estos estados definida en la ecuación 1.6.

Cruces de igual simetría espacial y de espín

Los puntos de cruce que tienen lugar entre estados de igual simetría espacial y de espín se conocen como puntos o hiperlíneas de Intersección Cónica (*Conical Intersection*, CI)^{50,59,60,61,62} y corresponden a la situación más favorable para que tenga lugar el proceso de conversión interna por el cual la energía se transfiere entre los dos estados implicados, ya que la probabilidad de transferencia va a ser mayor cuanto menor sea la diferencia energética, o *gap*, entre los estados.

Una CI tiene lugar cuando al representar la energía respecto a las F coordenadas nucleares se obtiene una intersección a lo largo de la hiperlínea de dimensión $F-2$, siendo $F = 3M-6$ los grados de libertad del sistema formado por M átomos. En la descripción topológica de las CIs, el espacio total de coordenadas (de dimensión F) se divide en: i) el subespacio de intersección, siendo su dimensión $F-2$, en el que los dos estados tienen la misma energía, y ii) subespacio ramificado (*branching subspace*), de dos dimensiones, en el que la degeneración desaparece. Estas dos dimensiones son:

- Vector diferencia de gradientes

$$x_1 = \frac{\partial(E_1 - E_2)}{\partial Q} \quad (1.9)$$

- Vector de acoplamiento no-adiabático

$$x_2 = \left\langle \Psi_1 \left| \frac{\partial \Psi_2}{\partial Q} \right. \right\rangle \quad (1.10)$$

donde Q representa las coordenadas nucleares del sistema.

Como puede verse en la figura 1.3, la representación de la energía de los dos estados respecto a los vectores x_1 y x_2 da lugar a una estructura de doble cono (o diábolo) en las PEHs correspondientes. Tanto x_1 como x_2 son combinaciones lineales de distancias y ángulos de enlace, siendo x_1 la dirección a lo largo de la cual las pendientes de las PEHs superior e inferior

son más pronunciadas y x_2 la dirección a lo largo de la cual el desplazamiento nuclear, que mezcla las funciones adiabáticas en el punto de intersección cónica, es el óptimo. La figura 1.3 también muestra que el espacio de intersección es una hiperlínea consistente en un número infinito de puntos de intersección cónica, denominada costura o *seam* de la intersección cónica. Cada uno de estos puntos cumple las condiciones de intersección cónica, sin embargo, tienen distinta energía, corresponden a geometrías diferentes del sistema considerado, y sus vectores x_1 y x_2 asociados no tienen por qué referirse a los mismos desplazamientos nucleares. Desde un punto de vista estático, se suele buscar la estructura y el emplazamiento energético del punto de más baja energía, aunque químicamente la geometría más interesante es la más accesible para la energía. Teniendo en cuenta que en un punto de la CI ni la superficie superior ni la inferior presentan gradiente cero, como lo sería para un verdadero punto estacionario, se requieren algoritmos especiales para optimizar la estructura de la CI. La mayoría de ellos se basan en seguir la proyección del gradiente en el subespacio de intersección $F-2$ ortogonal a x_1 y x_2 hasta que se anule.^{48,63}

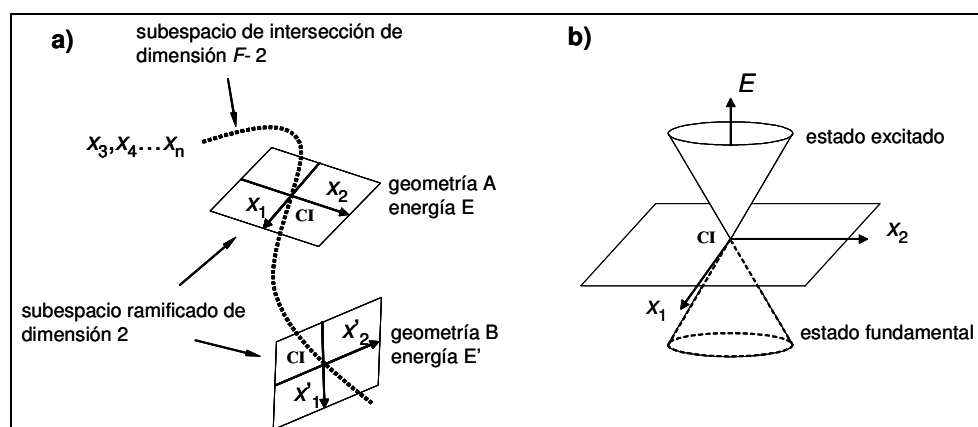


Figura 1.3 a) Espacio total de coordenadas y sus correspondientes subespacios relevantes para el cómputo de intersecciones cónicas. b) Representación de la energía de un sistema para dos estados (fundamental y excitado) respecto a x_1 (vector diferencia de gradientes) y x_2 (vector de acoplamiento no adiabático).

En la última década han sido muchos los estudios que han revelado el papel crucial que juegan las CIs en la fotofísica y fotoquímica no-adiabática de numerosos compuestos orgánicos.⁴⁸ En el caso de las reacciones fotoquímicas, la desactivación no radiativa por conversión interna, mediada a través de una CI, puede presentar distintas posibilidades de evolución en la PEH inferior. El tipo de topología que presente la CI va a ser la que controle la obtención de los productos fotoquímicos y el tiempo (o velocidad) de la reacción.

Cruces de igual simetría espacial y distinta multiplicidad

Las situaciones más comunes de cruces entre estados de la misma simetría espacial, pero distinta multiplicidad de espín, suceden entre estados singlete y triplete (*Singlet-Triplet Crossing*, STC).⁵⁰ A diferencia de las CIs, la máxima probabilidad de transición energética entre ambos estados tiene lugar no sólo en el punto de cruce sino en una región amplia de las PEHs, ya que en este caso hay que sopesar dos factores: el menor *gap* energético, ΔE_{fi} , y el máximo acoplamiento espín-órbita, SOC_{if} . Como se ha descrito anteriormente, este acoplamiento mezcla los estados de distinta multiplicidad de espín permitiendo la interacción entre ellos, cosa que de otro modo no sería posible si se tiene en cuenta la regla de selección de espín por dipolo eléctrico. La región de cruce con máxima relación $SOC_{if}/\Delta E_{fi}$ será la de máxima probabilidad de transición entre los dos estados.

Caminos de mínima energía

Hasta ahora se han descrito los puntos singulares (mínimos, SP, CI, STC) y la situación energética relativa de dichos puntos, como la topología en las PEHs de interés para el químico teórico, proporcionando una visión estática de la reactividad química. Sin embargo, se tiene que resaltar en este punto, que el número de grados de libertad de un sistema molecular crece rápidamente con el aumento de su complejidad. En consecuencia, las zonas

singulares de las correspondientes PEHs, de dimensiones muy grandes incluso para moléculas de modesto tamaño, van a ser muchas. La tarea de determinar teóricamente las regiones características de las PEHs, junto con la evaluación de los puntos necesarios para proporcionar una resolución suficiente se puede reducir si se localizan caminos de reacción dentro de las hipersuperficies. La evaluación de los puntos puede ser restringida entonces a aquellos que pertenezcan al dominio de dichos caminos. De entre las múltiples vías reactivas que pueden existir, cabe destacar el Camino de Mínima Energía (*Minimum Energy Path*, MEP) que sigue la dirección negativa del gradiente de la energía. En la Fotofísica y Fotoquímica Teóricas, el MEP resulta de gran interés para conocer el camino de mínima energía seguido por un sistema molecular tras la absorción de radiación, desde la disposición inicial Franck-Condon, hacia el mínimo del estado excitado. El procedimiento que se suele utilizar para la determinación del MEP que une dos puntos de una PEH, y el que se va a emplear en los estudios presentados en esta Tesis, corresponde al camino de máximo descenso (*steepest descendent path*),⁶⁴ el cual está basado en una modificación del algoritmo de optimización restringida de Anglada y Bofill⁶⁵ y sigue la aproximación de Müller-Brown.⁶⁶ En esta estrategia computacional cada paso requiere la minimización de las hipersuperficies de energía potencial, dentro de la sección de cruce de la hiperesfera en exploración, centrada en la geometría inicial y caracterizada por un radio predefinido (véase figura 1.4).

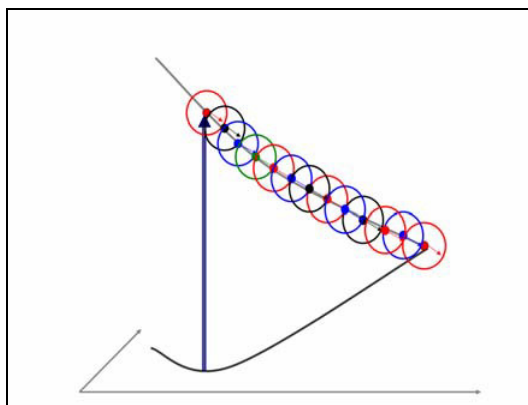


Figura 1.4. Representación esquemática del proceso de cálculo de un MEP.

La estructura optimizada se toma como centro de la nueva hiperesfera del mismo radio y el procedimiento se itera hasta que se alcanza el mínimo energético de la superficie. Se utilizan las coordenadas ponderadas en la masa, de forma equivalente al algoritmo de la Coordenada Intrínseca de Reacción (*Intrinsic Reaction Coordinate*, IRC).

Determinación estructural: el problema de la correlación diferencial

El principio variacional nos asegura que empleando un conjunto de funciones de base de mayor tamaño los resultados derivados de cálculos variacionales llevan a una energía más baja y, por tanto, más cercana al valor exacto.² El uso de una base más extensa no sólo repercute en la energía, sino que también permite definir una función de onda más flexible, proporcionando con ello una mejor descripción de la molécula. Por otra parte, tal como se ha visto anteriormente, existen métodos que no tienen en cuenta la correlación electrónica, unos que incluyen la correlación estática y otros la dinámica. Desafortunadamente, el aumento del conjunto de base electrónica y el empleo de métodos que traten con eficacia la correlación electrónica implica un aumento de coste computacional. Como consecuencia, muchas veces resulta necesario aplicar una mayor calidad metodológica en el cálculo de alguna de las propiedades, pero no en otras. Por ejemplo, es una práctica común el empleo de metodología de nivel medio para localizar las geometrías de los puntos singulares (mínimos, estados de transición, intersecciones cónicas...) y corregir después las energías mediante el empleo de métodos de alto nivel. A lo largo de la presente Tesis se va a seguir en repetidas ocasiones esta estrategia y, en concreto, el protocolo CASPT2//CASSCF, en el cual se tiene en cuenta la correlación estática a través del método CASSCF para optimizar la geometría del sistema y, a esa geometría, la energía es calculada aplicando un método de perturbaciones de segundo orden, CASPT2, a la función de onda multiconfiguracional CASSCF. La experiencia acumulada en el campo de la espectroscopia y fotoquímica teóricas haciendo uso de la estrategia CASPT2//CASSCF es amplia y se han resuelto problemas de gran

complejidad mediante su empleo.^{39,40,41,67} Sin embargo, hay que tener en cuenta que implícito en este protocolo y, en general, en la estrategia de usar distintos niveles metodológicos para el cálculo geométrico y energético, se encuentra el problema de los efectos de la correlación diferencial.⁴⁶ Así pues, en el protocolo CASPT2//CASSCF, los efectos de las energías de correlación estática y dinámica van a diferir dependiendo de la naturaleza del estado, afectando directamente con ello a la descripción de la hipersuperficie. La figura 1.5 resume los diferentes efectos de la correlación en la localización de cruces entre superficies con el protocolo CASPT2//CASSCF.

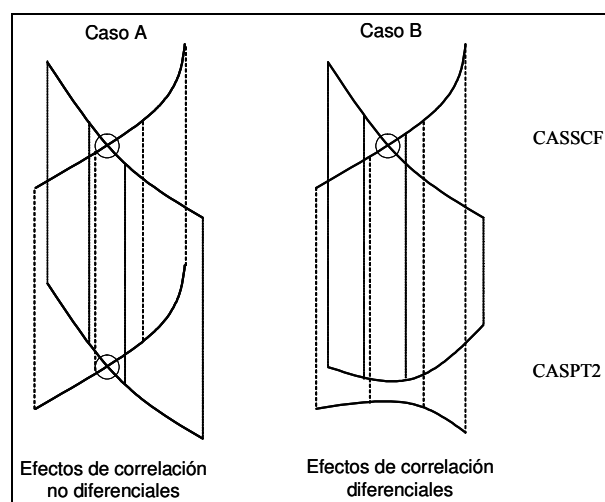


Figura 1.5. Distintos efectos de la correlación en el protocolo CASPT2//CASSCF.

En los casos donde las PEHs calculadas a nivel CASSCF y CASPT2//CASSCF se comporten aproximadamente paralelas (Caso A), las geometrías optimizadas CASSCF serán en general correctas, a pesar de haberse calculado a un nivel de teoría inferior. Esto significa que las contribuciones de la correlación dinámica son bastante regulares y similares para los dos estados en amplias regiones de las PEHs. En estos casos, la estrategia computacional en dos pasos CASPT2//CASSCF puede aplicarse con seguridad. Por otro lado, cuando la correlación dinámica es

marcadamente diferente para los estados considerados y varía significativamente a lo largo de las PEHs de interés, las optimizaciones de geometría deben llevarse a cabo al nivel más alto correlacionado. De otro modo, las contribuciones desiguales de la correlación dinámica pueden conducir a interacciones espurias, es decir, sin sentido físico, entre los dos estados electrónicos, pues el cruce puede estar situado en otra geometría o incluso no existir.

Error por superposición del conjunto de base

El Error por Superposición del Conjunto de Base (*Basis Set Superposition Error*, BSSE)⁶⁸ aparece debido al empleo de conjuntos de base finitos y, aunque siempre está presente, resulta crucial cuando se comparan moléculas con algunos de sus fragmentos o sistemas moleculares con sus entidades individuales. En el caso de dos moléculas interaccionando, se podría esperar que la energía de interacción o de enlace (*binding energy*, E_b) se correspondiese con la diferencia entre la energía de la supermolécula “dímero” y la energía de los monómeros a distancia infinita. Sin embargo, la energía calculada de esta manera sobreestima el valor real, debido a que en el dímero las funciones de base de un monómero contribuyen a la mejor descripción de la distribución electrónica del monómero adyacente, produciendo un efecto de pseudoextensión de base en el dímero.

Aunque existen varias técnicas para corregir este error, la forma más usual es utilizar el método *counterpoise* de Boys y Bernardi,⁶⁹ desarrollado para dimerizaciones atómicas o de monómeros rígidos. Esta técnica se basa en calcular la energía de los monómeros por separado con sólo sus funciones de base y luego restarle la energía de ellos mismos, también de forma individual, pero ahora en presencia de las funciones de base de todo el dímero. Esto permite cuantificar la estabilización extra en el dímero y, por tanto, corregir el error mediante la sustracción de estas cantidades a la energía total del dímero.

Para un dímero $A \cdots B$ la energía de enlace E_b no corregida por el BSSE viene dada por la ecuación

$$E_b(AB) = E_A + E_B - E_{AB}, \quad (1.11)$$

donde E_A y E_B son las energías de los fragmentos aislados A y B , respectivamente, a sus geometrías de equilibrio y E_{AB} representa la energía total del dímero AB a una geometría dada. La energía de interacción corregida mediante la técnica *counterpoise* ($CP-E_b$) viene determinada por la expresión

$$CP-E_b(AB) = E_b(AB) - [CP-BSSE(AB)], \quad (1.12)$$

donde

$$CP-BSSE(AB) = E_A(B, R=\infty) - E_A(B, R=AB) + E_B(A, R=\infty) - E_B(B, R=AB). \quad (1.13)$$

En $CP-BSSE(AB)$ la geometría de los monómeros se mantiene igual a la del dímero. Entonces, la notación $E_A(B, R=AB)$ indica que la energía de A ha sido calculada con los orbitales fantasma o *ghost* de B a la geometría de AB , mientras que en $E_A(B, R=\infty)$ los MOs fantasma de B se encuentran a una distancia infinita de A .

La inclusión del error por superposición del conjunto de base es crucial para llevar a cabo una descripción precisa de las energías de enlace en el caso del estudio de dimerizaciones.

CASPT2 y sus modificaciones *level-shift* e *IPEA-shift* al Hamiltoniano de orden cero

El uso de la teoría de perturbaciones para el tratamiento de la correlación electrónica en átomos y moléculas ha quedado bien establecido en la actualidad. Para sistemas en los cuales se requiere una correcta evaluación de la correlación tanto dinámica como estática, la teoría de perturbaciones a segundo orden sobre una función de referencia multiconfiguracional, CASPT2, resulta de gran precisión y productividad. El método CASPT2

está basado en una función de onda de referencia CASSCF, que considera todos los efectos de cuasidegeneración en la estructura electrónica (correlación estática o no dinámica), presentes en muchas situaciones, como por ejemplo, en estados excitados, roturas de enlace, estructuras de estados de transición, etc. Los efectos restantes de correlación dinámica son añadidos a través del método de perturbaciones, calculando la función de onda y la energía a primer y segundo orden, respectivamente, y empleando la función CASSCF como función de referencia de orden cero.

La aproximación CASPT2 ha sido empleada con éxito en multitud de problemas químicos: estructuras moleculares, espectros electrónicos, metales de transición y otros compuestos de elementos pesados, etc.^{39,40,41,67}

Sin embargo, en ciertas situaciones como son la presencia de cuasidegeneraciones accidentales o en procesos donde el número de electrones apareados cambia, la formulación estándar de la matriz de Fock en el Hamiltoniano de orden cero no resulta lo suficientemente precisa. En estos casos, se han llevado a cabo ciertas modificaciones en la formulación original, introduciendo parámetros de desplazamiento, como son el *level-shift* e *IPEA-shift*. Veamos a continuación con un poco de detalle en qué consiste cada una de estas modificaciones y en qué situaciones hay que tenerlas en cuenta.

Level-shift CASPT2

La aplicación del tratamiento de perturbaciones a la función de onda multiconfiguracional CASSCF que tiene lugar en la aproximación CASPT2, resulta en una función corregida a primer orden, la cual, tras normalizarse, puede expresarse como

$$|\Psi\rangle = c_0|o\rangle + c_1|1\rangle, \quad (1.14)$$

donde $|o\rangle$ es la función de orden cero, generada a partir del cálculo CASSCF y $|1\rangle$ es la función de primer orden normalizada y $c_0^2 + c_1^2 = 1$.⁷⁰

La magnitud c_0^2 corresponde al peso de la función de referencia en la

función $|\Psi\rangle$ y se utiliza como un criterio de calidad rápido y sencillo del tratamiento de perturbaciones realizado. Idealmente, el peso de la referencia debería ser cercano a la unidad para que la expansión de perturbación pueda converger rápidamente. No obstante, su valor numérico depende del número de electrones correlacionados, de tal modo que, a mayor tamaño del sistema molecular tratado, menor será el peso de referencia. En la práctica, con la finalidad de que todos los estados electrónicos considerados estén descritos con un mismo nivel de correlación, y así sean comparativos, se suele comprobar que todos ellos tengan un peso similar, utilizando el mismo espacio activo. Existen ciertas situaciones en las cuales el coeficiente c_1 presenta un valor proporcionalmente alto, lo que conlleva a un peso de referencia relativamente bajo. Se pueden distinguir dos situaciones distintas como causantes de este hecho. En primer lugar, cuando hay un grupo reducido de estados que poseen valores altos del coeficiente c_1 (estados intrusos) y que interaccionan con la referencia de forma pronunciada, con contribuciones a la energía de segundo orden de ≥ 0.1 au. Esto sucede cuando el espacio activo es deficiente y, por tanto, la estrategia a seguir para solventar el problema es reformular el tratamiento de perturbaciones, considerando como activos los orbitales implicados en los estados intrusos. En otras ocasiones, lo que sucede es que el peso de referencia de un estado excitado es relativamente bajo, en comparación con el correspondiente al estado fundamental, sin que se pueda responsabilizar de ello a unos pocos estados concretos de la función de primer orden, la cual suele presentar un gran número de contribuciones de escasa importancia. En este caso, la causa del valor bajo del peso de referencia es debido a cuasidegeneraciones accidentales, generalmente causadas por orbitales difusos en bases con gran flexibilidad, como los ANOs, implicando a un grupo grande de estados intrusos que interactúan débilmente con la referencia. Es aquí donde la técnica *Level-Shift* (LS) queda especialmente aconsejada para constatar la validez del tratamiento de perturbaciones realizado.

El método *level-shift* CASPT2 (LS-CASPT2)^{71,72} elimina de forma efectiva los estados intrusos mediante la adición de un parámetro de desplazamiento al Hamiltoniano de orden cero y posterior corrección de su efecto a la energía de segundo orden. El desplazamiento aplicado eliminará los estados intrusos, pero la energía de segundo orden dependerá del parámetro

introducido, el cual resulta ambiguo ya que no puede escogerse en base a fundamentos físicos. Esta ambigüedad se resuelve mediante una transformación de retroceso aproximada de la energía de segundo orden para obtener su valor sin desplazamiento, pero libre ahora de estados intrusos. En principio, es recomendable realizar varios cálculos empleando distintos valores del parámetro de desplazamiento (0.0, 0.1, 0.2, 0.3, 0.4 au) y comprobar la adecuada dependencia de los resultados LS-CASPT2 respecto a los valores del parámetro *level-shift*, ya que si sólo se realiza un cálculo, accidentalmente podrían aparecer singularidades, precisamente para ese valor del parámetro de desplazamiento, y el resultado LS-CASPT2 no sería fiable. Con la finalidad de solucionar este problema y entonces eliminar por completo las singularidades accidentales, Forsberg y col.⁷³ desarrollaron en 1997 la técnica del *level-shift* Imaginario (*Imaginary*, IMAG). Ésta va a ser la opción escogida en la presente Tesis en los cálculos que requieran de la técnica *level-shift*, dando lugar a resultados CASPT2(IMAG).

IPEA-shift CASPT2

La formulación estándar de la matriz de Fock en el Hamiltoniano de orden cero produce normalmente una estimación por debajo, de 2–5 kcal/mol por cada enlace formado.⁷⁴ Esto es debido a un error sistemático en el método. En cada proceso en el cual tiene lugar un cambio en el número de electrones apareados, tendrá lugar un error de este tamaño por cada par de electrones. La causa de este error se encuentra en el hecho de que el Hamiltoniano de orden cero en su formulación original proporciona una descripción descompensada para las estructuras electrónicas de capa cerrada en comparación con las estructuras con electrones desapareados. Ghigo y col.⁷⁵ presentaron en 2004 una modificación del Hamiltoniano de orden cero que corrige en gran medida este error sistemático y da lugar a mejoras considerables en las energías de disociación y de excitación así como en potenciales de ionización y afinidades electrónicas.

En esta técnica se introduce un desplazamiento o *shift* en el Hamiltoniano de orden cero, específico para cada tipo de funciones doblemente excitadas

empleadas en la construcción de la función de onda de primer orden. Según Ghigo y col.,⁷⁵ la idea básica reside en hacer que las energías de los orbitales activos desde los cuales se excita sean más de tipo energías de ionización, mientras que las energías de los orbitales activos a los cuales se excita sean más parecidas a las afinidades electrónicas. El parámetro de desplazamiento empleado tiene que determinarse a través de la comparación con cálculos precisos o resultados experimentales. Tras la realización de gran cantidad de cálculos de distintas propiedades y en distintos sistemas y su comparación con datos precisos, los autores sugieren un valor para el *IPEA-shift* de 0.25 au.

Software

El programa elegido para realizar todos los cálculos computacionales de la presente Tesis Doctoral ha sido MOLCAS,⁷⁶ paquete de software consistente en un conjunto de módulos diseñados principalmente para el tratamiento cuántico de sistemas bajo una aproximación multiconfiguracional, y que es especialmente adecuado para el estudio de estados excitados.



<http://www.molcas.org/>

Referencias

1. Merchán, M.; Serrano-Andrés, L. Ab Initio Methods for Excited States. In *Computational Photochemistry*; Olivucci, M., Ed.; Elsevier: Amsterdam, 2005; Vol. 16, Chap. 2, pp 35.
2. Szabo, A.; Ostlund, N. S. *Modern Quantum Chemistry. Introduction to Advanced Electronic Structure Theory*; Dover Publications Inc.: Mineola, New York, 1996.
3. Helgaker, T.; Jørgensen, P.; Olsen, J. *Molecular Electronic-Structure Theory*; John Wiley & Sons: Chichester, 2004.
4. Widmark, P.-O.; Roos B. O. *European Summerschool of Quantum Chemistry 2005. Books I, II, III*; Lund University: Lund, 2005.
5. Jensen, F. *Introduction to Computational Chemistry*; John Wiley & Sons: Chichester, 1999.
6. Hartree, D. R. *Proc. Cambridge Philos. Soc.* **1928**, *24*, 328.
7. Fock, V. A. *Z. Phys.* **1930**, *15*, 126.
8. Roothaan, C. C. J. *Rev. Mod. Phys.* **1951**, *23*, 69.
9. Pople, J. A.; Nesbet, R. K. *J. Chem. Phys.* **1954**, *22*, 571.
10. Roothaan, C. C. J. *Rev. Mod. Phys.* **1960**, *32*, 179.
11. Schlegel, H. B. *J. Phys. Chem.* **1988**, *92*, 3075.
12. Knowles, P. J.; Handy, N. C. *J. Phys. Chem.* **1988**, *92*, 3097.
13. Meckler, A. *J. Chem. Phys.* **1953**, *21*, 1750.
14. Roos, B. O.; Taylor, P. R.; Siegbahn, P. E. M. *Chem. Phys.* **1980**, *48*, 157.
15. Roos, B. O. *Int. J. Quantum Chem. Symp.* **1980**, *14*, 175.
16. Roos, B. O. The Complete Active Space Self-Consistent Field Method and its Applications in Electronic Structure Calculations. In *Advances in Chemical Physics; Ab Initio Methods in Quantum Chemistry-II*; Lawley, K. P., Ed.; John Wiley & Sons Ltd.: New York, 1987; Vol 69, pp 399-445.
17. Huron, B.; Rancurel, P.; Malrieu, J. P. *J. Chem. Phys.* **1973**, *58*, 5745.
18. Buenker, R. J.; Peyerimhoff, S. D. *Theor. Chim. Acta* **1974**, *35*, 33.
19. Møller C.; Plesset, M. S. *Phys. Rev.* **1934**, *46*, 618.
20. Andersson, K.; Malmqvist, P.-Å.; Roos, B. O.; Sadlej, A. J.; Wolinski, K. *J. Phys. Chem.* **1990**, *94*, 5483.
21. Andersson, K.; Malmqvist, P.-Å.; Roos, B. O. *J. Chem. Phys.* **1992**, *96*, 1218.
22. Čížek, J. *J. Chem. Phys.* **1966**, *45*, 4256.
23. Čížek, J. *Adv. Chem. Phys.* **1969**, *14*, 35.
24. Čížek, J.; Paldus, J. *Int. J. Quantum Chem.* **1971**, *5*, 359.

25. Bartlett, R. J. *J. Phys. Chem.* **1989**, *93*, 1697.
26. Lee, T. J.; Taylor, P. R. *Int. J. Quantum Chem. Symp.* **1989**, *23*, 199.
27. Lee, T. J.; Scuseria, G. E. In *Quantum Mechanical Electronic Structure Calculations with Chemical Accuracy*; Langhoff, S. R., Ed.; Kluwer Academic Publishers: Dordrecht, 1995; 47.
28. Daudey, J.-P.; Heully, J.-L.; Malrieu, J.-P. *J. Chem. Phys.* **1993**, *99*, 1240.
29. Slater, J. C. *Phys. Rev.* **1930**, *36*, 57.
30. Boys, S. F. *Proc. R. Soc. (London) A* **1950**, *200*, 542.
31. Heher, W. J.; Radom, L.; Schleyer, P. v. R.; Pople, J. A. *Ab Initio Molecular Orbital Theory*; John Wiley & Sons: New York, 1986.
32. Hehre, W. J.; Stewart, R. F.; Pople, J. A. *J. Chem. Phys.* **1969**, *51*, 2657.
33. Heher, W. J.; Ditchfield, R.; Pople, J. A. *J. Chem. Phys.* **1972**, *56*, 2257.
34. Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. *J. Chem. Phys.* **1980**, *72*, 650.
35. Dunning Jr, T. H. *J. Chem. Phys.* **1989**, *90*, 1007.
36. Wilson, A. K.; Van Mourik, T.; Dunning Jr, T. H. *J. Mol. Struct.* **1996**, *388*, 339.
37. Kendall, R. A., Dunning Jr, T. H.; Harrison, R. J. *J. Chem. Phys.* **1992**, *96*, 6796.
38. Almlöf, J.; Taylor, P. R. *J. Chem. Phys.* **1978**, *86*, 4070.
39. Merchán, M.; Serrano-Andrés; Fülcher, M. P.; Roos, B. O. Multiconfigurational Perturbation Theory Applied to Excited States of Organic Compounds In *Recent Advances in Multireference Methods*; Hirao, K., Ed.; World Scientific Publishing: Singapore, 1999, pp 161-195.
40. Roos, B. O.; Fülcher, M. P.; Malmqvist, P.-Å.; Merchán, M.; Serrano-Andrés, L. Theoretical Studies of Electronic Spectra of Organic Molecules. In *Quantum Mechanical Electronic Structure Calculations with Chemical Accuracy*; Langhoff, S. R., Ed.; Kluwer Academic Publishers: Dordrecht, 1995; pp 357-438.
41. Roos, B. O.; Fülcher, M. P.; Malmqvist, P.-Å.; Serrano-Andrés, L.; Pierloot, K.; Merchán, M. In *Advances in Chemical Physics: New Methods in Computational Quantum Mechanics*; Prigogine, I.; Rice, S. A.; John Wiley & Sons: New York, 1996; pp. 219.
42. Widmark, P.-O.; Malmqvist, P.-Å.; Roos, B. O. *Theor. Chim. Acta* **1990**, *77*, 291.
43. Widmark, P.-O.; Persson, B. J.; Roos, B. O. *Theor. Chim. Acta* **1991**, *79*, 291.
44. Pou-Américo, R.; Merchán, M.; Nebot-Gil, I.; Widmark, P.-O.; Roos, B. O. *Theor. Chim. Acta* **1995**, *92*, 149.

45. Pierloot, K.; Dumez, B.; Widmark, P.-O.; Roos, B. O. *Theor. Chim. Acta* **1995**, *90*, 87.
46. Serrano-Andrés, L.; Merchán, M. *J. Mol. Struct. THEOCHEM* **2005**, *729*, 99.
47. Turro, N. J. *Modern Molecular Photochemistry*; University Science Books: Sausalito, 1991.
48. Serrano-Andrés, L.; Merchán, M. Spectroscopy: Applications. In *Encyclopedia of Computational Chemistry*; Schlegel, P. V. R., Schreiner, P. R., Schaefer III, H. F., Jorgensen, W. L., Thiel, W., Glen, R. C., Eds.; Wiley: Chichester, 2004.
49. Robb, M. A.; Olivucci, M.; Bernardi, F. Photochemistry. In *Encyclopedia of Computational Chemistry*; Schlegel, P. V. R.; Schreiner, P. R.; Schaefer III, H. F.; Jorgensen, W. L.; Thiel, W.; Glen, R. C., Eds.; Wiley: Chichester, 1998.
50. Klessinger, M.; Michl, J. *Excited States and Photochemistry of Organic Molecules*; VCH Publishers: New York, 1995.
51. Peyerimhoff, S. Spectroscopy: Computational Methods. In *Encyclopedia of Computational Chemistry*; Schlegel, P. V. R.; Schreiner, P. R.; Schaefer III, H. F.; Jorgensen, W. L.; Thiel, W.; Glen, R. C., Eds.; Wiley: Chichester, 1998; Vol. 4, pp 2646 - 2664.
52. Serrano-Andrés, L.; Merchán, M. *Anales de Química* **2004**, *100*, 16.
53. Malmqvist, P.-Å. *Int. J. Quantum Chem.* **1986**, *30*, 479.
54. Malmqvist, P. -Å.; Roos, B. O. *Chem. Phys. Lett.* **1989**, *155*, 189.
55. Rodríguez, A.R.; Román, J. Z. *Espectroscopia*; Pearson Prentice Hall: Madrid, 2004.
56. Pilar, F. *Elementary Quantum Chemistry*; Dover Publications, Inc.: 2001.
57. Strickler, S. J.; Berg, R. A. *J. Chem. Phys.* **1962**, *37*, 814.
58. McHale, J. L. *Molecular Spectroscopy*; Prentice-Hall: Upper Saddle River, New Jersey, 1999.
59. Teller, E. *J. Phys. Chem.* **1937**, *41*, 109.
60. Klessinger, M. *Angew. Chem. Int. Ed. Engl.* **1995**, *34*, 549.
61. Bernardi, F.; Olivucci, M.; Robb, M. A. *Chem. Soc. Rev.* **1996**, *25*, 321.
62. Domcke, W.; Yarkony, D. R.; Köppel, H. *Conical Intersections: Electronic Structure, Dynamics and Spectroscopy*; World Scientific: Singapore, 2004.
63. Bearpark, M. J.; Robb, M. A.; Schlegel, H. B. *Chem. Phys. Lett.* **1994**, *223*, 269.
64. De Vico, L.; Olivucci, M.; Lindh, R. *J. Chem. Theory Comput.* **2005**, *1*, 1029.
65. Anglada, J. M.; Bofill, J. M. *J. Comput. Chem.* **1997**, *18*, 992.

66. Müller, K.; Brown, L. D. *Theor. Chim. Acta* **1979**, *53*, 75.
67. Merchán, M.; Serrano-Andrés, L.; González-Luque, R.; Roos, B. O.; Rubio, M. *J. Mol. Struct. THEOCHEM* **1998**, *463*, 201.
68. Davidson, E. R.; Chakravory, S. J. *Chem. Phys. Lett.* **1994**, *217*, 49.
69. Boys, S. F.; Bernardi, F. *Molecular Physics* **1970**, *19*, 553.
70. Andrés, J.; Beltrán, J. *Química Teórica y Computacional*; Universitat Jaume I: Castelló, 2000.
71. Roos, B. O.; Andersson, K. *Chem. Phys. Lett.* **1995**, *245*, 215.
72. Roos, B. O.; Andersson, K.; Fülcher, M. P.; Serrano-Andrés, L.; Pierloot, K.; Merchán, M.; Molina, V. *J. Mol. Struct. THEOCHEM* **1996**, *388*, 257.
73. Forsberg, N.; Malmqvist, P.-Å. *Chem. Phys. Lett.* **1997**, *274*, 196.
74. Andersson, K.; Roos, B. O. *Int. J. Quantum Chem.* **1993**, *45*, 591.
75. Ghigo, G.; Roos, B. O.; Malmqvist, P.-Å. *Chem. Phys. Lett.* **2004**, *396*, 142.
76. Andersson, K.; Barysz, M.; Bernhardsson, A.; Blomberg, M. R. A.; Carissan, Y.; Cooper, D. L.; Cossi, M.; Fülcher, M. P.; Gagliardi, L.; de Graaf, C.; Hess, B.; Hagberg, G.; Karlström, G.; Lindh, R.; Malmqvist, P.-Å.; Nakajima, T.; Neogrády, P.; Olsen, J.; Raab, J.; Roos, B. O.; Ryde, U.; Schimmelpfennig, B.; Schütz, M.; Seijo, L.; Serrano-Andrés, L.; Siegbahn, P. E. M.; Stålring, J.; Thorsteinsson, T.; Veryazov, V.; Widmark, P. –O. *MOLCAS, version 6.0-7.0*; Department of Theoretical Chemistry, Chemical Centre, University of Lund, P.O.B. 124, S-221 00 Lund: Sweden, 2004.

PARTE I

Luz visible y fotocromismo



... Y Dios dijo: "Hágase la luz". Y la luz se hizo.

Entonces dijo: "Que la tierra produzca vegetales, hierbas que den semilla y árboles frutales, que den sobre la tierra frutos de su misma especie con su semilla adentro". Y así sucedió.

Génesis 1 Capítulo 1 Antiguo Testamento de la Biblia.

2 Introducción y objetivos

Receptores de luz y fotomorfogénesis

Las plantas son bombardeadas por un sin número de señales procedentes no sólo por su entorno físico sino también de los animales, tanto amigos como enemigos. Como consecuencia de ello, han desarrollado un sistema muy sofisticado de receptores y vías de transducción de señales que generan las respuestas adecuadas en cada situación. Que la luz juega un importante papel en el desarrollo de la planta no es de extrañar, pues la capacidad de una planta para maximizar su productividad fotosintética depende de su capacidad para detectar, evaluar y responder a la calidad de la luz, así como a su intensidad y dirección. De este modo, el calendario de etapas del desarrollo de las plantas, como pueden ser la floración o la entrada en latencia, depende de un sistema de medición y respuesta a los cambios en la duración del día. Los fitocromos (*phytocrome*, phy) son una gran familia de fotorreceptores muy heterogéneos que median en una amplia gama de respuestas fotomorfogénicas al absorber luz (600-800 nm) en el rojo y en el rojo lejano.^{1,2} Por otro lado, existen dos clases distintas de fotorreceptores reguladores de los efectos de la luz UV-A y azul (320-500 nm): los criptocromos (*cryptochrome*, cry) y las fototropinas (*phototropin*, phot).³ Actualmente, se ha identificado un tercer tipo de receptor de luz azul, el fitocromo 3 (phy3), relacionado con fitocromos y fototropinas, y que incluye a proteínas que median en la proteólisis de los componentes asociados con el ciclo circadiano y la floración.⁴ En la figura 2.1 puede verse la representación esquemática de los cuatro tipos de fotorreceptores conocidos a día de hoy.⁴

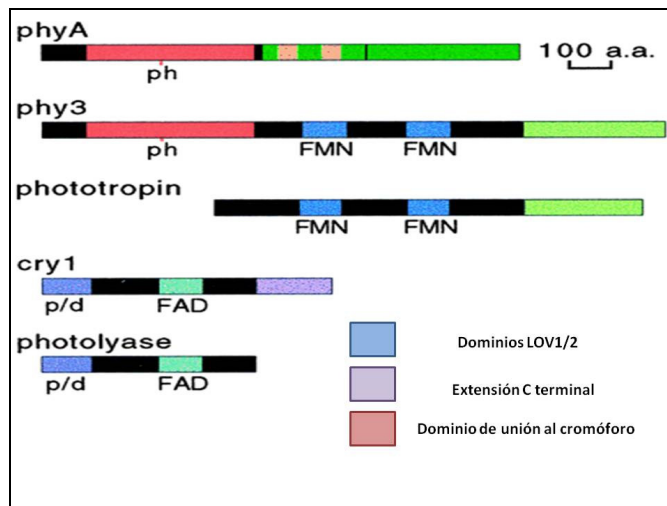


Figura 2.1. Los fotorreceptores, conocidos a día de hoy, responsables de la fotomorfogénesis en las plantas son:³ cinco fitocromos, un supercromo (o phy3), dos fototropinas y dos criptocromos. La fotoliasa, que aparece en la figura, es una enzima implicada en la reparación de dímeros de pirimidina y que no se clasifica como fotorreceptor propiamente dicho.

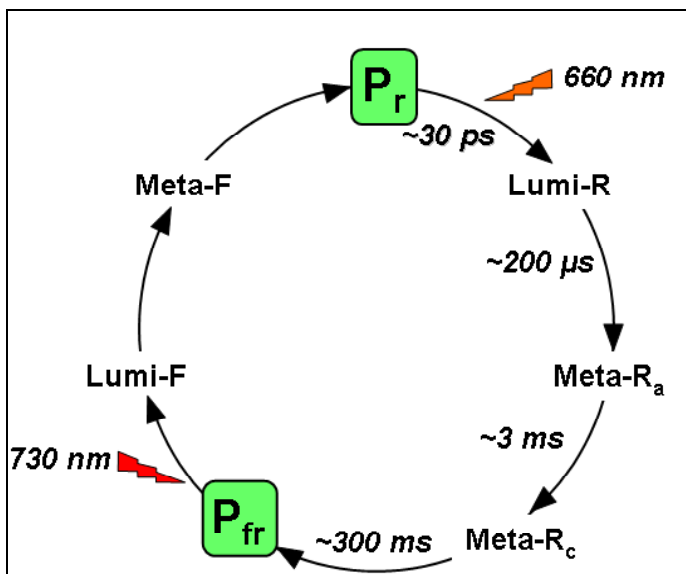
El progreso en la comprensión de los mecanismos básicos de respuesta de las plantas a la luz roja y roja lejana, fue espectacular tras el aislamiento y caracterización inicial de un fitocromo. Por el contrario, el progreso en la comprensión de los acontecimientos desencadenados por la luz azul se vio obstaculizado por la dificultad de identificar al/los cromóforo(s) y/o receptore(s) de luz azul. Las plantas contienen innumerables flavoproteínas y carotenoproteínas, que complicaban la identificación de los cromóforos que podrían funcionar como receptores de luz azul, mientras que los investigadores que estudiaban a los fitocromos no tenían una variedad tan desconcertante de candidatos. Por ello, en 1975 todavía no se había identificado ningún receptor de luz azul en las plantas y existía, en esa época, una considerable controversia en cuanto a cómo podría ser el cromóforo de un receptor de este tipo de luz. Algo menos polémica fue la idea (ahora sabemos que errónea) de que probablemente se trataba de un solo receptor y Gressel,⁵ quien lo acuñó con el término de criptocromo, ya

advirtió en contra de esta interpretación simplista que, sin embargo, persistió durante años.

Fitocromos

Fueron descubiertos por Borthwick et al.⁶ en el marco de una investigación sobre el efecto de la luz en la germinación de las semillas de lechuga. Estos investigadores observaron que las semillas que germinaban en oscuridad no llegaban al 20 %, mientras que, el porcentaje de germinación era máximo cuando irradiaban las semillas con un pulso de luz roja (R). Se halló, además, que la irradiación subsiguiente con un pulso de luz roja lejana (RL) anulaba el efecto inductor de la luz roja, evitando la germinación. Las irradiaciones alternas con luz R y RL (R, R + RL, R + RL + R, R + RL + R + RL, etc.) pusieron de manifiesto que el último color aplicado determinaba la germinación de las semillas, y que la luz roja constituía el factor estimulante del proceso, mientras que el factor inhibidor era la luz roja lejana. En busca de una explicación para tales fenómenos se propuso la existencia de un pigmento, al que llamaron fitocromo, que absorbía la luz roja. El fitocromo en cuestión, tras absorber luz roja, se convertía en una forma capaz de absorber radiación roja lejana, forma que tornaba a su condición inicial tras realizar dicha absorción. A principios de los años sesenta, esta hipótesis halló respaldo experimental con la purificación, a partir de extractos de plántulas de cereales, de una proteína soluble dotada de las características predichas. Los fitocromos, que también pueden encontrarse en algunas bacterias y hongos, están formados por una parte proteica y un sistema tetrapirrólico lineal, o cromóforo, que es el que realmente absorbe la luz a determinadas longitudes de onda. La energía de la luz absorbida por los fitocromos inicia un cambio estructural en su molécula que desencadena toda una cascada de señales transcripcionales; señales que regulan funciones tan importantes como el crecimiento de las plantas, la germinación de las semillas, o el movimiento y pigmentación en bacterias. Todo ello promovido por el hecho de que puedan intercambiarse sus dos formas, de ahí que se les denomine “foto-convertibles”, P_r y P_{fr} , tal como se muestra en el esquema 2.1,⁷ donde P_r absorbe luz roja y P_{fr} se encarga de la radiación roja lejana. Puesto que en los procesos fisiológicos la luz roja tiene un

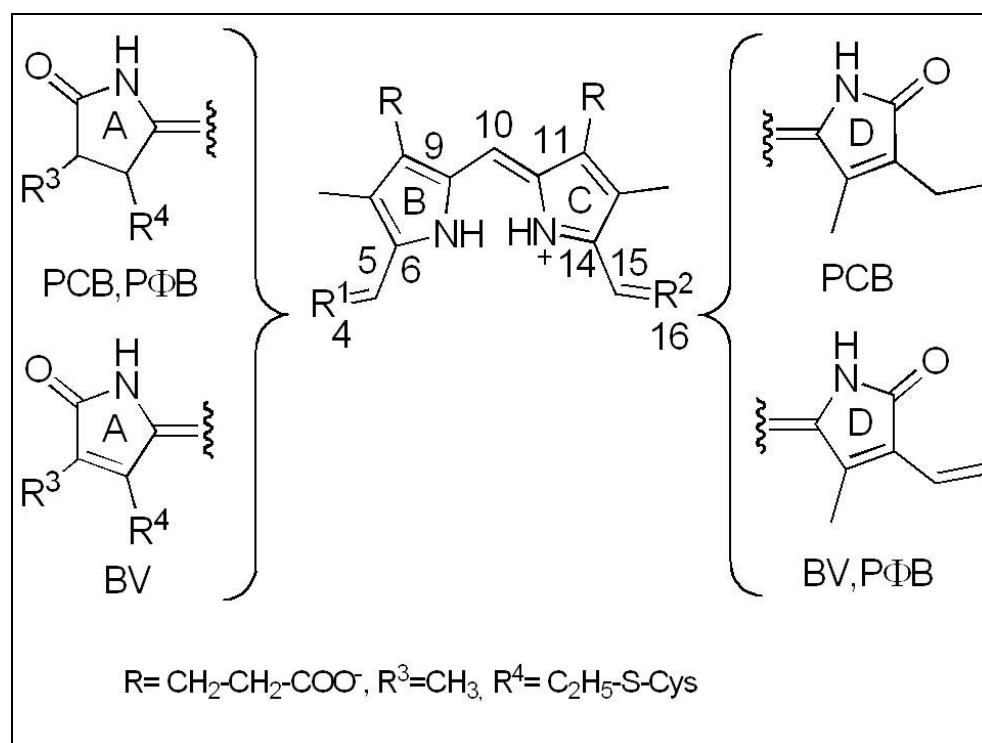
efecto activador y la roja lejana un efecto inhibitor, P_{fr} se considera la forma biológicamente activa del fitocromo y P_r la inactiva.



Esquema 2.1. Fotociclo de los fitocromos.

En el ciclo de activación-desactivación de la molécula, aparecen compuestos intermedios, Lumi-R, Lumi-F, Meta-R y Meta-F, con diferentes tiempos de excitación-desexcitación.

Se han encontrado tres cromóforos en los cinco fitocromos: la fitocromobilina (*phytochromobilin*, PΦB), la biliverdina (*biliverdin*, BV) y la ficocianobilina (*phycocyanobilin*, PCB), que son muy similares, pues comparten el sistema tetrapirrólico lineal conjugado, pero presentan diferencias en el primer y último anillo. En el esquema 2.2 se incluye la descripción de los tres sistemas dependiendo de los sustituyentes, R, R¹, R², R³ y R⁴.



Esquema 2.2. Cromóforos de los fitocromos: fitocromobilina, PΦB, biliverdina, BV, y ficocianobilina, PCB, formados por un sistema tetrapirrólico cuyos anillos se nombran como A, B, C y D. La unión del cromóforo a la parte proteica del fitocromo se realiza a través de un enlace tioéter entre su anillo A y un aminoácido cisteína de la proteína, representado en el esquema por R⁴.

Fototropinas

En 1988, Gallagher, Briggs y colaboradores⁸ informaron, por primera vez, de que la luz azul podría activar la fosforilación de una proteína asociada a la membrana plasmática de plántulas etioladas (decoloradas, o de crecimiento en oscuridad). Después de una amplia caracterización bioquímica, genética y fisiológica, se hizo patente que esta proteína no sólo era el fotorreceptor responsable del fototropismo, sino que también era la enzima de su propia fosforilación. Identificada originalmente a partir de un mutante (*non-phototropic hypocotyl 1*, *nph1*) de la *Arabidopsis thaliana*, sus

análisis genéticos concluyeron que contenía dos fototropinas, phot1 y phot2.^{9,10} Ambas se solapan parcialmente para regular el fototropismo del hipocotiledón en respuesta a la luz azul muy intensa,¹¹ mientras que con la luz azul débil sólo actúa el phot1.^{11,12,13} Además, phot1 y phot2 regulan la apertura de los estomas de las hojas, respuesta que permite la absorción de CO₂ en la fotosíntesis y controla la pérdida de agua por transpiración.¹⁴ Las fototropinas también controlan el movimiento de los cloroplastos en respuesta a distintas intensidades lumínicas.¹¹ Así, por ejemplo, bajo condiciones de alta intensidad los cloroplastos se alejan del lugar de la irradiación para prevenir que se dañe el aparato fotosintético con el exceso de luz.¹⁵ Además, están implicadas en la expansión tanto del cotiledón¹⁶ como de las hojas en Arabidopsis.¹⁷ Las fototropinas, pues, posibilitan que se desarrollen las respuestas fotosintéticas, y promueven el crecimiento de las plantas, en condiciones de pobreza lumínica.¹⁸

La estructura proteica de las fototropinas (véase figura 2.1), contiene dos dominios muy similares, con aproximadamente 110 aminoácidos, denominados LOV¹⁹ (porque regulan la respuesta a la luz, al oxígeno y/o al voltaje). Para formar la proteína activa, u holoproteína, cada uno de los dominios, LOV1 y LOV2, se une covalentemente a una molécula de flavin mononucleótido, FMN, que es el que actúa como cromóforo del receptor. En la figura 2.2 se presenta un modelo de la estructura terciaria de la proteína incorporando el cromóforo de FMN en su interior.

Las fototropinas pertenecen al grupo de las flavinas, grupo que engloba a todas aquellas sustancias que contienen en su molécula el anillo de isoaloxacina (véase figura 4.1). Estas flavinas se encuentran involucradas en la cadena de transporte de electrones mitocondrial, que relaciona la oxidación de los alimentos con la producción de adenosín trifosfato, ATP. La transferencia de dos electrones se puede realizar de forma secuencial o de forma simultánea y, por tanto, en cada uno de los estados redox (completamente oxidado, reducido en un electrón y completamente reducido) los compuestos flavínicos pueden existir en forma catiónica, neutra y aniónica, dependiendo del pH de la disolución. Como curiosidad, los humanos, y animales superiores, son incapaces de sintetizar el anillo de isoaloxacina de las flavinas y la fuente de estas sustancias en la dieta es la vitamina B₂, que se encuentra en la leche, la cerveza y en vegetales de hojas verdes.

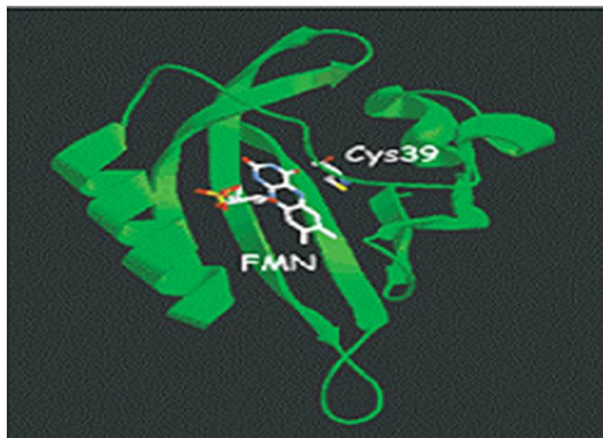
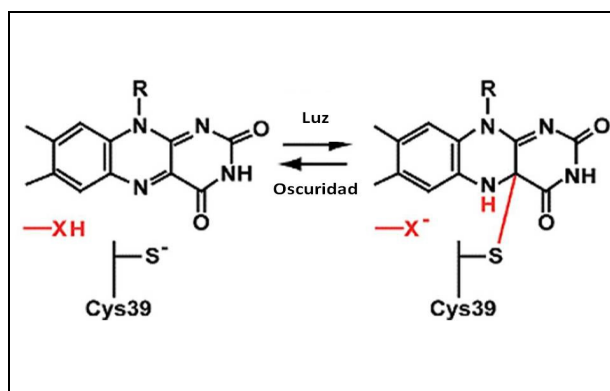


Figura 2.2. Estructura de la fototropina, fotorreceptor de luz azul, en la que puede observarse en su interior el cromóforo flavín mononucleotido, FMN. El modelo es el resultado de una modelización computacional a partir de parámetros obtenidos por resonancia magnética nuclear.²⁰

El responsable de la fotoquímica de las flavinas es el anillo de isoaloxacina del FMN completamente oxidado. Después de la absorción de la luz, se forma un aducto, entre la cisteína-39 de la proteína y el FMN, aducto que se descompone, en una escala de tiempo de minutos, en la oscuridad posterior^{21,22,23} (véase esquema 2.3).



Esquema 2.3. Molécula de FMN al someterse a un fotociclo luz-oscuridad.

Criptocromos

No fue hasta 1993 cuando Ahmad y Cashmore²⁴ descubrieron el criptocromo 1, cry1, en plantas de *Arabidopsis* y, posteriormente, el grupo de Cashmore identificó el criptocromo 2, cry2, que contiene una extensión C-terminal diferente de la del cry1. Hay evidencia de que los criptocromos están localizados en el núcleo pero, hasta la fecha, poco se conoce acerca de las consecuencias inmediatas de la fotoexcitación de cualquiera de los criptocromos, aunque se sabe que ambos están implicados en la inhibición del alargamiento del hipocotiledón. No se profundiza en su descripción por no ser motivo de estudio en la presente Tesis.

Objetivos

Se describen a continuación los objetivos de esta primera parte de la presente Tesis doctoral.

- ✓ Se pretende estudiar las propiedades espectroscópicas *in vacuo* del cromóforo de los fitocromos. El estudio realizado permitirá explicar de forma cualitativa la fotoquímica intrínseca de este cromóforo, ampliando los conocimientos que ya se disponen sobre él. Los resultados de esta parte de la investigación, realizados sobre un modelo ligeramente simplificado debido al gran tamaño molecular del sistema tetrapirrólico, se presentan en el capítulo 3.
- ✓ Se pretende estudiar los estados electrónicos excitados del anillo de la isoaloxacina mediante cálculos *ab initio* mecano-cuánticos CASPT2//CASSCF. Se presentará el panorama general de los espectros obtenidos (naturaleza de los estados e intensidad relativa de absorción de las principales bandas) y el mecanismo intrínseco para la población fotoinducida del estado triplete más bajo, principal responsable de la reactividad del anillo de isoaloxacina y, por tanto, de la activación del ciclo fotoquímico de las flavinas. Se presentan los resultados de esta parte de la investigación en el capítulo 4.

Referencias

1. Chen, M.; Chory, J.; Fankhauser, C. *Annu. Rev. Genet.* **2004**, *38*, 87.
2. Franklin, K. A.; Lerner, V. S.; Whitelam, G. C. *Int. J. Dev. Biol.* **2005**, *49*, 653.
3. Banerjee, R.; Batschauer, A. *Planta* **2005**, *220*, 498.
4. Briggs, W.R.; Olney, M. A. *Plant Physiology* **2001**, *125*, 85.
5. Gressel, J. *Photochemistry and Photobiology* **1979**, *30*, 749.
6. Hendricks, S. B.; Borthwick, H. A. The Physiological Functions of Phytochrome. In *Chemistry and Biochemistry of Plant Pigments*; Goodwin, T.W. Ed.; Academic Press, Inc. London Ltd: London, U.K., 1965; pp 405-436.
7. Rockwell, N. C.; Su, Y.-S.; Lagarias, J. C. *Annu. Rev. Plant Biol.* **2006**, *57*, 837
8. Gallagher, S.; Short, T. W.; Ray, P. M.; Pratt, L. H.; Briggs, W. R. *Proc. Natl. Acad. Sci. USA* **1988**, *85*, 8003.
9. Briggs, W. R.; Christie, J. M. *Trends Plant Sci.* **2002**, *7*, 204.
10. Celaya, R. B.; Liscum, E. *Photochem. Photobiol.* **2005**, *81*, 73.
11. Sakai, T.; Kagawa, T.; Kasahara, M.; Swartz, T. E.; Christie, J. M.; Briggs, W. S.; Wada, M.; Okada, K. *Proc. Natl. Acad. Sci. USA* **2001** *98*, 6969.
12. Liscum, E.; Briggs, W. R. *Plant Cell* **1995**, *7*, 473.
13. Sakai, T.; Wada, T.; Ishiguro, S.; Okada, K. *Plant Cell* **2000**, *12*, 225.
14. Kinoshita, T.; Doi, M.; Suetsugu, N.; Kagawa, T.; Wada, M.; Shimazaki, K. *Nature* ,**2001**, *414*, 656.
15. Kasahara, M.; Kagawa, T.; Oikawa, K.; Suetsugu, N.; Miyao, M.; Wada, M. *Nature* **2002**,*420*, 829.
16. Ohgishi, M.; Saji, K.; Okada, K.; Sakai, T. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 2223.
17. Sakamoto, K.; Briggs, W.R. *Plant Cell* **2002**, *14*, 1723.
18. Takemiya, A.; Inoue, S.; Doi, M.; Kinoshita, T.; Shimazaki, K. *Plant Cell* **2005**, *17*, 1120.
19. Huala, E.; Oeller, P.W.; Liscum, E.; Han, I. S.; Larsen, E.; Briggs, W. R. *Science* **1997**, *278*, 2120.
20. Harper, S. M.; Neil, L. C.; Gardner, K. H. *Science* **2003**, *301*, 1541.
21. Crosson, S.; Moffat, K. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 2995.
22. Crosson S, Moffat K. *Plant Cell* **2002**, *14*, 1067.

23. Fedorov, R.; Schlichting, I.; Hartmann, E.; Domratcheva, T.; Fuhrmann, M.; Hegemann, P. *Chlamydomonas reinhardtii*. *Biophys. J.* **2003**, *84*, 2474.
24. Ahmad, M.; Cashmore, A. R. *Nature* **1993**, *366*, 162.

3 Fitocromos: Rutas de desactivación

En el capítulo anterior se ha introducido el fotociclo de los fitocromos (véase esquema 2.1) y se ha puesto de manifiesto el hecho de que pueden intercambiarse las dos formas P_r y P_{fr} , donde P_r absorbe luz roja y P_{fr} roja lejana.¹ También se han descrito los tres cromóforos encontrados en los fitocromos, $P\Phi B$, BV y PCB , que comparten el sistema tetrapirrólico lineal conjugado, aunque con diferencias en el primer y último anillo (véase esquema 2.2). Gracias a los estudios experimentales realizados sobre estos cromóforos, es ampliamente conocida su fotofísica y fotoquímica, tanto inmersos en la proteína como en disolución.²⁻⁶ Los cambios en las propiedades espectrales P_r/P_{fr} se asocian a una fotoisomerización inicial ($Z \rightarrow E$)^{*} en el doble enlace $C_{15}=C_{16}$, que desemboca en un compuesto intermedio Lumi-R en el intervalo de unas pocas decenas de picosegundos,^{7,8} y con la isomerización posterior del enlace adyacente, $C_{14}-C_{15}$, aunque este último paso se ha puesto en duda. Algunos autores proponen una posible rotación del puente A-B,^{9,10} mientras que otros descartan cualquier isomerización que no sea la fotoquímica, es decir, la que involucra únicamente al enlace $C_{15}=C_{16}$.^{11,12} Después de la absorción de luz, el primer acontecimiento que tiene lugar se sabe que sigue la misma cinética compleja, tanto en la proteína como en disolución, con una componente rápida, alrededor de 2-3 ps, y una lenta, entre 30-40 ps, lo cual sugiere un mecanismo de desactivación similar en ambos medios. Es muy interesante resaltar que se ha constatado, a través de experimentos combinados de *pump-probe* vis/vis y vis/IR,⁷ que la componente lenta (30-40 ps) es la

* El convenio de nomenclatura de la Unión Internacional de Química Pura y Aplicada (*International Union of Pure and Applied Chemistry*, IUPAC) ha adoptado un sistema en el que a los sustituyentes de cada extremo del doble enlace se le asigna una prioridad en función de su número atómico. Si los sustituyentes de alta prioridad están en el mismo lado del doble enlace, al compuesto se le asigna la letra Z (del alemán, *zusammen*, juntos). Si están en lados opuestos del enlace, se le asigna la letra E (del alemán, *entgegen*, enfrentados). No siempre Z y cis (o bien, E y trans) son intercambiables.

fotorreactiva y conduce a la fotoisomerización del doble enlace $C_{15}=C_{16}$ (y formación de Lumi-R) con un rendimiento cuántico relativamente bajo ($\sim 9\%$), que encaja con el encontrado para la formación del P_{fr} ($\sim 8\%$).¹³ En estos experimentos se sugiere⁷ que la isomerización ocurre en un intervalo del orden del tiempo de vida media del estado excitado S_1 , para el que se ha detectado un tiempo de decaimiento incluso menor (0.7 ps), sin descartar un proceso con múltiples pasos, que se iniciaría en el estado excitado S_1 para acabar en S_0 . Que la fotoisomerización $Z \rightarrow E$, del enlace $C_{15}=C_{16}$, suceda en la hipersuperficie del estado excitado también se sustenta con estudios espectroscópicos Raman recientes.¹⁴ Estos resultados sugieren la formación inicial (en ~ 0.6 ps) de un estado excitado intermedio (I^*), seguido de la formación (en ~ 3 ps) de un estado excitado tardío, llamado fotoproducto, similar a Lumi-R (Lumi-R*), donde la isomerización $C_{15}=C_{16}$ estaría prácticamente finalizada.

Aunque se han realizado multitud de estudios experimentales, los primeros pasos en los que se produce el almacenamiento de la energía de los fotones, e iniciación del fotociclo, todavía no han sido aclarados. La fotoquímica computacional^{15,16} podría proporcionar las herramientas adecuadas para analizar los detalles que llevarían a comprender los mecanismos moleculares del proceso de fotoisomerización pero, debido a la gran complejidad del sistema, se han realizado pocos estudios computacionales.¹⁷⁻¹⁹

Con la finalidad de dilucidar tales procesos, hemos realizado los perfiles energéticos de los caminos de relajación del estado excitado, *in vacuo*, con cálculos *ab initio* CASPT2. Dada la complejidad de la molécula, fue necesario utilizar modelos simplificados de la misma. El modelo más simple consistió en la molécula de maleimida (1H-pyrrole-2,5-dione),[▼] que representa uno de los anillos del sistema tetrapirrólico del fitocromo. El segundo modelo del cromóforo consistió en el catión de cuatro anillos simplificado del PΦB, en el que R , R^3 y R^4 se sustituyeron por H (véase esquema 2.2). En el presente capítulo se presentan los resultados obtenidos con dicho modelo y publicados recientemente.²¹

▼ El estudio de la molécula de la maleimida, además de formar parte del entrenamiento propio que supone abordar un sistema nuevo, fue el trabajo de investigación desarrollado para la obtención del Diploma de Estudios Avanzados, DEA.²⁰

Detalles computacionales

Selección del modelo molecular PΦB

A pesar de que en los últimos años ha aumentado de forma exponencial la capacidad y velocidad de las máquinas de cálculo, y el número de ellas disponibles para cada usuario, a día de hoy aún es imposible tratar, computacionalmente, un sistema tan grande como el que nos ocupa. Es, por tanto, que en el modelo simplificado del cromóforo PΦB los grupos carboxi, de los anillos B y C, y los metilos y el tioéter, del anillo A, se han reemplazado por átomos de hidrógeno. Se ha optado por una geometría correspondiente a la conformación *ZZZasa*,[♦] tal y como ha sido propuesta en base a varios experimentos espectroscópicos,²² aunque, debemos decir que la conformación inicial del cromóforo en la proteína está todavía bajo discusión y que se proponen estructuras *ZZZasa* y *ZZZssa*, ambas bien argumentadas con experimentos de espectroscopia vibracional.^{1,4} Es incluso probable que el cromóforo adopte una geometría cíclica *ZZZsss*, tipo porfirina, cuando se encuentra libre en disolución.^{1,23} Conviene aclarar que los resultados presentados en esta Tesis pretenden dar una visión general del proceso de fotoisomerización, que no se ve especialmente afectado por la estabilidad relativa de los diferentes conformeros en el estado fundamental, más que estudiar en sí el giro de dicho enlace.

Selección y validación de los métodos computacionales

Para poder seleccionar y validar los métodos computacionales adecuados en cada caso con los que realizar las optimizaciones de las distintas geometrías y el cálculo puntual de las energías de los distintos estados, se ha utilizado toda una batería de métodos computacionales, tales como HF, MP2, CIS, CASSCF y CASPT2, y se han comparado los resultados obtenidos entre ellos.

[♦] La isomería *anti*, *a*, y *syn*, *s*, es una isomería similar a la isomería geométrica cis-trans, pero en vez de doble enlace C=C, tenemos otros dobles enlaces diferentes, como C=N; N=N o N=S. Se denomina *syn* al isómero cuyos grupos de mayor prioridad están juntos, y *anti* en el caso en que dichos grupos de mayor prioridad están opuestos. Fuente: Estereoquímica. Pedro Antonio García Ruiz. EDITUM. Secretariado de publicaciones, Universidad de Murcia, 1991.

Tabla 3.1. Distancias de enlaces, d (Å) y ángulos diedros, a (grados), para las estructuras optimizadas del estado fundamental y del primer estado excitado. La numeración de los átomos se ha realizado en concordancia con el esquema 2.2.

Parámetros Geométricos	Método					
	HF	MP2 (FC)	CASSCF(8,8) (Lewis)	CASSCF(8,8) (Lewis 2)	CASSCF(12,11) (Lewis 2)	CIS S ₁ (PLA)
$d(C_4-C_5)$	1.341	1.362	1.355	1.329	1.345	1.355
$d(C_5-C_6)$	1.454	1.438	1.433	1.473	1.473	1.431
$d(C_9-C_{10})$	1.380	1.394	1.342	1.451	1.441	1.394
$d(C_{10}-C_{11})$	1.390	1.397	1.441	1.352	1.354	1.394
$d(C_{14}-C_{15})$	1.458	1.435	1.474	1.431	1.442	1.427
$d(C_{15}-C_{16})$	1.339	1.368	1.345	1.356	1.355	1.366
$a(C_3-C_4-C_5-C_6)$	179	177	179	179	179	177
$a(C_4-C_5-C_6-C_7)$	-46	-45	-33	-60	-62	-35
$a(C_8-C_9-C_{10}-C_{11})$	170	165	178	145	148	168
$a(C_9-C_{10}-C_{11}-C_{12})$	168	163	150	179	177	167
$a(C_{13}-C_{14}-C_{15}-C_{16})$	-48	-44	-66	-32	-40	-28
$a(C_{14}-C_{15}-C_{16}-C_{17})$	178	174	179	176	179	174

Las optimizaciones de las geometrías, del estado fundamental y del primer estado excitado se realizaron con métodos MP2 y CIS, respectivamente, usando Gaussian03.²⁴ Posteriormente, para conseguir perfiles energéticos más precisos, todas las energías fueron recalculadas incluyendo efectos de correlación electrónica mediante cálculos CASPT2,²⁵ empleando MOLCAS 6.0.²⁶ A lo largo de todo el trabajo se utilizó un conjunto de bases 6-31G(d), siempre teniendo presente que un conjunto de bases mayor hubiera dado resultados más precisos pero que, sin embargo, es prácticamente imposible de emplear en un modelo tan grande como éste.

Optimizaciones de la geometría del estado fundamental

A la complicación de elegir el espacio activo, cuando se emplea el método CASSCF, se une la necesidad de reducción drástica que supone el no poder incluir el sistema π completo (28 electrones en 25 orbitales π) en esta molécula. Inicialmente se eligió un espacio activo π de 8 orbitales/8 electrones, CASSCF(8,8), para realizar las optimizaciones de las geometrías. Esto supone utilizar los correspondientes orbitales topológicamente equivalentes a los tres últimos orbitales π ocupados HF y los tres primeros orbitales π^* virtuales HF, que están situados sobre los anillos B y C, y los dos orbitales π/π^* que describen el doble enlace $C_{15}=C_{16}$, es decir, el biológicamente activo implicado en la isomerización. Desafortunadamente, la elección no estaba perfectamente equilibrada debido a que el espacio activo no está uniforme/simétricamente distribuido sobre el anillo tetrapirrólico; esto podría llevar a que, durante la optimización CASSCF de la geometría, algunas zonas estuvieran mejor descritas que otras y, por tanto, diera lugar a asimetrías no realistas de las estructuras optimizadas. Para evitar este problema, el espacio activo π se aumentó a 11 MOs/12 e^- , CASSCF(12,11), que está uniformemente distribuido en el sistema tetrapirrólico, con cuatro electrones en cuatro orbitales π/π^* , describiendo los dobles enlaces $C_4=C_5$ y $C_{15}=C_{16}$, más ocho electrones en siete orbitales π/π^* de la parte central (anilloB– C_{10} H–anilloC) del sistema P Φ B. Cabe resaltar que, mientras que con los métodos HF y MP2 se obtiene una única estructura estable, en forma de híbrido resonante

deslocalizado, con la carga positiva repartida igualmente por los anillos B y C, los cálculos CASSCF(8,8) y CASSCF(12,11) conducen a dos mínimos estables correspondientes a dos estructuras resonantes de Lewis, en las que la carga positiva se localiza: en una, en el anillo B, y en la otra, en el anillo C.

En la tabla 3.1 se compilan los parámetros geométricos del estado fundamental para las distintas estructuras obtenidas con diferentes métodos: HF, MP2 y CASSCF, así como las del estado excitado, con metodología CIS. A partir de ahora, a la estructura optimizada en el estado fundamental a nivel MP2, mínimo S_0 , la llamaremos **FC**, mientras que, a la estructura del estado excitado, mínimo S_1 , la denominaremos **PLA**, por tratarse de una estructura cuasi-plana.

Tabla 3.2. Energías absolutas CASPT2, en au, para las tres estructuras optimizadas (Lewis 1 y Lewis 2 CASSCF(8,8) y MP2) del estado fundamental.

Método	OPT CAS (Lewis 1)	OPT CAS (Lewis 2)	OPT MP2 (FC)
S_0 CASSCF	-1290.774659	-1290.767086	-1290.765780
S_0 CASPT2	-1294.698331	-1294.693540	-1294.722054

Para validar la geometría con la que empezar a trabajar, se realizaron cálculos CASPT2 sobre las tres estructuras obtenidas, MP2 y CASSCF. En la tabla 3.2 se detallan los valores de la energía absoluta que, básicamente, confirman que la estructura híbrida-resonante, obtenida tras la optimización MP2, tiene un mínimo de energía de, al menos, 15 kcal/mol por debajo del mínimo de las dos estructuras localizadas CASSCF (véase figura 3.1), confirmando que el mínimo en el estado fundamental es la estructura deslocalizada.

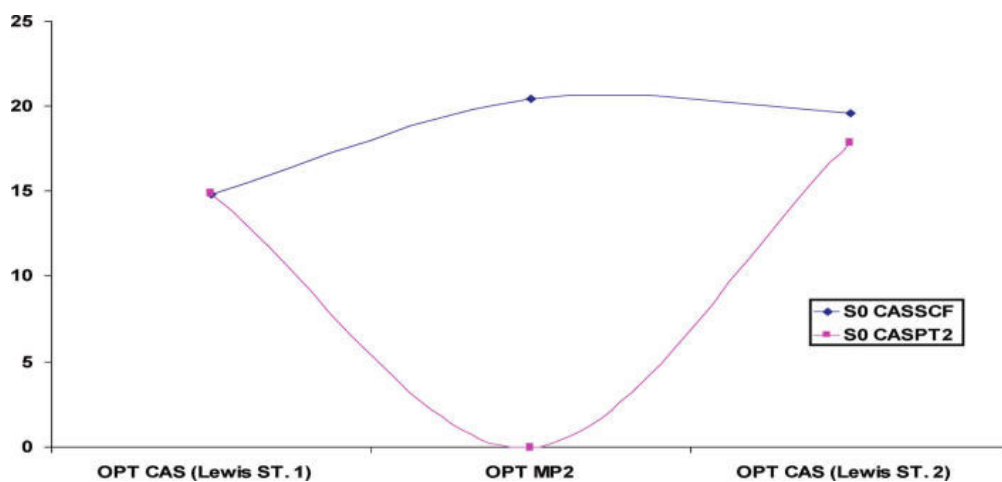


Figura 3.1. Perfil de energías (kcal/mol) CASSCF y CASPT2 conectando los mínimos del estado fundamental calculados a nivel CASSCF y MP2.

Optimización de la geometría del estado excitado

En base a los resultados obtenidos en la optimización de la geometría del estado fundamental, se pensó que, quizás, en este sistema concreto, el uso del método CASSCF no fuera adecuado para realizar la optimización de la geometría del estado excitado. Desafortunadamente, son pocos los métodos capaces de realizar dichas optimizaciones a un nivel altamente correlacionado²⁷ y, por supuesto, serían técnicamente impracticables en un sistema del tamaño molecular del cromóforo que estamos tratando, especialmente para localizar intersecciones cónicas. Se sabe que, en el caso de las bases del ADN/ARN, las optimizaciones con el método CIS desembocan directamente en una intersección cónica entre los estados S_0/S_1 de las bases, hecho confirmado usando metodologías de más alto nivel.²⁸ Así pues, inspirados en este hecho, y dado que el carácter del estado S_1 es una excitación casi pura HOMO \rightarrow LUMO, se seleccionó la aproximación CIS para la determinación preliminar de la geometría del estado excitado más bajo del modelo P Φ B. Recordemos que la descripción más simple de un estado excitado en el marco del modelo orbital simple, supone que un electrón se ha promocionado desde un orbital ocupado a uno desocupado.

Por tanto, la optimización de la geometría de los estados electrónicamente excitados descritos fundamentalmente por configuraciones simplemente excitadas, se pueden realizar cualitativamente mediante la aproximación CIS. La elección realizada se respaldó con calibraciones CASPT2, análogas a las realizadas para el estado fundamental, en las distintas estructuras optimizadas del modelo PΦB, mostrando que la geometría obtenida a nivel CIS, se comporta, en este caso, mejor que las optimizadas a nivel CASSCF, dando una energía CASPT2 más baja. En resumen, las diferentes estructuras de la hipersuperficie S_1 se han caracterizado a nivel CIS, como una estrategia computacional práctica y siempre después de un cuidadoso calibrado.

Como se muestra en la figura 3.2, el mínimo del estado fundamental en la región FC (que está ligeramente girado para minimizar las interacciones estéricas debido a los sustituyentes en los anillos) aparece como un híbrido resonante estable, cuya estructura es casi simétrica respecto al carbono central C_{10} . Estas propiedades se conservan en la estructura del estado excitado, PLA, donde ocurre una ligera restructuración del esqueleto de la molécula.

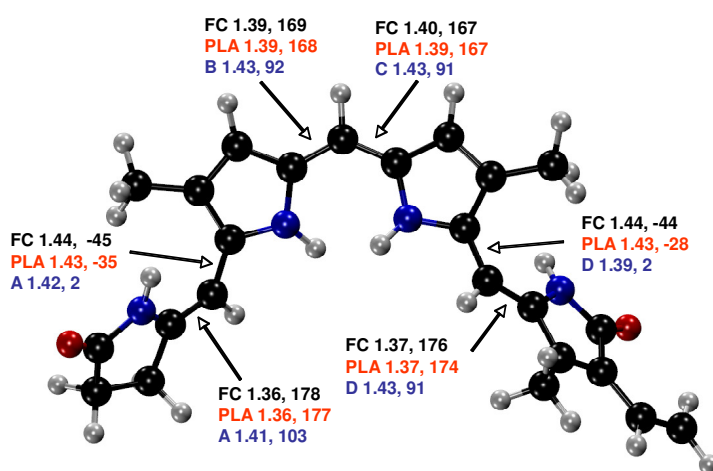


Figura 3.2. Parámetros geométricos relevantes (longitudes de enlace en Å y ángulos de giro en grados) calculados para los mínimos del estado fundamental (FC) y para el mínimo de los estados excitados (PLA, A, B, C, D).

La estructura PLA aparece como un estado transitorio localizado en una región muy plana de la hipersuperficie de energía potencial de S_1 , y en nuestro estudio presentamos cuatro posibles canales de isomerización para salir de él, cada uno de ellos involucrando el giro de uno de los enlaces exocíclicos siguientes: C_4-C_5 (canal A), C_9-C_{10} (canal B), $C_{10}-C_{11}$ (canal C) y $C_{15}-C_{16}$ (canal D).

Estimación de las energías del estado fundamental y estados excitados

Una vez finalizadas las optimizaciones de las geometrías del estado fundamental y del estado excitado, se realizaron cálculos puntuales CASPT2 para contabilizar la energía de correlación y poder cuantificar, de forma precisa, las diferencias energéticas entre los distintos estados. La elección del espacio activo para los cálculos CASPT2 se realizó analizando la convergencia de las propiedades espectroscópicas verticales al ir aumentando sistemáticamente el número de orbitales y de electrones activos (véase tabla 3.3). Un espacio activo con 10 electrones en 11 orbitales π/π^* , promediando 3 raíces, desembocó en resultados CASPT2//CASSCF(10,11) que pueden considerarse, en la práctica, convergidos dentro del error del método. En los cálculos CASPT2 se empleó el método estándar del Hamiltoniano de orden cero, donde se incluyó una corrección de 0.2 au en el *imaginary level-shift* para evitar la presencia de estados intrusos en el cálculo. El método CASSI²⁹ se empleó de la forma que está incorporado en MOLCAS-6.0²⁶ para calcular los momentos dipolares, que a su vez se emplearon, junto con las energías corregidas CASPT2, para computar la fuerza de oscilador, f . Los resultados revelan, de forma clara, que la excitación singlete HOMO \rightarrow LUMO, correspondiente al estado S_1 , es el estado poblado inicialmente, o estado brillante. En la figura 3.3 puede verse la descripción de los orbitales HOMO y LUMO, ambos deslocalizados en los dos anillos centrales (B y C) y en los dos enlaces exocíclicos (C_4-C_5 y $C_{15}-C_{16}$), sin afectar a los anillos externos (A y D).

Tabla 3.3. Energías totales (au) a nivel CASSCF y CASPT2, energías de excitación (eV) y fuerzas de oscilador correspondientes, f , del estado fundamental y los estados excitados más bajos energéticamente del modelo del cromóforo tetrapirrólico del PΦB, calculadas a la geometría del estado fundamental, optimizada a nivel MP2. En todos los casos se empleó el conjunto de base 6-31G(d).

Estado	CASSCF(au)	CASSCF(eV)	CASPT2(au)	CASPT2(eV)	f
8 MOs/8e	-1290.765071		-1294.726870		
S ₁	-1290.650008	3.13	-1294.662748	1.74	0.845
S ₂	-1290.614579	4.10	-1294.611532	3.14	0.039
10 MOs/8e	-1290.777459		-1294.725156		
S ₁	-1290.670506	2.91	-1294.658263	1.82	0.952
S ₂	-1290.635828	3.86	-1294.611759	3.08	0.009
12 MOs/8e	-1290.781276		-1294.724271		
S ₁	-1290.675183	2.89	-1294.658039	1.80	0.981
S ₂	-1290.641318	3.81	-1294.612075	3.05	0.012
10 MOs/10e	-1290.788850		-1294.724071		
S ₁	-1290.686892	2.77	-1294.648970	2.04	0.819
S ₂	-1290.644129	3.94	-1294.602483	3.31	0.054
11 MOs/10e	-1290.802833		-1294.723283		
S ₁	-1290.703329	2.70	-1294.652361	1.93	0.968
S ₂	-1290.665041	3.75	-1294.607431	3.15	0.026
12 MOs/10e	-1290.804178		-1294.722787		
S ₁	-1290.705417	2.67	-1294.652623	1.91	0.986
S ₂	-1290.667238	3.73	-1294.607659	3.13	0.028
12 MOs/12e	-1290.811828		-1294.720333		
S ₁	-1290.716310	2.60	-1294.652366	1.85	0.915
S ₂	-1290.675034	3.72	-1294.602042	3.21	0.050

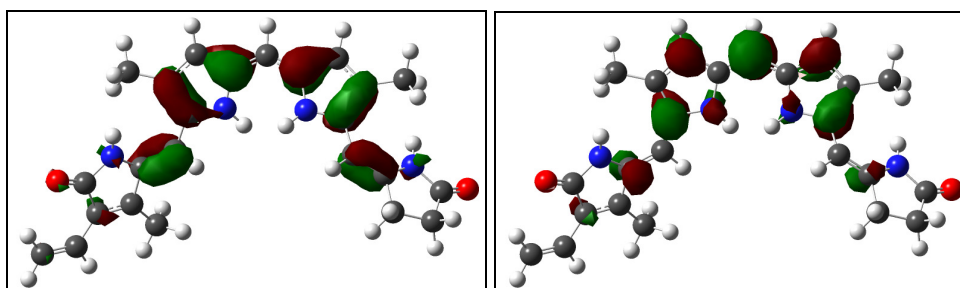


Figura 3.3. Orbitales (CIS/HF) HOMO (izquierda) y LUMO (derecha).

Absorción y emisión

Una vez caracterizada la transición principal en nuestro modelo del cromóforo PΦB, se procedió al análisis de las energías de absorción y emisión, cuyos resultados se muestran en la tabla 3.4. La energía de absorción vertical $S_0 \rightarrow S_1$ concuerda con los valores experimentales conocidos tanto en disolución³⁰ como en el fitocromo.³¹

Hay que tener en cuenta que la conformación del cromóforo en disolución (ZZZsss) puede diferir de la considerada aquí (ZZZasa), tal como se sugiere para el fitocromo.²² Sin embargo, desde el momento en que la excitación se localiza en el centro del cromóforo, estructuralmente protegido en los dos confórmeros, el efecto en las transiciones energéticas cabe que sea minoritario y encaja dentro del error del método de cálculo empleado.

El mínimo S_1 , PLA, se puede asignar, razonablemente, al estado emisor; en este punto es donde el sistema, después de la absorción, permanece unos instantes como tal antes de la redistribución energética interna de forma vibracional, y desde éste se pueblan los canales que conducen a la recuperación del estado fundamental. Obsérvese en la tabla 3.4 que la diferencia de energía vertical $S_1 \rightarrow S_0$, calculada a la geometría de PLA, encaja de forma plausible con las longitudes de onda observadas para el cromóforo en distintos medios.

Tabla 3.4. Longitudes de onda, λ , en nm, de absorción y emisión, calculadas *in vacuo* a nivel CASPT2. Se incluyen también los valores experimentales.

abs/em	PΦB		
	CASPT2 (<i>in vacuo</i>)	Experimental	
		disolvente ^a	proteína ^b
λ_{abs}	642	590	666
λ_{em}	673	680-690	685

^a El cromóforo de la dihidrobiliverdina en CH₃Cl difiere ligeramente del PΦB, ya que contiene un sustituyente metoxycarbonilmetil en C₂ y un grupo etilo, en lugar de un vinilo, en C₁₈.³⁰ Esto puede contribuir, en parte, a que el valor se encuentre desplazado hacia el azul.

^b Espectro realizado del PΦB en el fitocromo.³¹

Rutas de desactivación

Para realizar el estudio de cada uno de los cuatro posibles canales de isomerización A, B, C y D, a través de los cuales evoluciona la estructura PLA, se han calculado las energías, a nivel CASPT2//CASSCF, para un total de 20 puntos de la hipersuperficie de S₁. Los resultados se muestran en la tabla 3.5 y en las figuras 3.2 (donde se han compilado los parámetros geométricos más relevantes) y 3.4. Obsérvese que los mínimos de los estados excitados a los que conducen cada uno de los canales corresponden a las estructuras giradas en 103 (C₄-C₅), 92 (C₉-C₁₀), 91 (C₁₀-C₁₁) y 91 (C₁₅-C₁₆) grados, respectivamente. Tal como se muestra en la figura 3.4, de los cuatro posibles canales de isomerización, tres casi no presentan barrera: B, C y D, con energías de activación de 1.9, 1.5 y 0.7 kcal/mol, respectivamente, mientras que el cuarto canal, A, se encuentra desfavorecido, acabando en un punto de la hipersuperficie de S₁ que está 15.7 kcal/mol por encima de PLA, y que haría que quedara descartado como uno de los caminos de desactivación de la molécula.

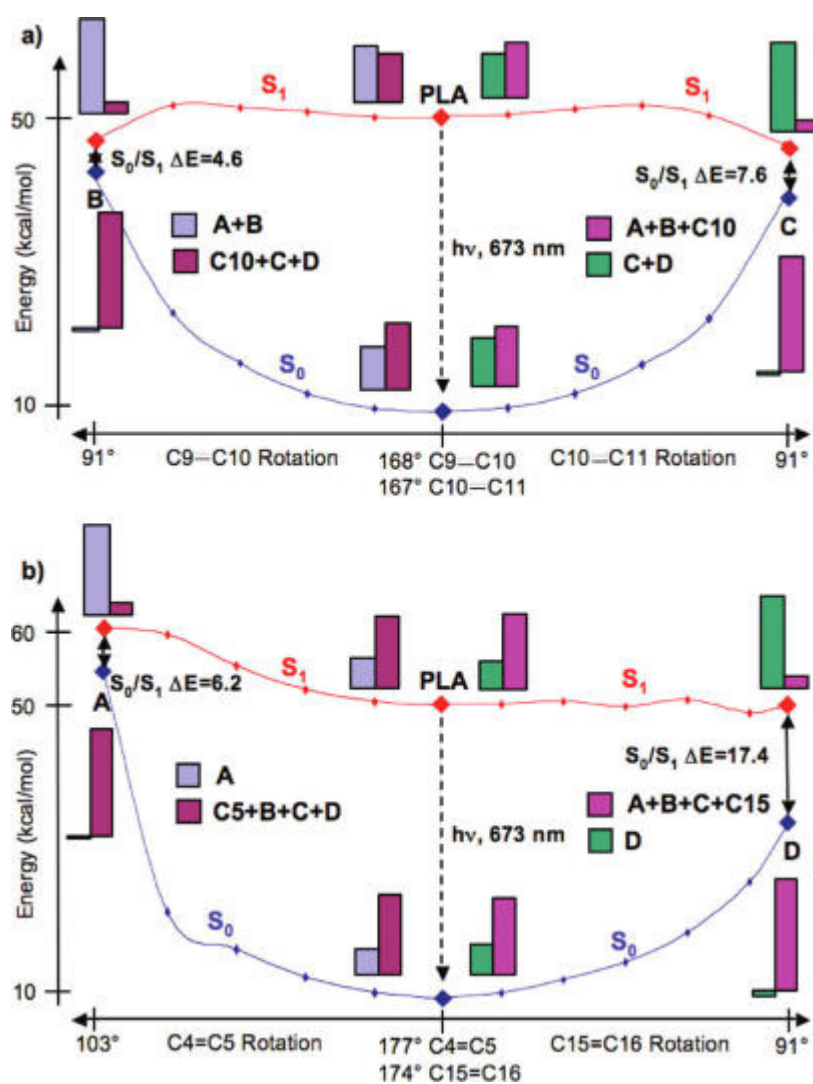


Figura 3.4. Cuatro caminos de relajación a través de distintas torsiones de enlace. Los puntos de inicio (PLA) y finales girados (parte a, caminos B y C; parte b, caminos A y D) son los mínimos optimizados a nivel CIS. Los perfiles energéticos se han recalculado a nivel CASPT2 en cada uno de los puntos a lo largo de los cuatro canales. La carga positiva sobre cada uno de los fragmentos que giran se muestra en los diagramas de barras de colores y pone de relieve la transferencia de carga del estado electrónico.

Tabla 3.5. Energías absolutas (en au) en los distintos puntos de los 4 canales de relajación (A, B, C y D) a lo largo la superficie de S_1 . El número al lado de cada letra indica el valor (en grados) del ángulo diedro en cada punto analizado en dichos canales. Entre paréntesis se indica el enlace girado en cada canal.

CASPT2	A_162	A_147	A_132	A_117	A_103 (C₄-C₅)
S_0	-1294.709678	-1294.706146	-1294.699647	-1294.691089	-1294.635409
S_1	-1294.642646	-1294.639859	-1294.634368	-1294.627115	-1294.625512
S_2	-1294.602455	-1294.601571	-1294.599381	-1294.596836	-1294.584149
CASPT2	B_153	B_138	B_123	B_108	B_92 (C₉-C₁₀)
S_0	-1294.710111	-1294.706748	-1294.699763	-1294.688112	-1294.655998
S_1	-1294.643230	-1294.641950	-1294.641090	-1294.640493	-1294.648741
S_2	-1294.599374	-1294.593576	-1294.583482	-1294.575195	-1294.564341
CASPT2	C_152	C_137	C_122	C_107	C_91 (C₁₀-C₁₁)
S_0	-1294.710243	-1294.706956	-1294.700304	-1294.689721	-1294.662187
S_1	-1294.643135	-1294.641793	-1294.641033	-1294.643309	-1294.650134
S_2	-1294.602065	-1294.599198	-1294.591434	-1294.579199	-1294.547982
CASPT2	D_159	D_144	D_129	D_114	D_91 (C₁₅-C₁₆)
S_0	-1294.709952	-1294.706872	-1294.702838	-1294.696020	-1294.671421
S_1	-1294.643381	-1294.642767	-1294.643943	-1294.642340	-1294.643760
S_2	-1294.602549	-1294.604066	-1294.613053	-1294.621229	-1294.622861

Por otro lado, la diferencia energética S_0-S_1 en el punto de giro de los caminos B y C es bastante pequeña, 4.6 y 7.6 kcal/mol, respectivamente, mientras que, sorprendentemente, en el mínimo girado del camino D (el biológicamente activo que incluye la isomerización del enlace $C_{15}=C_{16}$) es mucho más grande, 17.4 kcal/mol.

El análisis de población de Mulliken, obtenido con la función de onda CASSCF (véase el diagrama de barras en la figura 3.4), revela que no se produce transferencia de carga intramolecular significativa desde FC a PLA, y que en el transcurso de la excitación la carga permanece principalmente localizada en el centro del cromóforo en ambos puntos. Sin embargo, en el proceso de desactivación la carga positiva se mueve progresivamente hasta quedar localizada en la mitad girada más corta de las estructuras torsionadas al final del camino, y es aquí cuando aparece un estado girado de transferencia de carga intramolecular puro (*twisted intramolecular charge transfer*, TICT). Estos resultados excluirían, definitivamente, que esta molécula se pudiera estudiar con aproximaciones computacionales en las que no se tiene en cuenta, de forma precisa, la contribución de la transferencia de carga fotoinducida, como es el caso de métodos TD-DFT con funcionales estándar.³² En la información suplementaria del artículo original²¹ también se pueden consultar los resultados de los cálculos realizados con métodos DFT, tanto en la optimización de las estructuras como en el calibrado energético, pero que no se han incluido en el presente trabajo de Tesis por no ser relevantes para el estudio que nos compete.

Si consideramos la pequeña diferencia energética S_0-S_1 , una vez hecha la corrección CASPT2, en los mínimos girados B y C, se podría pensar que estos puntos están cercanos a, o realmente son, una CI; sin embargo, la diferencia energética en el mínimo D parece demasiado grande como para pensar en la existencia de una CI en ese punto. Estos datos llevan a plantear un esquema de desactivación con tres rutas de conversión internas competitivas entre sí, es decir, dos caminos de fotoisomerización, B y C, que llevarían a un decaimiento ultrarrápido, que conducirían directamente a una CI por mediación de las estructuras giradas que implican a los anillos B y C, y un tercer camino, a través del giro del anillo D, que conduciría a poblar el mínimo de un estado excitado que se encuentra más alejado de una CI, y que supone, en todo caso, una conversión interna más lenta.

Debemos tener en cuenta que la viabilidad de la componente más lenta, pero todavía ultrarrápida, del orden de los picosegundos, a través del camino D, depende de la existencia y accesibilidad a una CI desde el mínimo girado en S_1 . Aunque los resultados CASPT2//CIS que presentamos puedan parecer inconsistentes en este punto, para explicar un decaimiento en aproximadamente 30 ps, pensamos que la CI está cercana, tanto en energía como geometría, al mínimo girado D. El hecho de que la optimización de dicho mínimo ha sido realizada con la aproximación CIS es, posiblemente, la causa que ha llevado a una sobreestimación de la diferencia energética S_1-S_0 . Si nos basamos en que hay una similitud estructural entre el camino A y el D, y que el camino a través de A ha finalizado en una CI (véase figura 3.4), podríamos pensar que D debería haber acabado del mismo modo. Pensamos que, probablemente, no lo haya hecho porque la metodología CIS no incluye los efectos de correlación dinámica en la optimización y que, por tanto, infravalora la longitud del doble enlace en la geometría girada (la longitud del enlace $C_{15}=C_{16}$ es de 1.43 Å, véase figura 3.2) desembocando en un incremento artificial de la diferencia energética en esas geometrías. Para comprobarlo se optimizó la estructura del mínimo D a nivel CASSCF (10,11) obteniéndose una distancia enlace $C_{15}=C_{16}$ mayor (1.46 Å, véase figura 3.5). La diferencia energética S_1-S_0 con cálculos CASPT2 se reduce, en este caso, a tan solo 5.4 kcal/mol, lo que justificaría la presencia de una CI a través del canal D. Como se ha explicado en la sección de metodología del presente capítulo, los cálculos CASSCF no se han utilizado para realizar una exploración sistemática de la superficie por no ser, de forma general, los más adecuados en este caso.

El esquema que se sugiere en la presente Tesis (véase figura 3.6) representa un modelo general para la desactivación fotoquímica del cromóforo de los fitocromos. Nuestros resultados podrían explicar muchas de las propiedades observadas de forma experimental, tanto en disolución como en la proteína, a saber: el decaimiento ultrarrápido de la emisión (en picosegundos), su pequeño rendimiento cuántico ($\sim 10^{-3}$), el corto periodo de vida de S_1 y su decaimiento multiexponencial, con una componente rápida (2-3 ps)⁶ que quedaría explicada con la desactivación a través de los caminos B y C, y una componente lenta (>30 ps) que quedaría explicada con la desactivación a través del camino D y que está relacionada con la fotoisomerización⁷ del

enlace $C_{15}=C_{16}$. De acuerdo con nuestros resultados CASPT2//CIS podemos sugerir que hay una componente lenta del decaimiento a través del camino D y una componente rápida a través de los caminos B y C; que el estado excitado intermedio, I^* , observado recientemente,¹⁴ y que se forma en un primer instante en 0.6 ps, correspondería a PLA, donde tiene lugar la distorsión inicial del grupo $C_{14}-C_{15}=C_{16}$, mientras que el estado excitado intermedio, Lumi-R*,¹⁴ formado más tarde, en 3 ps, correspondería al mínimo girado obtenido por optimización de D en S_1 , y que sería el precursor del fotoproducto Lumi-R.

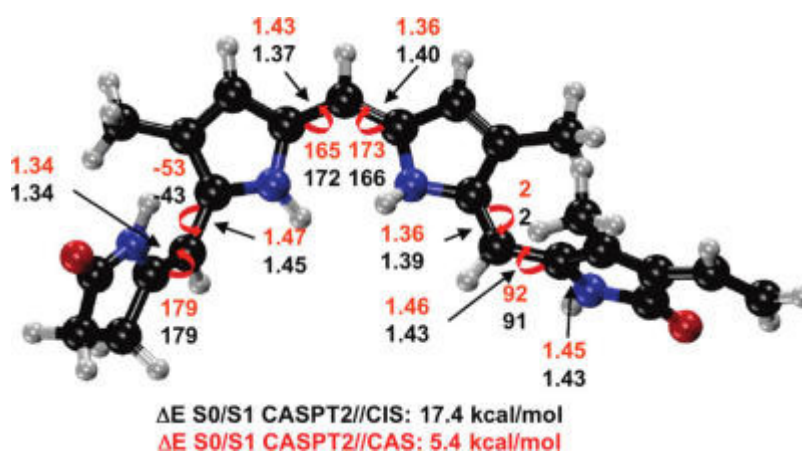


Figura 3.5. Parámetros geométricos (enlaces en Å y ángulos en grados) para la geometría optimizada en S_1 en el punto girado en D, tanto para el nivel de cálculo CIS (valores negros) como para CASSCF (valores rojos). La diferencia energética S_0-S_1 CASPT2 también queda reflejada.

Contrariamente a los datos de las estructuras cristalinas de los dominios del cromóforo del fitocromo,¹¹ que indican que la isomerización completa de C_9-C_{10} y $C_{10}-C_{11}$ (a lo largo de los caminos B y C), no pueden ocurrir en el interior de la proteína, sugerimos que dichos giros pueden iniciarse en S_1 (cuando todavía no hay una demanda excesiva de espacio para la torsión); estos giros iniciarían la componente rápida de conversión interna hacia los canales de isomerización que, sin embargo, al ir evolucionando se ven

abortados debido a una demanda de volumen mayor y que, por tanto, no pueden acabar en S_0 .

Por otro lado, a través del camino D la molécula tiene suficiente espacio como para finalizar en S_0 , pero con una vida media más larga y con un rendimiento cuántico más pequeño ($\sim 9\%$).⁷ La completa validación de esta hipótesis se encuentra fuera de los objetivos de este trabajo, pues requiere un mapeo de los canales de fotoisomerización dentro de la proteína usando, por ejemplo, metodología QM/MM.^{33,34}

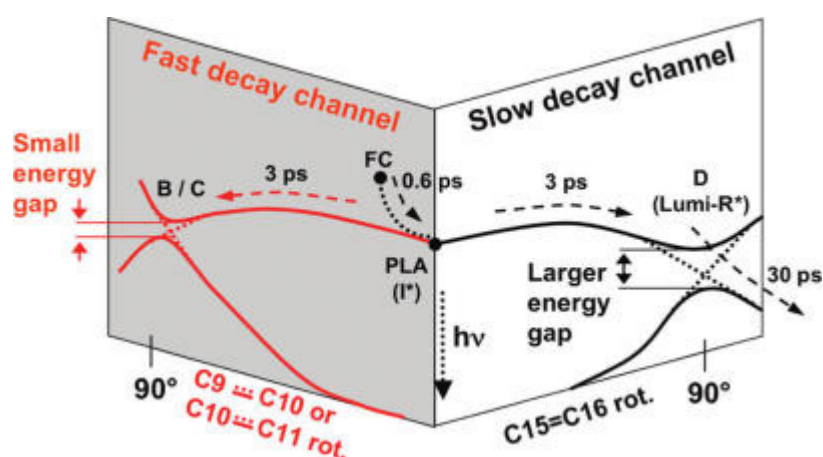


Figura 3.6. Representación esquemática de los distintos caminos competitivos de fotorreacción en los cromóforos de los fitocromos de acuerdo con los resultados CASPT2//CIS. Los tiempos de decaimiento experimentales se han asignado de forma tentativa.

Conclusiones del capítulo 3

- ✓ En el presente trabajo hemos estudiado las propiedades espectroscópicas del cromóforo de los fitocromos sobre un modelo del mismo, y hemos comprobado que son reproducibles y están en acuerdo con los datos experimentales de la bibliografía. Todo esto lleva a dar por buenos tanto la metodología como el modelo utilizados en el estudio de la desactivación fotoquímica de dicho cromóforo.
- ✓ El estudio presenta, además, un esquema de la fotoquímica intrínseca del cromóforo de los fitocromos que permite explicar de forma cualitativa su comportamiento, observado tanto en disolución como en el interior de la proteína.
- ✓ Con nuestro estudio añadimos, a los conocimientos que ya se disponen sobre la fotofísica/fotoquímica de este cromóforo, que existen rutas de isomerización competitivas, casi sin barrera, que conducen a la formación de unas TICT-CIs, de estructura girada, en las que se ha transferido carga desde una parte de la molécula a la otra. Desde dichas CIs se produce el decaimiento no radiativo ultrarrápido multicanal, o multiexponencial.
- ✓ De acuerdo con nuestros resultados CASPT2//CIS, PLA correspondería a un estado intermedio excitado temprano (I^*) y D a uno tardío (en la escala de los 3 ps de formación) correspondiente al estado intermedio excitado (Lumi- R^*).
- ✓ El decaimiento desde D se asignaría a una componente de desactivación más lenta (alrededor de 30 ps), mientras que los caminos inactivos biológicamente, B y C, conducirían a la desactivación rápida (3 ps) y principal depleción del estado excitado.

- ✓ Del mismo modo explicaría el menor rendimiento cuántico observado a través de D y la disminución en la intensidad de la señal asociada con las correspondientes señales vibracionales.

Referencias

1. Rockwell, N. C.; Su, Y.-S.; Lagarias, J. C. *Annu. ReV. Plant Biol.* **2006**, *57*, 7.
2. Schaffner, K.; Braslavsky, S. E.; Holzwarth, A. R. Protein Environment and Photophysics and Photochemistry of the Prosthetic Chromophores of Biliproteins. In *Frontiers in Supramolecular Organic Chemistry and Photochemistry*; Schneider, H.-J., Dürr, H. Eds.; VCH Verlagsgesellschaft: Weinheim, Germany, 1991; pp 421-452.
3. Holzwarth, A. R.; Venuti, E.; Braslavsky, S. E.; Schaffner, K. *Biochim. Biophys. Acta* **1992**, *1140*, 59.
4. Gärtner, W.; Braslavsky, S. E. The Phytochromes: Spectroscopy and Function. In *Photoreceptors and Light Signaling*; Batschauer, A. Ed.; Royal Society of Chemistry: Cambridge, U.K., 2004; Vol. 3, pp 136-180.
5. Falk, H. *The Chemistry of Linear Oligopyrroles and Bile Pigments*; Springer-Verlag: New York, 1989.
6. Rentsch, S.; Hermann, G.; Bischoff, M.; Strehlow, D.; Rentsch, M. *Photochem. Photobiol.* **1997**, *66*, 585.
7. Schumann, C.; Groß, R.; Michael, N.; Lamparter, T.; Diller, R. *ChemPhysChem* **2007**, *8*, 1657.
8. Müller, M. G.; Lindner, I.; Martin, I.; Gärtner, W.; Holzwarth, A. R. *Biophys. J.* **2008**, *94*, 4370.
9. Mroginski, M. A.; Murgida, D. H.; Hildebrandt, P. *Acc. Chem. Res.* **2007**, *40*, 258.
10. Thor, J. J. V.; Ronayne, K. L.; Towrie, M. *J. Am. Chem. Soc.* **2007**, *129*, 126.
11. Wagner, J. R.; Brunzelle, J. S.; Forest, K. T.; Vierstra, R. D. *Nature* **2005**, *438*, 325.
12. Inomata, K.; Hammam, M. A. S.; Kinoshita, H.; Murata, Y.; Khawn, H.; Noack, S.; Michael, N.; Lamparter, T. *J. Biol. Chem.* **2005**, *280*, 24491.
13. Hughes, J.; Lamparter, T.; Mittmann, F.; Hartmann, E.; Gärtner, W.; Wilde, A.; Borner, T. *Nature* **1997**, *386*, 663.
14. Dasgupta, J.; Frontiera, R. R.; Taylor, K. C.; Lagarias, J. C.; Mathies, R. A. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 1784.
15. Olivucci, M. *Computational Photochemistry*; Elsevier: Amsterdam, 2005; Vol. 16.
16. Domcke, W.; Yarkony, D.; Köppel, H. *Conical Intersections: Electronic Structure, Dynamics & Spectroscopy*; World Scientific: Singapore, 2004.
17. Durbeej, B.; Eriksson, L. A. *Phys. Chem. Chem. Phys.* **2006**, *8*, 4053.

18. Borg, O. A.; Durbeej, B. *Phys. Chem. Chem. Phys.* **2008**, *10*, 2528.
19. Durbeej, B. *Phys. Chem. Chem. Phys.* **2009**, *11*, 1354.
20. Climent, T.; González-Luque, R.; Merchán, M. *J. Phys. Chem. A* **2003**, *107*, 6995.
21. Altoe, P.; Climent T.; De Fusco, G.; Stenta, M.; Bottoni, A.; Serrano-Andrés, L.; Merchán, M.; Orlandi, G.; Garavelli, M. *J. Phys. Chem. B* **2009**, *113*, 15067.
22. Kneip, C.; Hildebrandt, P.; Schlamann, W.; Braslavsky, S. E.; Mark, F.; Schaffner, K. *Biochemistry* **1999**, *38*, 15185.
23. Knipp, B.; Müller, M.; Metzler-Nolte, N.; Balaban, T. S.; Braslavsky, S. E.; Schaffner, K. *Helv. Chim. Acta* **1998**, *81*, 881.
24. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03*, revision C.02; Gaussian, Inc.: Wallingford, CT, 2004.
25. Andersson, K.; Malmqvist, P.-Å.; Roos, B. O. *J. Chem. Phys.* **1992**, *96*, 1218.
26. Andersson, K.; Barysz, M.; Bernhardsson, A.; Blomberg, M. R. A.; Carissan, Y.; Cooper, D. L.; Cossi, M.; Fülscher, M. P.; Gagliardi, L.; de Graaf, C.; Hess, B.; Hagberg, G.; Karlström, G.; Lindh, R.; Malmqvist, P.-Å.; Nakajima, T.; Neogrády, P.; Olsen, J.; Raab, J.; Roos, B. O.; Ryde, U.; Schimmelpfennig, B.; Schütz, M.; Seijo, L.; Serrano-Andrés, L.; Siegbahn, P. E. M.; Stålring, J.; Thorsteinsson, T.; Veryazov, V.; Widmark, P. O. *MOLCAS 6.0*; Department of Theoretical Chemistry, Chemical Centre, University of Lund: Lund, Sweden, 2004.
27. Merchán, M.; Serrano-Andrés, L. *THEOCHEM* **2005**, *729*, 99.

28. Merchán, M.; González-Luque, R.; Climent, T.; Serrano-Andrés, L.; Rodríguez, E.; Reguero, M.; Pelaez, D. *J. Phys. Chem. B* **2006**, *110*, 26471.
29. Malmqvist, P.-Å.; Roos, B. O. *Chem. Phys. Lett.* **1989**, *155*, 189.
30. Ditto, M.; Brunner, H.; Lippitsch, M. E. *Chem. Phys. Lett.* **1991**, *185*, 61.
31. Büchler, R.; Hermann, G.; Lap, D. V.; Rentsch, S. *Chem. Phys. Lett.* **1995**, *233*, 514.
32. Dreuw, A.; Head-Gordon, M. *J. Am. Chem. Soc.* **2004**, *126*, 4007.
33. Altoe, P.; Stenta, M.; Bottoni, A.; Garavelli, M. *Theor. Chem. Acc.* **2007**, *118*, 219.
34. Case, D. A.; Cheatham, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. *J. Comput. Chem.* **2005**, *26*, 1668.

4 Isoaloxacina, el anillo básico de las flavinas

Tal como se ha introducido en el capítulo 2, el responsable de la fotoquímica de las flavinas es el anillo de isoaloxacina del FMN completamente oxidado.^{1,2,3} Después de la absorción de la luz, se forma un aducto cromóforo-proteína que se descompone en una escala de tiempo de minutos en la oscuridad posterior (véase esquema 2.3).⁴ Además, está totalmente aceptado entre la comunidad científica que, después de la absorción de la luz, el principal paso del fotociclo del FMN consiste en la población del estado excitado singlete más bajo, S_1 , para posteriormente experimentar un rápido decaimiento al estado excitado triplete más bajo, T_1 , a través de un mecanismo de cruce entre sistemas, ISC.⁵⁻⁹ La transferencia fotoinducida de electrones al triplete de las flavinas, que se comportan como aceptores, ha sido ampliamente analizada por medio de diferentes técnicas experimentales,^{10,11} y en este capítulo presentamos el análisis detallado del proceso no radiativo singlete-triplete dentro del marco de la fotoquímica adiabática. En particular, al iniciar la investigación sobre este fotociclo nos planteamos la cuestión clave de si la relajación del estado triplete consiste en un acoplamiento directo entre los estados singlete y triplete más bajos o se produce a través de un tercer estado excitado. Para dar respuesta a esta cuestión, se estudió, con cálculos químico-cuánticos *ab initio*, el mecanismo detallado de la fotoquímica de la isoaloxacina (benzo[g]pteridina-2, 4 (3H, 10H)-diona) (véase figura 4.1), que es el anillo básico de las flavinas. La elección de este anillo se fundamenta en que los espectros electrónicos de la lumiflavina, la 8-metil-isoaloxacina y la riboflavina muestran las mismas características, pudiendo ser descritos como típicos de isoaloxacinas¹² en la

región de energías hasta 5.5 eV, en que el espectro de absorción de estos sistemas se compone principalmente de tres bandas. Por otro lado, no se han encontrado diferencias significativas en los espectros de polarización de la lumiflavina, la riboflavina y el FMN en relación con la molécula original, la forma oxidada del anillo de la isoaloxacina¹³ y, por tanto, se puede esperar que éste sea el responsable de la fotoquímica de las flavinas.

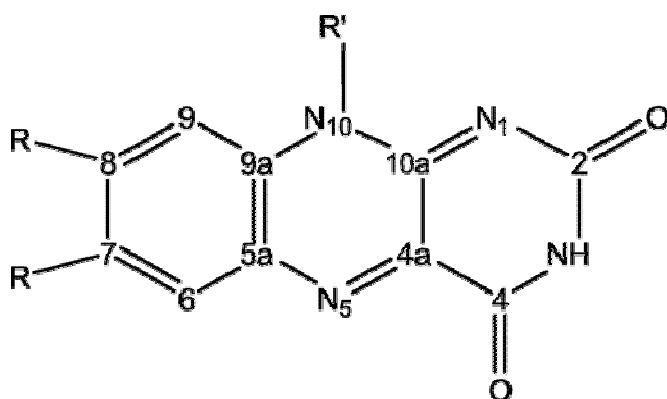


Figura 4.1. Estructura y numeración de los átomos del anillo de isoaloxacina, cromóforo de las flavinas.

Se presenta, en este capítulo, la investigación teórica sobre los estados excitados de valencia de la molécula de isoaloxacina que han sido caracterizados mediante el método CASPT2//CASSCF,¹⁴ que ha demostrado su eficacia en la interpretación de los espectros electrónicos en un gran número de compuestos orgánicos¹⁵⁻²³ y de cromóforos de distintas biomoléculas²⁴⁻²⁷ (estas últimas estudiadas por el grupo de Química Cuántica del Estado Excitado de la Universidad de Valencia, QCEXVAL). El presente trabajo incluye la determinación de las energías de transición vertical, orígenes de la banda y máximos de emisión relacionados con los estados excitados de valencia singlete y triplete más bajos del anillo de isoaloxacina. Además, se presentan los resultados obtenidos del análisis realizado de la relajación electrónica del sistema excitado a lo largo de la hipersuperficie de S_1 que conduce a la población del estado triplete más bajo T_1 , mediado por el estado triplete del T_N .

Detalles computacionales

Las optimizaciones de la geometría del estado fundamental y de los estados excitados de valencia más bajos se llevaron a cabo a nivel CASSCF.²⁸ El espacio activo π de valencia del anillo de isoaloxacina incluye 18 electrones y 16 orbitales moleculares, CASSCF(18,16). El espacio activo para el cálculo de los estados de carácter $\pi\pi^*$, después de una cuidadosa calibración, se redujo a 14 electrones distribuidos en 13 MOs π de valencia, CASSCF(14,13), mientras que para los estados $n\pi^*$ el espacio activo fue ampliado con los MOs correspondientes a los pares solitarios, manteniendo inactivo el MO π con el número de ocupación más alto del espacio CASSCF(14,13) y pasando al espacio secundario el MO π con el número de ocupación más bajo en dicha función, dando finalmente una función de onda CASSCF(16,13). Los resultados que se presentan corresponde, en todos los casos, a cálculos realizados con el conjunto de bases 6-31G(d), ya que la inclusión de funciones de polarización en los átomos de hidrógeno, mediante el uso de bases 6-31G(dp), tiene, en este caso, un efecto menor sobre el cálculo de las energías electrónicas de transición. Los cálculos exploratorios de las diferentes regiones de las hipersuperficies, del estado fundamental y estados excitados, a partir de geometrías muy distorsionadas del anillo de isoaloxacina acababan siempre en estructuras planas, incluso con bases ANO-S. Es por tanto que, de acuerdo con nuestra investigación, concluimos que la molécula se mantiene plana durante la reacción fotoquímica. Las energías de transición vertical se calcularon, pues, dentro de las limitaciones de simetría C_s . Los MOs de los estados excitados se obtuvieron de cálculos CASSCF promedio, que incluyen todos los estados de interés dentro de una simetría dada: ocho raíces para los estados $^1A'$, dos raíces para los estados $^{1,3}A''$ incluyendo la participación del MO del par solitario (se empleó el mismo tratamiento para los ubicados en los átomos de oxígeno o de nitrógeno, pero en cálculos independientes), y una raíz para el estado más bajo $^3A'$. El resto de efectos electrónicos de correlación se añadieron al realizar cálculos CASPT2.¹⁴ Además, el efecto de la interacción débil de los estados intrusos se minimizó mediante el uso de la técnica *level-shift*.²⁹ El cálculo de los TDM se realizó con el método

CASSI,^{30,31} y éstos, a su vez, se emplearon, junto con las energías CASPT2, para computar la fuerza de oscilador, f , así como para calcular los tiempos de vida media radiativa usando la aproximación de Strickler-Berg.^{32,33} Desde la geometría FC se realizó un MEP-CASSCF a lo largo de S_1 , empleando el algoritmo descrito en la referencia 34 y calculando para cada punto la energía corregida CASPT2. El SOC se calculó de la forma que se describe en la referencia 35, empleándose en todo momento el paquete de cálculo MOLCAS 6.0.^{36,37}

Estructuras de equilibrio

En la figura 4.2 se muestran los parámetros geométricos, optimizados a nivel CASSCF, del estado fundamental, $(gs)_{\min}$, y de los estados excitados singlete y triplete más bajos de carácter $\pi\pi^*$, $^1(\pi\pi^*)_{\min}$ y $^3(\pi\pi^*)_{\min}$. En ella también se han incluido, los resultados para el estado triplete $^3(n_N\pi^*)$, estado especialmente relevante para la racionalización del comportamiento fotoquímico de la molécula. Las coordenadas cartesianas de las estructuras optimizadas se pueden encontrar en la información suplementaria del artículo publicado con los resultados que se presentan en este capítulo.³⁸ La mayoría de las distancias de enlace calculadas para el estado fundamental se desvían en menos de 0.020 Å de los datos cristalográficos disponibles.³⁹ Como la isoaloxacina tautomeriza inmediatamente a aloxacina por transferencia del protón del N_{10} al N_1 , el derivado más simple con el que poder comparar adecuadamente se corresponde con la 10-metil-isoaloxacina.³⁹ La concordancia entre los resultados CASSCF obtenidos para el estado fundamental y los datos de rayos X es satisfactoria y, por tanto, se puede deducir que el sustituyente metilo desempeña un papel menor en la determinación de la estructura de equilibrio de los anillos fusionados del sistema molecular; situación que se espera similar para las propiedades espectroscópicas.

Los principales cambios que tienen lugar en la geometría de las estructuras $^1(\pi\pi^*)_{\min}$ y $^3(\pi\pi^*)_{\min}$, con respecto al estado fundamental $(gs)_{\min}$, implican un aumento de las distancias de enlace N_1-C_{10a} , $C_{4a}-N_5$ y, en menor medida, $C_{5a}-C_{9a}$, mientras que se observa una disminución de la distancia de enlace

C_{4a}-C_{10a}. Los estados excitados ¹(ππ*) y ³(ππ*) se describen, principalmente, por la promoción de un electrón desde el HOMO al LUMO. Como puede observarse en la figura 4.3 el HOMO se localiza, principalmente, sobre el átomo de nitrógeno N₁ y en la región de unión C_{5a}-C_{9a}. En el caso del LUMO, la densidad electrónica se sitúa, principalmente, en la región del enlace C_{4a}-C_{10a} y del átomo de nitrógeno N₅. Para el estado ³(n_Nπ*) la transición electrónica se produce, fundamentalmente, desde los pares solitarios que se encuentran en los átomos de nitrógeno hacia el LUMO.

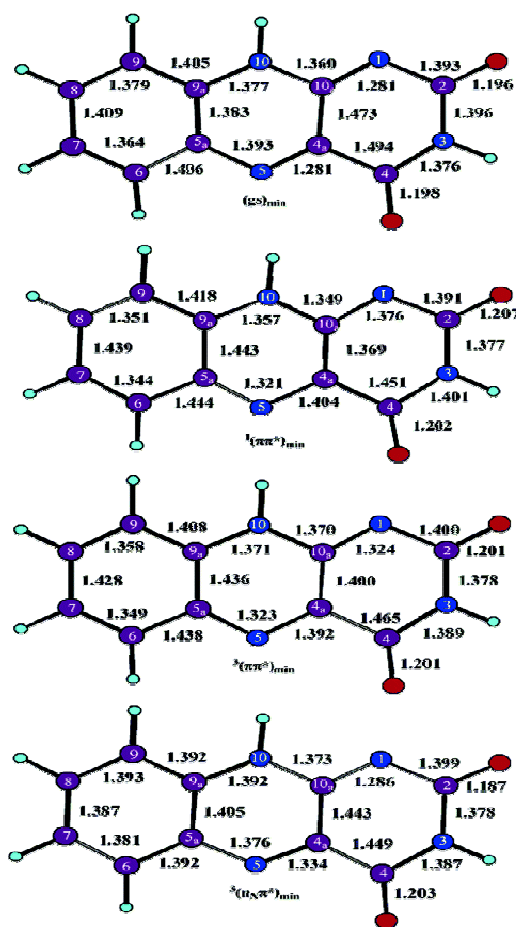


Figura 4.2. Parámetros geométricos de las estructuras de los estados fundamental y excitados más bajos optimizadas a nivel CASSCF/6-31G(d)

Tabla 4.1. Energías de excitación, experimentales y calculadas a nivel CASSCF y CASPT2, fuerzas de oscilador, f , correspondientes a las transiciones energéticas y momentos dipolares, μ , de los estados excitados de la molécula de isoaloxacina.

Estado ^a	Energía de excitación (eV)		f	μ (D)	Experimental ^b
	CASSCF	CASPT2			
Estado fundamental					
1 ¹ A'				2.4	
Estados excitados singlete					
2 ¹ A'($\pi \rightarrow \pi^*$)	4.15	3.09	0.239	9.7	2.85
1 ¹ A''($n_{N-} \rightarrow \pi^*$)	4.79	3.34	0.007	5.4	
2 ¹ A''($n_{O-} \rightarrow \pi^*$)	5.15	3.75	0.001	1.9	
3 ¹ A'($\pi \rightarrow \pi^*$)	5.47	4.28	0.158	10.1	3.76
3 ¹ A''($n_{O+} \rightarrow \pi^*$)	5.62	4.43	0.000	2.6	
4 ¹ A''($n_{N+} \rightarrow \pi^*$)	5.83	4.66	0.001	5.0	
4 ¹ A'($\pi \rightarrow \pi^*$)	6.27	4.69	0.104	7.8	
5 ¹ A'($\pi \rightarrow \pi^*$)	5.88	4.90	0.035	7.9	
6 ¹ A'($\pi \rightarrow \pi^*$)	6.78	4.92	0.058	1.8	
7 ¹ A'($\pi \rightarrow \pi^*$)	7.16	5.00	0.337	7.1	4.68–4.77
8 ¹ A'($\pi \rightarrow \pi^*$)	7.48	5.37	0.641	5.5	
Estados excitados triplete					
1 ³ A'($\pi \rightarrow \pi^*$)	3.14	2.52		6.0	
1 ³ A''($n_{N-} \rightarrow \pi^*$)	4.30	2.97		6.3	
2 ³ A''($n_{O-} \rightarrow \pi^*$)	4.99	3.70		2.2	
3 ³ A''($n_{O+} \rightarrow \pi^*$)	5.47	4.39		2.8	
4 ³ A''($n_{N+} \rightarrow \pi^*$)	5.57	4.53		5.1	

^a El carácter y orden de los estados es el obtenido con el método CASPT2.

^b Máximos de banda del espectro de absorción de la 8-metil-isoaloxacina en etanol a 77 K. Ref:12.

Las geometrías optimizadas se utilizaron para la caracterización de los espectros electrónicos verticales (absorción y emisión) y transiciones electrónicas no verticales.

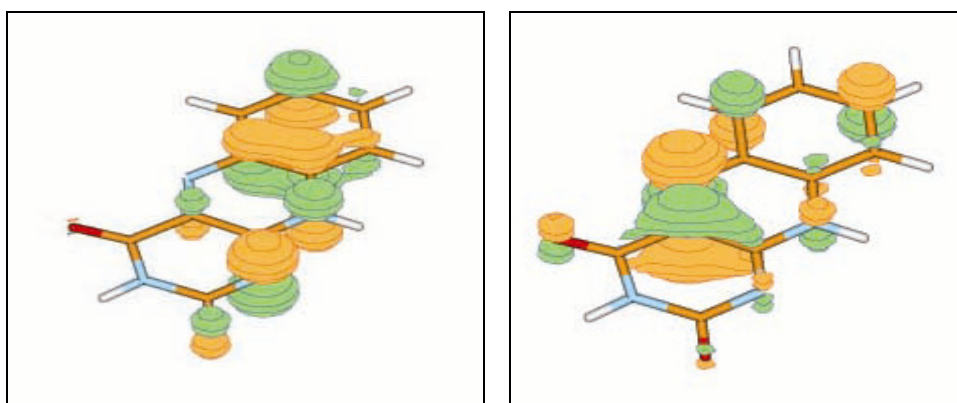


Figura 4.3. HOMO (izquierda) LUMO (derecha)

Espectros Verticales

En la Tabla 4.1 se enumeran las energías de excitación vertical, calculadas a nivel CASSCF y CASPT2, junto con las fuerzas de oscilador de las transiciones electrónicas relacionadas, momentos dipolares de los estados, así como una selección de datos experimentales. Las asignaciones que identifican a los estados excitados se hallan entre paréntesis en la primera columna de la tabla, donde n_{N-} y n_{N+} representan las combinaciones lineales de los pares solitarios, menos y más, que pertenecen a los átomos N_1 y N_5 . Lo mismo se aplica para los dos átomos de oxígeno: n_{O-} y n_{O+} . A partir de aquí, nos referiremos a los estados singlete, $2^1A'(\pi \rightarrow \pi^*)$, y triplete, $1^3A'(\pi \rightarrow \pi^*)$, como S_1 y T_1 , respectivamente y al estado $1^3A''(n_{N-} \rightarrow \pi^*)$ como T_N . Como se verá más adelante, este estado podría estar especialmente involucrado en la desactivación del estado S_1 y posterior población del estado triplete más bajo T_1 .

Encontramos cinco transiciones electrónicas de valencia de carácter $\pi\pi^*$, con una fuerza de oscilador significativa, que pueden relacionarse con las bandas principales del espectro electrónico registrado hasta 5.50 eV. Sólo

una transición se encuentra en el intervalo del visible, a 3.09 eV, correspondiente a la absorción vertical desde el estado fundamental al primer estado excitado singlete. Esta transición $S_0 \rightarrow S_1$ (HL) tiene una fuerza de oscilador asociada de aproximadamente 0.2. Como el momento dipolar de S_1 (9.7 D) es mucho mayor que el del estado fundamental (2.4 D), en disolventes polares se puede esperar que ocurra un desplazamiento hacia el rojo en la transición de menor energía. Esto es consistente con los espectros de absorción, realizados por Sun et al.,¹² en derivados de la metil-isoaloxacina, en etanol a 77 K, donde se observa un pico máximo a 2.85 eV en la 8-metil-isoaloxacina y a 2.76 eV en lumiflavina, 3-metil-lumiflavina y riboflavina. El estado excitado triplete más bajo de tipo $n\pi^*$, T_N , se calcula que se encuentran verticalmente a 3.34 eV, es decir, cerca de la transición más baja. Como es de esperar, la intensidad de la transición se prevé que será sumamente débil (la fuerza de oscilador es 0.007). Sin embargo, debido a la proximidad de estados $n\pi^*$ y $\pi\pi^*$, éstos pueden interactuar mediante acoplamiento vibrónico y la banda resultante podría estar relacionada con el hombro observado en la parte alta de la región del espectro de menor energía, registrado en etanol.¹² La segunda banda de energía registrada experimentalmente se puede relacionar con la transición de naturaleza $\pi\pi^*$ calculada, *in vacuo*, a 4.28 eV y con una fuerza de oscilador ligeramente inferior, 0.158, que la obtenida para la transición más baja. El resultado está ligeramente desplazado hacia el azul con respecto al máximo de la banda correspondiente al espectro observado en disolución, 3.62-3.76 eV (en función del derivado metílico de que se trate).¹² Un estudio de espectroscopia de absorción y fluorescencia, realizado en diferentes isoaloxacinas,⁴⁰ concluyó que el aumento en la polaridad del disolvente tiene sólo una pequeña influencia en el máximo de la banda de absorción de más baja energía, mientras que el máximo de la segunda banda presenta un desplazamiento batocrómico pronunciado. De hecho, el momento dipolar calculado para el estado $3^1A'$ es mayor que el del estado $2^1A'$ y, por tanto, en disolventes polares debería ocurrir un desplazamiento mayor del máximo de la segunda banda hacia el rojo. Dado que los dos estados $n\pi^*$ de los pares solitarios del grupo carbonilo se encuentran cercanos energéticamente, también puede ocurrir una interacción con el estado $2^1A'(\pi \rightarrow \pi^*)$, resultando en una banda compleja, donde el máximo no necesariamente tendría que

coincidir con la transición vertical aquí calculada, lo que es más una regla que una excepción en cuanto a la espectroscopia del grupo carbonilo.²¹ Debido a la gran diferencia entre las geometrías del estado fundamental y del excitado, las características espectroscópicas relacionadas con el grupo CO se encuentran altamente influidas por los estados de vibración, y como resultado el máximo de la banda no coincide con la transición vertical, tal como se supone por el principio de FC.^{19,41}

La tercera banda de energía de las isoaloxacinas en etanol tiene un máximo de banda en el intervalo de energías de 4.7-4.8 eV.¹² De 4.7 a 5.4 eV se han calculado cinco transiciones electrónicas que podrían contribuir al espectro observado. La mayoría de los estados tienen momento dipolar más alto que el estado fundamental y, en consecuencia, las transiciones calculadas se espera que queden un poco desplazadas al azul con respecto a los datos experimentales en etanol, comportamiento que está siendo una constante en el sistema en estudio, sobre todo en las transiciones con mayores fuerzas de oscilador. Concluimos, pues, que la tercera banda del espectro de energía es el resultado de varias transiciones electrónicas $\pi\pi^*$, en particular de aquellas que tienen lugar desde el estado fundamental a los estados excitados $4^1A'$ (4.69 eV), $7^1A'$ (5.00 eV) y $8^1A'$ (5.37 eV), que tiene fuerzas de oscilador significativas (véase tabla 4.1) y, por tanto, constituyen los principales contribuyentes a la banda. En las funciones de onda CASSCF de los estados excitados $4^1A'$ y $7^1A'$ tienen una participación dominante la excitación electrónica HOMO \rightarrow LUMO + 1, mientras que, los dos estados $5^1A'$ y $6^1A'$ se describen principalmente por la promoción de un electrón desde el HOMO-2 al LUMO y la configuración descrita por la promoción de dos electrones desde el HOMO al LUMO, en el primero, y desde el HOMO-3 al LUMO en el segundo. Como se puede observar en la tabla 4.1, una serie de estados del tipo $n\pi^*$ se intercalan entre las principales transiciones de valencia. Los estados $1^1A''(n_{N-}\rightarrow\pi^*)$, $2^1A''(n_{O-}\rightarrow\pi^*)$, $3^1A''(n_{O+}\rightarrow\pi^*)$ y $4^1A''(n_{N+}\rightarrow\pi^*)$ se encuentran verticalmente a 3.34, 3.75, 4.43 y 4.66 eV, respectivamente. Las fuerzas de oscilador asociadas a esas transiciones son insignificantes y, por tanto, no tienen especial relevancia en relación con el espectro de absorción. Una discusión más amplia respaldando las asignaciones propuestas, y que incluye el análisis de las

direcciones de polarización calculadas para las principales transiciones, se puede consultar en la publicación correspondiente.³⁸

El orden y la naturaleza de los cinco estados triplete de valencia más bajos es el mismo que el obtenido para los estados excitados singlete. El espectro singlete-triplete, calculado a la geometría del estado fundamental de la isoaloxacina, sitúa al estado triplete más bajo $1^3A'(\pi \rightarrow \pi^*)$ a 2.52 eV. A continuación, a 2.97 eV, se encuentra el estado $1^3A''(n_{N-} \rightarrow \pi^*)$. Como se puede observar en la tabla 4.1, el resto de estados tripletes se encuentran dentro del intervalo de energía de 3.70-4.53 eV. Como era de esperar, por su integral de intercambio, la mayor separación singlete-triplete se encuentra entre los estados más bajos $\pi\pi^*$ (0.57 eV). Normalmente, esta integral es relativamente más grande en estados $\pi\pi^*$ que en $\sigma\pi^*$ y, por tanto, la separación singlete-triplete se vuelve más pronunciada entre los estados con carácter $\pi\pi^*$. A excepción de la pareja $1,3A''(n_{N-} \rightarrow \pi^*)$, que tiene una diferencia de energía singlete-triplete de 0.37 eV, el resto de diferencias triplete-singlete se encuentran como mucho en los 0.1 eV.

Energías de emisión y de transición no vertical

El conocimiento de las hipersuperficies de energía del estado fundamental y de los estados excitados más bajos de la isoaloxacina es esencial para la determinación de sus propiedades espectroscópicas, que son particularmente relevantes en un amplio grupo de cromóforos biológicamente activos. La comprensión del comportamiento espectroscópico de estos compuestos requiere, no sólo el cálculo exacto de los estados verticales sino también la caracterización de las energías de emisión y de transición no vertical. El estudio incluye (véase tabla 4.2) tanto la determinación de las transiciones verticales como la adiabática, relacionadas con las bandas de absorción y emisión de los tres estados excitados de valencia: $\pi\pi^*$ y $n\pi^*$, singletes y tripletes.

Los resultados obtenidos predicen, para la transición singlete-singlete más baja (del tipo $\pi\pi^*$), un origen banda de fluorescencia a 2.69 eV y una emisión vertical a 2.04 eV. Estos resultados se pueden relacionar con el espectro de la 8-metil-isoaloxacina en etanol,¹² que muestra orígenes de

banda de emisión cerca de los 2.7 eV y un máximo de fluorescencia entre 2.5 y 2.3 eV. La vida media de radiación calculada, 15 ns, también está de acuerdo con los datos experimentales para la vida de fluorescencia de los derivados de la metil-isoaloxacina;¹² 1 ns en el caso de la 3,10-dimetil-isoaloxacina en acetonitrilo.⁴⁰

Tabla 4.2. Energías de transición vertical, origen de banda y de emisión (en eV), y tiempos de vida media de los estados excitados singletes más bajos de la isoaloxacina.

Estado	Transición vertical		Origen de banda		Emisión		τ_{rad}
	CASSCF	CASPT2	CASSCF	CASPT2	CASSCF	CASPT2	
Estados Singlete							
$\pi\pi^*$	4.15	3.09	3.20	2.69	2.21	2.04	15 ns
$n_{\text{N}}\pi^*$	4.79	3.34	3.82	3.33	2.65	2.61	467 ns
$n_{\text{O}}\pi^*$	5.15	3.75	3.52	3.16	2.19	2.33	6458 ns
Estados Triplete							
$\pi\pi^*$	3.14	2.52	2.41	2.03	1.73	1.75	116 ms
$n_{\text{N}}\pi^*$	4.30	2.97	3.50	2.73	3.00	2.37	
$n_{\text{O}}\pi^*$	4.99	3.70	3.40	3.10	2.20	2.34	

El estado triplete fosforescente de las isoaloxacinas ha sido asignado como de tipo $^3(\pi,\pi^*)$, en base a los tiempos de vida relativamente largos y a la tendencia de polarización negativa de fosforescencia en la región 0-0 con respecto a las transiciones más bajas tipo $\pi\pi^*$.¹² Nuestros resultados están de acuerdo con ello, y la predicción de que el estado triplete más bajo tiene carácter $^3(\pi,\pi^*)$, con un origen de la banda fosforescencia a 2.03 eV y una emisión vertical a 1.75 eV, están en concordancia con los valores publicados para la 8-metil-isoaloxacina en etanol,¹² alrededor de 2.2 y 2.0 eV, respectivamente.

Formación del estado triplete

Tras ser irradiada con luz UV, la isoaloxacina, como ya hemos analizado, se promueve al estado S_1 , con una energía de transición vertical, de 3.09 eV y con una fuerza de oscilador de alrededor de 0.2. Para analizar la evolución del decaimiento de este estado, se realizó un MEP CASSCF a lo largo de la coordenada de reacción de S_1 . Para cada punto del MEP se calcularon, posteriormente, las energías CASPT2//CASSCF de los correspondientes estados excitados singlete y triplete. En la figura 4.4 se muestra el MEP- S_1 y como puede observarse, el primer punto de cruce aparece cerca de la región FC, donde se lleva a cabo un cruce singlete-triplete $(S_1/T_N)_{STC}$, mientras que el estado T_1 se encuentra siempre por debajo de S_1 .

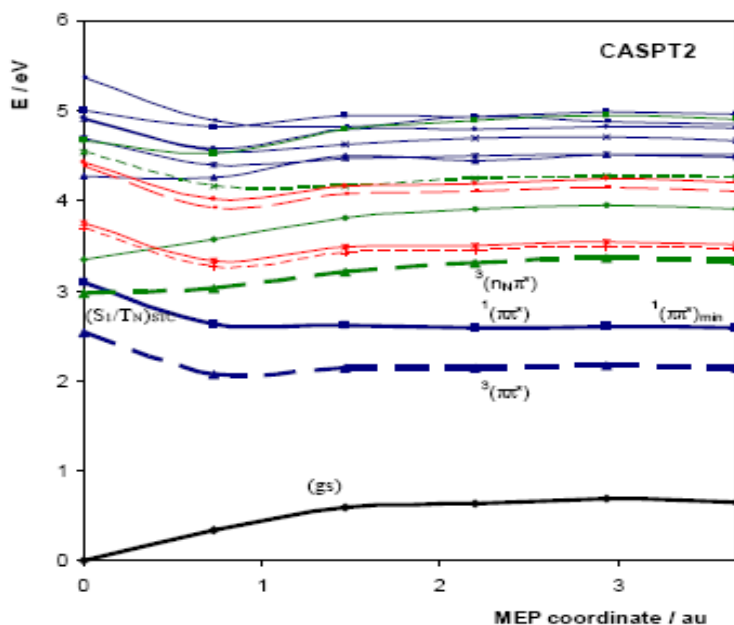


Figura 4.4. Evolución CASPT2 del estado fundamental y estados excitados a lo largo del MEP- S_1 .

En las inmediaciones del $(S_1/T_N)_{STC}$, el mecanismo de ISC está favorecido por la proximidad de los dos estados energéticos y un SOC relativamente

grande. El SOC electrónico calculado en esta región tiene un valor de alrededor de 10 cm^{-1} .

Por tanto, se puede prever que tras la absorción de un fotón, el anillo isoaloxacina, que ha sido promovido al estado excitado singlete más bajo, S_1 , se desactiva, en menos tiempo que el tiempo de resolución experimental, a un intermedio relacionado con T_N . La población de este estado se produce antes de que se alcance el mínimo de S_1 , lo cual es consistente con el hecho de que el rendimiento cuántico de fluorescencia de los compuestos flavínicos es más bajo que el rendimiento cuántico de cruce entre sistemas y de formación del triplete.⁴² Los disolventes polares parece que aumentan los rendimientos cuánticos de ISC y disminuyen el de fluorescencia en isoaloxacinas.^{12,40} A causa de su naturaleza, S_1 y T_N estarían, en un medio polar, estabilizado y desestabilizado, respectivamente, con respecto a los valores *in vacuo*. Esto daría lugar a una disminución de la diferencia de energía vertical, que favorecería que se diera, con mayor probabilidad, un cruce singlete-triplete en la región FC, explicando las tendencias observadas experimentalmente.

En la figura 4.5.a se muestra un esquema con la representación de los principales procesos fotoquímicos que tienen lugar a lo largo de la hipersuperficie de S_1 . En dicha figura se muestran las estructuras de la isoaloxacina con los orbitales π y el par solitario del N_5 involucrados en el cambio de espín. Un análisis comparativo de población de Mulliken, de las estructuras $^1(\pi, \pi^*)_{\text{min}}$ y $^3(n_N, \pi^*)_{\text{min}}$, nos lleva a concluir que el N_5 es el centro directamente implicado en ese cambio de espín.

A partir del $(S_1/T_N)_{\text{STC}}$ se ha estudiado la evolución del estado $^3(n_N, \pi^*)$, siguiendo una estrategia de cálculo similar. Se realizó el MEP a lo largo del estado $^3(n_N, \pi^*)$, T_N (véase figura 4.6) y, posteriormente, a las correspondientes coordenadas del MEP del T_N se calcularon las energías CASPT2 de los estados singlete y triplete. Los resultados de esta parte de la investigación se resumen en la figura 4.5.b.

Una vez el estado T_N ha sido poblado, se lleva a cabo una rápida conversión interna desde este estado a T_1 , mediada por la intersección cónica $(T_N/T_1)_{\text{CI}}$, que se encuentra a 2.5 kcal/mol (0.11 eV) por encima de $^1(\pi, \pi^*)_{\text{min}}$, siendo ésta la estructura de referencia elegida para representar las energías de la figura 4.5.b.

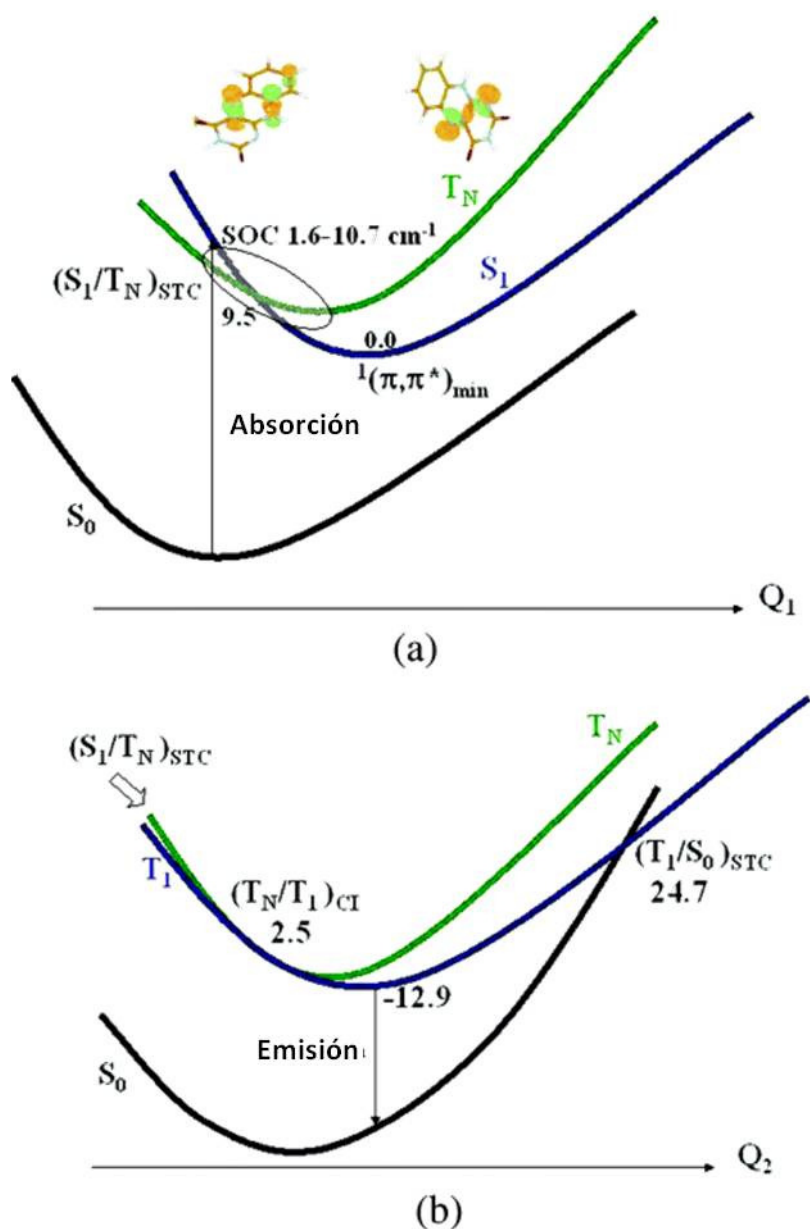


Figura 4.5. Representación esquemática de la fotoquímica de la isoaloxacina después de la absorción de luz y promoción vertical al estado S_1 . (a) Población del estado T_N , cerca de la región FC, por un mecanismo de cruce entre sistemas. (b) Conversión interna hacia el estado triplete, T_1 , mediado por una cónica $(T_N/T_1)_{CI}$ que lleva, en última instancia, a la estructura ${}^3(\pi\pi^*)_{min}$. Las energías se dan en $\text{kcal}\cdot\text{mol}^{-1}$. Q_1 y Q_2 representan las coordenadas de reacción.

El MEP de T_N conduce, en última instancia, a la estructura de equilibrio del estado T_1 , ${}^3(\pi,\pi^*)_{\min}$, localizada 12.9 kcal/mol (0.56 eV) por debajo de la estructura de equilibrio del estado excitado ${}^1(\pi,\pi^*)_{\min}$.

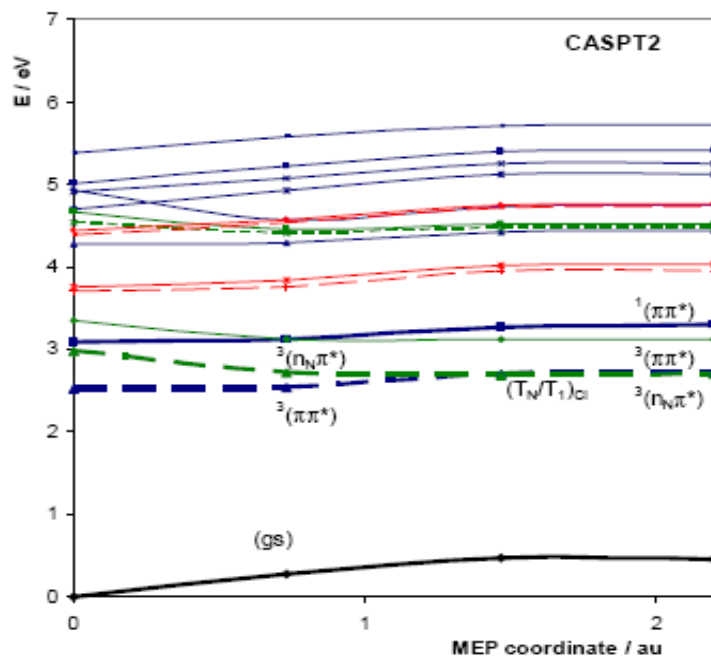


Figura 4.6. Evolución CASPT2 del estado fundamental y estados excitados a lo largo del MEP- T_N desde el cruce $(S_1/T_N)_{STC}$.

Desde el estado relajado T_1 tres eventos fotoquímicos diferentes pueden tener lugar: la emisión radiativa (fosforescencia), la reactividad del estado T_1 de la isoaloxacina con diferentes residuos de la proteína, y el decaimiento no radiativo al estado fundamental a través de un $(T_1/S_0)_{STC}$.

Con el fin de explorar esta última posibilidad, se obtuvo la estructura del cruce $(T_1/S_0)_{STC}$. Con el sistema localizado en el punto ${}^3(\pi,\pi^*)_{\min}$ de la hipersuperficie, la barrera de energía, calculada a nivel CASPT2, que se debe superar para acceder al cruce $(T_1/S_0)_{STC}$ es de 37.6 kcal/mol. Por tanto, la desactivación el estado T_1 de la isoaloxacina al estado fundamental, de forma no radiativa a través del cruce STC (T_1/S_0) , depende, en última

instancia, del exceso de energía vibracional del sistema en este punto; de lo contrario la molécula solamente emite. Como se observa en la tabla 4.2, el tiempo de vida radiativo del estado T_1 se encuentra en el orden de 116 ms, consistente con la naturaleza del estado. Dentro de la fototropina se sabe que, después de la formación de T_1 , el FMN reacciona con el átomo de azufre de la cisteína-39 para formar un aducto, pero para comprender mejor este proceso, se requiere realizar estudios de reactividad del estado T_1 de la isoaloxacina en el interior de la proteína con metodología, por ejemplo QM/MM, que queda fuera de los objetivos del presente trabajo de Tesis.

Conclusiones del capítulo 4

- ✓ Se han estudiado los estados electrónicos excitados del anillo de la isoaloxacina mediante cálculos *ab initio* mecano-cuánticos CASPT2//CASSCF.
- ✓ Las energías de excitación vertical se calcularon a nivel CASPT2 utilizando la geometría optimizada CASSCF del estado fundamental. El estudio engloba a los estados excitados singlete hasta 5.5 eV. Se han caracterizado once estados singlete y cinco estados triplete.
- ✓ El panorama general de los espectros obtenidos muestra un acuerdo satisfactorio con los datos experimentales disponibles en cuanto a: la naturaleza de los estados, las direcciones de polarización de las transiciones y la intensidad relativa de absorción de las principales bandas. Las principales bandas del espectro de absorción, *in vacuo*, se computan a 3.09, 4.28, 4.69, 5.00 y 5.37 eV, sólo ligeramente desplazadas hacia el azul con respecto a los datos experimentales en disolución.
- ✓ La fosforescencia en el anillo de isoaloxacina es debida a una transición de carácter $\pi\pi^*$ calculada a 2.52 eV. El origen de banda se encuentra a 2.03 eV.
- ✓ Hemos analizado con detalle la posibilidad de que el estado triplete más bajo, T_1 , de la isoaloxacina se pueble a lo largo del camino de relajación del estado singlete más bajo, S_1 . Para ello, se han calculado las energías potenciales de las distintas hipersuperficies y los correspondientes acoplamientos espín-órbita entre los estados implicados. El análisis teórico revela que el mecanismo intrínseco para la población fotoinducida de T_1 tiene dos etapas. En primer lugar, en la región cercana a la transición vertical de S_0 - S_1 , se produce un STC singlete-triplete entre S_1 y un estado triplete de tipo $n\pi^*$. En segundo lugar, desde este punto de cruce, y siguiendo la

hipersuperficie del triplete $n\pi^*$ en su camino de relajación, se produce una CI que implica la intersección de los dos estados triplete más bajos y que conduce directamente al mínimo del estado T_1 , principal responsable de la reactividad del anillo de isoaloxacina y, por tanto, de la activación del ciclo fotoquímico de las flavinas.

- ✓ Finalmente, por los presentes resultados *in vacuo*, se puede concluir que la formación del estado triplete, T_1 , es una propiedad intrínseca de la molécula de isoaloxacina. A pesar de que el efecto del entorno proteico aumenta la tasa de ISC,⁴² y que es claramente necesario para la formación de los aductos, el mismo cromóforo se convierte en el actor principal en la fase inicial del fotociclo, mostrando una capacidad implícita para poblar de forma muy eficiente el estado triplete más bajo.

Referencias

1. Voet, D.; Voet, J. *Biochemistry*; John Wiley and Sons, Inc.: New York, 1990.
2. Briggs, W. R.; Beck, C. F.; Cashmore, A. R.; Christie, J. M.; Hughes, J.; Jarillo, J. A.; Kagawa, T.; Kanegae, H.; Liscum, E.; Nagatanni, A.; Okada, K.; Salomon, M.; Rüdiger, R.; Sakai, T.; Zakano, M.; Wada, M.; Watson, J. C. *Plant Cell* **2001**, *13*, 993.
3. Dürr, H.; Bouas-Laurent, H. *Photochromism Molecules and Systems*; Elsevier Science: New York, 2003.
4. Christie, J. M.; Salomon, M.; Nozue, K.; Wada, M.; Briggs, W. R. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 8779.
5. Swartz, T. E.; Corchnoy, S. B.; Christie, J. M.; Lewis, J. W.; Szundi, I.; Briggs, W. R.; Bogomolni, R. A. *J. Biol. Chem.* **2001**, *276*, 36493.
6. Neiss, C.; Saalfrank, P. *Photochem. Photobiol.* **2003**, *77*, 101.
7. Kennis, J. T. M.; Crosson, S.; Gauden, M.; Stokkum, I. H. M.; Moffat, K.; van Grondelle, R. *Biochemistry* **2003**, *42*, 3385.
8. Kottke, T.; Heberle, J.; Hehn, D.; Dick, B.; Hegemann, P. *Biophys. J.* **2003**, *84*, 1192.
9. Scheleicher, E.; Kowalczyk, R. M.; Kay, C. W. M.; Hegemann, P.; Bacher, A.; Fischer, M.; Bittl, R.; Richter, G.; Weber, S. *J. Am. Chem. Soc.* **2004**, *126*, 11067.
10. Crovetto, L.; Braslavsky, E. *J. Phys. Chem. A* **2006**, *110*, 7307 and references cited therein.
11. Kowalczyk, R. M.; Schleicher, E.; Bittl, R.; Weber, S. *J. Am. Chem. Soc.* **2004**, *126*, 11393.
12. Sun, M.; Moore, T. A.; Song, P.-S. *J. Am. Chem. Soc.* **1972**, *94*, 5.
13. Johansson, L. B.-A.; Davidsson, A.; Lindblom, G.; Naqvi, K. R. *Biochemistry* **1979**, *18*, 4249.
14. Andersson, K.; Malmqvist, P.-Å.; Roos, B. O. *J. Chem. Phys.* **1992**, *96*, 1218.
15. Roos, B. O.; Fülcher, M. P.; Malmqvist, P.-Å.; Merchán, M.; Serrano-Andrés, L. Theoretical Studies of Electronic Spectra of Organic Molecules. In *Quantum Mechanical Electronic Structure Calculations with Chemical Accuracy*; Langhoff, S. R. Ed.; Kluwer Academic Publishers: Dordrecht: The Netherlands, 1995; pp 357-438.
16. Roos, B. O.; Andersson, K.; Fülcher, M. P.; Malmqvist, P.-Å.; Serrano-Andrés, L.; Pierloot, K.; Merchán, M. *Adv. Chem. Phys.* **1996**, *93*, 219.

17. Merchán, M.; Serrano-Andrés, L.; Fülcher, M. P.; Roos, B. O. Multiconfigurational Perturbation Theory Applied to Excited States of Organic Compounds; In *Recent Advances in Multireference Methods*; Hirao, K., Ed.; World Scientific Publishing Company: Amsterdam, 1999; Vol. 4, pp 161-195.
18. Merchán, M.; Serrano-Andrés, L. An Initio Methods for excited states. In *Computational Photochemistry*; Olivucci, M. Ed; Elsevier: Amsterdam, 2005; pp 35-92.
19. Serrano-Andrés, L.; Merchán, M. Spectroscopy: Applications. In *Encyclopedia of Computational Chemistry*; Schleyer, P. v. R., Schreiner, P. R., Schaefer, H. F., III, Jorgensen, W. L., Thiel, W., Glen, R. C., Eds.; Wiley: Chichester, U.K. 2004.
20. (a) Serrano-Andrés, L.; Merchán, M.; Nebot-Gil, I.; Roos, B. O.; Fülcher, M. *J. Am. Chem. Soc.* **1993**, *115*, 6184. (b) Roos, B. O.; Malmqvist, P.-Å.; Molina, V.; Serrano-Andrés, L.; Merchán, M. *J. Chem. Phys.* **2002**, *116*, 7526. (c) Serrano-Andrés, L.; Fülcher, M. P.; Roos, B.O.; Merchán, M. *J. Chem. Phys.* **1996**, *100*, 6484.
21. (a) Merchán, M.; Roos, B. O. *Theor. Chim. Acta* **1995**, *92*, 227. (b) Merchán, M.; Roos, B. O.; McDiarmid, R.; Xing, X. *J. Chem. Phys.* **1996**, *104*, 1791.
22. (a) Pou-Amérigo, R.; Merchán, M.; Ortí, E. *J. Chem. Phys.* **1999**, *110*, 9536. (b) Pou-Amérigo, R.; Serrano-Andrés, L.; Merchán, M.; Ortí, E.; Forsberg, N. *J. Am. Chem. Soc.* **2000**, *122*, 6067.
23. Climent, T.; González-Luque, R.; Merchán, M. *J. Phys. Chem. A* **2003**, *107*, 6995.
24. González-Luque, R.; Garavelli, M.; Bernardi, F.; Merchán, M.; Robb, M. A.; Olivucci, M. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 9379.
25. Molina, V.; Merchán, M. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 4299.
26. Merchán M.; Serrano-Andrés, L. *J. Am. Chem. Soc.* **2003**, *125*, 8108.
27. Serrano-Andrés, L.; Merchán, M.; Borin, A. C. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 8691.
28. For reviews of the CASSCF methods, see different contributions in: *Ab Initio Methods in Quantum Chemistry II*; Lawley, K. P., Ed.; Wiley: New York, 1987.
29. Forsberg, N.; Malmqvist, P.-Å. *Chem. Phys. Lett.* **1997**, *274*, 196.
30. Malmqvist, P.-Å. *Int. J. Quantum Chem.* **1986**, *30*, 479.
31. Malmqvist, P.-Å.; Roos, B. O. *Chem. Phys. Lett.* **1989**, *155*, 189.
32. Strickler, S. J.; Berg, R. A. *J. Chem. Phys.* **1962**, *37*, 814.

33. Rubio-Pons, O.; Serrano-Andrés, L.; Merchán, M. *J. Phys. Chem.A* **2001**, *105*, 9664.
34. De Vico, L.; Olivucci, M.; Lindh, R. *J. Chem. Theory Comput.* **2005**, *1*, 1029.
35. Merchán, M.; Serrano-Andrés, L.; Robb, M. A.; Blancafort, L. *J. Am. Chem. Soc.* **2005**, *127*, 1820.
36. Andersson, K.; Barysz, M.; Bernhardsson, A.; Blomberg, M. R. A.; Carissan, Y.; Cooper, D. L.; Cossi, M.; Fülcher, M. P.; Gagliardi, L.; de Graaf, C.; Hess, B.; Hagberg, G.; Karlström, G.; Lindh, R.; Malmqvist, P.-Å.; Nakajima, T.; Neogrády, P.; Olsen, J.; Raab, J.; Roos, B. O.; Ryde, U.; Schimmelpfennig, B.; Schütz, M.; Seijo, L.; Serrano-Andrés, L.; Siegbahn, P. E. M.; Stålring, J.; Thorsteinsson, T.; Veryazov, V.; Widmark, P. O. *MOLCAS 6.0*; Department of Theoretical Chemistry, Chemical Centre, University of Lund: Lund, Sweden, 2004.
37. Veryazov, V.; Widmark, P.-O.; Serrano-Andrés, L.; Lindh, R.; Roos, B. O. *Int. J. Quantum. Chem.* **2004**, *100*, 626.
38. Climent, T.; González-Luque, R.; Merchán, M.; Serrano-Andrés, L. *J. Phys.Chem.* **2006**, *110*, 13584.
39. Wang, M.; Fritchie, J., Jr. *Acta Crystallogr.* **1973**, *B29*, 2040.
40. Visser, A. J. W. G.; Müller, F. *Helv. Chim. Acta* **1979**, *62*, 593.
41. Serrano-Andrés, L.; Forsberg, N.; Malmqvist, P.-Å. *J. Chem. Phys.* **1998**, *108*, 7202.
42. Shüttrigkeit, T. A.; Kompa, C. K.; Salomon, M.; Rüdiger, W.; Michel-Beyerle, M. E. *Chem. Phys.* **2003**, *294*, 501.

PARTE II

Luz ultravioleta y mutaciones genéticas



Nuestra idea era estéticamente elegante. Una estructura tan bella tenía que existir.

James D. Watson.

The Double Helix: A Personal Account of the Discovery of the Structure of DNA, 1968.

5

Introducción y objetivos

Los ácidos nucleicos: Descripción y estructura

Los ácidos nucleicos, ácido desoxirribonucleico (ADN) y ácido ribonucleico (ARN), son polímeros constituidos por largas cadenas en las que se repite una unidad básica formada por un azúcar, un grupo fosfato y una base nitrogenada.

Las bases nitrogenadas del ADN/ARN, también llamadas **bases nucleicas o nucleobases**, son estructuras heterocíclicas y aromáticas derivadas de la pirimidina (bases pirimidínicas) o de la purina (bases púricas). En la figura 5.1 se muestran las tres bases pirimidínicas más comunes uracilo (*uracil*, U), timina (*thymine*, T) y citosina (*cytosine*, C), cuyo esqueleto molecular consiste en un anillo de seis átomos, cuatro de carbono y dos de nitrógeno, y las dos bases púricas adenina (*adenine*, A) y guanina (*guanine*, G), constituidas por un sistema de dos anillos fusionados pirimidina–imidazol. Las bases nucleicas A, G y C se encuentran tanto en el ADN como en el ARN, mientras que la T se encuentra principalmente en el ADN y el U en el ARN.

Las cinco moléculas presentan distribuciones de carga localizada en determinados átomos. Concretamente, los grupos –NH son centros de carga positiva, mientras que los átomos de nitrógeno del anillo y los pares solitarios de electrones del oxígeno de los grupos –C=O son centros de carga negativa, lo que contribuye a que se establezcan uniones entre las bases a través de enlaces de hidrógeno; uniones fundamentales en el mantenimiento de la estructura del ADN/ARN.

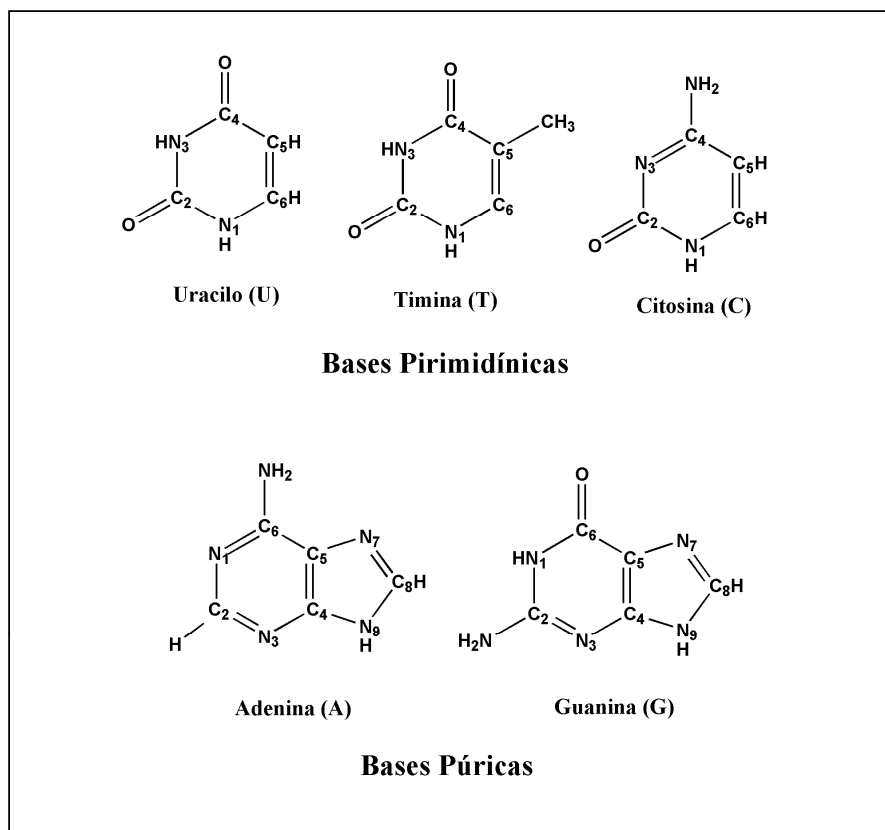


Figura 5.1. Bases pirimidínicas y púricas del ADN/ARN. El convenio de nomenclatura de la Unión Internacional de Química Pura y Aplicada (*International Union of Pure and Applied Chemistry, IUPAC*)¹ establece que sólo se numeran los átomos del anillo aromático, mientras que los sustituyentes adoptan el mismo número que el átomo al que están unidos.

La unión de una base nucleica con un azúcar del tipo pentosa, la D-ribosa en el ARN y la 2'-desoxi-D-ribosa en el ADN, forma un **nucleósido** (véase figura 5.2), cuyo nombre viene determinado por la base nitrogenada en cuestión (adenosina, guanosina, citidina, timidina y uridina).¹

La pentosa se presenta en forma de anillo pentagonal cerrado, o β -furanosa, cuya estructura puede adoptar básicamente dos conformaciones de anillo no rígido: una de tipo sobre (*envelope*) y otra de tipo torsionada (*twist*). Dichas

conformaciones fuera del plano son energéticamente mucho más estables que la forma coplanar, siendo, entre ellas, la forma tipo sobre un poco más estable que la torsionada.¹ Los ángulos diedros internos del anillo de azúcar, tales como el de pseudorrotación (*phase*) y el de distorsión (*puckering*), son algunos de los parámetros empleados para definir los distintos conformeros.²

El enlace entre la base y el azúcar es del tipo N-β-glucosídico entre el C_{1'} de la pentosa y el N₁, en las pirimidinas, o el N₉, en las purinas, formado con eliminación de agua.

La unión, a través del C_{5'}, del azúcar, de un nucleósido con una molécula de fosfato da lugar a un **nucleótido** (véase figura 5.2), cuya nomenclatura¹ es la correspondiente al nucleósido del que procede (adenilato, guanilato, citidilato, timidilato o uridilato). En este caso, se ve ampliado el número de posibles conformaciones, pues a los movimientos del anillo de azúcar se une la rotación alrededor de su enlace con el grupo fosfato.³

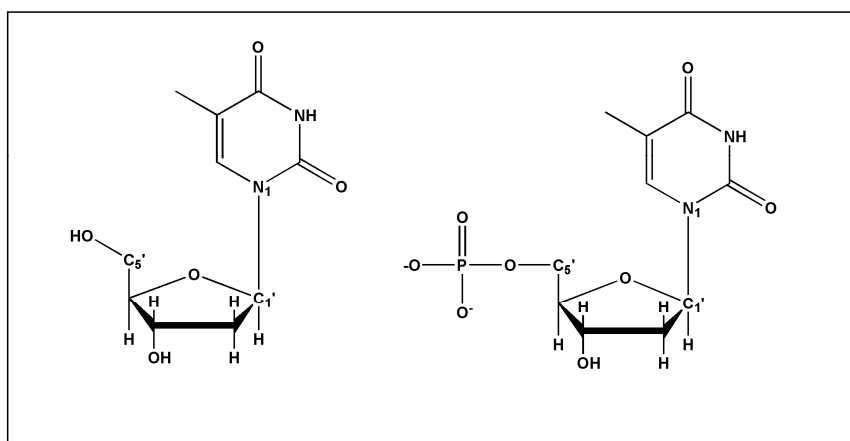


Figura 5.2. Nucleósido (izquierda) y nucleótido (derecha) de timina, donde puede observarse la unión de la base (a través del N₁) con el azúcar (a través del C_{1'}) para formar el nucleósido y la unión del fosfato al azúcar (a través del C_{5'}) para formar el nucleótido.

La unión de los distintos nucleótidos para formar el polinucleótido se produce a través de un enlace covalente fosfodiéster. En él, el grupo

hidroxilo, en posición 5', de uno de los nucleótidos queda unido al grupo hidroxilo, en posición 3', del nucleótido siguiente. Lo que distingue a un nucleótido de otro es únicamente la base nitrogenada y, por ello, la secuencia del ácido nucleico, ADN y ARN, se especifica nombrando sólo la secuencia de sus bases.

Estructura del ADN

En el año 1953, James Watson y Francis Crick publicaron en la revista Nature⁴ uno de los descubrimientos más trascendentales del siglo XX, la estructura de doble hélice del ADN. Propuesta por ellos, pero fruto del esfuerzo de numerosas investigaciones anteriores,^{6,7,8,9,10,11} Watson y Crick tuvieron la genialidad de encajar el puzle que les llevaría a recibir el Premio Nobel de Fisiología o Medicina unos años más tarde. Se basaron, fundamentalmente, en la información de las imágenes de difracción de rayos X proporcionadas por Rosalind Franklin y Maurice Wilkins,⁹ y en el análisis químico realizado por Edwin Chargaff¹² que mostraba, entre otras conclusiones, que en todos los ADN celulares, independientemente de la especie, el número de residuos de adenosina era igual al de residuos de timidina (es decir A=T) y el número de residuos de guanosina era igual al número de residuos de citidina (G=C). A partir de estas relaciones, dedujeron que la suma de los residuos de purina es igual a la suma de los residuos de pirimidina.¹² La conclusión a la que llegaron Watson y Crick fue que el ADN consta de dos cadenas de polinucleótidos enrolladas, cuya estructura de doble hélice se forma al unirse las bases de cadenas opuestas en planos perpendiculares al eje de la hélice, mientras que el esqueleto azúcar-fosfato forma parte del exterior de la misma, tal como se muestra en el esquema de la figura 5.3. Su estudio mostraba que los pares de bases unidos por enlaces de hidrógeno, G con C y A con T, eran los que mejor encajaban en la estructura, proporcionando, de esta forma, una explicación a los análisis de Chargaff. La adenina se aparea con la timina a través de la formación de un par de puentes de hidrógeno (simbolizado como A=T) y la guanina con la citosina a través de tres (G≡C). Los enlaces que unen de forma horizontal las dos cadenas del ADN son enlaces débiles de hidrógeno

entre las bases, pero son las fuerzas de mayor importancia pues se encargan de mantener y estabilizar las estructuras tridimensionales de los ácidos nucleicos, permitiendo, a su vez, que las dos cadenas se puedan separar y abrir como una cremallera. Se ha observado que, cuanto mayor sea la relación de pares de bases $G\equiv C$ con respecto a los de $A=T$, mayor es la dificultad para separar las hebras apareadas del ADN.

Las cadenas de ADN se estabilizan también de forma vertical por el apilamiento, o *stacking*, de las bases. Esto es debido al solapamiento parcial de las nubes electrónicas π de las bases apiladas, más o menos intensa según la secuencia de bases implicadas.

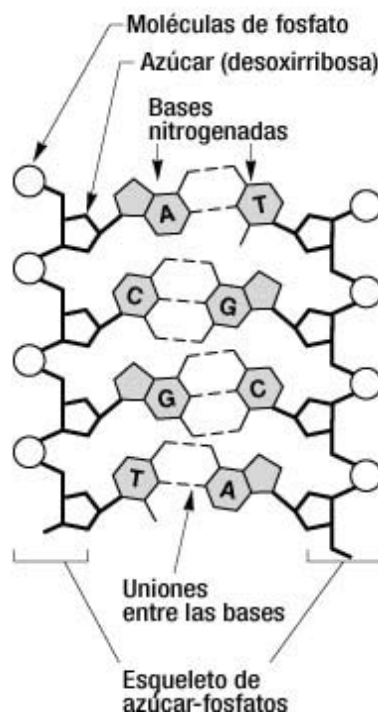


Figura 5.3. Esquema básico de apareamiento de las bases nitrogenadas y esqueleto azúcar-fosfato para formar una hélice de ADN.

La estructura propuesta por Watson y Crick, o forma ADN-B, es la más común en la naturaleza, aunque no es la única, ya que pueden existir otras estructuras secundarias¹³ (véase figura 5.4), entre ellas el ADN-A,^{14,15} que se trata de una doble hélice dextrógira, como el ADN-B. En el caso del ADN-B la estructura helicoidal se repite cada 10 pares de bases, encontrándose las bases adyacentes separadas por una distancia vertical de 3.4 Å, mientras que en el ADN-A hay 11 pares de bases en cada vuelta de hélice.

Otra estructura es el ADN-Z,¹⁶ una doble hélice levógira donde los planos de las bases están en zig-zag, y que se suele dar cuando se suceden secuencias de nucleótidos en las que alternan residuos de citosina y guanina o 5-metil-citosina y guanina.

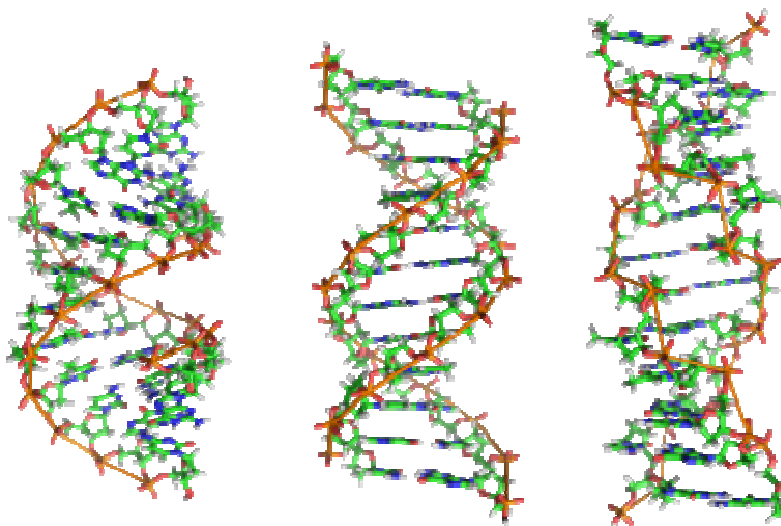


Figura 5.4. De izquierda a derecha se muestran las estructuras del ADN-A, B y Z.

La principal característica del modelo de Watson y Crick es la de la complementariedad de las dos hebras de la doble hélice de ADN. Mucho antes de disponer de pruebas experimentales a su favor, ambos investigadores comprendieron que este hecho podía ser relevante en la

replicación de la estructura, dado que al separarse las dos hebras se podrían sintetizar, de forma natural, nuevas hebras totalmente complementarias a cada una de ellas. Además, en la secuencia de las bases podría residir su importancia como molécula portadora de la información genética, pues cada hebra preexistente haría de molde para dirigir la síntesis de la hebra complementaria.

Actualmente, se sabe que la disposición secuencial de las cuatro bases del ADN contiene toda la información necesaria para la formación de las secuencias de las proteínas y de los ARN. A cada uno de los segmentos de ADN que “codifica” la información para la síntesis de una proteína o de un ARN se le denomina gen. Una célula ordinaria contiene muchos miles de genes y, por tanto, no es sorprendente que las moléculas de ADN sean muy largas. Una vez formada la doble hélice de ADN, ésta se pliega para quedar almacenada en el interior de las células, en formas y tamaños que varían según se trate de organismos procariotas o eucariotas. En los procariotas, y en orgánulos celulares como mitocondrias y cloroplastos, se pliega adoptando una forma de superhélice circular y desnuda, es decir, no asociada a proteínas. En cambio, en los organismos eucariotas el empaquetamiento es más complejo y compacto, y necesita estar unido a proteínas para dar lugar a la cromatina, cuyo último grado de condensación constituye los cromosomas.

Estructura del ARN

Del mismo modo que el ADN, el ARN está presente tanto en las células procariotas como en las eucariotas. Se trata de un polímero formado por la repetición de una unidad básica constituida por un azúcar (D-ribosa), un grupo fosfato y una base nitrogenada. Sus bases púricas son las mismas que en el ADN pero, en las bases pirimidínicas, el uracilo substituye a la timina. Aunque la existencia de diferentes pentosas y la presencia de uracilo en el ARN y de timina en el ADN, son los dos aspectos que distinguen el ADN del ARN, son las pentosas las que definen la identidad de un ácido nucleico. Si el ácido nucleico contiene 2'-desoxi-D-ribosa, es ADN por definición, aunque pueda contener unos cuantos uracilos. Del mismo modo, si el ácido

nucleico contiene D-ribosa, es ARN, independientemente de las bases que lo componen.

En la célula se encuentran varias clases de ARN con distintas funciones. Así los ARN ribosómicos (ARNr) son componentes de los ribosomas, complejos que llevan a cabo la síntesis de proteínas. Los ARN mensajeros (ARNm) actúan de intermediarios, transportando la información desde un gen, o unos pocos genes, hasta el ribosoma donde se sintetizan las proteínas. Los ARN de transferencia (ARNt) son moléculas adaptadoras que traducen con fidelidad la información contenida en el ARNm a secuencias específicas de aminoácidos. También se conoce una amplia variedad de ARN con funciones especiales, incluyendo algunos con actividad enzimática denominados ribozimas. Las diversas, y a menudo complejas, funciones de los ARN son reflejo de una diversidad de estructuras mucho más rica que la observada en el ADN.

A diferencia del ADN, las moléculas de ARN son de cadena simple, donde las monohebras tienden a adoptar una conformación helicoidal dextrógira dominada por las interacciones de apilamiento de bases. La fuerza de la interacción es mayor entre dos purinas que entre una purina y una pirimidina o entre dos pirimidinas. De hecho, la interacción purina-purina es tan fuerte que una pirimidina que se encuentre entre dos purinas suele ser desplazada de la estructura de bases apiladas para que las dos purinas puedan interactuar, formándose protuberancias y bucles internos. Esto hace que las estructuras tridimensionales de muchos ARN sean unas estructuras complejas y únicas.

El ARN puede formar pares de bases con regiones complementarias del propio ARN o de un ADN. En la figura 5.5, izquierda, se muestra un esquema del plegado donde se observa que la complementariedad de bases se produce al aparearse la adenina con el uracilo ($A=U$) y la guanina con la citosina ($G\equiv C$).¹⁷ En la figura 5.5, derecha, se muestra la estructura de uno de los tipos de ARN, concretamente el de transferencia, ARNt, que posee más de un 50% de sus bases apareadas en cuatro brazos con estructura de doble hélice. En disolución, estos brazos están plegados en forma de "L" compacta debido a la estabilización que se produce por apareamientos de Watson y Crick convencionales ($A=U$, $C\equiv G$), y a las interacciones entre las bases de dos o más nucleótidos.¹⁸

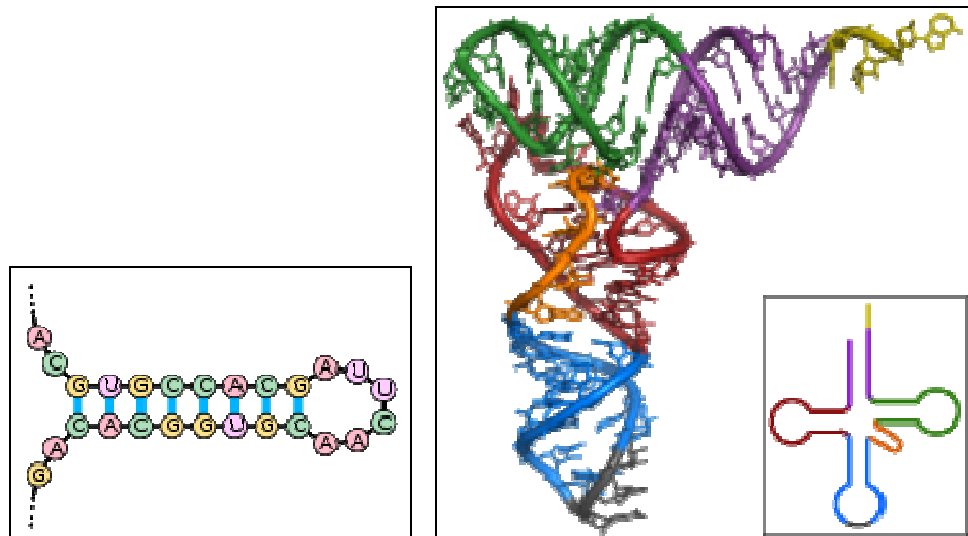


Figura 5.5. A la izquierda, estructura secundaria del ARN. A la derecha, estructura terciaria del ARN de transferencia (ARNt).

Implicación de la radiación UV en la formación de biodímeros

Nuestra estructura genética sufre alteraciones espontáneas, normalmente muy lentas pero fisiológicamente significativas, pues dan lugar a cambios permanentes en la información codificada en el ADN. Dichas alteraciones se denominan mutaciones, siendo la acumulación de éstas las que se han relacionado con procesos de envejecimiento y cáncer.

Uno de los procesos lentos de mutación espontánea del ADN es la desaminación de una base nucleotídica como puede ser la desaminación de la citosina para dar lugar a uracilo (véase figura 5.6), que se produce en aproximadamente uno de cada 107 residuos de citosina cada 24 horas. Ello equivale, en promedio, a unas 100 mutaciones espontáneas por día en una

célula de mamífero.♦

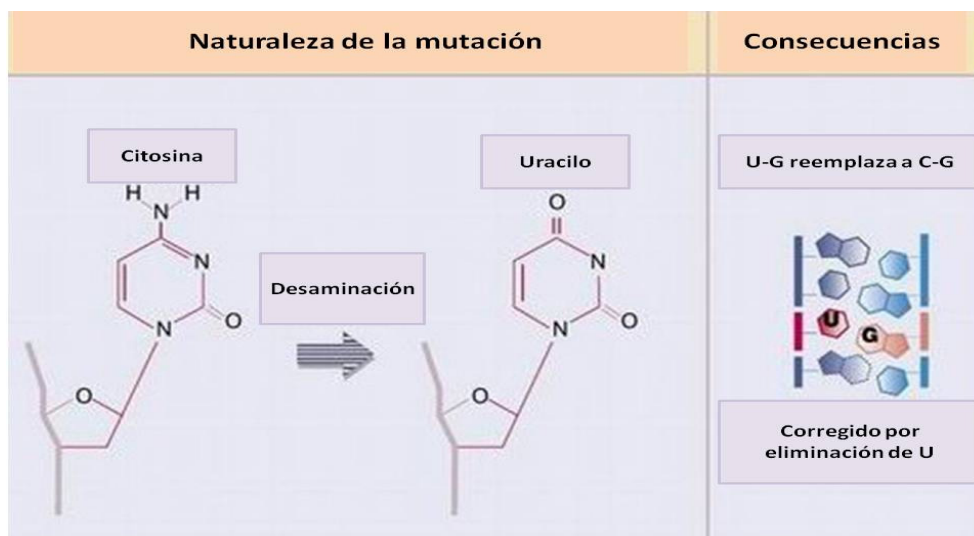


Figura 5.6. Ejemplo de uno de los muchos daños espontáneos que sufre el ADN. En este caso la desaminación de la citosina dando lugar a uracilo.

El ADN también puede sufrir daños causados por reactivos químicos que pueden no ser peligrosos en sí mismos, pero su metabolización por las células los convierte en productos que sí los son.

Otras reacciones que pueden producir lesiones en el ADN están favorecidas por ciertos tipos de radiación. Radiaciones ionizantes, como los rayos gamma y los rayos X, que nos llegan a través de la radiación cósmica y exámenes médicos o dentales; o por radiaciones no ionizantes, siendo la radiación UV uno de los agentes responsable de las mutaciones que

♦ La lenta reacción de la desaminación de la citosina parece inocua a primera vista, pero es, casi con toda seguridad, la causa de que el ADN contenga timina en lugar de uracilo. El producto de la desaminación de la citosina (uracilo) se reconoce rápidamente como extraño y se elimina por un sistema de reparación enzimático. Si el ADN contuviera normalmente uracilo, el reconocimiento de los uracilos procedentes de la desaminación de la citosina sería más difícil y los uracilos no reparados darían lugar a cambios permanentes en la secuencia al aparearse con adeninas en la replicación. Fuente: Lehninger. Principios de Bioquímica. D. Nelson y M. Cox. Quinta edición. 2009.

desembocan en distintos tipos de cáncer inducidos por la luz solar. Estas mutaciones se producen a pesar de que la estructura intrínseca de las nucleobases parece haber sido seleccionada con cuidado para que constituyan sistemas con la mínima respuesta posible a la radiación.[♥]

El espectro de la radiación UV^{21,22} se subdivide, en función de la longitud de onda, en tres segmentos designados como UV-A (400–320 nm), UV-B (320–295 nm) y UV-C (295–100 nm). La radiación UV solar que llega a la superficie terrestre está compuesta básicamente por UV-A y UV-B, pues la UV-C queda filtrada en la atmósfera, principalmente por la capa de ozono.

La radiación UV, bien sea tipo B, en condiciones naturales, o tipo C, en circunstancias artificiales (lámparas germicidas, por ejemplo), puede producir un gran número de reacciones en el material genético celular que, en la mayoría de los casos, corresponden a dimerizaciones, o formaciones de aducto,[▲] y que afectan sólo a las bases pirimidínicas adyacentes en una misma cadena de ADN/ARN^{21,22,23,24,25} para formar bipyrimidinas.^{26,27}

Entre estos fotoproductos diméricos, cuyas estructuras aparecen esquematizadas en las figuras 5.7 y 8.1, se encuentran frecuentemente dos tipos:

- Los dímeros de ciclobutilpirimida (*Cyclobutane Pyrimidine Dimers*, CPD, CBP, CBPyr, Pyr\leftrightarrowPyr), en los cuales el aducto se forma por fotocicloaddición [2+2] involucrando a los dobles enlaces C=C de las pirimidinas (véase figura 5.7.a).
- Los fotoproductos (6-4), pirimidina-(6-4)-[pirimidín-2'-ona], en los cuales la cicloaddición tiene lugar entre los dobles enlaces de los C=C y C=O de moléculas adyacentes (véase figura 5.7.b).

De estos dos tipos de fotoproductos, los que se producen *in vivo*, con un mayor rendimiento, son los de tipo CPDs.^{21,23}

[♥] Se ha propuesto^{19,20} que las cinco bases nucleicas se desactivan, una vez excitadas por la radiación UV, a través de un mecanismo ultrarrápido, tal como se verá en el capítulo 6, que disipa la energía absorbida antes de que se produzcan fotorreacciones indeseables.

[▲] El término aducto surge de la combinación de las palabras adición y producto. El aducto AB está formado por la unión directa de dos moléculas A y B, sin que se produzcan cambios estructurales en su topología. Los productos de una ciclación son aductos en los que dos enlaces π se pierden para formar dos enlaces σ , como por ejemplo, la condensación de dos grupos etileno para formar un anillo de ciclobutano. Fuente: <http://es.wikipedia.org/>

A pesar de que las bases pirimidínicas pueden reaccionar para formar CPDs en diversas combinaciones, la combinación TT, es la que se observa con mayor frecuencia entre residuos adyacentes en la misma hebra de ADN mientras que la lesión CC resulta ser la más mutagénica.²⁴

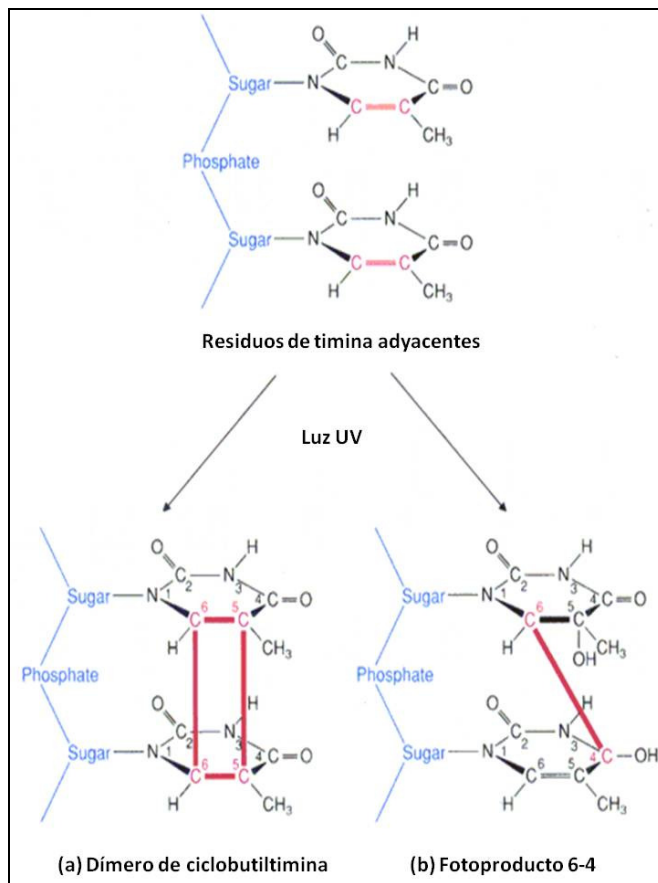


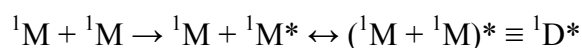
Figura 5.7. Forma esquemática de dos tipos de dímero de timina que pueden formarse en la doble cadena de ADN por acción de la radiación UV. En (a) puede verse el dímero de ciclobutilpirimidina de timina (*Cyclobutane Thymine Dimer*, CTD). En (b) puede verse un fotoproducto pirimidina–(6-4)–pirimidona, o fotoproducto 6-4, aducto simbolizado por T(64)T.

Excímeros, las moléculas responsables de la formación de fotoproductos: definición y caracterización

El término excímero^{17,28,29} hace referencia a un dímero (D) entre moléculas idénticas cuyo estado electrónico excitado es enlazante, es decir, en él los monómeros (M) se encuentran unidos, mientras que su estado fundamental es disociativo, y por tanto, sus monómeros, en ausencia de restricciones externas, se encuentran separados.

En el caso de un heterodímero los dos monómeros interaccionantes son de distinta naturaleza química, y el complejo formado se conoce como exciplejo.^{30,31}

El excímero constituye una entidad física, y no química, puesto que sólo existe en el estado excitado, y al disiparse la energía el sistema revierte al par de monómeros separados en su estado fundamental. En la práctica, puede ser visto como un dímero en el cuál uno de los monómeros está excitado electrónicamente y disocia inmediatamente tras la desexcitación:



Se da el caso de que en algunas moléculas aromáticas la interacción de tipo excímero es tan fuerte que a partir del excímero, ${}^1D^*$, se forma un fotodímero estable, fenómeno conocido como fotodimerización.³²

Cuando se habla de excímero, sin hacer ningún tipo de especificación, se está haciendo referencia al excímero singlete, ${}^1D^*$, es decir, al estado excitado singlete enlazante más bajo en energía. Sin embargo, la colisión entre una molécula en estado excitado triplete, ${}^3M^*$, y otra no excitada, 1M , puede formar un excímero triplete, ${}^3D^*$, dando lugar a un estado excitado triplete, en el que los monómeros están enlazados:



Los estudios experimentales han mostrado que la capacidad de formar excímeros es común en la mayoría de las moléculas aromáticas planas. Por tanto, no es de extrañar que la naturaleza, un tanto inusual, de este tipo de

enlace molecular sea de gran interés para la comprensión de los procesos fotoquímicos y fotofísicos en los compuestos aromáticos.

El hecho de que se forme un excímero supone una vía competitiva al decaimiento radiativo del monómero y, por ello, la fluorescencia detectada experimentalmente se verá disminuida respecto a la del monómero. Las moléculas aromáticas que forman excímeros presentan además una banda de emisión por fluorescencia desestructurada y desplazada hacia frecuencias más bajas respecto a la fluorescencia del monómero; mientras que sus espectros de absorción no se ven afectados.

La presencia de excímeros en sistemas biológicos, o bioexcímeros, ha sido sugerida en el ADN en varios experimentos,^{33,34,35,36,37} y su formación e importancia en la reactividad de los ácidos nucleicos se va a demostrar en los estudios realizados en la presente Tesis, siempre desde un punto de vista teórico.

Objetivos

Se describen a continuación los objetivos de esta segunda parte de la presente Tesis doctoral.

- ✓ Presentar una visión unificada para explicar los principales sucesos fotoquímicos en las tres bases pirimidínicas del ADN/ARN, uracilo, timina y citosina, que justifiquen la gran fotoestabilidad de estos sistemas moleculares tras ser irradiados con luz UV. Se pretende descubrir la presencia o no de barreras energéticas a lo largo de su camino de desactivación, y para ello se describen de forma precisa los perfiles de energía de dichos caminos de reacción. Se presentan los resultados de esta parte de la investigación en el capítulo 6.
- ✓ Se sabe que los estados tripletes de los sistemas moleculares son intermedios frecuentes en muchas reacciones fotoinducidas, ya que el hecho de poseer un marcado carácter birradicalario y unos tiempos de vida media relativamente largos, les hace ser especies candidatas proclives a interactuar con otros compuestos. Los estados tripletes de las bases del ADN/ARN no son una excepción y participan activamente en las fotorreacciones promovidas por la luz UV. Por ello, se pretende explorar el mecanismo de población eficiente del estado triplete más bajo –altamente implicado en los procesos fotoquímicos de formación de dímeros ciclobutánicos– y presentar, si es posible, un esquema unificado, desde el punto de vista químico–cuántico, en las tres bases pirimidínicas. Se presentan los resultados de esta parte de la investigación en el capítulo 7.
- ✓ A pesar de la gran fotoestabilidad demostrada por las bases nucleicas, que las ha hecho las perfectas candidatas a formar parte del material genético a lo largo de la evolución, y a pesar de los mecanismos de autorreparación enzimática permanente, que mantienen la estructura genética protegida del ataque de la radiación,

sabemos que se producen mutaciones genéticas inducidas por la luz solar, que llegan incluso a desembocar en lesiones graves de dicho material genético. Muchas de estas lesiones corresponden a dimerizaciones de las bases adyacentes pirimidínicas en el ADN/ARN empaquetado, del tipo ciclobutilpirimidinas. Nuestro objetivo es el de profundizar en los mecanismos de formación de los dímeros de ciclobutilpirimidinas en el estado triplete para los tres homodímeros de las bases nucleicas, Uracilo–Uracilo, Timina–Timina y Citosina–Citosina. Se presentan los resultados de esta parte de la investigación en el capítulo 8.

Referencias

1. IUPAC-IUB Joint Commission on Biochemical Nomenclature *Eur. J. Biochem.* **1983**, *131*, 9.
2. Hall, L. D. *Chem. Ind.* **1963**, 950.
3. Altona, C.; Sundaralingam, M. *J. Am. Chem. Soc.* **1972**, *94*, 8205.
4. Seeman, N. C.; Rosenberg, J. M.; Suddath, F. L.; Kim, J. J. P.; Rich, A. J. *Mol. Biol.* **1976**, *104*, 142.
5. Watson, J. D.; Crick, F. H. C. *Nature.* **1953**, *171*, 737.
6. Levene, P. J. *Biol Chem.* **1919**, *40*, 415.
7. Astbury, W. T. *Symp. Soc. Exp. Bio.* **1947**, *1*, 66.
8. Avery, O; MacLeod, C; McCarty, M. *J Exp Med*, **1944**, *79*, 137.
9. Franklin, R. E. *Nature.* **1953**, *171*, 740.
10. Wilkins, M. H. F.; Stokes, A. R.; Wilson, H. R. *Nature* **1953**, *171*, 738.
11. Dahm, R. *Hum Genet.* **2008**, *122*, 565.
12. Chargaff, E. *Biochim. et Biophys. Acta* **1952**, *9*, 402.
13. Ghosh, A.; Bansal, M. *Acta Crystallogr D Biol Crystallogr.* **2003**, *59*, 620.
14. Wahl, M.; Sundaralingam, M. *Biopolymers* **1997**, *44*, 45.
15. Lu, X. J.; Shakked, Z.; Olson, W. K. *J. Mol. Biol.* **2000**, *300*, 819.
16. Rothenburg, S.; Koch-Nolte, F.; Haag, F. *Immunol Rev.* **2001**, *184*, 286.
17. Devlin, T. M. *Bioquímica*, 4ª ed; Reverté: Barcelona, 2004.
18. Lee, J.C.; Gutell, R.R. *J. Mol. Biol.* **2004**, *344*, 1225.
19. Serrano-Andrés, L; Merchán, M. "*Vida y Luz: Una Perspectiva Químico-Cuántica*". *Anales de Química*, **2004**, *100*, 16.
20. Serrano-Andrés, L.; Merchán, M. *J. Photochemistry and Photobiology C: Photochemistry Reviews.* **2009**, *10*, 21.
21. Friedberg, E. C.; Walker, G. C.; Siede, W.; Wood, R. D.; Schultz, R. A. *DNA Repair and Mutagenesis*; Ellenberger, T. ed.; ASM Press: Washington DC, 2006.
22. Björn, L. O.; McKenzie, R.L. *Photobiology. The Science of Life and Light*; Björn L. O. ed.; Springer: New York, 2008, 503.
23. Cadet, J.; Vigny, P. *Bioorganic Photochemistry*; Morrison, H. Ed.; John Wiley & Sons: New York, 1990, Vol. 1, pp 1-272.
24. Douki, T.; Cadet, J. *Biochemistry*, **2001**, *40*, 2495.
25. Schuch, A. P.; Menck, C. F. M. *J Photochem Photobiol B Biol.* **2010**, *99*, 111.
26. Delaney, S.; Barton, J. K. *J. Org. Chem.* **2003**, *68*, 6475.

27. Rajski, S. R.; Jackson, B. A.; Barton, J. K. *Mutat. Res.* **2000**, 447, 49.
28. Birks, J. B. *Rep. Prog. Phys.* **1975**, 38, 903.
29. Stevens, B.; Hutton, E. *Nature*, **1960**, 186, 1045.
30. Birks, J. B. *The Exciplex*; Academic Press: New York, San Francisco, London, 1975.
31. Birks, J. B. *Molecular Luminescence*; Lim, E. C. ed; Benjamin Inc.: New York, 1969, 907.
32. Calas, R.; Lalande, R. *Bull. Soc. Chim. France.* **1959**, 763 y 770.
33. Plessow, R.; Brockhinke, A.; Eimer, W.; Kohse-Höinghaus, K. *J. Phys. Chem. B* **2000**, 104, 3695.
34. Danilov, V. I.; Slyusarchuk, O. N.; Alderfer, J. L.; Stewart, J. J. P.; Callis, P. R. *Photochem. Photobiol.* **1994**, 59, 125.
35. Crespo-Hernández, C. E.; Cohen, B.; Kohler, B. *Nature* **2005**, 436, 1141.
36. Kwok, W.-M.; Ma, C.; Phillips, D. L. *J. Am. Chem. Soc.* **2006**, 128, 11894.
37. Eisinger, J.; Guéron, M.; Shulman, R. G.; Yamane, T. *Proc. Natl. Acad. Sci. USA.* **1966**, 55, 1015.

6 Decaimiento ultrarrápido de las bases pirimidínicas

Como ya se ha mencionado en el capítulo 5, la estructura intrínseca de las bases nucleicas del ADN/ARN parece haber sido seleccionada con cuidado para que constituyan sistemas con la mínima respuesta posible a la radiación.[♦] Después de la absorción de luz UV, las nucleobases presentan una gran fotoestabilidad que las protege del daño causado por la radiación y les hace ser los constituyentes óptimos del código genético.¹

Desde que por primera vez se publicó la desactivación por fluorescencia en los monómeros de las nucleobases del ADN a temperatura ambiente² se han producido avances en las técnicas experimentales que han mejorado nuestro conocimiento en la dinámica del ADN fotoexcitado. Por ejemplo, experimentos realizados en las bases aisladas han detectado que existen canales de decaimiento en intervalos de tiempo comprendidos entre los femtosegundos, fs, y picosegundos, ps.^{1,3,4,5} La ausencia de efectos acusados cuando los mismos experimentos se realizan en disolución y la detección de caminos ultrarrápidos similares en los nucleósidos y nucleótidos, sugiere que los tiempos de vida ultracortos son propiedades moleculares intrínsecas de las bases nucleicas, íntimamente relacionadas con los procesos dinámicos no adiabáticos que conducen a una relajación no radiativa hacia el estado fundamental.¹

En el capítulo 1 se ha descrito que los decaimientos no radiativos de los estados electrónicos ocurren en procesos de conversión interna, cuya eficiencia está asociada con la presencia de cruces entre las diferentes PHEs

[♦] Véase referencias 29 y 30 del capítulo 5.

en regiones o uniones de puntos de CIs que se comportan como embudos. En ellos, la probabilidad de que se produzcan saltos no adiabáticos, no radiativos, es alta⁶. La interpretación teórica de las evidencias experimentales se lleva a cabo con descripciones químico-cuánticas, basadas en cálculos de obtención de MEPs, optimización de geometrías y localización de CIs, que permiten establecer los principales aspectos del proceso fotoquímico.

Los rendimientos cuánticos de fluorescencia extremadamente bajos encontrados en las nucleobases aisladas,^{2,7,8,9} junto con la energía de relajación ultrarrápida, detectada experimentalmente en el orden de los femtosegundos, sugieren que el sistema encuentra un camino muy eficiente para desexcitar el estado singlete $\pi\pi^*$, S_1 , que ha sido el inicialmente poblado por absorción directa, y que regresa al estado fundamental, S_0 , a través de una CI que involucra a ambos estados, S_1 y S_0 .

Por otro lado, estudios químico cuánticos *ab initio* en las PEHs de los monómeros de las nucleobases de pirimidina¹⁰⁻¹⁶ han descrito CIs representando cruces de la PEH del estado S_1 , tanto de naturaleza $\pi\pi^*$ como $n\pi^*$, con la PEH del estado S_0 , e incluso CIs involucrando tres estados en la citosina¹⁷ y el uracilo.¹⁸ Pero el problema fotoquímico no puede ser entendido solamente describiendo las intersecciones cónicas entre los distintos estados a distintos niveles de cálculo. Hoy en día, se está haciendo evidente que se pueden encontrar un gran número de CIs para sistemas multidimensionales en regiones de PEHs con estructuras de estado electrónico complejo donde la degeneración es frecuente.^{19,20} Por tanto, tan importante como obtener la energía y localización de una CI, es determinar cuan accesible ésta puede ser a lo largo del camino de relajación energética; en este caso desde el estado $\pi\pi^*$ inicialmente poblado. La clave se encuentra, pues, en las barreras energéticas encontradas a lo largo de la evolución de los distintos estados implicados.

En un estudio CASPT2 del tautómero 9H de la molécula de adenina²¹ se encontró que el MEP a lo largo del estado $^1(\pi\pi^* \text{ HL})$, partiendo de la geometría del estado fundamental FC, evoluciona, sin barrera de energía, hacia la intersección cónica $(gs/\pi\pi^*)_{CI}$, involucrando la torsión del enlace N_1-C_2 y el enlace perpendicular C_2-H en el anillo de seis átomos (véase

figura 5.1). En estudios similares, la desactivación de la adenina se encontró también sin barrera,²²⁻²⁵ incluso cuando en lugar de MEPs se realizaron optimizaciones de geometría e interpolaciones lineales entre la geometría FC y la estructura de la intersección cónica.²²⁻²⁴ En estudios de las bases pirimidínicas realizados con pares guanina–citosina,²⁶ así como en las bases independientes citosina,^{11,14} timina¹³ y uracilo,¹⁵ se identificaron CIs similares consistentes en una torsión del enlace etilénico C₅–C₆. Pero no hay evidencia firme y común de cuáles son los perfiles del camino para acceder a la CI y no hay indicaciones claras del papel que juegan en la fotofísica del sistema el mínimo del estado S₁ y otras intersecciones cónicas, (gs/ππ*)_{CI} y (gs/nπ*)_{CI}, encontradas como cruces de las PHEs (S₀/S₁)_{CI},^{10,12} donde el origen de la débil fluorescencia medida para el sistema en medio acuoso^{1,2,7,8,9,16} debería tenerse en cuenta en el esquema.

En el presente capítulo de esta Tesis se han realizado cálculos multiconfiguracionales CASPT2//CASSCF de los estados singletes más bajos de las nucleobases pirimidínicas, uracilo, timina y citosina, para elaborar un modelo unificado que explique la forma de poblar y desactivar el estado excitado singlete más bajo ¹(ππ* HL) HOMO → LUMO en las tres nucleobases, así como la característica más representativa de su fotoquímica, es decir, el decaimiento ultrarrápido responsable de su fotoestabilidad.

Detalles computacionales

Todos los cálculos se han llevado a cabo usando el protocolo CASPT2//CASSCF en el que la optimización de geometrías, cruces de hipersuperficies y MEPs se han realizado a nivel CASSCF, y los cálculos de energía electrónica a nivel CASPT2.²⁷⁻³⁰ Se han probado diferentes espacios activos para asegurar la convergencia y confianza de los resultados. Los resultados usados en la discusión para sustentar las conclusiones se basan en optimizaciones de geometrías de los mínimos de los estados, caminos de mínima energía, y búsqueda de intersecciones cónicas entre el estado fundamental y los estados excitados ππ* más bajos energéticamente a nivel

CASSCF (10,8), con un espacio activo que incluye todos los electrones y orbitales moleculares $\pi\pi^*$. En el caso del uracilo, los puntos estacionarios (mínimos, puntos del MEP y STCs) se obtuvieron a nivel CASSCF, excepto para el mínimo del estado $^1(\pi\pi^*)$ que se realizó a nivel CASPT2. La estabilidad de los resultados obtenidos fue corroborada al ir incrementando los espacios activos, tal como se explicará en el texto. En particular, los resultados finales incluyen caminos de mínima energía calculados al añadir tres orbitales π^* de correlación extras, CASSCF(10,11) al espacio π de valencia. Las determinaciones de la geometría optimizada de los estados $^1(n\pi^*)$, donde n_O describe al MO del par solitario localizado en el átomo de O_4 (véase figura 5.1) se llevaron a cabo usando una función de onda CASSCF(12,9), que añade el correspondiente par solitario al espacio activo π completo (10,8).

En los cálculos finales CASPT2, para cada una de las estructuras optimizadas a nivel CASSCF, se usó el conjunto de base ANO-S C, N, O [3s2p1d]/H [2s1d]³¹ y un espacio activo de 14 electrones distribuidos en 10 orbitales que incluye los pares solitarios, quedando descrito de la siguiente forma: CASPT2//CASSCF(14,10)/ANO-S C, N, O [3s2p1d]/H [2s1d]. En los cálculos CASPT2 se empleó el Hamiltoniano de orden cero estándar,³² y para evitar la presencia de estados intrusos se incluyó una corrección de 0.2 au en el *imaginary level-shift*. La exactitud en esta estrategia computacional ha sido demostrada previamente en diferentes estudios.^{10,21,28-30,33-37} El paquete de programa MOLCAS 6.0³⁸ es el que se empleó en todos los cálculos realizados.

Modelo unificado sin barrera energética

En la figura 6.1 se resumen mediante un esquema los resultados CASPT2 obtenidos para las tres bases pirimidínicas del ADN/ARN. La química fotoinducida más relevante para las nucleobases aisladas empieza con la absorción de energía del sistema que le conduce al estado singlete $^1(\pi\pi^* \text{ HL})$, calculado verticalmente a 5.02 (en el caso del uracilo), 4.89 (T) y 4.41 eV (C), con una fuerza de oscilador cerca de 0.2 (0.09 para C), en

concordancia con los datos teóricos previos y experimentales en fase gas, 5.1 (U), 4.8 (T) y 4.6 eV (C).^{1,39,40} La absorción directa de los estados próximos $n\pi^*$ se encuentra esencialmente prohibida, por tanto, dichos estados se poblarán sólo a través de los procesos de conversión interna que suceden a lo largo del camino de relajación del estado $^1(\pi\pi^* \text{ HL})$, no esperándose que jueguen un papel importante en el proceso fotoquímico primario.

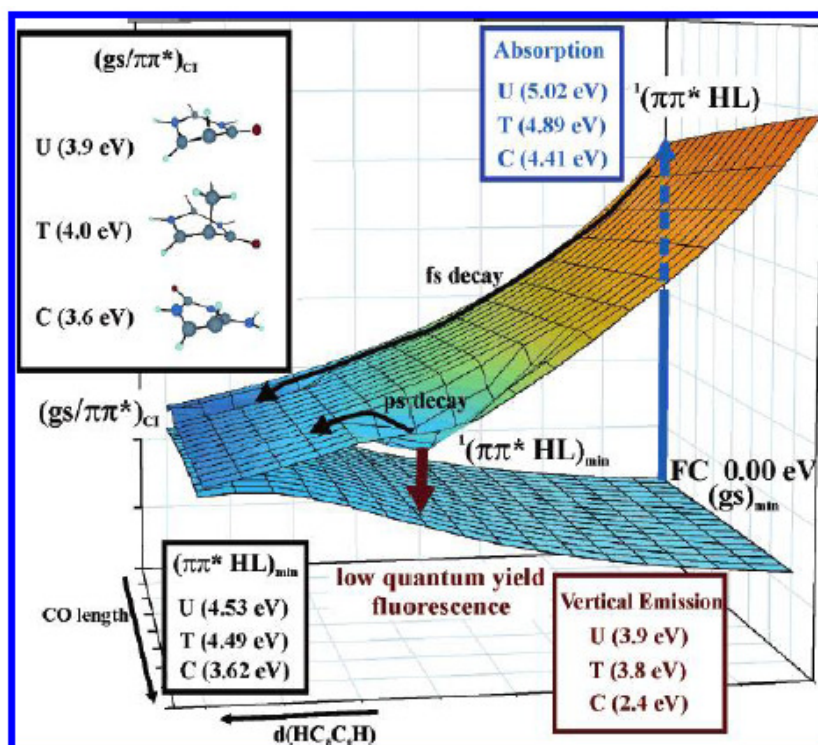


Figura 6.1. Esquema global de la fotoquímica del uracilo (U), timina (T) y citosina (C) sugerido por cálculos CASPT2.

Después de la fotoexcitación, el camino de relajación a lo largo del estado $^1(\pi\pi^* \text{ HL})$ en los tres sistemas conduce, sin barrera energética, a una intersección cónica $(gs/\pi\pi^*)_{CI}$ del tipo eteno, en la que los átomos de

hidrógeno (o de carbono en T) unidos al C₅ adquieren un ángulo cercano a 120°. Esta CI se encuentra a 3.9 (U), 4.0 (T) y 3.6 eV (C) adiabáticamente desde el mínimo del estado fundamental. La mayor parte de la energía absorbida decaerá, por tanto, no radiativamente al estado fundamental a través del embudo de la intersección cónica, $(gs/\pi\pi^*)_{CI}$, en un proceso de relajación ultrarrápido, que asignamos al tiempo del decaimiento medidos experimentalmente en 130 fs (C), en 150 fs (T) y en 160 fs (C).⁴

A continuación se detalla cómo se ha obtenido el perfil de decaimiento de cada una de las bases.

Uracilo

A día de hoy, solamente los métodos multiconfiguracionales son capaces de describir correctamente las intersecciones cónicas y las barreras energéticas encontradas en los caminos de desexcitación de los estados excitados de las nucleobases. Sin embargo, hay que tener presente que la utilización de estos métodos no garantiza la correcta descripción de los diferentes caminos que puede experimentar el sistema, pues su descripción depende de la cantidad de correlación energética aportada por cada uno. Por ejemplo, en el U una optimización geométrica CASSCF/6–31G(d,p) (espacio activo π de valencia) del estado $^1(\pi\pi^* HL)$ realizada a partir de la geometría FC, lleva directamente a una estructura plana. Sin embargo, cuando la misma optimización geométrica se realiza tras la adición de tres orbitales de correlación activos extra, se alcanza directamente la CI del tipo eteno, $(gs/\pi\pi^*)_{CI}$.

La figura 6.2 muestra la evolución del estado fundamental y los dos estados excitados singlete más bajos de la molécula de U desde la geometría FC a lo largo del MEP $^1(\pi\pi^* HL)$ calculado a nivel CASSCF(10,11)/ANO–S, con un espacio activo que incluye la adición de tres orbitales extra de correlación requeridos para describir adecuadamente el carácter “*steepest descent*” del MEP (cuando el cálculo se realizó empleando bases 6–31 G(d) se llegó a la misma conclusión). A cada punto del MEP se calcularon las energías CASPT2//CASSCF(14,10)/ANO–S para los estados de interés.

El perfil del camino de desactivación del estado espectroscópico $^1(\pi\pi^* \text{ HL})$ (adiabáticamente S_1) en el uracilo se considera, pues, sin barrera energética hacia la intersección cónica tipo eteno $(gs/\pi\pi^*)_{CI}$, cónica relacionada con el decaimiento medido en femtosegundos y que representa el camino más efectivo para la desactivación energética en las nucleobases.

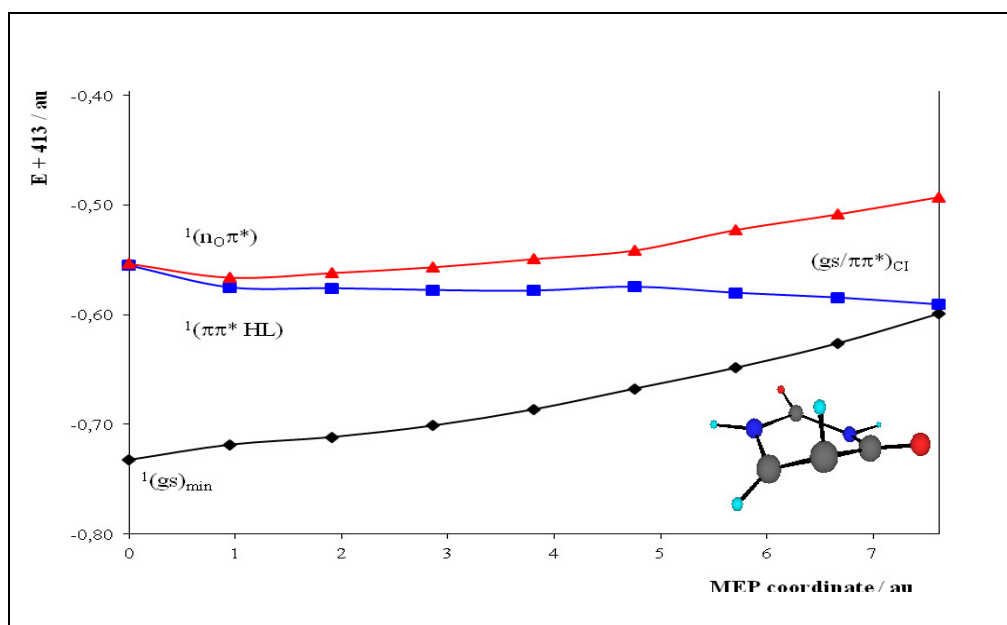


Figura 6.2. Evolución del estado fundamental y estados excitados más bajos a lo largo del MEP $^1(\pi\pi^* \text{ HL})$ desde la geometría FC del uracilo.

Timina

En el caso de la timina, tal como ocurre en el uracilo, una optimización geométrica CASSCF/6–31G(d,p) del estado $^1(\pi\pi^* \text{ HL})$ realizada tomando como punto de partida la geometría FC, con espacios activos reducidos, hacía imposible que el sistema evolucionara hacia la cónica tipo eteno, desembocando, sin embargo, en un mínimo plano. En este sentido, la adición de orbitales de correlación extra mejora el nivel de tratamiento

CASSCF, llevando a resultados coincidentes con los obtenidos con métodos más correlacionados como es el MRCI.¹⁵

La mejora de la base monoeléctronica también apunta en la misma dirección en la timina. Las optimizaciones de geometría CASSCF para el estado más bajo $^1(\pi\pi^* \text{ HL})$ usando el espacio activo π de valencia y bases 6–31G(d) ó 6–31G(d,p) conducen directamente desde la región FC al mínimo plano del estado $^1(\pi\pi^* \text{ HL})$, mínimo localizado en una región totalmente diferente de la hypersuperficie. Sin embargo, la optimización con bases ANO–S lleva a la intersección cónica del tipo eteno. Está demostrado que estas bases tienen en cuenta de mejor forma la energía de correlación, proporcionando mejores resultados que las bases segmentadas con el mismo número de funciones contraídas.^{30,37,41}

Ambas mejoras, el incremento de la calidad del conjunto de base o la adición de orbitales de correlación extra al espacio activo, compensa parcialmente la falta de correlación dinámica del método CASSCF π de valencia.

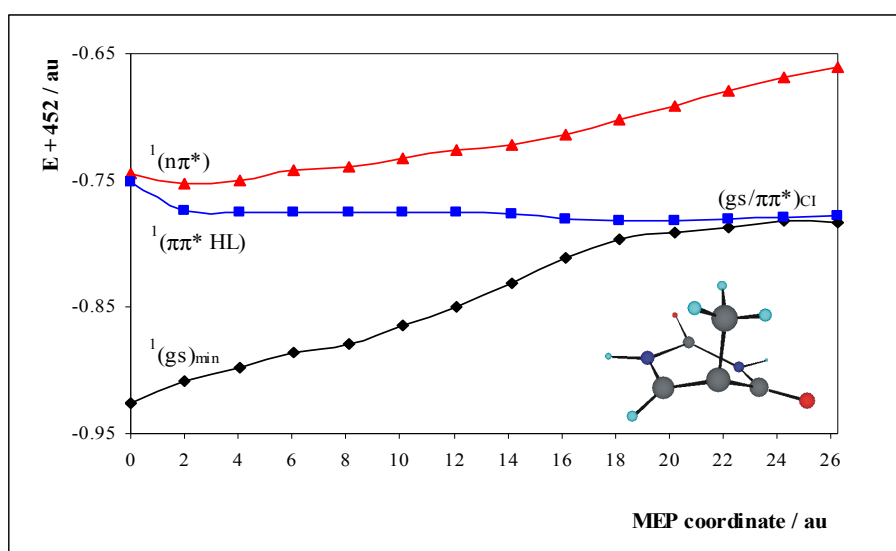


Figura 6.3. Evolución de los estados fundamental y singletes excitados más bajos a lo largo del MEP $^1(\pi\pi^* \text{ HL})$ desde la geometría FC de la timina.

Para asegurar un perfil sin barrera en la relajación de la timina, se analizó la evolución de los estados fundamental y singletes excitados más bajos a lo largo del MEP $^1(\pi\pi^* \text{ HL})$, desde la geometría de FC hasta la CI tipo eteno, evolución que se muestra en la figura 6.3. Como puede observarse el MEP obtenido muestra un perfil sin barrera desde FC hasta la $(gs/\pi\pi^*)_{CI}$.

Citosina

Los problemas encontrados en la descripción de los posibles caminos de relajación en la citosina son más serios. En un trabajo previo,³⁴ mostrábamos que el MEP calculado a nivel CASSCF (espacio π de valencia) para la citosina, discurre directamente desde la geometría FC hasta el mínimo plano $^1(\pi\pi^* \text{ HL})_{\text{min}}$, y que las optimizaciones energéticas añadiendo orbitales de correlación extra, o mejorando la base, no cambiaba el resultado, al contrario de lo ocurre en el uracilo y la timina.

Se nos plantea, pues, la pregunta de si la citosina es diferente a las otras bases pirimidínicas y de si, en este caso, existe realmente una barrera energética en el camino desde FC a la CI tipo eteno. La respuesta es, que probablemente no. A diferencia de lo que ocurre para los otros sistemas, los dos puntos críticos, la cónica tipo eteno $(gs/\pi\pi^*)_{CI}$ y el mínimo plano $^1(\pi\pi^* \text{ HL})_{\text{min}}$, están en la citosina casi degenerados (véase figura 6.1), con valores de 3.62 y 3.6 eV, respectivamente. Este hecho puede afectar a la descripción de las posibles barreras, que depende altamente de efectos casi imperceptibles relacionados con el nivel de correlación usado en la optimización. Las evidencias apuntando en dicha dirección quedaron claras cuando se realizó una Interpolación Lineal de Coordenadas Internas (*Linear Interpolation in Internal Coordinates*, LIIC) entre las geometrías optimizadas FC y $(gs/\pi\pi^*)_{CI}$ de la citosina. La figura 6.4 muestra la evolución de su estado singlete excitado $\pi\pi^*$ más bajo desde la geometría FC hasta la intersección cónica tipo eteno $(gs/\pi\pi^*)_{CI}$ a lo largo del LIIC calculado a nivel CASPT2//CASSCF(14,10)/ANO-S.

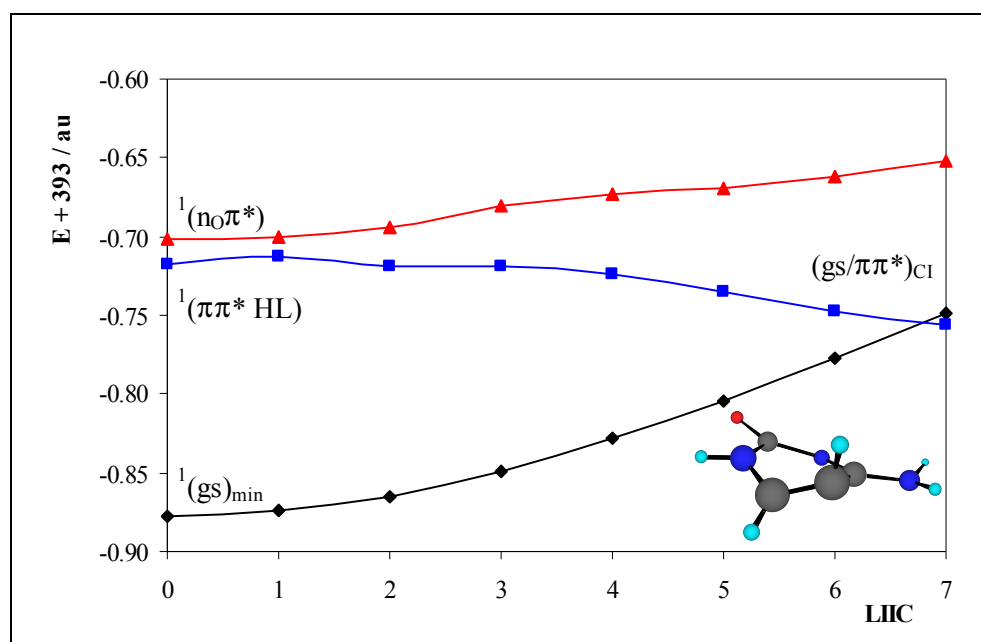


Figura 6.4. Evolución del estado singlete $\pi\pi^*$ más bajo en la citosina desde la geometría FC hasta la geometría de la intersección cónica tipo eteno, $(gs/\pi\pi^*)_{CI}$, al realizar la interpolación lineal en coordenadas internas, LIIC.

Las geometrías, FC inicial y $(gs/\pi\pi^*)_{CI}$ final, se determinaron en cálculos separados. Como puede observarse, el perfil de energía obtenido a lo largo del estado $^1(\pi\pi^* HL)$ en dicho proceso LIIC tiene una pequeña barrera de 2.5 kcal/mol a nivel CASPT2 al principio de la interpolación, en la que el ángulo diedro HC_5C_6H (véase figura 5.1) tiene un valor de 16° . Si se aumentan las bases a ANO-S (C, N, O [4s3p1d]/H[2s1p])³¹ y ANO-L (C, N, O [4s3p2d]/H[3s2p])⁴² también se encuentran barreras en la misma sección inicial de la interpolación, pero de menor energía, 1.0 y 1.1 kcal/mol a nivel CASPT2, respectivamente. Estas barreras, sin embargo, son mucho mayores en cálculos CASSCF, aunque nunca superiores a 6.6 kcal/mol al incrementar los espacios activos. Esto es indicativo de lo difícil que es describir el camino desde FC a la cónica tipo eteno a nivel CASSCF, ya que el camino hacia el mínimo plano sería el elegido por el sistema, quedando patente la necesidad de usar correlación dinámica en la optimización de la citosina.

Las barreras encontradas en las interpolaciones lineales, en particular cuando se usa un conjunto de puntos muy ajustado, representan un límite superior de la verdadera barrera existente. Por tanto, en la citosina se puede concluir también que el camino descrito es sin barrera energética y que la falta de correlación dinámica en la optimización impide una descripción más precisa.

Para comprobar este aspecto, se hicieron interpolaciones en sistemas demostrados previamente “sin barrera”, tales como el uracilo y la timina, y se calcularon, con ambas moléculas, los caminos LIIC (véase figuras 6.5 y 6.6) al mismo nivel de cálculo que el empleado para realizar los MEPs (véase figuras 6.2 y 6.3).

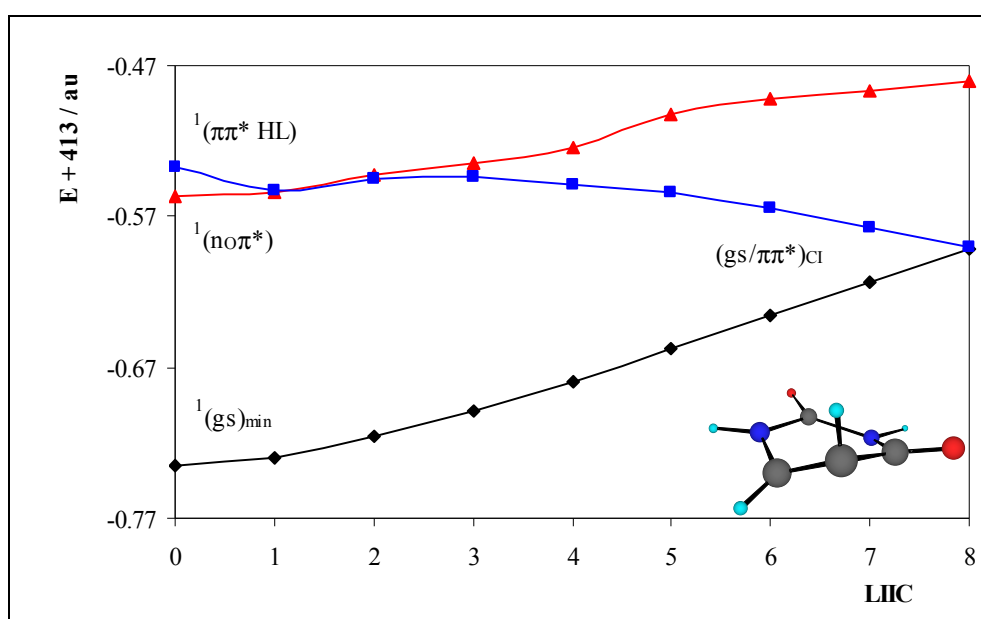


Figura 6.5. Evolución de los estados excitados singletes más bajos en el uracilo, entre las estructuras FC y la CI tipo eteno, $(\text{gs}/\pi\pi^*)_{\text{CI}}$, al realizar la interpolación lineal en coordenadas internas, LIIC.

Las figuras 6.5 y 6.6 muestran la evolución del estado singlete excitado $\pi\pi^*$ más bajo $^1(\pi\pi^* \text{ HL})$ en el uracilo y la timina, respectivamente, desde la

geometría FC hasta la geometría de la intersección cónica $(gs/\pi\pi^*)_{CI}$ a lo largo del LIIC, calculadas a nivel CASPT2//CASSCF(14,10)/ANO-S. El perfil obtenido muestra en el uracilo una barrera energética de cerca de 5 kcal/mol para un ángulo diedro HC_5C_6H cercano a 30° y en la timina una barrera de cerca de 12 kcal/mol con un ángulo de 16° . Estas barreras energéticas pueden asociarse con el carácter “sin barrera” de los MEPs realizados en ambos sistemas. Como la barrera energética derivada del LIIC en la citosina es mucho más baja que la obtenida para el uracilo y la timina, se puede asumir, casi con total garantía, que el carácter del camino de relajación para la citosina es también sin barrera energética.

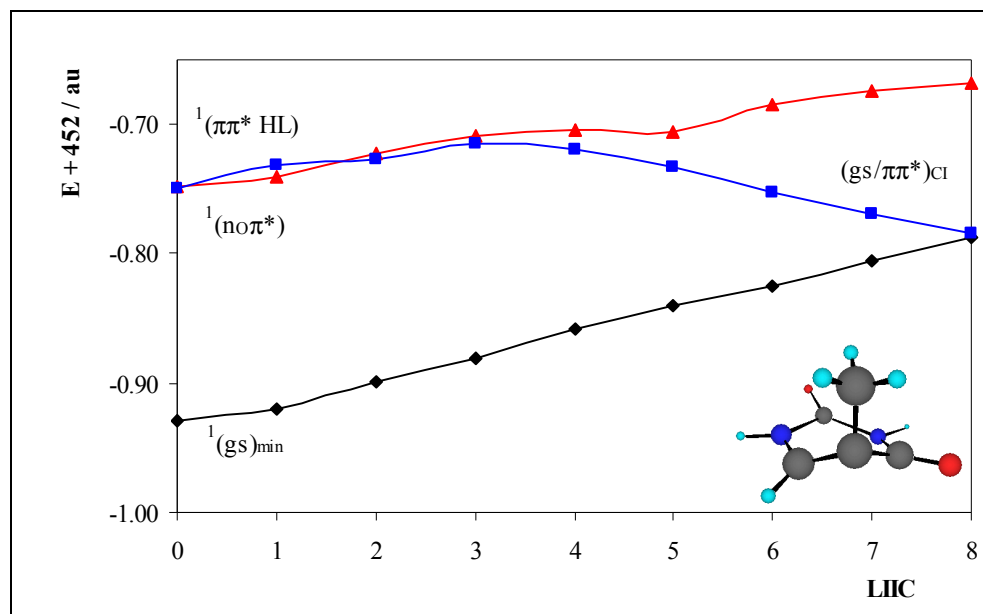


Figura 6.6. Evolución de los estados excitados singletes más bajos en la timina, entre las estructuras FC y la CI tipo eteno, $(gs/\pi\pi^*)_{CI}$, al realizar la interpolación lineal en coordenadas internas, LIIC.

Camino alternativo de relajación energética

El hecho de que las optimizaciones geométricas o los MEPs de cada uno de los sistemas conduzcan a una CI de tipo eteno no implica, sin embargo, que el mínimo plano no exista, sino que, simplemente, el MEP tipo *steepest-descent*, que describe el principal camino fotoquímico desde la estructura FC, no lo incluye. La contribución de dicho mínimo plano $^1(\pi\pi^* HL)_{\min}$, que debería considerarse como un mínimo local en la hipersuperficie S_1 , no es relevante para la desactivación ultrarrápida de cada uno de los sistemas, ya que se encuentra en una región de la hipersuperficie lejos del camino principal de desactivación, que involucra distorsiones fuera del plano hacia la cónica de tipo eteno. Aun así, la menor población que alcanza el mínimo plano puede considerarse responsable de la emisión residual observada en disolución, así como de los tiempos de decaimiento elevados.^{10,12} En este sentido, no podemos olvidar, especialmente en la citosina, que el camino hacia el mínimo plano $^1(\pi\pi^* HL)_{\min}$ juega un papel competitivo en la relajación energética. De hecho, tal como se ha sugerido en estudios previos,¹⁰ los decaimientos más largos (~ 1 ps) descritos por experimentos de haz molecular⁴ podrían estar relacionados con la desactivación del estado $^1(\pi\pi^*)$ y con las barreras energéticas desde su mínimo hacia diferentes CIs, no sólo la eténica sino también la que incluye el alargamiento del enlace $C_2=O$.^{10,34} Para analizar el papel de dichos decaimientos secundarios, se ha realizado una determinación precisa de esas barreras para las tres moléculas al nivel de cálculo más alto posible.

Por un lado, se ha obtenido que el mínimo plano $^1(\pi\pi^* HL)_{\min}$ para los tres sistemas se sitúa adiabáticamente a 4.53 (U), 4.49 (T) y 3.62 eV (C) y, por otro, las emisiones verticales son de 3.9 (U), 3.8 (T) y 2.4 eV (C). Estos mínimos pueden ser considerados como la causa del bajo rendimiento cuántico de fluorescencia ($\Phi_F \sim 10^{-4}$) observado en disolución,^{1,2,7,8,9,16,43} pues el origen de banda de fluorescencia en medio acuoso se ha medido cerca de 4.6 (U) y 4.5 eV (T)¹⁶ y sus respectivos máximos en la banda de emisión tienen un valor cercano a 3.9 eV tanto para U como para T.¹⁶ Es por tanto que, los resultados de los cálculos teóricos se pueden relacionar claramente con los datos experimentales. Para la citosina, los resultados

calculados, tanto adiabáticamente como en el caso de la emisión vertical son, en cierto modo, más bajos que la banda de fluorescencia máxima medida en agua, ~ 3.8 eV.¹² La asignación de la banda de emisión observada con el mínimo plano, $^1(\pi\pi^* \text{HL})_{\text{min}}$, se ha demostrado en estudios teóricos previos,^{10,11} donde el tiempo de vida media calculado de 33 ns se ajusta al estimado experimentalmente, 40 ns.¹⁰ Además, la estructura obtenida en el mínimo de la citosina es bastante diferente a las de las otras dos nucleobases. La longitud del enlace $\text{C}_2=\text{O}$ (véase figura 5.1) aumenta de 1.198 Å, en la estructura del estado fundamental en los tres sistemas, a 1.203 (U), 1.202 (T) y 1.353 Å (C), en la estructura del $^1(\pi\pi^* \text{HL})_{\text{min}}$. En la citosina, el mínimo del estado $^1(\pi\pi^* \text{HL})_{\text{min}}$, que se encuentra isoenergético con la CI tipo eteno, tiene una estructura con una distancia de enlace $\text{C}_2=\text{O}$ más larga, lo cual refleja una interacción con el estado excitado $n\text{O}\pi^*$ cercano.^{10,34} Para el uracilo y la timina los mínimos $^1(\pi\pi^* \text{HL})_{\text{min}}$ se localizan adiabáticamente 0.5–0.6 eV por encima de la cónica tipo eteno y mantienen un carácter puramente $\pi\pi^*$. En ambas nucleobases, los mínimos $n\pi^*$ en las hipersuperficies, incluso cuando se encuentran más bajos en energía, al menos en fase gas donde están menos perturbados (se estima un corrimiento al azul de 0.5 eV en agua),⁴⁶ jugarán, seguramente, un papel menor en las propiedades de emisión, teniendo lugar, en estos sistemas, esencialmente desde el mínimo plano $^1(\pi\pi^* \text{HL})_{\text{min}}$. De hecho, los tiempos de vida media radiativos, τ_{rad} , para U y T, obtenidos por medio de la relación de Strickler–Berg^{44,45} son de ~ 9 ns, un valor que está de acuerdo con la estimación experimental 7–8 ns,¹⁶ y que apunta a un estado $\pi\pi^*$ como responsable de la emisión.

Conclusiones del capítulo 6

- ✓ El presente capítulo presenta un modelo unificado que explica los principales sucesos fotoquímicos observados en las tres bases pirimidínicas del ADN/ARN, uracilo, timina y citosina, después de ser irradiadas con luz UV.
- ✓ Se ha estudiado el estado singlete más bajo, $^1(\pi\pi^* \text{ HL})$ HOMO \rightarrow LUMO, de las tres nucleobases, empleado métodos químico-cuánticos CASPT2//CASSCF y conjunto de base tipo ANO.
- ✓ El decaimiento ultrarrápido, medido en distintos medios para los tres sistemas aislados, se puede relacionar con el camino de desactivación directo, sin ningún tipo de barrera, del estado singlete $^1(\pi\pi^* \text{ HL})$ hacia una intersección cónica fuera del plano de tipo eteno, $^1(\text{gs}/\pi\pi^*)_{\text{Cl}}$.
- ✓ La presencia o no de una barrera a lo largo del camino y la descripción precisa de los perfiles de los diferentes caminos de reacción se han mostrado como dependientes del nivel de cálculo empleado. La adición de orbitales de correlación extra en el espacio activo o la mejora de la base monoeléctronica usando, por ejemplo, funciones tipo ANO conduce a un tratamiento más equilibrado.
- ✓ Se propone en este capítulo que la causa del bajo rendimiento cuántico de fluorescencia ($\Phi_{\text{F}} \sim 10^{-4}$), medido en disolución en las tres nucleobases, se puede asociar a la presencia de un mínimo plano $^1(\pi\pi^*)$ en la hipersuperficie de S_1 . Dicho mínimo es la fuente de la emisión débil que se ha registrado en estos sistemas, aunque tal estructura plana no es de esperar que participe en la desactivación ultrarrápida de cada una de las nucleobases.

Referencias

1. Crespo-Hernández, C. E.; Cohen, B.; Hare, P. M.; Kohler, B. *Chem. Rev.* **2004**, *104*, 1977.
2. Daniels, M.; Hauswirth, W. W. *Science* **1971**, *171*, 675.
3. Samoylova, E.; Lippert, H.; Ullrich, S.; Hertel, I. V.; Radloff, W.; Schultz, T. *J. Am. Chem. Soc.* **2004**, *127*, 1782.
4. Canuel, C.; Mons, M.; Piuze, F.; Tardivel, B.; Dimicoli, I.; Elhanine, M. *J. Chem. Phys.* **2005**, *122*, 74316.
5. Kuimova, M. K.; Dyer, J.; George, M. W.; Grills, D. C.; Kelly, J. M.; Matousek, P.; Parker, A. W.; Sun, X. Z.; Towrie, M.; Whelan, A. M. *Chem. Commun.* **2005**, 1182.
6. Olivucci, M., Ed. *Computational Photochemistry*; Elsevier: Amsterdam, 2005.
7. Hauswirth, W.; Daniels, M. *Chem. Phys. Lett.* **1971**, *10*, 140.
8. Callis, P. R. *Chem. Phys. Lett.* **1979**, *61*, 563.
9. Callis, P. R. *Ann. Rev. Phys. Chem.* **1983**, *34*, 329.
10. Merchán, M.; Serrano-Andrés, L. *J. Am. Chem. Soc.* **2003**, *125*, 8108.
11. Tomic, K.; Tatchen, J.; Marian, C. M. *J. Phys. Chem. A* **2005**, *109*, 8410.
12. Blancafort, L.; Cohen, B.; Hare, P. M.; Kohler, B.; Robb, M. A. *J. Phys. Chem. A* **2005**, *109*, 4431.
13. Zgierski, M. Z.; Patchkovskii, S.; Fujiwara, T.; Lim, E. C. *J. Phys. Chem. A* **2005**, *109*, 9384.
14. Zgierski, M. Z.; Patchkovskii, S.; Lim, E. C. *J. Chem. Phys.* **2005**, *123*, 81101.
15. Matsika, S. *J. Phys. Chem. A* **2004**, *108*, 7584.
16. Gustavsson, T.; Banyasz, A.; Lazzarotto, E.; Markovitsi, D.; Scalamani, G.; Frisch, M. J.; Barone, V.; Improta, R. *J. Am. Chem. Soc.* **2006**, *128*, 607.
17. Blancafort, L.; Robb, M. A. *J. Phys. Chem. A* **2004**, *108*, 10609.
18. Matsika, S. *J. Phys. Chem. A* **2005**, *109*, 7538.
19. Truhlar, D. G.; Mead, C. A. *Phys. Rev. A* **2003**, *68*, 32501.
20. Dreuw, A.; Worth, G. A.; Cederbaum, L. S.; Head-Gordon, M. *J. Phys. Chem. B* **2004**, *108*, 19049.
21. Serrano-Andrés, L.; Merchán, M.; Borin, A. C. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 8691.
22. Marian, C. M. *J. Chem. Phys.* **2005**, *122*, 104314.

23. Perun, S.; Sobolewski, A. L.; Domcke, W. *J. Am. Chem. Soc.* **2005**, *127*, 6257.
24. Chen, H.; Li, S. H. *J. Phys. Chem. A* **2005**, *109*, 8443.
25. Blancafort, L. *J. Am. Chem. Soc.* **2006**, *128*, 210.
26. Sobolewski, A. L.; Domcke, W. *Phys. Chem. Chem. Phys.* **2004**, *6*, 2763.
27. Andersson, K.; Malmqvist, P.-Å.; Roos, B. O. *J. Chem. Phys.* **1992**, *96*, 1218.
28. Roos, B. O.; Fülcher, M. P.; Malmqvist, P.-Å.; Serrano-Andrés, L.; Pierloot, K.; Merchán, M. *Adv. Chem. Phys.* **1996**, *93*, 219.
29. Merchán, M.; Serrano-Andrés, L.; Fülcher, M. P.; Roos, B. O. Multiconfigurational Perturbation Theory Applied to Excited States of Organic Compounds; In *Recent Advances in Multireference Methods*; Hirao, K., Ed.; World Scientific Publishing Company: Amsterdam, 1999; Vol. 4, pp 161-195.
30. Merchán, M.; Serrano-Andrés, L. An Initio Methods for excited states. In *Computational Photochemistry*; Olivucci, M. Ed; Elsevier: Amsterdam, 2005; pp 35-92.
31. Pierloot, K.; Dumez, B.; Widmark, P.-O.; Roos, B. O. *Theor. Chim. Acta* **1995**, *90*, 87.
32. Forsberg, J.; Malmqvist, P.-Å. *Chem. Phys. Lett.* **1997**, *274*, 196.
33. Serrano-Andrés, L.; Fülcher, M. P. *J. Am. Chem. Soc.* **1998**, *120*, 10912.
34. Merchán, M.; Serrano-Andrés, L.; Robb, M. A.; Blancafort, L. *J. Am. Chem. Soc.* **2005**, *127*, 1820.
35. González-Luque, R.; Garavelli, M.; Bernardi, F.; Merchán, M.; Robb, M. A.; Olivucci, M. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 9379.
36. Serrano-Andrés, L.; Merchán, M.; Lindh, R. *J. Chem. Phys.* **2005**, *122*, 104107
37. Serrano-Andrés, L.; Merchán, M. Spectroscopy: Applications. In *Encyclopedia of Computational Chemistry*; Schleyer, P. v. R., Schreiner, P. R., Schaefer, H. F., III, Jorgensen, W. L., Thiel, W., Glen, R. C., Eds.; Wiley: Chichester, U.K. 2004.
38. Andersson, K.; Barysz, M.; Bernhardsson, A.; Blomberg, M. R. A.; Carissan, Y.; Cooper, D. L.; Cossi, M.; Fülcher, M. P.; Gagliardi, L.; de Graaf, C.; Hess, B.; Hagberg, G.; Karlström, G.; Lindh, R.; Malmqvist, P.-Å.; Nakajima, T.; Neogrády, P.; Olsen, J.; Raab, J.; Roos, B. O.; Ryde, U.; Schimmelpfennig, B.; Schütz, M.; Seijo, L.; Serrano-Andrés, L.; Siegbahn, P. E. M.; Stålring, J.; Thorsteinsson, T.; Veryazov, V.; Widmark, P. O.

- MOLCAS 6.0*; Department of Theoretical Chemistry, Chemical Centre, University of Lund: Lund, Sweden, 2004.
39. Fülcher, M. P.; Roos, B. O. *J. Am. Chem. Soc.* **1995**, *117*, 2089.
 40. Lorentzon, J.; Fülcher, M. P.; Roos, B. O. *J. Am. Chem. Soc.* **1995**, *117*, 9265.
 41. Serrano-Andrés, L.; Merchán, M. *J. Mol. Struct. Theochem* **2005**, *729*, 99.
 42. Widmark, P. O.; Malmqvist, P.-A.; Roos, B. O. *Theor. Chim. Acta* **1990**, *77*, 291.
 43. Vigny, P.; Duquesne, M. *Photochem. Photobiol.* **1974**, *20*, 15.
 44. Strickler, S. J.; Berg, R. A. *J. Chem. Phys.* **1962**, *37*, 814.
 45. Rubio-Pons, O.; Serrano-Andrés, L.; Merchán, M. *J. Phys. Chem. A* **2001**, *105*, 9664.
 46. Zazza, C.; Amadei, A.; Sanna, N.; Grandi, A.; Chillemi, G.; Di, Nola, A.; D'Abramo, M.; Aschi, M. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1385.
 47. Shukla, M. K.; Mishra, P. C. *Chem. Phys.* **1999**, *240*, 319.

7 Población del estado triplete en las bases pirimidínicas

En el capítulo anterior se ha presentado un modelo unificado para explicar el decaimiento ultrarrápido de las bases pirimidínicas del ADN/ARN, uracilo, timina, y citosina. En él se propone que éste se debe a la rápida conversión interna, descrita a través de un camino de mínima energía, libre de barreras, desde el estado excitado singlete más bajo, $^1(\pi\pi^* \text{ HL})$, hasta una intersección cónica con el estado fundamental;¹ cónica que ha sido descrita también por diferentes autores a diferentes niveles de cálculo.^{2,3,4,5}

Sin embargo, y a pesar de los cortos periodos de vida media asociados al estado excitado singlete más bajo, se observa que en el ADN se forman fotodímeros de pirimidina (véase figuras 5.7 y 8.1), tanto en disolución como en estado sólido.⁶ Desde los años 60, la determinación del mecanismo de formación de dímeros ciclobutánicos ha sido objeto de numerosos estudios.^{6,7,8,9,10,11} Uno de los modelos propuestos para explicar la fotodimerización de las nucleobases pirimidínicas en disolución implica la fotoexcitación de una molécula al estado singlete más bajo, seguida de la población del estado triplete a través de un mecanismo de cruce entre sistemas y posterior reacción de éste con una segunda molécula en su estado fundamental.^{6,10,11,12}

Los estados triplete de los sistemas moleculares son intermedios frecuentes en muchas reacciones fotoinducidas, ya que el hecho de poseer un marcado carácter birradicalario y tiempos de vida media relativamente largos, les hace ser especies candidatas proclives a interactuar con otros compuestos.¹³ Los estados triplete de las bases púricas y pirimidínicas del ADN/ARN no son una excepción y éstos participan activamente en

fotorreacciones desencadenadas por la luz UV, como por ejemplo en la de formación de aductos fototerapéuticos fármaco-nucleobase¹⁴ o en la fotodimerización de las bases pirimidínicas, considerada ésta como una de las mutaciones del material genético más frecuentes.^{6,7,15,16} En particular, la presencia del triplete de uracilo en disolución ha sido observado usando diferentes técnicas experimentales directas, como la de *laser flash photolysis*,¹⁷ o por técnicas indirectas de detección de la intervención del triplete, como son el uso de *triplet quenchers* y de fotosensibilizadores externos.^{6,7,17,18,19} Con ellas se ha puesto de manifiesto, además, el hecho de que el valor del rendimiento cuántico del cruce entre sistemas, Φ_{ISC} , es altamente dependiente de la longitud de onda de excitación en la zona de más baja energía de absorción. Así pues, en el caso del uracilo, el valor se incrementa de 1.4×10^{-3} , a 280 nm (4.43 eV), a 1.6×10^{-2} , a 230 nm (5.39 eV).^{6,7} Por último, con dichas técnicas se ha demostrado que el espectro de fosforescencia del ADN a bajas temperaturas está formado por dos componentes básicas que son debidas principalmente a la timina y, en menor medida, a la adenina.^{20,21,22}

Pero, a pesar de que ya se han recabado y revisado los datos de fosforescencia de cada una de las nucleobases del ADN y de sus diferentes derivados en distintos medios y condiciones,^{6,23,24,25} incluyendo estudios que emplean fotosensibilizadores externos,^{18,19,26} todavía, a día de hoy, no se conocen los mecanismos específicos por los cuales se puebla el triplete en cada una de ellas. Aún quedan por dilucidar aspectos tales como el diferente “destino” de sus estados triplete, que explicaría por ejemplo la prevalencia de las dos bases mencionadas, timina y adenina, en el espectro de fosforescencia del ADN, o la implicación del estado triplete en el proceso de relajación rápida de las nucleobases, en particular de la timina.²⁷

La población del estado triplete se puede producir desde otras especies que se encuentran en estado triplete, o por un ISC eficiente desde estados singlete de la molécula inicialmente fotoexcitada. Como ya se ha comentado en el capítulo 1, las transiciones no radiativas entre estados de la misma multiplicidad, que conducen a una conversión interna, son más eficientes si tienen lugar en las cercanías de las regiones de intersección cónica.^{28,29,30} En el caso de interacción entre estados de diferente multiplicidad, el cálculo de las tasas de ISC cumple con la regla de oro de Fermi, que relaciona la fuerza

de la interacción con los efectos de acoplamiento espín–órbita vibrónico y la densidad de los estados FC. Los estudios de Marian y colaboradores^{31,32,33} han demostrado que la eficiencia del proceso de ISC se basa en una serie de pequeños efectos, incluyendo una gran densidad de estados vibracionales y el que se produzca un correcto solapamiento de las funciones de onda vibracionales que, al final, intensifican los efectos SOC. La disminución de la separación energética entre los estados singlete y triplete, y en particular la presencia de degeneraciones singlete–tripleto, o regiones de cruce, especialmente cuando están relacionadas con la existencia de modos vibracionales fuera del plano, es un buen indicador de una alta densidad de estados y es, por tanto, concebible que los cruces singlete–tripleto jueguen un papel importante en el incremento de las tasas de transferencia de población por ISC.¹³

El objetivo del presente capítulo es proporcionar un mecanismo de población del estado tripleto más bajo de las tres bases pirimidínicas del ADN/ARN, que ayude a racionalizar los datos observados experimentalmente en estas moléculas. En definitiva, pretende ayudar a entender:

- (i) cómo se puede llevar a cabo, *in vacuo*, la población del estado tripleto más bajo para las tres bases,
- (ii) por qué los estados tripleto de la timina prevalecen en el espectro de fosforescencia del ADN pudiéndose, por tanto, esperar un rendimiento cuántico de formación, Φ_{ISC} , mayor que para las otras nucleobases.
- (iii) cuál es la base molecular para la dependencia observada del Φ_{ISC} con la longitud de onda.⁶

Detalles computacionales

Los cálculos realizados para llevar a cabo este proyecto incluyen optimización de geometrías y obtención de los MEPs, CIs y STCs, a nivel CASSCF, seguidas de cálculos de energías CASPT2 de cada uno de los puntos obtenidos. También se han realizado cálculos de los valores de SOC y los momentos dipolares de transición, (*Transition Dipolar Moment*, TDM)

y se han estimado los tiempos de vida radiativa por medio de la relación Strickler–Berg.^{34,35} Para cada sistema se empleó el mismo conjunto de base monoeléctronica y el mismo espacio activo que se indica en el capítulo anterior en el estudio de la fotoquímica de los estados singlete. El resultado final queda descrito como CASPT2//CASSCF(14,10), con un espacio activo de 14 electrones distribuidos en 10 orbitales, en donde se han incluido todos los orbitales de valencia $\pi\pi^*$ (10 electrones distribuidos en 8 orbitales activos) más los pares solitarios. En los procesos de optimización se usaron los espacios activos, que habían demostrado que funcionaban correctamente en cálculos previos;^{1,36,37,38,39} en ellos, y en los artículos publicados con los presentes resultados,^{40,41} pueden encontrarse todos los detalles técnicos de carácter computacional. Conviene aclarar que las tablas y gráficas del presente capítulo no se han modificado en la discusión de resultados para que aparezcan tal como están publicadas. Así el primer estado excitado singlete, S_1 , se nombra como ($\pi\pi^*$ La) o como ($\pi\pi^*$ HL) dependiendo de la base nucleica descrita.

Mecanismos de población del estado triplete en las bases pirimidínicas

A continuación se describen los mecanismos de población del estado triplete más bajo energéticamente en cada una de las bases. La estrategia empleada consiste en obtener los MEPs que conducen a las bases desde el primer paso del proceso fotoquímico, es decir, después de la absorción de la luz UV, hacia las regiones degeneradas de STC, para desembocar, finalmente, en el estado excitado triplete energéticamente más bajo y reactivo, calculando los términos SOC entre los estados más relevantes.

Teniendo en cuenta que los cálculos se han realizado *in vacuo*, sin las consideraciones explícitas de los efectos del disolvente, los resultados pueden considerarse como una característica molecular inherente a las nucleobases, que pueden diferir de lo esperado por la exposición al entorno, bien sea en disolución o en fase sólida, *in vitro* o *in vivo*, pero que sirven como modelo para conseguir dar una visión general de la fotoquímica del ADN/ARN, tarea altamente complicada.

Uracilo

En la tabla 7.1 se compilan los valores de las transiciones verticales, origen de banda, fuerzas de oscilador y tiempos de vida media radiativos, calculados a nivel CASSCF y CASPT2, de los estados singlete y triplete más bajos del uracilo. A no ser que se especifique lo contrario, los resultados empleados en la discusión son valores CASPT2.

Tabla 7.1. Propiedades espectroscópicas calculadas para los estados excitados singlete y triplete más bajos del uracilo a nivel CASPT2//CASSCF(14/10)/6–31G(d,p).

Estado	Transición Vertical (eV)		Origen de Banda (T_e , eV)		
	CASSCF	CASPT2 ^a	CASSCF	CASPT2	τ_{rad}^b
¹ ($n_O\pi^*$) ^c	5.18	4.93 (0.001)	4.07	4.03	3051 ns
¹ ($\pi\pi^*$)	6.82	5.18 (0.196)	6.30 ^d	4.48 ^d	7 ns
¹ ($\pi\pi^*$)	7.29	6.18 (0.073)			
³ ($\pi\pi^*$)	3.98	3.80	3.16	3.15	135 ms
³ ($n_O\pi^*$) ^c	4.87	4.71	3.81	3.91	
³ ($\pi\pi^*$)	5.76	5.33			

^a Fuerza de oscilador entre paréntesis.

^b Calculado usando la aproximación de Strickler–Berg.

^c Incluye básicamente al O_4 .

^d Mínimo optimizado a nivel CASPT2.

Como puede observarse, en la región FC, la transición al estado excitado S_1 , ¹($\pi\pi^*$ HL), con una energía de 5.18 eV, es la que tiene la mayor fuerza de oscilador, con un valor de 0.196, siendo éste el estado poblado tras la absorción de la radiación y que se encuentra 0.25 eV por encima del estado ¹ $n_O\pi^*$.

La figura 7.1 muestra la evolución del estado fundamental y los excitados singlete y triplete más bajos de la molécula uracilo, en el MEP calculado, partiendo de la geometría FC, a lo largo hipersuperficie de S_1 $^1(\pi\pi^*$ HL), que conduce, como ya se explicó en el capítulo 6, por un camino sin barrera hacia la CI con el estado fundamental $(gs/{}^1\pi\pi^*)_{CI}$ de tipo eteno.¹ Al analizar la figura 7.1, se pone de manifiesto que hay dos regiones diferentes de STC, que pueden alcanzarse fácilmente a lo largo del principal camino de desactivación energética a través del MEP- S_1 , siendo éstos los posibles canales de ISC hacia el estado triplete más bajo.

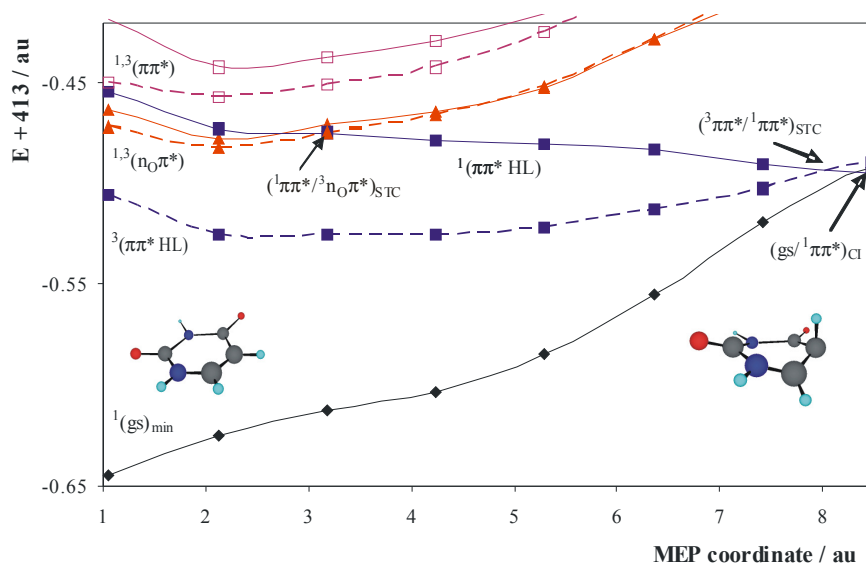


Figura 7.1. Evolución de los estados fundamental y excitados singlete y triplete más bajos, del uracilo desde la geometría de FC a lo largo del MEP $^1(\pi\pi^*$ HL) realizado con cálculos CASPT2//CASSCF(14,10)/6-31G(d,p).

Entre el segundo y tercer punto del MEP, el estado $^1(\pi\pi^*$ HL), se cruza tanto con el singlete como con el triplete $n\pi^*$. Aparte de un posible cruce interno hacia el singlete $^1(n_o\pi^*)$, a través de su correspondiente cónica, esta región puede ser responsable del primer proceso de ISC que tiene lugar en el

uracilo a altas energías (4.6 eV), en el que el estado $^3(n_0\pi^*)$ se podría poblar a partir del estado singlete activado, $^1(\pi\pi^* \text{ HL})$.

Este primer cruce singlete–tripleto, $(^1\pi\pi^*/^3n_0\pi^*)_{\text{STC}}$, del estado S_1 con el estado tripleto $^3(n_0\pi^*)$ se produce cerca del tercer punto del MEP, mientras que el segundo cruce, con el tripleto T_1 $^3(\pi\pi^*)$, se establece en las proximidades de la cónica con el estado fundamental, $(g_s/^1\pi\pi^*)_{\text{CI}}$. Por tanto, cabe pensar que ambos STC contribuirán a la despoblación del estado S_1 , en su camino hacia el tripleto más bajo T_1 .

El primer STC, se sitúa ~ 0.6 eV por debajo de la energía de transición vertical correspondiente al estado $^1(\pi\pi^* \text{ HL})$ en la región FC. Como es de esperar, debido a la diferente naturaleza de los estados, en la región de cruce hay un elevado SOC, $\sim 25 \text{ cm}^{-1}$, entre los estados $^3(n_0\pi^*)$ y $^1(\pi\pi^* \text{ HL})$, sustentando la eficiente transferencia de población desde el estado excitado singlete poblado inicialmente hacia el estado tripleto.

Para analizar la evolución del tripleto $^3(n_0\pi^*)$, hacia el tripleto más bajo $^3(\pi\pi^*)$, se ha calculado el MEP a lo largo de la hipersuperficie del tripleto $^3(n_0\pi^*)$ desde el cruce $(^1\pi\pi^*/^3n_0\pi^*)_{\text{STC}}$, tal como se muestra en la figura 7.1.a. En ella se muestra que la evolución de dicho estado lleva a un mínimo, $^3(n_0\pi^*)_{\text{min}}$, situado adiabáticamente a 3.91 eV del estado fundamental. Sólo ligeramente por encima, a 0.01 eV, hemos localizado una intersección cónica conectando los estados tripleto $^3(n_0\pi^*)$ y $^3(\pi\pi^*)$ que favorece el intercambio de población hacia el estado más bajo, T_1 $^3(\pi\pi^*)$. Desde $(^3n_0\pi^*/^3\pi\pi^*)_{\text{CI}}$ es posible seguir la hipersuperficie de T_1 hasta alcanzar su mínimo, $^3(\pi\pi^*)_{\text{min}}$. Como no ha sido posible seguir el MEP directamente hacia el mínimo, hemos recurrido a hacer una interpolación lineal en coordenadas internas, LIIC, conectando $(^3n_0\pi^*/^3\pi\pi^*)_{\text{CI}}$ y $^3(\pi\pi^*)_{\text{min}}$ (cuyos resultados se muestran en la figura 7.1.b). Encontramos una pequeña barrera de cerca de 0.02 eV (0.50 kcal/mol) en la parte inicial del LIIC que, a pesar de que probablemente es lo que hace que no se pueda realizar el MEP, es tan pequeña como para restarle importancia en la evolución del estado $^3(\pi\pi^*)$ hacia su mínimo.

Como puede observarse en la figura 7.1.a, las PEHs de los estados singlete y tripleto $n\pi^*$ están siempre muy próximas y cerca de la cónica $(^3n_0\pi^*/^3\pi\pi^*)_{\text{CI}}$ hemos encontrado el cruce $(^1n_0\pi^*/^3\pi\pi^*)_{\text{STC}}$. En el caso de que algo de

población pasara al estado $^1(n_0\pi^*)$, a través de su cruce con $^1(\pi\pi^* \text{ HL})$ (véase figura 7.1), su evolución hacia el estado triplete más bajo sería extremadamente favorable porque el valor de SOC calculado se incrementa en la zona $(^1n_0\pi^*/^3\pi\pi^*)_{\text{STC}}$ a 40 cm^{-1} .

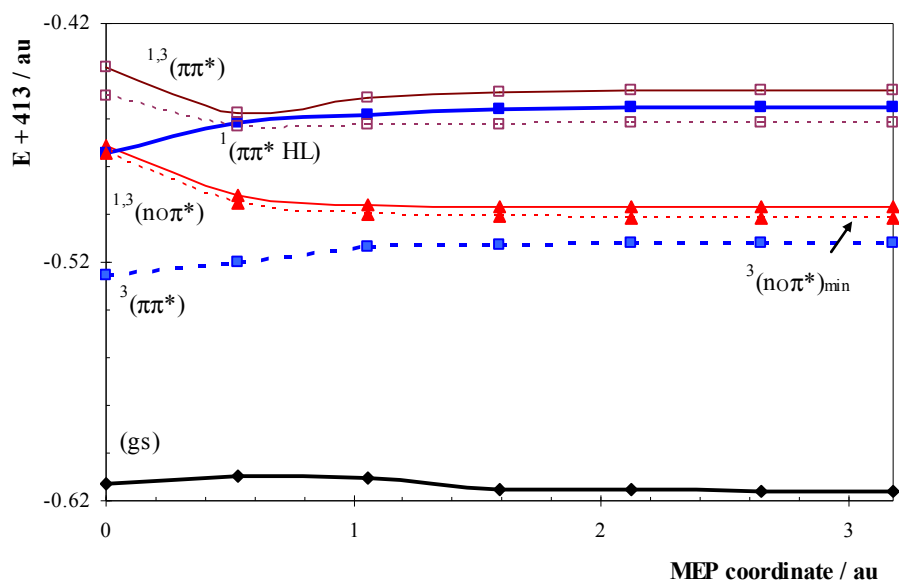


Figura 7.1.a. Evolución del MEP a lo largo de la hypersuperficie de $(^3n_0\pi^*)$ desde el cruce singlete–triplete $(^3n_0\pi^*/^1\pi\pi^*)_{\text{STC}}$ hasta alcanzar el mínimo $(^3n_0\pi^*)_{\text{min}}$ en el uracilo, realizado con cálculos CASPT2//CASSCF(14,10)/6–31G(p,d).

Los mecanismos anteriores descritos para la población del triplete más bajo son similares a los encontrados en otros cromóforos biológicos, tales como la isoaloxacina⁴⁴ y el psoraleno,⁴⁵ pero éstos no son la única posibilidad para poblar el triplete más bajo, T_1 $^3(\pi\pi^*)$, desde el estado inicialmente excitado $^1(\pi\pi^* \text{ HL})$. Tal como se muestra en la figura 7.1, cerca de la intersección cónica de tipo eteno, $(\text{gs}/^1\pi\pi^*)_{\text{CI}}$, que se produce entre el estado fundamental y el estado excitado singlete $^1(\pi\pi^* \text{ HL})$, hay un cruce, $(^3\pi\pi^*/^1\pi\pi^*)_{\text{STC}}$, que conecta los estados excitados triplete y singlete $\pi\pi^*$ más bajos. El cruce se encuentra a una energía, calculada adiabáticamente

desde el mínimo del estado fundamental, de 4.2 eV. La estructura obtenida en ese punto ha sufrido las mismas distorsiones que conducen a la molécula desde la región FC a la cónica $(gs/{}^1\pi\pi^*)_{CI}$, es decir, alargamiento y giro del doble enlace $C_5=C_6$.

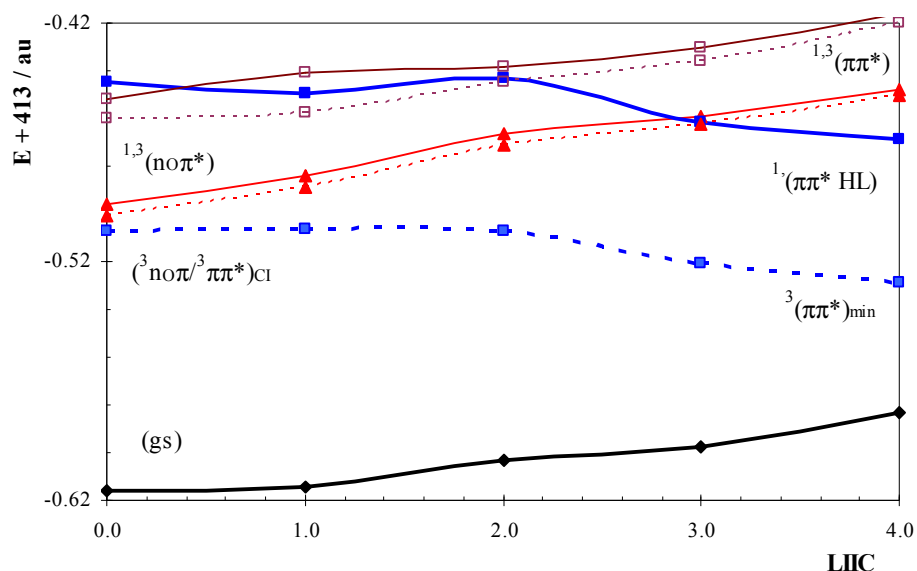


Figura 7.1.b. Interpolación lineal en coordenadas internas (LIIC) desde la intersección cónica $({}^3no\pi^3/{}^3\pi\pi^*)_{CI}$ al mínimo $({}^3\pi\pi^*)_{min}$ en el uracilo, realizado con cálculos CASPT2//CASSCF(14,10)/6–31G(p,d).

A diferencia de los STC en los que se ven involucrados los estados $n\pi^*$, el valor calculado de SOC en este caso es bajo, 1 cm^{-1} y, desde el punto de vista de cruce entre sistemas, podría esperarse, en cierta forma, menos eficiente. Sin embargo, los tres mecanismos pueden contribuir a la población general del estado triplete más bajo, que sucede con un rendimiento de cruce entre sistemas bajo, sin duda alguna debido al carácter ultrarrápido por el que se disipa la energía a lo largo del singlete.^{1,42,43}

Como aspecto final de la evolución a lo largo del triplete en el uracilo, hemos localizado el cruce singlete–triplete que conecta el $({}^3\pi\pi^*)$ y el estado fundamental y realizado el MEP que conduce desde tal cruce STC hacia el

$^3(\pi\pi^*)_{\min}$ (véase figura 7.1.c). El cruce se sitúa cerca de 3.4 eV del mínimo del estado fundamental, lo que significa que hay una barrera de 0.21 eV (4.7 kcal/mol) para alcanzar $(gs/{}^3\pi\pi^*)_{STC}$ desde el $^3(\pi\pi^*)_{\min}$. El SOC electrónico calculado es bajo, $\sim 1 \text{ cm}^{-1}$, lo cual predice para el triplete una vida media larga y una lenta relajación, teniendo, por tanto, predisposición para reaccionar.

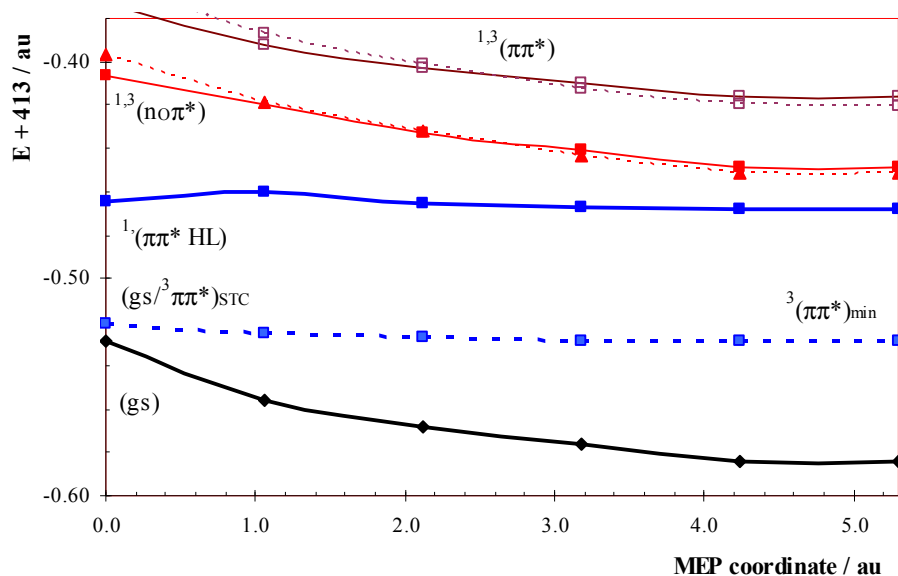


Figura 7.1.c. Evolución del camino de mínima energía (MEP) desde el cruce singlete–triplete $(gs/{}^3\pi\pi^*)_{STC}$ hasta el mínimo $({}^3\pi\pi^*)_{\min}$ realizado con cálculos CASPT2//CASSCF(14,10)/6–31G(p,d), en el uracilo.

En la figura 7.2 se muestra el esquema de la fotoquímica del uracilo, en el que se presentan los tres mecanismos propuestos que explican el proceso de población del estado triplete del uracilo, T_1 , a través de tres cruces entre sistemas.

La dependencia observada de la longitud de onda sobre el rendimiento cuántico del cruce entre sistemas, en el uracilo, puede ser entendida por la posición de las regiones de cruce STC. A energías de excitación altas, por

ejemplo, 5.39 eV, el Φ_{ISC} medido,^{6,7} (1.6×10^{-2}), es mayor que a 4.43 eV, (1.4×10^{-3}), probablemente porque en el primer caso los tres mecanismos de ISC descritos, con regiones de cruce STC cerca de 4.6 (los conducidos a través de los estados $n_0\pi^*$) y de 4.2 eV (el conducido a través del estado $\pi\pi^*$), pueden estar activos; mientras que, en el segundo caso, el estado $^1(\pi\pi^* \text{ HL})$ es poblado por debajo de la región de cruce STC y sólo el ISC conducido por estados $\pi\pi^*$ puede contribuir al proceso general.

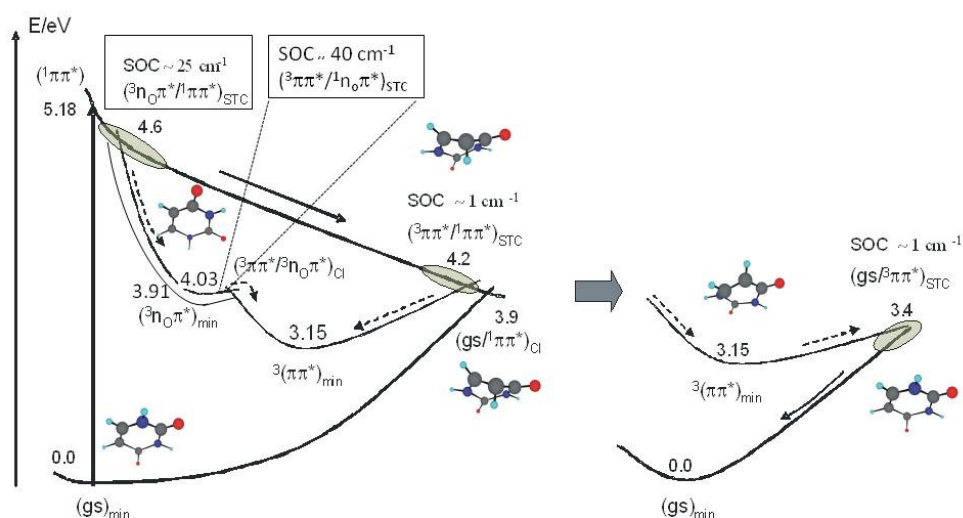


Figura 7.2. Esquema de la fotoquímica del uracilo, basado en los resultados de cálculos CASPT2, focalizado en la población del estado triplete más bajo.

En cualquier caso, estos mecanismos conducen a la población del triplete más bajo y el MEP a lo largo de $(^3\pi\pi^*)$ conduce al mínimo de dicho estado $^3(\pi\pi^*)_{\text{min}}$, en el que la molécula presenta un anillo distorsionado, con una deformación de 45° en el ángulo diedro formado por los átomos $\text{C}_2\text{N}_1\text{C}_6\text{C}_5$ y un aumento en la longitud del enlace C_5C_6 (1.503 \AA), con cierto carácter birradicalario que le confiere alta reactividad. El mínimo se sitúa a 3.15 eV,

calculado adiabáticamente desde el mínimo optimizado del estado fundamental, valor que puede compararse con los 3.3 eV estimados para la localización del estado triplete del mononucleótido de uracilo en disolución acuosa a temperatura ambiente,¹⁹ que es consistente con determinaciones teóricas previas de 3.2² y 3.1 eV.⁴⁶

Timina

La estructura del triplete más bajo de la timina es muy similar a la del uracilo, por tanto, los mecanismos fotofísicos propuestos para poblar dicho triplete en esta base pirimidínica, son semejantes a los ya explicados para aquélla.

En la tabla 7.2 se compilan las transiciones verticales, origen de banda, fuerzas de oscilador y tiempos de vida media radiativos, calculados a nivel CASSCF y CASPT2, de los estados singlete y triplete más bajos de la timina. Como puede observarse, en la región FC, el estado más bajo $^1(n_o\pi^*)$ (relacionado principalmente con el O_4) se encuentra situado 0.12 eV por debajo del estado espectroscópico $^1(\pi\pi^*L_a)$ HOMO \rightarrow LUMO, el cual se encuentra a 4.89 eV por encima del mínimo del estado fundamental, con una fuerza de oscilador de 0.167.

La figura 7.3 muestra la evolución del estado fundamental y los excitados singlete y triplete más bajos de la molécula timina, en el MEP calculado, partiendo de la geometría FC, a lo largo hipersuperficie de S_1 $^1(\pi\pi^*L_a)$. Al analizar dicha figura se pone de manifiesto que hay dos regiones diferentes de STC, que pueden alcanzarse fácilmente a lo largo del principal camino de desactivación energética, a través del MEP del $^1(\pi\pi^*L_a)$, partiendo de la geometría FC, siendo éstos los posibles canales de ISC hacia el estado triplete más bajo.

Nada más iniciarse dicho MEP, el estado $^1(\pi\pi^*L_a)$ se cruza tanto con el singlete como con el triplete $n\pi^*$. Tal como se ha descrito en el uracilo, estos cruces conducen a la población del estado triplete más bajo de la timina en las regiones de altas energías (4.8 eV).

Siguiendo la misma estrategia que para el uracilo, se han realizado varios MEPs conectando los distintos puntos críticos (véase figuras 7.3.a–c).

El MEP calculado desde la geometría del cruce (${}^1\pi\pi^*/{}^3n_O\pi^*$)_{STC}, a lo largo de la PEH del estado ${}^3(n_O\pi^*)$, lleva a la molécula hacia el mínimo de este estado (véase figura 7.3.a), en cuyas proximidades hemos encontrado una intersección cónica con el triplete más bajo, (${}^3n_O\pi^*/{}^3\pi\pi^*$)_{CL}, situada cerca de 3.9 eV respecto al mínimo del estado fundamental. A lo largo de este MEP, puede observarse que las PEHs del singlete y triplete $n\pi^*$ están siempre muy próximas, y cerca de la CI, hemos encontrado el cruce (${}^1n_O\pi^*/{}^3\pi\pi^*$)_{STC}. En el caso de que algo de población pasara al estado ${}^1(n_O\pi^*)$, a través del cruce con ${}^1(\pi\pi^*$ HL), (de hecho, se ha publicado un camino de desactivación a través de este intermedio⁴⁷) el paso energético hacia el estado triplete más bajo sería extremadamente favorable porque el valor de SOC calculado se incrementa en la zona (${}^1n_O\pi^*/{}^3\pi\pi^*$)_{STC} a 61 cm^{-1} .

Tabla 7.2. Propiedades calculadas para los estados excitados singlete y triplete más bajos de la timina.^a

Estado	Transición Vertical (eV)		Origen de Banda (T _e , eV)		τ_{rad}^c
	CASSCF	CASPT2 ^b	CASSCF	CASPT2	
${}^1(n_O\pi^*)^d$	5.41	4.77 (0.004)	4.23	4.05	2501 ns
${}^1(\pi\pi^*L_a)$	6.52	4.89 (0.167)	6.07	4.49	9 ns
${}^1(\pi\pi^*)$	7.36	5.94 (0.114)			
${}^3(\pi\pi^*L_a)$	3.95	3.59	2.99	2.87	17 ms
${}^3(n_O\pi^*)^d$	5.21	4.75	3.84	3.93	
${}^3(\pi\pi^*)$	5.86	5.14			

^a Véase también Ref. 37

^b Fuerzas de oscilador entre paréntesis.

^c Calculado usando la aproximación de Strickler–Berg.

^d Incluye básicamente al O₄ (en posición *orto* respecto al grupo metil).

Es posible confirmar nuestras propuestas acerca de la efectividad de este tipo de mecanismo gracias al estudio de Etinski y col.³¹ en el que se ha establecido la eficiencia del canal de ISC ($^1n_o\pi^*/^3\pi\pi^*$)_{STC} calculando los factores FC vibracionales y las tasas de ISC. Del mismo modo que en el uracilo, una región de STC ($^3\pi\pi^*/^1\pi\pi^*$) se encuentra en la región de baja energía, cerca del final del MEP $^1(\pi\pi^* HL)$ (véase figura 7.3), en este caso a 4.0 eV. Los valores de SOC calculados en los diferentes puntos del camino varían de 5 a 8 cm^{-1} .

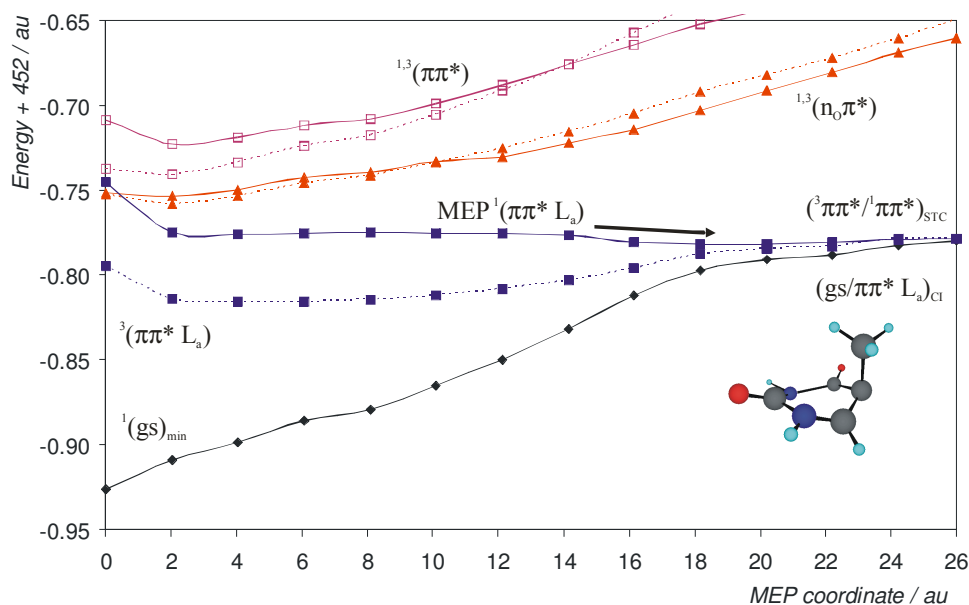


Figura 7.3. Evolución de los estados fundamental y excitados singlete y triplete más bajos para la timina desde la geometría FC a lo largo del MEP $^1(\pi\pi^*L_a)$.

Como aspecto final del estudio en la timina, hemos seguido la evolución de la hipersuperficie del triplete $^3(\pi\pi^*)$, desde la geometría ($^3n_o\pi^*/^3\pi\pi^*$)_{CI} a lo largo del MEP $^3(\pi\pi^*)$ (véase figura 7.3.b) para llegar a $^3(\pi\pi^*)_{\text{min}}$ y hemos localizado el cruce singlete–triplete que conecta dicho mínimo, $^3(\pi\pi^*)_{\text{min}}$, con el estado fundamental (véase figura 7.3.c). El cruce se encuentra cerca de 3.0 eV del mínimo del estado fundamental, lo cual implica una barrera de

0.13 eV (3.0 kcal/mol) para alcanzar $(gs/{}^3\pi\pi^*)_{STC}$ desde ${}^3(\pi\pi^*)_{min}$, y donde la molécula recupera su planaridad.

En la figura 7.4 se muestra el esquema con los tres mecanismos propuestos para explicar el proceso de población del estado triplete de la timina, T_1 , a través de tres cruces entre sistemas.

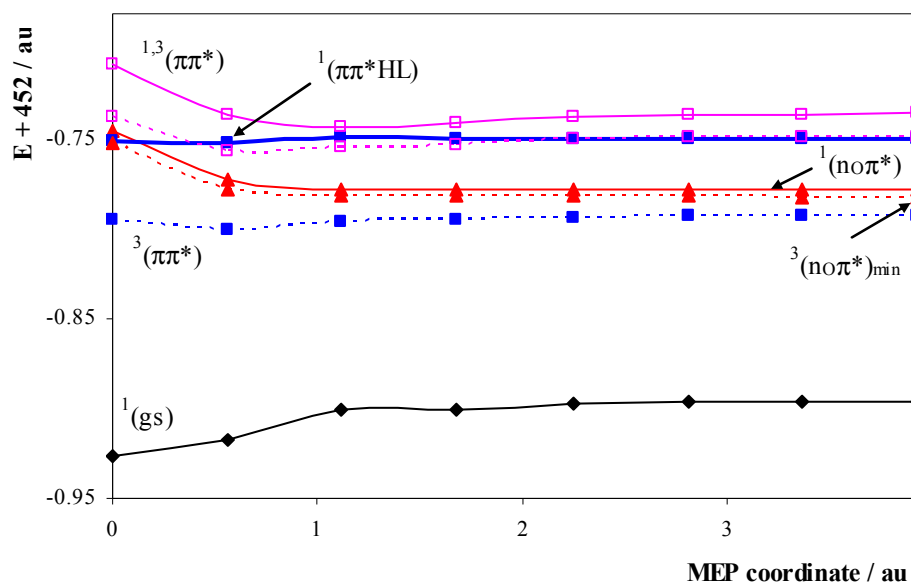


Figura 7.3.a. Evolución de los estados singlete y triplete más bajos de la timina desde la geometría $({}^3n_0\pi^*/{}^1\pi\pi^*HL)_{STC}$ a lo largo del MEP ${}^3(n_0\pi^*)$.

En el mínimo, ${}^3(\pi\pi^*L_a)$, la molécula muestra una estructura distorsionada, con una deformación en su anillo que incluye un ángulo diedro de 44° en $C_2N_1C_6C_5$ y un aumento en la longitud del enlace C_5C_6 (1.494 \AA) que le imprime un cierto carácter birradicalario. El mínimo se sitúa a 2.87 eV, calculado adiabáticamente desde el mínimo optimizado del estado fundamental, por debajo del estimado para el triplete del mononucleótido de la timina, 3.2 eV en disolución acuosa, a temperatura ambiente,¹⁹ y consistente con cálculos teóricos previos 2.8–3.0 eV.⁴⁶

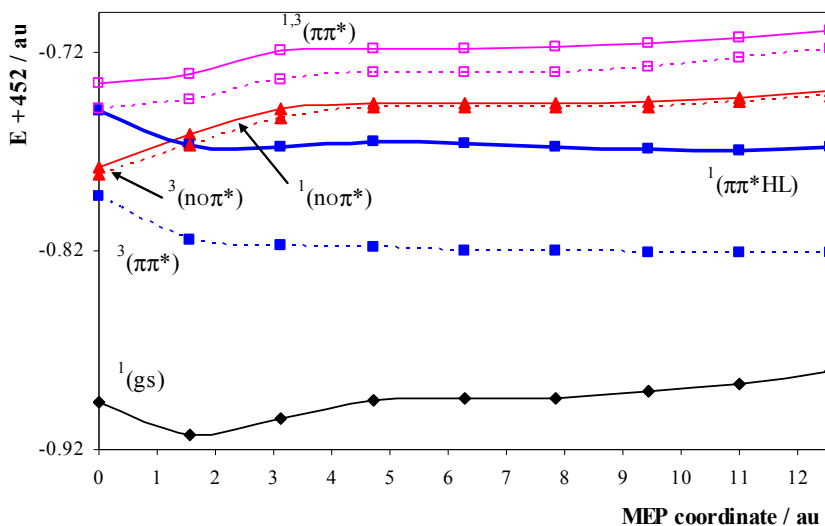


Figura 7.3.b. Evolución de los estados singlete y triplete más bajos de la timina desde la geometría $({}^3n_0\pi^*/{}^3\pi\pi^*)_{CI}$ a lo largo del MEP ${}^3(\pi\pi^*)$.

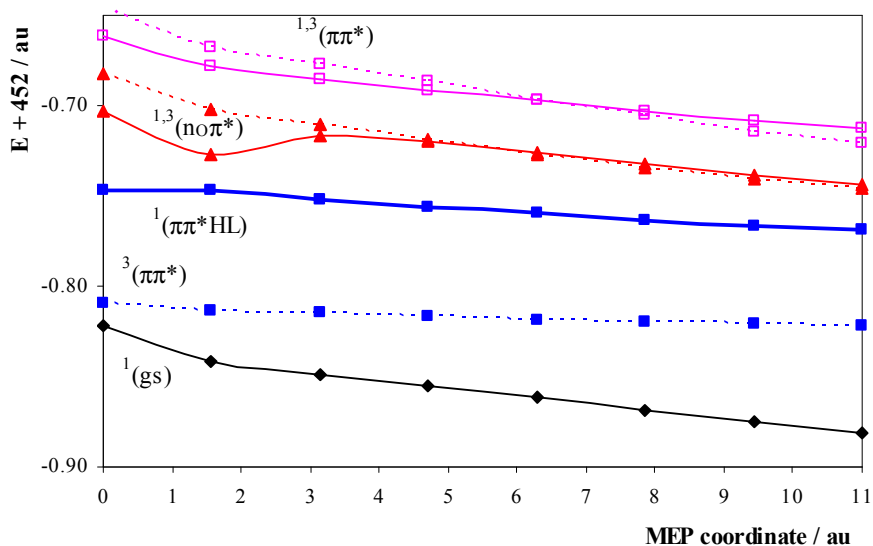


Figura 7.3.c. Evolución de los estados singlete y triplete más bajos de la timina desde la geometría $({}^1gs/{}^3\pi\pi^*)_{STC}$ a lo largo del MEP ${}^3(\pi\pi^*)$.

El presente modelo permite comprender la razón, desde un punto de vista molecular, por la que en las bases nucleicas existe una dependencia del rendimiento cuántico del ISC con la longitud de onda, seguramente causada por las dos regiones de interacción STC y su accesibilidad una vez se ha producido la excitación inicial del sistema. En el caso de la timina, el valor aumenta desde 3.9×10^{-3} , a 280 nm (4.43 eV), donde sólo el canal de más baja energía se puede alcanzar, hasta 5.2×10^{-2} , a 240 nm (5.17 eV),^{6,48} donde los otros dos canales descritos son accesibles.

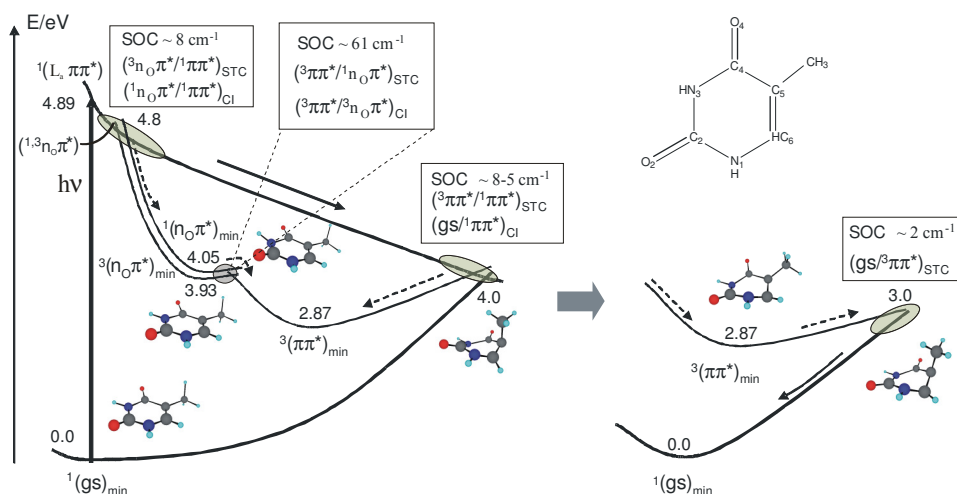


Figura 7.4. Esquema basado en resultados CASPT2, de la fotoquímica de la timina centrados en la población del triplete más bajo. El estado $^1\pi\pi^*$ representa al estado ($^1\pi\pi^*$ L_a) o ($^1\pi\pi^*$ HL).

Citosina

Con respecto a la citosina, los valores de la tabla 7.3 ayudan a entender, y también en cierta manera a predecir, el comportamiento fotofísico de su triplete. Ésta tiene un estado singlete más bajo $^1(\pi\pi^*$ HL) cuya interacción inicial con el estado $n\pi^*$, situado más alto energéticamente no será fuerte ni en la región de FC ni a lo largo del camino de desactivación del $^1(\pi\pi^*$ HL), como puede observarse en la figura 7.5.

Tabla 7.3. Propiedades calculadas para los estados excitados más bajos singlete y triplete de la citosina.

Estado	Transición Vertical (eV)		Origen de Banda (T_e , eV)		
	CASSCF	CASPT2 ^a	CASSCF	CASPT2	τ_{rad}^b
¹ (L _a $\pi\pi^*$)	5.22	4.41 (0.069)	4.14	3.62	30 ns
¹ (n _O π^*)	5.23	4.95 (0.001)	3.68	3.72	1200 ns
¹ (n _N π^*)	5.59	5.06 (0.003)			
¹ (L _b $\pi\pi^*$)	6.17	5.89 (0.106)			
³ (L _a $\pi\pi^*$)	3.64	3.53	2.85	2.98	437 ms
³ ($\pi\pi^*$)	4.87	4.45			
³ (n _O π^*)	5.13	4.63	3.49	3.66	
³ (n _N π^*)	5.31	4.94			

^a Fuerzas de oscilador entre paréntesis.

^b Calculados usando la aproximación de Strickler–Berg.

Como se ha visto en el capítulo 6, la relajación del singlete en la citosina es más compleja que en el resto de bases pirimidínicas. La presencia de un estado mínimo plano a 3.62 eV, casi isoenergético con la cónica tipo eteno ($gs/\pi\pi^*$)_{CI}, genera varios caminos de decaimiento competitivos que han sido ampliamente analizados.^{1,36,49} Es, por tanto, que hay más posibilidades de que ocurran diferentes procesos de ISC, pero siempre a bajas, y no a altas energías, como ocurre en la timina y en el uracilo. En particular, en las figuras 6.4 y 7.5 (en esta última con más detalle) se muestra el resultado de una interpolación lineal de las coordenadas internas, LIIC, en el camino desde la región FC hacia la CI tipo eteno con el estado fundamental. Ya quedó explicado en el capítulo anterior que la barrera a lo largo del estado ¹($\pi\pi^*$ HL), calculada en 2.5 kcal mol⁻¹ como límite superior, es lo suficientemente pequeña para que pueda considerarse el camino sin barreras.

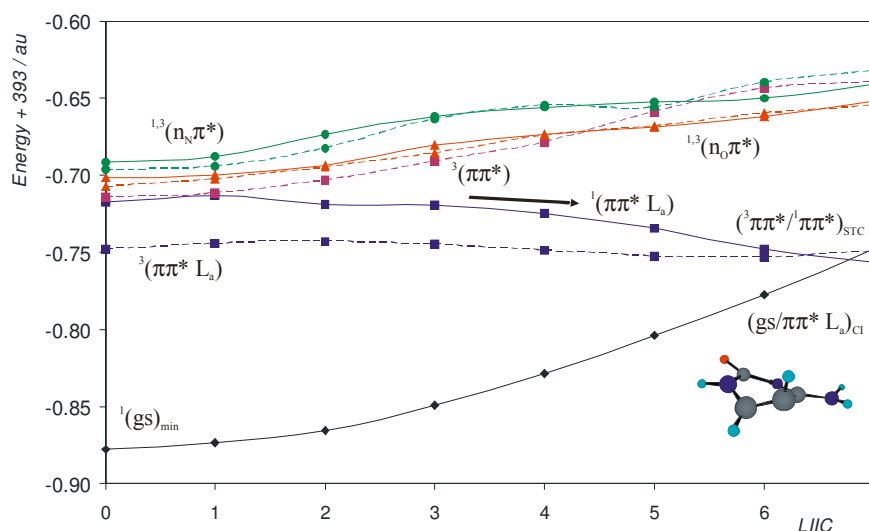


Figura 7.5. Evolución de los estados fundamental y excitados singlete y triplete más bajos para la citosina desde la geometría FC hasta la cónica $(gs/\pi\pi^*L_a)_{CI}$ al realizar la interpolación lineal en coordenadas internas, LIIC.

En un estudio previo,³⁶ hemos analizado los procesos de ISC que tienen lugar en regiones de baja energía. En la figura 7.6 se resume de forma esquemática el proceso fotofísico por el que se puebla el estado triplete de la molécula de citosina obtenido por medio de cálculos CASPT2 en la estructura de su estado excitado.

De la misma forma que en las otras nucleobases, a 3.6 eV tiene lugar un ISC entre los estados más bajos $\pi\pi^*$ cerca de la CI, con un valor de SOC de 6 cm^{-1} , pero a diferencia de las otras dos bases pirimidínicas, donde se han hallado tres canales básicos de población del triplete, dos a altas energías y otro a bajas energías, en la citosina parece que sólo el canal de baja energía es accesible. Esto podría ayudar a explicar la ausencia de contribución de la citosina en la fosforescencia del ADN^{20,21,22} así como los rendimientos cuánticos de fosforescencia, generalmente bajos, obtenidos para la citosina y sus derivados, comparados con las otras nucleobases.²⁵

Cabe indicar que el espectro de fosforescencia del ARN se compone principalmente de las propiedades individuales de los sistemas π de los

ribonucleótidos y en menor medida de las señales triplete de los grupos de adenosina y de centros de naturaleza desconocida que podrían ser asignadas a uracilos.⁵¹

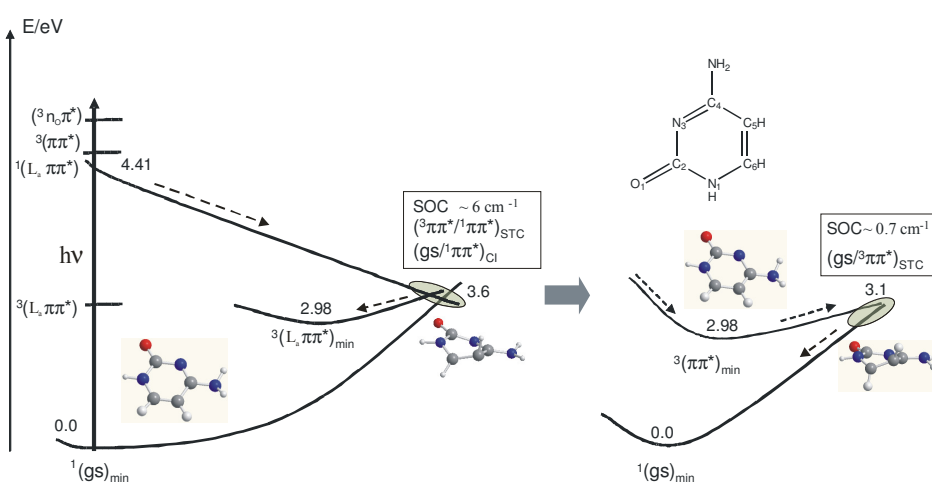


Figura 7.6. Esquema, basado en resultados CASPT2, de la fotoquímica de la citosina centrada en la población del triplete más bajo.

Conclusiones del capítulo 7

- ✓ Para poder establecer un mecanismo general de población del triplete de las tres bases pirimidínicas del ADN/ARN, uracilo, timina, y citosina, se han realizado cálculos químico-cuánticos *ab initio* del tipo multiconfiguracional CASPT2/CASSCF en las PEHs de sus estados más bajos singlete y triplete.
- ✓ La propuesta permite racionalizar las propiedades del estado triplete de los componentes del ADN/ARN, en particular, las medidas de mayores rendimientos cuánticos de fosforescencia que de fluorescencia en los sistemas individuales,^{6,23} la dependencia de la longitud de onda observada en la formación del triplete^{6,48} o la prevalencia de la contribución de la timina en las señales de fosforescencia del ADN a baja temperatura.^{20,21,22}
- ✓ Hemos analizado la accesibilidad de los canales de ISC para la población del estado triplete más bajo a lo largo del principal camino de decaimiento del estado singlete inicialmente poblado.
- ✓ Para poder determinar, sin ambigüedad, si las regiones propuestas de cruce singlete-triplete cumplen con los requisitos necesarios se ha analizado si las energías singlete-triplete están cercanas, si hay una alta densidad de estados vibracionales, con valores elevados de contribución vibracional a los acoplamientos espín-orbital y regiones donde la población se quede atrapada durante un tiempo suficiente para permitir que el proceso de ISC tenga lugar en competición con decaimientos por conversión interna.
- ✓ En la timina y el uracilo, nuestros resultados indican que tres regiones de STC pueden ser fácilmente alcanzadas a través del principal camino de desactivación desde el singlete, dos de ellas localizadas a altas energías, cruces mediados por la presencia de los estados singlete y triplete más bajos $n\pi^*$ y una tercera región, a baja energía, cerca del final del principal MEP-S₁. Estas tres regiones se

han propuesto como posibles canales de STC y, al menos aquellas relacionados con el STC $^1n\pi^* \rightarrow ^3\pi\pi^*$, parece que están confirmadas.³¹

- ✓ Adicionalmente, los resultados sugieren que la dependencia del rendimiento cuántico de formación del triplete observado en el uracilo y la timina podría estar relacionada con la activación de los tres canales de ISC (tanto a altas y bajas energías de excitación) o sólo de uno (a bajas energías).
- ✓ La citosina, por otro lado, en la que el estado excitado singlete, $^1(\pi\pi^*)$, se encuentra por debajo de los estados del tipo $n\pi^*$, no muestra mecanismos de ISC mediados por el $n\pi^*$ en regiones cercanas al inicio del principal MEP, y sólo tiene un canal de ISC eficaz a bajas energías, cerca de la CI del singlete, característica común a las tres bases. Por tanto, puede entenderse que la citosina contribuye en menor medida a la fosforescencia del ADN.^{20,21,22}

Referencias

1. Merchán, M.; González-Luque, R.; Climent, T.; Serrano-Andrés, L.; Rodríguez, E.; Reguero, M.; Peláez, D. *J. Phys. Chem. B* **2006**, *110*, 26471.
2. Marian, C. M.; Schneider, F.; Kleinschmidt, M.; *J. Tatchen, Eur. Phys. J. D* **2002**, *20*, 357.
3. Matsika, S. *J. Phys. Chem. A* **2004**, *108*, 7584.
4. Zgierski, M. Z.; Patchkovskii, S.; Fujiwara, T.; Lim, E. C. *J. Phys. Chem. A* **2005**, *109*, 9384.
5. Gustavsson, T.; Bányász, A.; Lazzarotto, E.; Markovitsi, D.; Scalamani, G.; Frisch, M. J.; Barone, V.; Improta, R. *J. Am. Chem. Soc.* **2006**, *128*, 607.
6. Cadet, J.; Vigny, P. The Photochemistry of Nucleic Acids. In *Bioorganic Photochemistry*, Morrison, H., Ed., John Wiley & Sons: New York, 1990, Vol. 1, pp 1-272.
7. Brown, I. H.; Johns, H. E. *Photochem. Photobiol.* **1968**, *8*, 273.
8. Beukers, R.; Ijlstra, J.; Berends, W.; *Rev. Trav. Chim. Pays. Bas.* **1960**, *79*, 101.
9. Danilov, V. I.; Slyusarchuk, O. N.; Alderfer, J. L.; Stewart, J. J.; Callis, P. R.; *Photochem. Photobiol.* **1994**, *59*, 125.
10. Leszczynski, J., Ed., *Computacional Molecular Biology*; Elsevier: Amsterdam, 1999.
11. Ericsson, L. A., Ed., Theoretical Biochemistry. Processes and Properties. In *Biological Systems*, Elsevier, Amsterdam, 2001.
12. Sztumpf-Kulikowska, E.; Shugar, D. Boag, J. W. *Photochem. Photobiol.* **1967**, *6* 41.
13. Klessinger, M. Triplet Photoreactions; Structural Dependence of Spin-Orbit Coupling and Intersystem Crossing in Organic Biradicals. In *Theoretical Organic Chemistry – Theoretical and Computational Chemistry*, Párkányi, C., Ed., Elsevier, Amsterdam, 1998, 581-610.
14. Serrano-Pérez, J. J.; Merchán, M.; Serrano-Andrés, L. *J. Phys. Chem. B.* **2008**, *112*, 14002.
15. Schreier, W. J.; Schrader, T. E.; Koller, F. O.; Gilch, P.; Crespo-Hernández, C. E.; Swaminathan, V. N.; Carell, T.; Zinth, W.; Kohler, B. *Science* **2007**, *315*, 625.
16. Roca-Sanjuán, D.; Olaso-González, G.; González-Ramírez, I.; Serrano-Andrés, L.; Merchán, M. *J. Am. Chem. Soc.* **2008**, *130*, 10768.
17. Whillans, D. W.; Johns, H. E. *Photochem. Photobiol.* **1969**, *6*, 323.

18. Gut, I. G.; Wood, P. D.; Redmond, R. W. *J. Am. Chem. Soc.* **1996**, *118*, 2366.
19. Wood, P. D.; Redmond, R. W. *J. Am. Chem. Soc.* **1996**, *118*, 4256.
20. Imakubo, K. *J. Phys. Soc. Jpn.* **1968**, *24*, 1124.
21. Szerenyi, P.; Dearman, H. H. *Chem. Phys. Lett.* **1972**, *15*, 81.
22. Arce, R.; Rodríguez, G. *J. Photochem.* **1986**, *33*, 89.
23. Gueron, M.; Eisinger, J.; Lamola, A. A. Excited States of Nucleic Acids. In *Basic Principles in Nucleic Acid Chemistry*, Tso, P. O. P., Ed.; Academic Press: New York, 1974, Vol. 1, pp 311-398.
24. Daniels, M. In *Photochemistry and Photobiology of Nucleic Acids*, Wang, S. Y., Ed., Academic Press: New York, 1976, Vol. 1, pp 23-108.
25. Görner, H. *J. Photochem. Photobiol. B: Biol.* **1990**, *5*, 359.
26. Bosca, F.; Lhiaubet-Vallet, V.; Cuquerella, M. C.; Castell, J. V.; Miranda, M. A. *J. Am. Chem. Soc.* **2006**, *128*, 6318.
27. Kang, H.; Lee, K. T.; Jung, B.; Ko, Y. J.; Kim, S. K. *J. Am. Chem. Soc.* **2002**, *124*, 12958.
28. Klessinger, M.; Michl, J. *Excited States and Photochemistry of Organic Molecules*, VCH Publishers, Inc.: New York, 1995.
29. Olivucci, M. *Computational Photochemistry*, Elsevier: Amsterdam, 2005.
30. Serrano-Andrés, L.; Merchán, M.; Borin, A. C. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 8691.
31. Etinski, M.; Fleig, T.; Marian, C. M. *J. Phys. Chem. A*, **2009**, *113*, 11809.
32. Tatchen, J.; Gilka, N.; Marian, C. M. *Phys. Chem. Chem. Phys.* **2007**, *9*, 5209.
33. Salzmann, S.; Tatchen, J.; Marian, C. M. *J. Photochem. Photobio. A* **2008**, *198*, 221.
34. Strickler, S. J.; Berg, R. A. *J. Chem. Phys.* **1962**, *37*, 814.
35. Rubio-Pons, O.; Serrano-Andrés, L.; Merchán, M. *J. Phys. Chem. A* **2001**, *105*, 9664.
36. Merchán, M.; Serrano-Andrés, L.; Robb, M. A.; Blancafort, L. *J. Am. Chem. Soc.* **2005**, *127*, 1820.
37. Serrano-Pérez, J. J.; González-Luque, R.; Merchán, M.; Serrano-Andrés, L. *J. Phys. Chem. B* **2007**, *111*, 11880.
38. Serrano-Andrés, L.; Merchán, M.; Borin, A. C. *Chem. Eur. J.* **2006**, *12*, 6559.
39. Serrano-Andrés, L.; Merchán, M.; Borin, A. C. *J. Am. Chem. Soc.* **2008**, *130*, 2473.
40. Climent, T.; González-Luque, R.; Merchán, M.; Serrano-Andrés, L. *Chem. Phys. Lett.* **2007**, *441*, 327.

41. González-Luque, R.; Climent, T.; González-Ramírez, I.; Merchán, M.; Serrano-Andrés, L. *J Chem Theor Comp.* **2010**, *6*, 2103.
42. Crespo-Hernández, C. E.; Cohen, B.; Hare, P. M.; Kohler, B. *Chem. Rev.* **2004**, *104*, 1977.
43. Canuel, C.; Mons, M.; Pluzzi, F.; Tardivel, B.; Dimicoli, I.; Elhanine, M. *J. Chem. Phys.* **2005**, *122*, 74316.
44. Climent, T.; González-Luque, R.; Merchán, M.; Serrano-Andrés, L. *J. Phys. Chem. A* **2006**, *110*, 13584.
45. Serrano-Pérez, J. J.; Merchán, M.; Serrano-Andrés, L. *Chem. Phys. Lett.* **2007**, *434*, 107.
46. Nguyen, M. T.; Zhang, R.; Nam, P. -C.; Ceulemans, A. *J. Phys. Chem. A* **2004**, *108*, 6554.
47. Hare, P. M.; Crespo-Hernández, C. E.; Kohler, B. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 435.
48. Nikogosyan, D. N.; Letokhov, L. S. *Riv. Nuov. Cim.* **1983**, *6*, 1.
49. Blancafort, L. *Photochem. Photobiol.* **2007**, *83*, 603.
50. Merchán, M.; Serrano-Andrés, L. *J. Am. Chem. Soc.* **2003**, *125*, 8108.
51. Kudrya, V. Y.; Yashchuk, V. M.; Levchenko, S. M.; Melnik, V. I.; Zaika, L. A.; Govorun, D. M. *Mol. Cryst. Liq. Cryst.* **2008**, *497*, 425.

8 Formación de ciclobutanos del ADN/ARN

Como ya se ha mencionado a lo largo de la presente Tesis, una de las evidencias más notorias de la fotosensibilidad del material genético a la acción de la luz UV es la formación de dímeros ciclobutánicos pirimidínicos (véase figuras 5.7 y 8.1) entre pares de bases nucleicas de pirimidinas adyacentes en las cadenas de ADN/ARN,¹ aductos que incluso pueden llegar a desencadenar procesos de fotomutagénesis y carcinogénesis.²

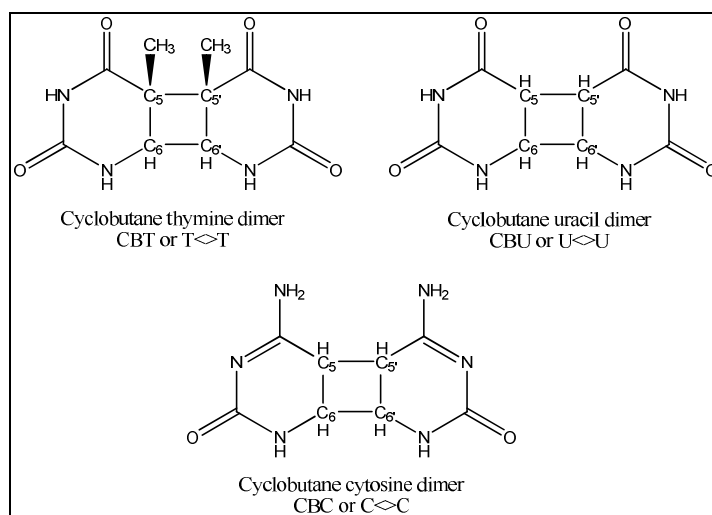


Figura 8.1. Estructuras y nomenclatura en inglés de los fotoproductos ciclobutánicos (CPD, CBPyr, Pyr\rightleftharpoonsPyr) de las tres bases pirimidínicas del ADN/ARN.

Actualmente está claramente establecido que la secuencia de bases en la cadena de ADN/ARN, y las orientaciones relativas entre ellas, son factores esenciales para la estabilidad, dinámica y reactividad del material genético, ya que aquellas orientaciones que maximizan el solapamiento entre las nubes π de las nucleobases empaquetadas favorecen la formación de excímeros o exciplejos,^{3,4,5,6} es decir, homodímeros o heterodímeros, respectivamente, asociados en un estado electrónico excitado y disociados en el estado fundamental (véase capítulo 5).⁷

La posibilidad de formación de dichos excímeros en el ADN surge como consecuencia de la disposición de los pares de bases, A-T y G-C, situados en el interior del esqueleto formado por las dos cadenas de azúcar-fosfato de la doble hélice. En ésta, una vuelta incluye diez pares de bases y tiene una altura de 34 Å, quedando la distancia interplanar entre pares de bases vecinas alrededor de 3.4 Å, valor que aparece frecuentemente en cristales orgánicos tipo excímero.⁷ Las geometrías que se encuentran, en el momento de producirse la irradiación con luz UV, alrededor de una estructura empaquetada cara a cara se pueden considerar como las mejores candidatas para ser precursoras del fotodímero, pues una vez formado el excímero todo apunta a que su evolución hacia la formación del CBPyr es inmediata. De hecho, en el modelo presentado por Schreier y colaboradores,⁸ se propone que el grado de flexibilidad del ADN-A, relativamente pequeño comparado con el ADN-B, es lo que le impide alcanzar la orientación adecuada para reaccionar, justificando de este modo la gran resistencia del ADN-A a formar dímeros Pyr \leftrightarrow Pyr. Como muestran dichos autores en el estudio de la timina con técnicas de espectroscopia de femtosegundos, la dimerización (fotorreacción ultrarrápida en la que los dímeros CBT se forman completamente en aproximadamente un picosegundo después de la irradiación con luz UV) sólo ocurre en los residuos de bases que están en disposición reactiva en el momento de la excitación, pues la tasa de formación de dímeros con pares de timina que se encuentran orientadas favorablemente es mucho más rápida que la tasa de cambio de orientación, situación que se puede extrapolar a las otras bases nucleicas del ADN/ARN.⁸

Por otra parte, determinaciones teóricas, realizadas con cálculos CASPT2, han aportado un mecanismo para la fotorreacción en los dímeros de timina y citosina que estaría mediado por la presencia de una estructura generada a través de una zona de intersección cónica entre dos estados degenerados energéticamente, el singlete excitado más bajo, S_1 , y el estado fundamental, S_0 .^{3,4,5,6} En dichos estudios químico-cuánticos se ha caracterizado, en el estado singlete, la reacción de fotocicloaddición [2+2] que lleva a la formación de los dímeros ciclobutánicos de timina, CBT, y citosina, CBC^{3,4,5,6} y se ha comprobado que las conformaciones paralelas cara a cara del tipo *cis-syn* son las más fotorreactivas, de acuerdo con los altos rendimientos obtenidos en los fotoproductos.^{1,2,3,8} En ellos se muestra que existen caminos de relajación sin barreras que conducen al estado excitado singlete de las bases nucleicas, inicialmente poblado, hasta dicha cónica (S_1/S_0)_{Cl}, en cuya estructura los enlaces etilénicos C_5-C_6 y $C_5'-C_6'$ se encuentran paralelos (tipo paralelogramo) y alargados (tipo *shearing*).

Aparte de la reacción fotoinducida no adiabática que tiene lugar en el singlete, el proceso de formación de fotodímeros también es concebible a través del estado triplete,⁹ que puede ser poblado desde el estado singlete por diferentes mecanismos de ISC, tal como ha sido ampliamente explicado en el capítulo anterior y en los artículos correspondientes.^{10,11,12,13,14}

Experimentalmente la participación de los tripletes en la fotoquímica del ADN/ARN no tiene controversia¹⁰ y los aductos CBPyr se forman en disolución bajo condiciones llamadas de “fotosensibilización triplete”,¹ es decir, se forman después de la fotosensibilización de las bases nucleicas, por una transferencia de energía triplete-triplete (*Triplet-Triplet Energy Transfer*, TTET) desde dadores favorables como, por ejemplo, cetonas o derivados ftalimidínicos.^{1,2,15,16,17,18} Todo ello gracias a la habilidad de las bases nucleicas del ADN/ARN empaquetado a comportarse comoceptoras bajo condiciones favorables de fotosensibilización del triplete. Es importante recalcar que, para que una transferencia de energía tenga lugar de forma eficiente: el dador debe absorber en la región de interés, tener un estado triplete más alto en energía que el del aceptor y se debe producir un ISC eficiente. Así, por ejemplo, estudios experimentales realizados sobre la formación de dímeros de CBT a través de una TTET proponen que para que un fotosensibilizador dado pueda llegar a producir una fotolesión

relacionada con aductos de la timina la energía de su estado triplete ha de estar por encima de 2.6-2.8 eV,^{18,19} considerado como el intervalo de energía del estado triplete de la timina en el ADN. Gracias a un fotosensibilizador exógeno (o endógeno si ocurre entre diferentes nucleobases del ADN/ARN) se podría poblar el estado triplete de un monómero de la base nucleica, y éste interaccionar atractivamente con el estado fundamental de otro monómero, formando un excímero que evolucionaría hacia una estructura intermedia precursora de su dímero CBPyr. En la figura 8.2 se muestra un esquema, basado en resultados CASPT2, de la formación inducida de CBC a partir de dos moléculas de citosina que en una organización cara a cara forman el excímero de la nucleobase. Como puede observarse, el proceso sucede en el singlete evolucionando el sistema a través de una cónica, $(S_1/S_0)_{Cl}$,[♦] o en el triplete, que desemboca en la formación de un intermedio triplete, (*Step Wise Intermediate*, ³(SWI)), en el que se ha formado un enlace covalente entre los carbonos C₆-C_{6'} de las bases adyacentes. Dicha estructura intermedia tiene carácter birradical, con electrones no apareados y densidad de espín en los dos átomos de carbono, C₅ y C_{5'}, que no han formado enlace.³ Se sabe que para que se produzca un cruce singlete-triplete birradical degenerado en una cicloadición [2+2] entre enlaces etilénicos, se requiere una distribución de carga asimétrica en los centros de los carbonos³ y esto ocurre en las bases nucleicas empaquetadas, donde los átomos C₆ son más electrón deficientes que los átomos C₅, pues la afinidad electrónica del nitrógeno vecino al C₆ es mayor que la del C₅ en el anillo pirimidínico. Los átomos de carbono C₆-C_{6'} sustituidos compensan la falta de densidad electrónica acercándose a los otros monómeros y formando un nuevo enlace cuya distancia, en el caso de la citosina, es de 1.669 Å, mientras que los monómeros no enlazados permanecen a ~2.8 Å (véase figura 8.2). Este intermedio evoluciona hacia la formación del dímero correspondiente CBPyr y, por tanto, su caracterización debería proporcionar información altamente relevante acerca del proceso fotorreactivo.

♦ El mecanismo de formación del CBC y CBT a través del singlete ha sido ampliamente analizado en las Tesis de la Dra. Gloria Olaso-González (*Estudio teórico de procesos fotoinducidos en sistemas biológicos*, 2009. Universitat de València) y el Dr. Daniel Roca-Sanjuán (*Bases moleculares de la Fotorreactividad y Transferencia de carga en el ADN*, 2009. Universitat de València).

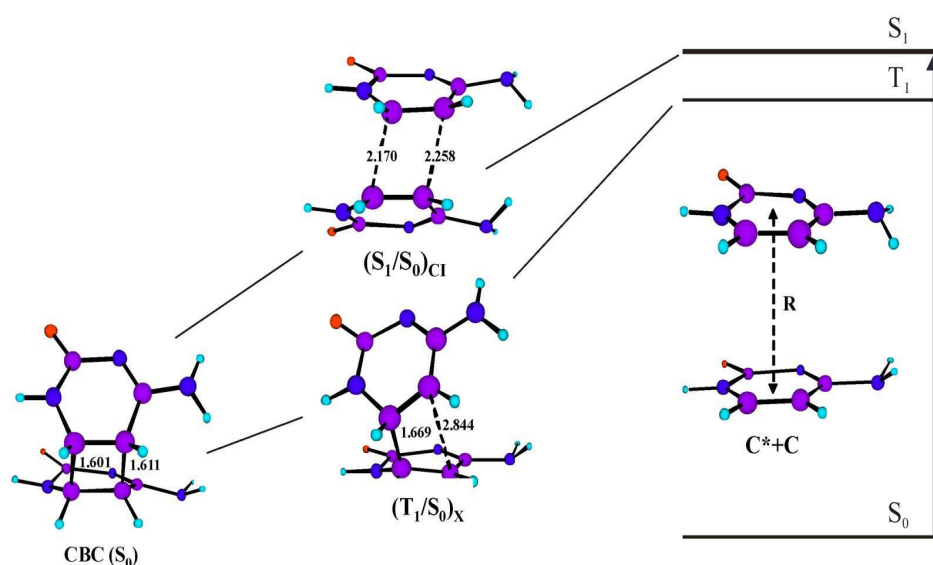


Figura 8.2. Esquema basado en cálculos CASPT2 de la formación fotoinducida de CBC desde el excímero de citosina empaquetado π , ($C^* + C$), a lo largo del singlete y triplete.

A continuación se presenta el estudio realizado sobre el mecanismo de fotodimerización, a través del estado triplete en el uracilo. El estudio analiza la formación del excímero UU, la obtención del intermedio triplete $^3(SWI)$ y la formación del aducto dímero CBU. Se completa el estudio comparando los resultados del uracilo con los de las otras dos bases pirimidínicas analizadas por el grupo QCEXVAL.^{3,4}

Detalles computacionales

Las optimizaciones de la geometría del estado fundamental de los dímeros $\text{Pyr} \langle \rangle \text{Pyr}$ y excímeros triplete se han llevado a cabo con cálculos CASSCF incluyendo 16 electrones π distribuidos en 14 MOs π . El MO π totalmente

simétrico de cada una de las pirimidinas se mantuvo inactivo. Para poder mimetizar la interacción de las bases en el ADN de la forma más real posible las optimizaciones se llevaron a cabo con las restricciones de simetría C_s . Una vez optimizadas las geometrías, las energías se calcularon sin restricciones de simetría, C_1 , requisito indispensable para poder describir correctamente las transiciones electrónicas entre las dos bases en el límite asintótico. En este caso, dos MOs π extra se mantuvieron también inactivos dado que el número de ocupación de sus correspondientes orbitales naturales era prácticamente 2.0. Por tanto, se usó una función de onda CASSCF con un espacio activo de 12 electrones π en 12 MOs π , CASSCF(12,12), siempre corrigiendo el efecto del error de superposición de base (*Basis Set Superposition Error*, BSSE) usando el procedimiento de counterpoise.²⁰ Las CI y los STC se estimaron a partir de los puntos de cruce de mínima energía (*Minimum energy crossing point*, MECP) en las PEHs de los sistemas de dos bases nucleicas. El conjunto de funciones de base empleado fue ANO-S C,N,O[3s2p1d]/H[2s1p]. Todos los cálculos se realizaron utilizando el paquete de software químico cuántico MOLCAS-7.²¹

Excímeros de uracilo

Siguiendo la misma estrategia que la empleada en el estudio teórico de los excímeros de citosina³ y timina,⁴ se analizaron las PECs del estado fundamental y de los estados excitados más bajos del homodímero de uracilo a lo largo de la distancia intermolecular, R , de los dos monómeros π empaquetados (véase las figuras 8.3 y 8.4) de la estructura tipo sándwich cara a cara porque, como se ha mencionado, es la que tiene en cuenta el máximo solape entre las nubes π y, por ende, representa la conformación más favorable para estabilizar un excímero.

Como puede observarse, comparando las figuras 8.3 y 8.4, la ausencia de BSSE produce una sobreestimación de las energías de enlace de los estados. Con la corrección BSSE-CASPT2 (12,12), el estado fundamental pasa a ser disociativo, mientras que el estado excitado más bajo permanece enlazado, con unos mínimos S_1 y T_1 situados a $R(C_5-C_5') \sim 3 \text{ \AA}$.

En las PCEs se ha monitorizado la energía en función de la distancia intermolecular C_5-C_5' , por ser ésta particularmente relevante en la formación posterior de los fotodímeros de ciclobutano. En ambas figuras, las energías están calculadas respecto al estado fundamental de las dos moléculas de uracilo separadas alrededor de 20 au. La absorción $S_0 \rightarrow S_1$ a esta distancia corresponde a la absorción del monómero, con un valor calculado para dicha transición de 4.83 eV, en concordancia con las evidencias experimentales en fase gas y resultados teóricos previos (véase capítulo 6).

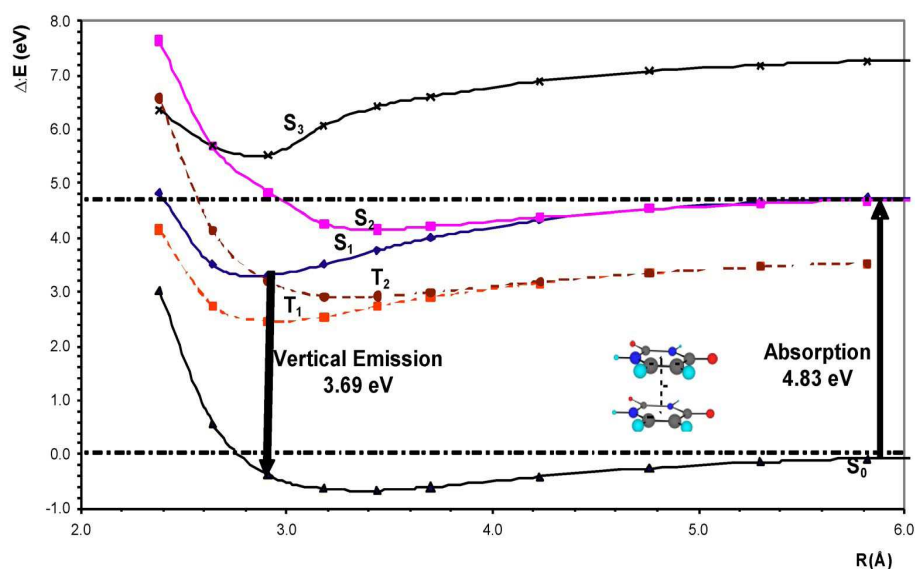


Figura. 8.3. PECs CASPT2(12,12)/ANO-S $C_5N_5O[3s2p1d]/H[2s1p]$ construidas con respecto a la distancia intermolecular ($R = C_5-C_5'$) de dos moléculas de uracilo empaquetadas cara a cara.

Como puede observarse, en el límite asintótico tanto S_1 y S_2 , como T_1 y T_2 , están degenerados, situación relacionada con $U + U^*$ y $U^* + U$, donde U y U^* representan el estado fundamental y el estado excitado singlete o triplete más bajo, respectivamente. El estado S_3 no se tendrá en cuenta en la discusión por corresponder a un estado excitado del monómero de mucha mayor energía.

Cuando se incluye en el cálculo la corrección del BSSE (véase figura 8.4) la emisión vertical desde el mínimo del estado S_1 , a 3.70 eV, permanece

prácticamente igual al resultado sin BSSE, a 3.69 eV, y con una diferencia de tan solo 0.3 eV de la emisión vertical calculada para el monómero de uracilo (véase capítulo 6). A la distancia intermolecular de ~ 2.9 Å los estados S_1 y T_2 están involucrados en un cruce singlete-triplete y, por tanto, T_2 podría ser poblado a través de un mecanismo de ISC y posteriormente poblar T_1 a través de una CI.

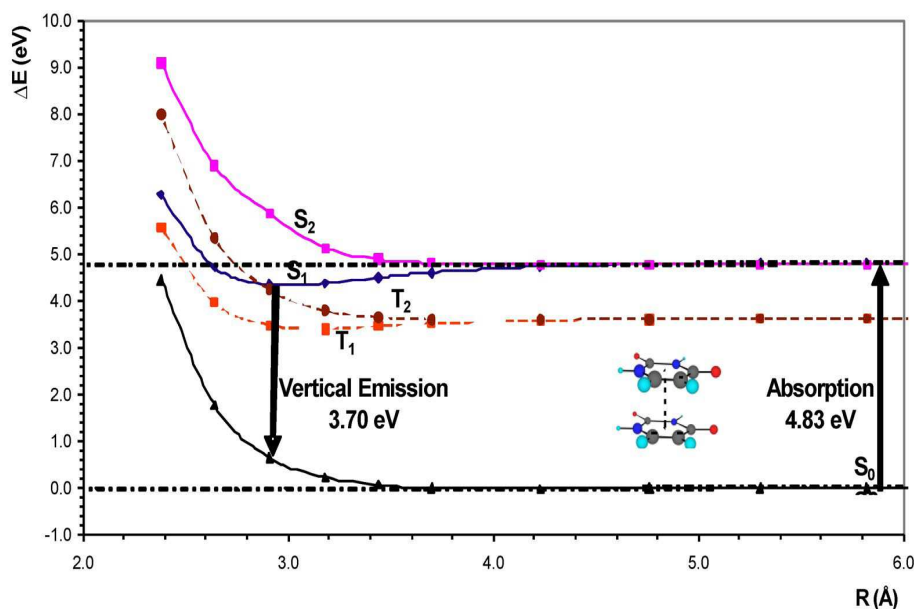


Figura. 8.4. PECs CASPT2(12,12)/ANO-S C,N,O[3s2p1d]/H[2s1p] con corrección BSSE construidas con respecto a la distancia intermolecular ($R = C_5-C_5'$) de dos moléculas de uracilo empaquetadas cara a cara.

Formación de ciclobutiluracilo a través del intermedio triplete

Tal como se ha mencionado, la posibilidad de formación de excímeros permite al sistema evolucionar hacia su dimerización y es, por tanto, que a partir del excímero formado se realizó el estudio sobre el posible camino de fotodimerización del uracilo.

La optimización del excímero triplete de uracilo conduce a un mínimo estable, $^3(\text{SWI})$, (véase figura 8.5), siendo esta estructura una buena candidata como precursora en la formación del dímero CBU.

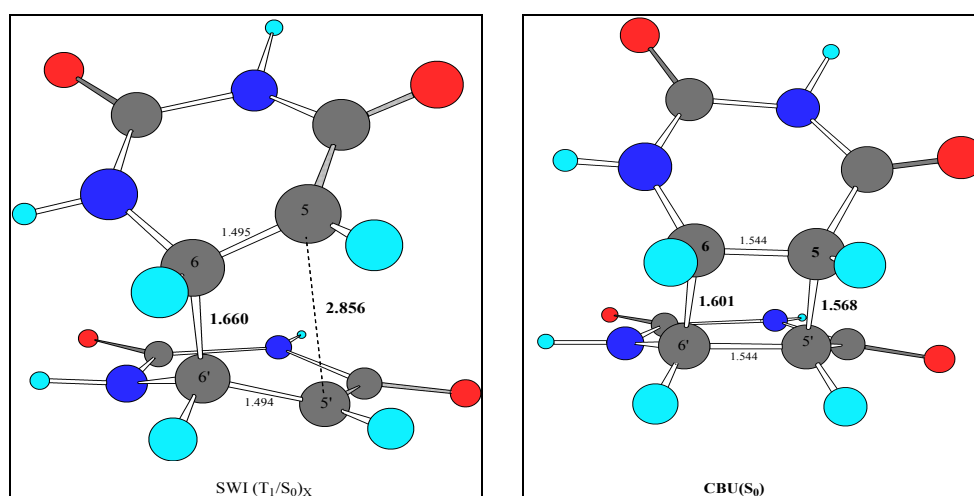


Figura 8.5. Estructura del intermedio $^3(\text{SWI})$ (izquierda) y $\text{CBU}(\text{S}_0)$ (derecha) del uracilo.

La estructura intermedia obtenida se caracteriza porque en ella se ha formado un enlace covalente simple entre ambos monómeros de uracilo, con una distancia $\text{C}_6\text{-C}_{6'}$ calculada de 1.660 Å. Por contra, la distancia de los carbonos no enlazados $\text{C}_5\text{-C}_{5'}$, permanece alargada alrededor de 2.856 Å, localizándose la densidad de espín básicamente en estos átomos no enlazados, C_5 y $\text{C}_{5'}$.

El estado triplete del intermedio queda situado a 2.47 eV por encima del estado fundamental de las dos moléculas de uracilo, diferencia energética que puede ser considerada como un límite inferior para la formación del triplete del uracilo en el ARN, y de forma general puede relacionarse con el umbral necesario para alcanzar el estado triplete del uracilo a partir de fotosensibilizadores exógenos o, tal como se ha mencionado, por un mecanismo de ISC y posterior CI. Una vez poblado el estado triplete del monómero, el excímero evoluciona a través del intermedio $^3(\text{SWI})$ hacia la

formación del dímero, alcanzando la estructura del estado fundamental del dímero de uracilo CBU(S_0) (véase figura 8.5) en la cual los enlaces etilénicos son prácticamente paralelos, siendo las distancias C_5-C_5' y C_6-C_6' de 1.568 y 1.601 Å, respectivamente.

Para analizar la evolución del intermedio $^3(SWI)$ hacia el fotodímero CBU se calculó el MEP a lo largo de la hipersuperficie del triplete T_1 (véase figura 8.6).

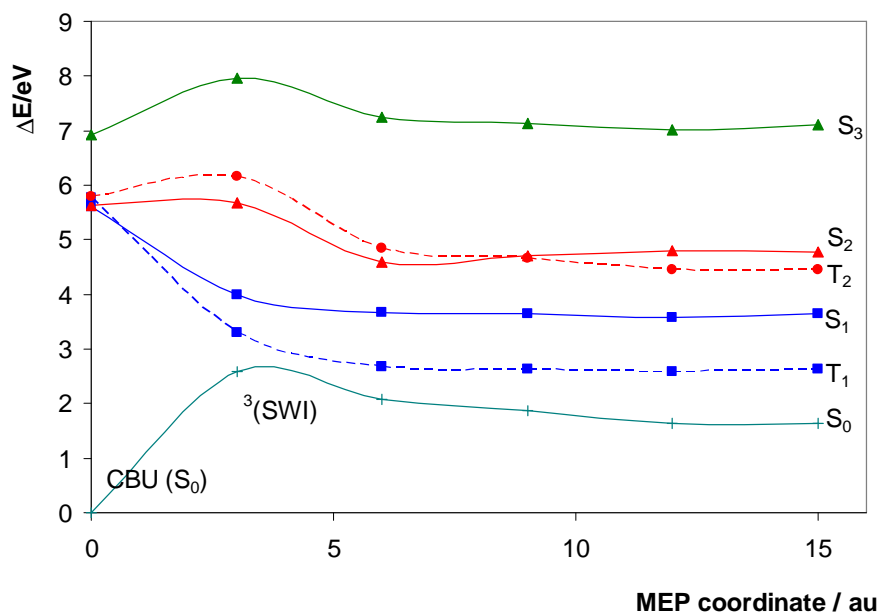


Figura 8.6. Estados excitados triplete y singlete más bajos del dímero de uracilo, calculados a nivel CASPT2//CASSCF a lo largo del MEP del estado T_1 realizado computacionalmente desde la geometría del estado fundamental del dímero de ciclobutiluracilo.

Dado que el triplete del CBU tiene una energía superior al mínimo $^3(SWI)$, para realizar el MEP- T_1 se partió de la geometría del estado fundamental del dímero de ciclobutiluracilo CBU(S_0). El resultado obtenido muestra que la estructura del intermedio $^3(SWI)$ conecta el estado triplete más bajo del dímero CBU con el mínimo del excímero triplete. Cabe destacar que esta

estructura intermedia se encuentra en una región de la hipersuperficie donde el estado fundamental del dímero aparece degenerado con el estado triplete, es decir, coinciden en un cruce singlete-triplete, $(T_1/S_0)_x$ y, consecuentemente, el decaimiento al estado fundamental del homodímero se ve particularmente favorecido. La estructura intermedia $^3(\text{SWI})$ es, pues, la mejor candidata para disparar la formación del fotoaducto ciclobutánico después del ISC y evolucionar en la región de la PEH del estado fundamental del aducto CBU. El hecho de que no haya barrera implicaría que el camino es posible en ambos sentidos y, por tanto, se pueda producir reversibilidad en la formación de los monómeros de uracilo a partir del dímero CBU.

Estos resultados en el uracilo junto con los obtenidos por nuestro grupo en la timina y citosina permiten proponer un esquema para la formación de los dímeros de ciclobutil pirimidina (véase figura 8.7) a través del estado triplete en los pares de bases aislados de ADN/ARN.

Fotodimerización de las bases nucleicas pirimidínicas

El comportamiento de los homodímeros de las tres bases pirimidínicas naturales es muy similar respecto de sus propiedades de enlace, a pesar de que la formación del excímero de timina debería estar estéricamente impedido por la presencia del grupo metilo, pero en cualquier caso, la conclusión es que los tres homodímeros dan lugar a excímeros. En la tabla 8.1 se muestran las energías del estado triplete más bajo, tanto del monómero como del dímero, de las tres bases pirimidínicas del ADN/ARN. En el monómero, las tres bases nucleicas aisladas tienen señales experimentales del triplete cercanas a 3.5-3.6 eV.^{13,22,23} Los cálculos computacionales predicen, tanto para las transiciones verticales singlete-triplete como para las adiabáticas (origen de banda), que el uracilo se queda por encima energéticamente, seguido por la citosina y la timina (véase figura 8.7).

Como se puede observar, en las tres bases el intermedio $^3(\text{SWI})$ es un mínimo real en la PEH del estado T_1 y también corresponde a una estructura de cruce singlete-triplete $(T_1/S_0)_x$ que, como ya se ha mencionado, será la que mediará en la formación del fotoproducto CBPyr.

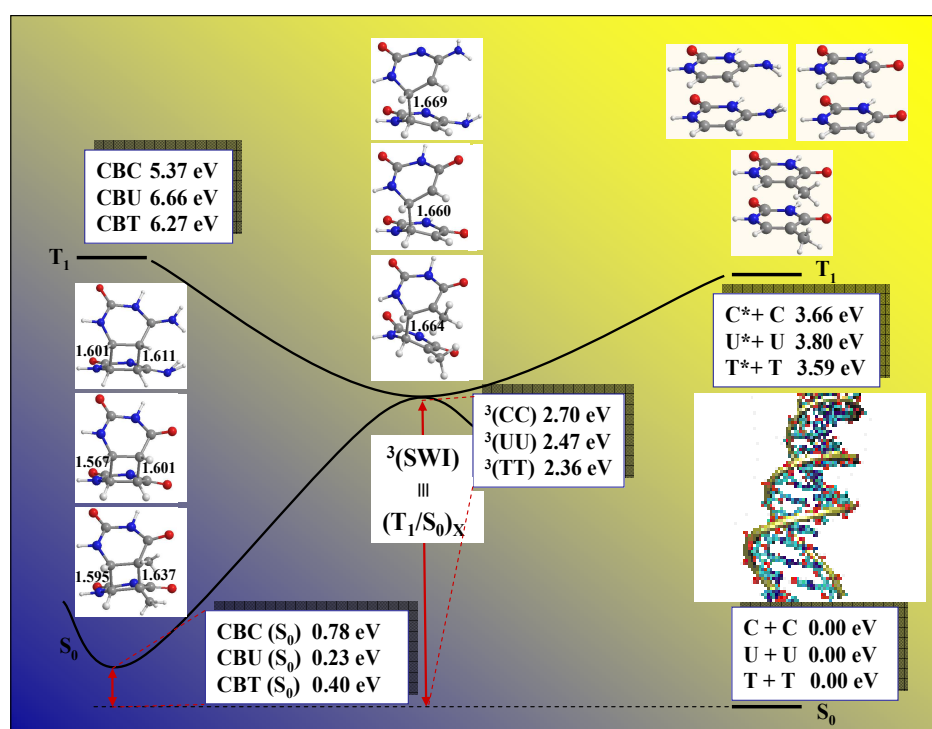


Figura 8.7. Esquema basado en resultados CASPT2 de la formación de los dímeros de ciclobutil pirimidina (izquierda) a través del triplete en los pares de bases aislados de ADN/ARN. Centro: intermedio singlete-triplete, $^3(\text{SWI})$. Derecha: pares de bases empaquetadas por separado. Las distancias de enlace (Å) se indican por su relevancia en: $C_6-C_{6'}$ en $^3(\text{SWI})$; $C_6-C_{6'}$ y $C_5-C_{5'}$ en los aductos.

En la localización del intermedio $^3(\text{SWI})$ se ha encontrado que el más bajo en energía es el del homodímero de timina (2.36 eV) seguido por el de uracilo (2.47 eV) y citosina (2.70 eV), que puede relacionarse con el límite mínimo de energía necesaria para poblar el estado triplete en el ADN/ARN. Este orden de energía se ha obtenido en experimentos en disoluciones acuosas^{18,19} y en medios rígidos,²⁴ y se ha encontrado que los aductos CBT y CBU dimerizan más rápidamente y más eficientemente que los CBC en presencia de acetona y acetofenona.¹

Los cálculos CASPT2 corroboran los datos experimentales ya que predicen que la producción de fotoaductos CBT y CBU, en el estado triplete, tiene

lugar a similares energías y condiciones, pero, sin embargo, se requieren energías superiores para que se forme el CBC, el cual puede estar también menos favorecido debido a la formación de excímeros muy estables en la citosina, tal como ocurre para el estado singlete.³

Así pues, dado que la eficacia de la formación del triplete es mayor en la timina que en la citosina,^{11,14} es posible racionalizar las preferencias por los sitios TT, y en menor extensión por los CT, TC y CC, para generar fotoproductos ciclobutánicos.

En resumen, una vez se ha propuesto la presencia de un precursor excímero singlete, que favorece la conversión interna ultrarrápida a través de una cónica $(S_0/S_1)_{CI}$ ^{3,4,5,6,8} como la fuente más importante de formación de dímeros CBPyr de las bases nucleicas, tanto en polímeros como en disoluciones congeladas,¹ el esquema que presentamos trata de racionalizar el mecanismo para la fotosensibilización de fotoaductos triplete-triplete desde un dador apropiado, proceso de transferencia de energía que se espera sea más favorable en disolución.¹ Proponemos, pues, que la estructura intermedia ³(SWI) es la mejor candidata para disparar la formación del fotoaducto ciclobutánico después del ISC y evolucionar en la región de la PEH del estado fundamental del aducto CBPyr, proceso que se ha sugerido tiene una eficiencia de cuasi la unidad.²⁵ La eficiencia real del proceso de decaimiento a lo largo del triplete también se basará en el aumento de los valores de acoplamiento espín-órbita. En este sentido, el disolvente, o en general el medio, se espera que juegue un papel crucial en el proceso de ISC.²⁶ La determinación de la naturaleza y localización relativa de los intermedios tripletes en los exciplejos de las diferentes nucleobases puede ayudar a racionalizar la fototoxicidad de potenciales fotosensibilizadores a consecuencia de la energía de sus estados tripletes. Del mismo modo se puede utilizar para diseñar derivados de la nucleobase adecuada, o fármacos fototerapéuticos, que sirvan para disminuir e incluso impedir la formación de dímeros de pirimidina.

Tabla 8.1. Energías experimentales y teóricas CASPT2 (eV) para el estado excitado más bajo, T_1 , de las bases pirimidínicas del ADN/ARN.^a

	Monómero		Dímero	Monómero
	E_{VA} ^b	T_e ^c	T_e (T_1/S_0) _X ^d	Exp ^e
Citosina	3.66	2.98	2.70	3.50
Uracilo	3.80	3.15	2.47	3.65
Timina	3.59	2.87	2.36	3.60

^a Todas las energías están referidas al estado fundamental singlete de dos bases aisladas. Véase Figura 8.7.

^b Energía de excitación vertical singlete-triplete (E_{VA}) en el mínimo del estado fundamental FC del monómero (S_0).

^c Energía adiabática de mínimo a mínimo de los monómeros S_0-T_1 , (T_e): origen de banda de excitación electrónica.

^d Energía adiabática de mínimo a mínimo de los homodímeros S_0-T_1 , correspondiente al cruce singlete-triplete más bajo (T_1/S_0)_X y la especie intermedia ³(SWI).

^e Banda máxima experimental del monomero singlete-triplete, corresponde aproximadamente a E_{VA} . Refs. 13,22 y 23.

Conclusiones del capítulo 8

- ✓ Se ha estudiado a nivel teórico CASPT2 el proceso de formación fotoinducido de los dímeros ciclobutánicos pirimidínicos en el estado excitado triplete.
- ✓ Se propone un mecanismo a través del triplete del homodímero para los pares de citosina, timina y uracilo en los que se incluye una estructura intermedia, $^3(\text{SWI})$, cruce singlete-triplete, de carácter birradical y que representa la conformación más estable del triplete de las bases nucleicas en el ADN.
- ✓ La eficiencia del mecanismo está modulada por dos factores:
 - la efectividad del proceso de transferencia de energía triplete-triplete desde una molécula fotosensibilizadora dadora, que está relacionada con la posición relativa de las tres bases pirimidínicas, requiriéndose menor energía en los pares de timina y uracilo que en los de citosina y
 - el proceso de cruce entre sistemas hacia el estado fundamental del fotoproducto. Cabe destacar que esta estructura intermedia $^3(\text{SWI})$ se encuentra en una región de la hipersuperficie donde el estado fundamental del dímero aparece degenerado con el estado triplete, es decir, coinciden en un cruce singlete-triplete, $(T_1/S_0)_x$ y, consecuentemente, el decaimiento al estado fundamental del homodímero se ve particularmente favorecido.
- ✓ Los rendimientos de formación más bajos observados en los dímeros de citosina, CBC, comparados con los de la timina, CBT, se puede relacionar con el hecho de que el intermedio triplete $^3(\text{SWI})$, mediador en la formación del fotoaducto a través de un proceso de transferencia de energía, está más alto y, por tanto, menos accesible en la citosina.

Referencias

1. Cadet, J.; Vigny, P. The Photochemistry of Nucleic Acids. In *Bioorganic Photochemistry*; Morrison, H., Ed.; John Wiley & Sons: New York, 1990; Vol. 1, pp 1-272.
2. Douki, T.; Cadet, J. *Biochemistry*, **2001**, *40*, 2495.
3. Roca-Sanjuán, D.; Olasso-González, G.; González-Ramírez, I.; Serrano-Andrés, L.; Merchán, M. *J. Am. Chem. Soc.* **2008**, *130*, 10768.
4. Serrano-Pérez, J. J.; González-Ramírez, I.; Coto, P. B.; Merchán, M.; Serrano-Andrés, L. *J. Phys. Chem. B.* **2008**, *112*, 14096.
5. Boggio-Pasqua, M.; Groenhof, G.; Schäfer, L. V.; Grubmüller, H.; Robb, M. A. *J. Am. Chem. Soc.* **2007**, *129*, 10996.
6. Blancafort, L.; Migani, A. *J. Am. Chem. Soc.* **2007**, *129*, 14540.
7. Birks, J. B. *Rep. Prog. Phys.* **1975**, *38*, 903
8. Schreier, W. J.; Schrader, T. E.; Soller, F. O.; Gilch, P.; Crespo-Hernández, C. E.; Swaminathan, V. N.; Carell, T.; Zinth, W.; Kohler, B. *Science* **2007**, *315*, 625.
9. Kwok, W. M.; Ma, C.; Phillips, D. L. *J. Am. Chem. Soc.* **2008**, *130*, 5131.
10. Hare, P. M.; Crespo-Hernandez, C. E.; Kohler, B. *Proc. Natl. Acad. Sci. USA.* **2007**, *104*, 435.
11. Merchán, M.; Serrano-Andrés, L.; Robb, M. A.; Blancafort, L. *J. Am. Chem. Soc.* **2005**, *127*, 1820.
12. Climent, T.; González-Luque, R.; Merchán, M.; Serrano-Andrés, L. *Chem. Phys. Lett.* **2007**, *441*, 327.
13. Etinski, M.; Fleig, T.; Marian, C. M. *J. Phys. Chem. A* **2009**, *113*, 11809.
14. Serrano-Pérez, J. J.; González-Luque, R.; Merchán, M.; Serrano-Andrés, L. *J. Phys. Chem. B.* **2007**, *111*, 11880.
15. Gut, I. G.; Wood, P. D.; Redmond, R. W. *J. Am. Chem. Soc.* **1996**, *118*, 2366.
16. Cadet, J.; Sage, E.; Douki, T. *Mutat. Res.* **2005**, *571*, 3.
17. Nguyen, M. T.; Zhang, R.; Nam, P.-C.; Ceulemans, A. *J. Phys. Chem. A* **2004**, *108*, 6554.
18. Bosca, F.; Lhiaubet-Vallet, V.; Cuquerella, M. C.; Castell, J. V.; Miranda, M. A. *J. Am. Chem. Soc.* **2006**, *128*, 6318-6319.
19. Albin, A.; Monti, S. *Chem. Soc. Rev.* **2003**, *32*, 238.
20. Boys, S. F.; Bernardi, F. *Mol. Phys.* **2002**, *100*, 65.
21. Aquilante, F.; De Vico, L.; Ferré, N.; Ghigo, G.; Malmqvist, P.-Å.; Pedersen, T.; Pitonak, M.; Reiher, M.; Roos, B. O.; Serrano-Andrés, L.;

- Urban, M.; Veryazov, V.; Lindh, R. MOLCAS 7: The Next Generation. *J. Comp. Chem.* **2010**, *31*, 224.
22. Abouaf, R.; Pommier, J.; Dunet, H. *Chem. Phys. Lett.* **2003**, *381*, 486.
23. Abouaf, R.; Pommier, J.; Dunet, H.; Quan, P.; Nam, P.- C.; Nguyen, M. T. *J. Chem. Phys.* **2004**, *121*, 11668.
24. Wood, P. D.; Redmond, R. W. *J. Am. Chem. Soc.* **1996**, *118*, 4256.
25. Marguet, S.; Markowitsi, D. *J. Am. Chem. Soc.* **2005**, *127*, 5780.
26. Hare, P. M.; Crespo-Hernández, C. E.; Kohler, B. *J. Phys. Chem. B* **2006**, *110*, 18641.

Glosario de siglas y abreviaturas

A

A	<i>Adenine</i>	Adenina
Abs	<i>Absorption</i>	Absorción
ADN	<i>Deoxyribonucleic Acid</i>	Ácido Desoxirribonucleico
ANO	<i>Atomic Natural Orbital</i>	Orbital Natural Atómico
AO	<i>Atomic Orbital</i>	Orbital Atómico
ARN	<i>Ribonucleic Acid</i>	Ácido Ribonucleico
ATP	<i>Adenosine 5'-Triphosphate</i>	Adenosina 5'-Trifosfato
au	<i>Atomic Unit</i>	Unidad Atómica

B

BSSE	<i>Basis Set Superposition Error</i>	Error por Superposición del Conjunto de Base
BV	<i>Biliverdin</i>	Biliverdina

C

C	<i>Cytosine</i>	Citosina
CASSCF	<i>Complete Active Space Self-Consistent Field</i>	Espacio Activo Completo Campo Autoconsistente
CASSI	<i>Complete Active Space State Interaction</i>	Espacio Activo Completo Interacción de Estados
CASPT2	<i>Complete Active Space Perturbation Theory to Second Order</i>	Espacio Activo Completo Teoría de Perturbaciones hasta Segundo Orden
CBC	<i>Cyclobutane Cytosine</i>	Ciclobutilcitosina
CBPyr (CBP)	<i>Cyclobutane Pyrimidine</i>	Ciclobutilpirimida
CBT	<i>Cyclobutane Thymine</i>	Ciclobutiltimina
CBU	<i>Cyclobutane Uracil</i>	Ciclobutiluracilo
CC	<i>Coupled Cluster</i>	Coupled Cluster
CI	<i>Configuration Interaction/Conical Intersection</i>	Interacción de Configuraciones/ Intersección Cónica

Glosario de siglas y abreviaturas

CPD	<i>Cyclobutane Pyrimidine Dimer</i>	Dímero de Ciclobutil Pirimida
CTD	<i>Cyclobutane Thymine Dimer</i>	Ciclobutilpirimidina de Timina
Cry	<i>Cryptochrome</i>	Criptocromo
D		
D	<i>Debye</i>	Debye
DFT	<i>Density Functional Theory</i>	Teoría del Funcional de la
DNA	<i>Deoxyribonucleic Acid</i>	Ácido Desoxirribonucleico
E		
E	<i>Energy</i>	Energía
Eb	<i>Energy binding</i>	Energía de enlace
ET	<i>Energy Transfer</i>	Transferencia de Energía
E _{VA}	<i>Vertical Absorption Energy</i>	Energía de Absorción
E _{VE}	<i>Vertical Emission Energy</i>	Energía de Emisión
eV	<i>electronvolt</i>	electronvoltio
F		
F	<i>Fluoresecence</i>	Fluorescencia
f	<i>Oscillator strength</i>	Fuerza de oscilador
FCI	<i>Full Configuration Interaction</i>	Interacción de Configuraciones Completa
FMN	<i>Flavin Mononucleotide</i>	Flavin Mononucleótido
fs	<i>femtosecond</i>	femtosegundo
G		
G	<i>Guanine</i>	Guanina
H		
HF	<i>Hartree-Fock</i>	Hartree-Fock
HM	<i>HOMO-LUMO</i>	HOMO-LUMO
HOMO	<i>Highest Occupied Molecular Orbital</i>	Orbital Molecular Ocupado Más Alto
I		
IC	<i>Internal Conversion</i>	Conversión Interna
IMAG	<i>Imaginary</i>	Imaginario
IPEA	<i>Ionization Potential Electron Affinity</i>	Potencial de Ionización Afinidad Electrónica
IRC	<i>Intrinsic Reactive Coordinate</i>	Coordenada Intrínseca de Reacción
ISC	<i>Intersystem Crossing</i>	Cruce Entre Sistemas
IUPAC	<i>International Union of Pure and Applied Chemistry</i>	Unión Internacional de Química Pura y Aplicada

Glosario de siglas y abreviaturas

IVR	<i>Applied Chemistry Intramolecular Vibrational Relaxation</i>	Química Pura y Aplicada Relajación Vibracional Intramolecular
J		
K		
Kcal	<i>Kilocalorie</i>	Kilocaloría
L		
LE	<i>Locally Excited</i>	Excitado Localmente
LIIC	<i>Linear Interpolation in Internal Coordinates</i>	Interpolación Lineal de Coordenadas Internas
LOV	<i>Light, Oxygen and Voltage</i>	Luz, Oxígeno y Voltaje
LS	<i>Level-Shift</i>	Desplazamiento de Nivel
LUMO	<i>Lowest Unoccupied Molecular Orbital</i>	Orbital Molecular más Bajo Desocupado
M		
MBPT	<i>Many-Body Perturbation Theory</i>	Teoría de Perturbaciones Multipartícula
MCSCF	<i>Multi-Configurational Self- Consistent Field</i>	Multiconfiguracional de Campo Autoconsistente
MEP	<i>Minimum Energy Path</i>	Camino de Mínima Energía
MM	<i>Molecular Mechanics</i>	Mecánica Molecular
MO	<i>Molecular Orbital</i>	Orbital Molecular
MP	<i>Møller-Plesset</i>	Møller-Plesset
MP2	<i>Møller-Plesset to Second order</i>	Møller-Plesset hasta Segundo orden
MRCI	<i>Multi-Reference Configuration Interaction</i>	Interacción de Configuraciones Multirreferencial
N		
O		
P		
P	<i>Phosphorescence</i>	Fosforescencia
PCB	<i>Phycocyanobilin</i>	Ficocianobilina
PEC	<i>Potential Energy Curve</i>	Curva de Energía Potencial
PEH	<i>Potential Energy Hypersurface</i>	Hipersuperficie de Energía Potencial

Glosario de siglas y abreviaturas

PES	<i>Potential Energy Surface</i>	Superficie de Energía Potencial
Phot	<i>Phototropin</i>	Fototropina
Phy	<i>Phycocrome</i>	Fitocromo
Pyr	<i>Pyrimidine</i>	Pirimidina
PΦB	<i>Phytochromobilin</i>	Fitocromobilina
Q		
QCEXVAL	<i>Quantum Chemistry of the Excited State Universitat de València</i>	Química Cuántica del Estado Excitado Universitat de València
QM	<i>Quantum Methods</i>	Métodos Cuánticos
R		
RNA	<i>Ribonucleic Acid</i>	Ácido Ribonucleico
S		
SCF	<i>Self-Consistent Field</i>	Campo Autoconsistente
SOC	<i>Spin-Orbit Coupling</i>	Acoplamiento Espín-Órbita
STC	<i>Singlet-Triplet Crossing</i>	Cruce Singlete Triplete
SWI	<i>Step-Wise Intermediate</i>	Intermediario
T		
T	<i>Thymine / Triple / Twist</i>	Timina / Triples / Torsión
TD-DFT	<i>Time-Dependent Density Functional Theory</i>	Teoría del Funcional de la Densidad Dependiente del Tiempo
TDM	<i>Transition Dipolar Moment</i>	Momentos Dipolares de Transición
TICT	<i>Twisted Intramolecular Charge Transfer</i>	Estado Girado de Transferencia de Carga Intramolecular
TS	<i>Transition State</i>	Estado de Transición
TT-EET	<i>Triplet Triplet Excitation Energy Transfer</i>	Transferencia de Energía de Excitación Triplete-Triplete
U		
U	<i>Uracil</i>	Uracilo
UV	<i>Ultraviolet</i>	Ultravioleta
UV-Vis	<i>Ultraviolet-Visible</i>	Ultravioleta-Visible
V		
VEA	<i>Vertical Electron Affinity</i>	Afinidad Electrónica Vertical
Z		
ZPVE	<i>Zero-Point Vibrational Energy</i>	Energía Vibracional de Punto Cero

A la memoria de Luis

En septiembre de 2010 a la edad de 44 años fallecía nuestro querido Luis Serrano Andrés.

Fue una muerte totalmente inesperada que impactó en la vida del grupo, no sólo profesional sino también personalmente. Persona de carácter bonachón, dispuesto siempre a ayudar en todo, con su toque especial de inteligente ironía, le hacían el compañero perfecto.

Luis fue un gran profesor y un excelente orador, que tenía la habilidad de atraer el interés de la audiencia desde el primer minuto de su charla hasta el final. Solía aderezar sus exposiciones con numerosas anécdotas, datos históricos y experiencias personales que las hacían más interesantes. Viajero de curiosidad insaciable, Luis era culto hasta la exageración, de mente preclara y abierta, que conocía de todo, daba igual lo que uno se imagine, desde la última noticia científica hasta los más pequeños detalles de una dinastía del antiguo Egipto.

En la última década dirigió junto a Manuela Merchán el grupo QCEXVAL (*Quantum Chemistry of the Excited State of Valencia*) y su intensa dedicación fue crucial para que el grupo fuera conocido internacionalmente. Su trabajo de excelencia obtuvo reconocimiento nacional e internacional con los contratos, becas y premios recibidos. Su curriculum con más de 125 artículos es impresionante. En el último año había sido habilitado para ser catedrático de Química Física, posición a la que, desafortunadamente, ya no pudo acceder. Varias de sus contribuciones están siendo un clásico en el campo de la Química Cuántica con un índice de impacto notablemente alto. Sus artículos han sido citados más de 3700 veces. Como bien dijo Manuela,¹

¹ In Remembrance of Dr. Luis Serrano-Andrés by Manuela Merchán In Valencia, October 27th, 2010.

todos los detalles se pueden encontrar en las bases de datos, pero lo que no se puede encontrar es su extraordinario esfuerzo por llevarlo todo adelante. Luis dejó muchos colaboradores y amigos que le echan de menos, y numerosos fueron los correos recibidos de condolencia, entre ellos me gustaría destacar el escrito por el profesor Ignacio Nebot Gil:

Ya sé que es una cita fácil, pero nada más conocer la terrible noticia se me ha venido a la memoria la dedicatoria y los versos finales de la “Elegía” de Miguel Hernández:

“En Orihuela, su pueblo y el mío,
se me ha muerto como del rayo Ramón Sijé,
con quien tanto quería”

“ ...

A las aladas almas de las rosas
del almendro de nata te requiero,
que tenemos que hablar de muchas cosas,
compañero del alma, compañero.”

Requiescat in pace.

Publicaciones

Publicación 1..... 205

Theoretical Analysis of the Excited States in Maleimide

Climent, T.; González-Luque, R.; Merchán, M.
J. Phys. Chem. A **2003**, *07*, 6995.

Publicación 2..... 217

Unified Model for the Ultrafast Decay of Pyrimidine Nucleobases

Merchán, M.; González-Luque, R.; Climent, T.; Serrano-Andrés, L.;
Rodríguez, E.; Reguero, M.; Pelaez, D.
J. Phys. Chem. B **2006**, *110*, 26471.

Publicación 3..... 225

Theoretical Insight into the Spectroscopy and Photochemistry of Isoalloxazine, the Flavin Core Ring

Climent, T.; González-Luque, R.; Merchán, M.; Serrano-Andrés, L.
J. Phys. Chem. **2006**, *110*, 13584.

Publicación 4..... 235

On the intrinsic population of the lowest triplet state of uracil

Climent, T.; González-Luque, R.; Merchán, M.; Serrano-Andrés, L.
Chem. Phys. Lett. **2007**, *441*, 327.

Publicación 5..... 243**Deciphering Intrinsic Deactivation/Isomerization Routes in a Phytochrome Chromophore Model**

Altoe, P.; Climent T.; De Fusco, G.; Stenta, M.; Bottoni, A.; Serrano Andrés, L.; Merchán, M.; Orlandi, G.; Garavelli, M.
J. Phys. Chem. B **2009**, *113*, 15067.

Publicación 6..... 253**The role of pyrimidine nucleobase excimers in DNA photophysics and photoreactivity**

González-Ramírez, I.; Climent, T.; Serrano-Pérez, J.; González-Luque, R.; Merchán, M.; Serrano-Andrés, L.
Pure Appl. Chem. **2009**, *81*, 1695.

Publicación 7..... 267**Singlet-Triplet States Interaction Regions in DNA/RNA Nucleobase Hypersurfaces**

González-Luque, R.; Climent, T.; González-Ramírez, I.; Merchán, M.; Serrano-Andrés, L.
J. Chem. Theor. Comp. **2010**, *6*, 2103.

Publicación 8..... 281**Cyclobutane Pyrimidine Photodimerization of DNA/RNA Nucleobases in the Triplet State**

Climent, T.; González-Ramírez, I.; González-Luque, R.; Merchán, M.; Serrano-Andrés, L.
J. Phys. Chem. Lett. **2010**, *1*, 2072.

Publicación 9..... 289**On the photoproduction of DNA/RNA cyclobutane pyrimidine dimers**

González-Ramírez, I.; Roca-Sanjuán, D.; Climent, T.; Serrano-Pérez, J.J.; Merchán, M.; Serrano-Andrés, L.
Theoretical Chemistry Accounts. **2010**, *128*, 705.

Publicación 1

Theoretical Analysis of the Excited States in Maleimide

Climent, T.; González-Luque, R.; Merchán, M.

J. Phys. Chem. A **2003**, *107*, 6995.

Theoretical Analysis of the Excited States in Maleimide

Teresa Climent, Remedios González-Luque, and Manuela Merchán*

Departamento de Química Física, Instituto de Ciencia Molecular, Universitat de Valencia, Dr. Moliner 50, Burjassot, ES-46100 Valencia, Spain

Received: December 9, 2002; In Final Form: May 20, 2003

The electronic excited states of maleimide have been studied using multiconfigurational second-order perturbation theory in its multistate formulation (MS-CASPT2) and extended atomic natural orbital (ANO) basis sets. The calculation of the singlet–singlet and singlet–triplet transition energies comprises 10 valence singlet excited states, four valence triplet states, and the singlet 3s, 3p, and 3d members of the Rydberg series converging to the first four ionization limits. The main features of the absorption spectrum correspond to the $1^1A_1 \rightarrow 1^1A_2(n\pi^*)$, $1^1A_1 \rightarrow 1^1B_2(\pi\pi^*)$, and $1^1A_1 \rightarrow 2^1B_2(\pi\pi^*)$ transitions computed at 3.29, 4.44, and 5.59 eV, respectively. The latter corresponds to the most intense band. The lowest Rydberg state is found to be at 5.98 eV in the high-energy side of the main band. A number of additional features are predicted in the higher-energy region of the spectrum, and our assignments follow the same trends as those offered earlier by Robin for *N*-methyl maleimide. The lowest triplet state is computed of $n\pi^*$ character in vacuo. However, arguments are given to rationalize the observed phosphorescence of maleimide in solution as being mainly of $\pi\pi^*$ nature.

1. Introduction

Cyclic imides are of fundamental importance in the realms of photochemistry and biochemistry, as well as commonly used reagents in organic and polymer chemistry. From a rapid literature review of the past few years, one can easily conclude that maleimides, apart from their classical function as photo-initiators,¹ are frequently used for a variety of purposes. Among them, a few selected applications are antitumor agents research,² study of membrane protein structure and function,³ and development of new diagnostic devices and photochromic molecules.⁴ It is also worth recalling the polyimides, polymers of recognized interest to the electronic industry because of their advantageous thermal and mechanical properties and low dielectric constant.⁵ All of these applications rely on the well-known behavior of maleimides as efficient cross-linker units. Nature also exploits this ability to design chromoproteins. For instance, open-chain tetrapyrroles serve as chromophores in light sensor proteins of plants, phytochromes (P). The conversion of the physiological P-inactive form into the P-active form is photochromic and behaves as a light-driven biological switch.⁶ Research on the photocycle of phytochromes is currently under way in our group accompanying recent work on chromophores related to rhodopsins⁷ and photoactive yellow protein (PYP).⁸ In the three cases, the key process involves a *cis* \rightarrow *trans* and *trans* \rightarrow *cis* double-bond photoisomerization. The view of maleimides as a building block of linear tetrapyrroles has actually been our main motivation to undertake the present study.

Despite the relevance of maleimides, to the best of our knowledge, there are no *ab initio* results published on the excited states of the simplest member of the family, maleimide (1*H*-pyrrole-2,5-dione). The electronic spectra of maleimides are, however, known since long ago.^{9–13} In the energy region up to 6.5 eV, the recorded absorption spectrum can be described as composed of three bands in increasing order of intensity and transition energy. That the lowest-energy transition in the singlet manifold has $n\pi^*$ character with a very weak intensity was

established in 1971 by Seliskar and McGlynn.⁹ In the absorption spectrum of maleimide in EPA mixed solvent at 77 K, the feature at 3.33 eV was considered by these authors as the probable 0–0 band of this transition.⁹ The second- and third-energy bands are assumed to involve excited states of $\pi\pi^*$ nature, which is confirmed in the present work. Among the available experimental data, we have chosen to compare to the maleimide data in solution using water as solvent (band maxima at 4.46 and 5.72 eV)¹³ and the *N*-methyl maleimide vapor-phase data (transitions at 4.34 and 5.53 eV).^{9,12} In the absorption spectrum of maleimide in water, the feature at 3.60 eV can be related to the $n\pi^*$ transition, which really looks like a shoulder of the second-energy band.¹³ Ideally, our results for maleimide computed in vacuo should be compared with the corresponding observed spectrum in the vapor (not recorded so far). Therefore, we have to bear in mind two effects in the current comparison, as appropriate: solvent effects in a polar medium such as water or the methyl substitution or both. Nevertheless, taking into account the experimental information, the substitution effect could be regarded of minor importance. Thus, the theoretical findings will be preferably contrasted to the vapor data for *N*-methyl maleimide. It especially holds true for the Rydberg excited states because of their collapse in solution. A number of Rydberg transitions and additional valence excited states have been identified in the vapor (*N*-methyl maleimide).^{9,12} Information about the triplet states comes from the phosphorescence spectra of maleimides in EPA solvent at 77 K.⁹ It was concluded that the $T_1 \rightarrow S_0$ transition in maleimide and its *N*-alkyl derivatives involves a triplet state of $n\pi^*$ nature. As shall be discussed below, the situation is more complex, and it is likely that the observed phosphorescence in solution comes mainly from a triplet state of $\pi\pi^*$ character, although the lowest two triplet states computed in vacuo are both of $n\pi^*$ type.

During the past decade, we have devoted considerable effort to characterize theoretically the excited states of a large number of fundamental organic compounds.^{14–16} Maleimide is a basic

heterocyclic polyene of intrinsic importance.¹² Previous works performed on related molecules such as pyrrole¹⁷ and imidazole¹⁸ give insight on the complexity of the problem. In addition, the theoretical spectroscopy of the carbonyl group addressed in different studies (formaldehyde,¹⁹ acetone,²⁰ *p*-benzoquinone²¹) points out the inherent difficulties implied in a dione. We report here a comprehensive theoretical research study of the excited states of maleimide, both of valence and Rydberg nature. We shall find that the previous descriptions are largely vindicated, but the situation is shown to be even more complicated than previously suspected. The excited states of the system have been characterized by using multiconfigurational second-order perturbation theory through the CASPT2 method.^{22,23} In addition, the indirect interaction of the resulting states has been taken into account within the framework of the multistate CASPT2 (MS-CASPT2) method.²⁴ The successful performance of the CASPT2 approach to the interpretation of the electronic spectra of a large number of organic compounds is well established.^{14–16} The study performed on maleimide includes determination of vertical excitation energies, band origin, and emission maximum for the electronic transitions involving the low-lying valence excited states of the singlet and triplet manifold. We believe that the present results give a more complete picture for a better understanding, on theoretical grounds, of the spectroscopic behavior of maleimides and provide confident assignments. They serve both to complement earlier studies and to bring a firm theoretical foundation to elucidate the photochemistry of maleimide.

The paper is organized as follows. Computational details are described in the next section. After that, the equilibrium geometry of the ground and low-lying excited states are considered. Results and analysis on the computed singlet \rightarrow singlet and singlet \rightarrow triplet absorption spectra and 0–0 transition are subsequently considered, together with comparison to previous findings. Our conclusions are summarized in the last section.

2. Theoretical Methods and Computational Details

Generally contracted basis sets of atomic natural orbitals type (ANO) obtained from the C,N,O(14s9p4d)/H(8s4p) primitive sets with the C,N,O[4s3p1d]/H[2s1p] contraction scheme²⁵ were used for the geometry optimization of the ground and low-lying valence excited states at the CASSCF level. For the geometry determination of the ground and low-lying excited states of $\pi\pi^*$ nature, the active space comprised the eight π -valence electrons distributed among the seven π -valence molecular orbitals (MOs). The active space was enlarged with the corresponding lone pair MO, CASSCF(8 MOs/10 electrons), for the description of the respective $n\pi^*$ states.

In the computation of the excitation energies, the basis set was supplemented with a 1s1p1d set of Rydberg functions (contracted from a set of 8s8p8d primitives). They were built following the protocol described elsewhere.¹⁴ Because of the relatively close spacing of the highest four occupied MOs (see Figure 1), the 3s-, 3p-, and 3d-Rydberg members converging to the lowest four ionization potentials are expected to overlap. In addition, the valence excited states would be interleaved among the Rydberg states. The low-lying Rydberg states of interest are described mainly by a one-electron promotion from the natural orbitals of the type 3b₁(π_1), 2b₁(π_2), 9b₂(lp₁), and 12a₁(lp₂) to the corresponding Rydberg orbital. Accordingly, the 1s1p1d Rydberg functions were placed at the averaged charge centroid of the ¹B₁, ²B₁, ¹B₂, and ¹A₁ states of the maleimide cation determined at the π -CASSCF level (plus lone

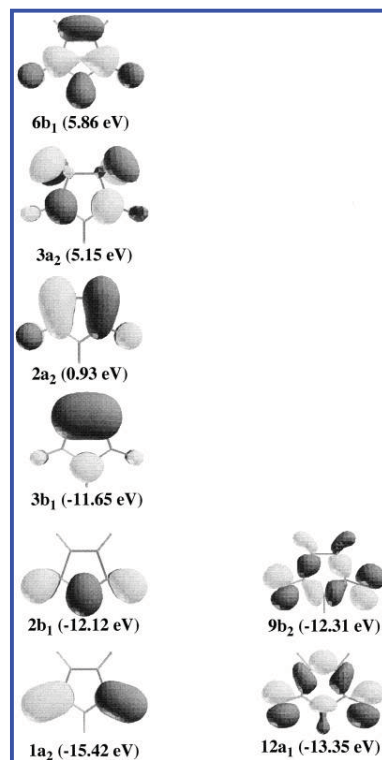


Figure 1. Highest five occupied and lowest three unoccupied canonical MOs and SCF orbital energies computed with the ANO-type C,N,O-[4s3p1d]/H[2s1p] basis set at the ground-state π -CASSCF equilibrium geometry.

pair lp₁ or lp₂). The final basis set consists of a total of 150 functions (559 primitives). The spectroscopic results reported here were computed including Rydberg functions.

Multiconfigurational wave functions were initially determined at the CASSCF level.²⁶ For the vertical excitation energies, the carbon, nitrogen, and oxygen 1s electrons were kept frozen in the form determined by the ground-state closed-shell Hartree–Fock wave function. The total SCF energy for the ground state (¹A₁) of maleimide was calculated to be $-357.534\,785$ au. The eight π electrons were always chosen to be active. The π -valence active space is denoted by (0;4;0;3), where the entries indicate the number of active orbitals belonging to the irreducible representations a₁, b₁, b₂, and a₂, respectively, of the point group C_{2v}. The active space was extended to include the Rydberg orbitals of symmetry a₁(3s, 3p_z, 3d_{z²}, 3d_{x²-y²}), b₁(3p_x, 3d_{xz}), b₂(3p_y, 3d_{yz}), and a₂(3d_{xy}), as appropriate. The active spaces used and the states computed are listed in Table 1.

Later, the relative vertical energy of each excited state is referenced to the ground-state energy computed with the same active space. The MOs for the excited states were obtained from average CASSCF computations, where the averaging includes all states of interest within a given symmetry. The number of states included in the state average CASSCF calculations is compiled in Table 1, together with the number of configurations involved in the CASSCF wave function.

TABLE 1: CASSCF Wave Functions Employed To Compute the Excited States of Maleimide with the ANO-type Basis Set C₄N₂O[4s3p1d]/H[2s1p] + 1s1p1d Rydberg Functions

wave function ^a	(N _e) ^b	states	N _{configurations} ^c	N _{states} ^d
Singlet Excited States Involving Ip ₁ (Inactive: 8;0;5;0)				
CASSCF (4;4;1;3)	(10)	¹ B ₂ (Ip ₁ →3s, 3p _z , 3d _z ² , 3d _{x²-y²})	42336	7
CASSCF (0;6;1;3)	(10)	¹ A ₂ (Ip ₁ →π*, 3p _z , 3d _z ²); ¹ A ₂ (Ip ₁ , π→π*π*)	4424	4
CASSCF (0;4;3;3)	(10)	¹ A ₁ (Ip ₁ →3p _y , 3d _{yz})	4984	6
CASSCF (0;4;1;4)	(10)	¹ B ₁ (Ip ₁ →π*, 3d _{xy})	1176	2
Singlet Excited States Involving Ip ₂ (Inactive: 7;0;6;0)				
CASSCF (5;4;0;3)	(10)	¹ A ₁ (Ip ₂ →3s, 3p _z , 3d _z ² , 3d _{x²-y²})	43030	7
CASSCF (1;6;0;3)	(10)	¹ B ₁ (Ip ₂ →π*, 3p _z , 3d _z ²)	4424	3
CASSCF (1;4;2;3)	(10)	¹ B ₂ (Ip ₂ →3p _y , 3d _{yz})	4840	5
CASSCF (1;4;0;4)	(10)	¹ A ₂ (Ip ₂ →π*, 3d _{xy})	1176	2
Singlet Excited States of ππ* Nature (Inactive: 8;0;6;0)				
CASSCF (4;4;0;3)	(8)	¹ B ₁ (π ₁ →3s, 3p _z , 3d _z ² , 3d _{x²-y²})	8176	8
		¹ B ₁ (π ₂ →3s, 3p _z , 3d _z ² , 3d _{x²-y²})		
CASSCF (0;6;0;3)	(8)	¹ A ₁ (ππ*); ¹ A ₁ (π ₁ →3p _z , 3d _z ²)	2688	7
		¹ A ₁ (ππ*); ¹ A ₁ (π ₂ →3p _z , 3d _z ²)		
CASSCF (0;4;2;3)	(8)	¹ A ₂ (π ₁ →3p _y , 3d _{yz}); ¹ A ₂ (π ₂ →3p _y , 3d _{yz})	1280	4
CASSCF (0;4;0;4)	(8)	¹ B ₂ (ππ*); ¹ B ₂ (ππ*); ¹ B ₂ (ππ*)		
		¹ B ₂ (π ₁ →3d _{xy} , π ₂ →3d _{xy})	864	5

^a Within parentheses are the number of orbitals of the symmetries a₁, b₁, b₂, and a₂, respectively, of the point group C_{2v}. Number of frozen orbitals = (4;0;3;0). Number of occupied SCF MOs = (12;3;9;1). ^b Number of active electrons. ^c Number of configurations in the CASSCF wave function. ^d States included in the average CASSCF wave function.

To take into account the remaining correlation effects, the CASSCF wave functions were employed as reference functions in a single-state second-order perturbation CASPT2 treatment.^{22,23} In addition, the effect of weakly interacting intruder states was minimized by using the so-called imaginary level-shift technique.²⁷ After careful calibration calculations, the shift of 0.2 au was selected to compile the final results. Even in cases where it was really not needed, the computation has been carried out at the same level shift because the energy results are not altered (at most 0.05 eV) and it renders consistency to the different comparisons performed (excitation energies, ionization potentials, singlet–triplet splitting, and so on). The coupling of the CASSCF wave functions via dynamic correlation was dealt with by using the extended multistate CASPT2 approach, the MS-CASPT2 method.²⁴ An effective Hamiltonian matrix is constructed where the diagonal elements correspond to the CASPT2 energies and the off-diagonal elements introduce the coupling to second order in the dynamic correlation energy. In this way, all states of a given symmetry can be treated simultaneously with the correlation effects on the reference functions included, and the possible erratic valence–Rydberg mixing can be removed yielding an effective separation of the computed states, which can be clearly identified as valence and Rydberg excited states. For a case of strong valence–Rydberg mixing and its successful solution at the MS-CASPT2 level, the reader is referred to the recent study on the anti conformer of *n*-tetrasilane.²⁸

The CASSCF state interaction (CASSI) method^{29,30} was used to calculate the transition dipole moments from the perturbation modified CAS (PMCAS) reference functions (the model states),²⁴ i.e., linear combinations of all CAS states involved in the MS-CASPT2 calculation. In the formula for the oscillator strength, the PMCAS-CI transition moment and the energy difference obtained in the MS-CASPT2 computation were used.

All calculations were performed with the MOLCAS-5 quantum-chemistry software.³¹

3. Results and Discussion

As a preliminary step toward the theoretical understanding of the electronic spectrum of maleimide, the geometry of the ground state and low-lying singlet and triplet excited states are first considered. The properties of the excited states computed

vertically are next analyzed. After that, the nonvertical transition energies are discussed. Furthermore, the present findings will be compared with the available experimental data and previous theoretical descriptions obtained at the semiempirical level.

3.1. Equilibrium Structures. Geometry optimizations for the ground state and low-lying excited states of both ππ* and nπ* character were carried out at the π-CASSCF level (ππ*) and at the same level enlarging the active space with the lone pair of the oxygen atoms (nπ*). The optimized geometric parameters for the ground state and singlet and triplet excited states are depicted in Figure 2. Optimizations were constrained to C_{2v} symmetry with the molecule placed in the yz plane, z being the C₂ axis passing along the N–H bond.

The bond lengths computed for the ground state are similar to the gas-phase molecular structure determined by electron diffraction (ED)³² (*r*(C=C) = 1.344 Å, *r*(C–C) = 1.508 Å, *r*(N–C) = 1.409 Å, *r*(C=O) = 1.206 Å), consistent with a C_{2v} symmetry model. In the ED study, the N–H and C–H bond lengths were assumed to be 1.025 and 1.096 Å, respectively. Both are somewhat too long compared to the π-CASSCF results and other theoretical determinations.^{32,33} Therefore, the fact that the larger deviation between the computed geometric parameters and the experimental data occurs in the N–C bond distance (1.383 vs 1.409 Å) could be related to the assumptions made in the ED research with a N–H distance fixed at a too long value. Similar comments are appropriate to rationalize the deviation on the computed CCH angle of 122.3° (cf. Figure 2) with respect to the experimentally derived datum (114.7°).³² Even when the known sensitivity of the carbonyl bond distance upon the level of theory is taken into consideration,^{19,20} the π-CASSCF trends are consistent with previous structure determination of ground-state maleimide performed at the ab initio and semiempirical levels.³² Therefore, it might be worth redoing the ED analysis employing the theoretical results instead of the above-mentioned assumed values.

From the nature of the canonical MOs determined of the ground state (see Figure 1), four low-lying singlet and triplet excited states can be predicted, as actual computation confirms. The minus and plus linear combinations of the in-plane lone-pair of the oxygen atoms lead to the 9b₂(Ip₁) and 12a₁(Ip₂) MOs. The one-electron promotion from Ip₁ to the lowest unoccupied

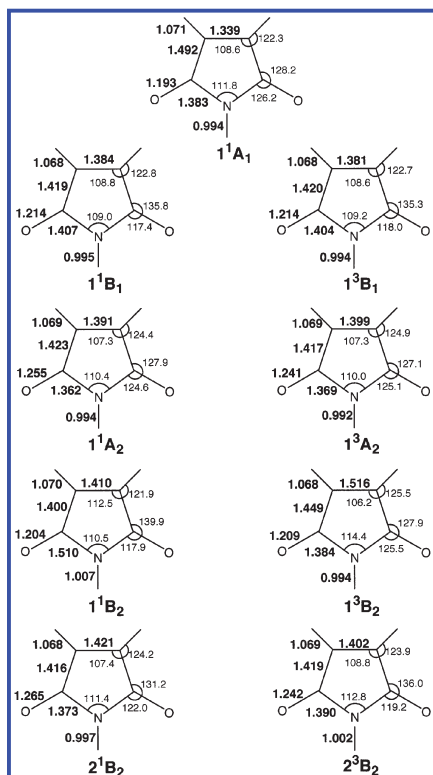


Figure 2. Geometries optimized at the π -CASSCF level (1^1A_1 , 1^3B_2 , and 2^1B_2 states) and with the active space enlarged with the lone pairs MOs of the oxygen atoms (1^1B_1 and 1^3A_2) as appropriate. Bond lengths given in Å and angles in deg.

molecular orbital (LUMO), $2a_2$, gives rise to the 1^1B_1 and 1^3B_1 states with the two electrons singlet- or triplet-coupled, respectively. Similarly, the $1^1A_2(lp_2 \rightarrow \pi^*)$ and $1^3A_2(lp_2 \rightarrow \pi^*)$ states can also be expected, where π^* represents a LUMO-like molecular orbital. There are also two singlet and two triplet states of B_2 symmetry as plausible candidates to appear in the low-energy region. They involve charge transfer from the highest occupied molecular orbital (HOMO) and $HOMO - 1$ into LUMO. Because there is electron reorganization within the π system in these states with respect to the ground state, they shall be labeled $\pi\pi^*$. The orbital energies between the HOMO ($3b_1$) and the $HOMO - 1$ ($2b_1$) differ in only 0.47 eV. Therefore, it is not surprising that they can be mixed in the CASSCF procedure, in accordance with the variational principle, to yield natural orbitals rotated with respect to the canonical MOs. Indeed, analysis of the homologous CASSCF natural orbitals (denoted by prime hereafter) reveals that they are mainly localized in the nitrogen atom ($2b_1'$) and the other in the C=C region ($3b_1'$). In terms of the natural orbitals, the two lowest $\pi\pi^*$ states are described by the corresponding singly excited configurations from a b_1' -type to the LUMO-like natural orbital. In all the cases, the LUMO-like natural orbital maintains its topology as the canonical LUMO ($2a_2$) shown in Figure 1. Analysis of the obtained excited-state geometric trends in terms

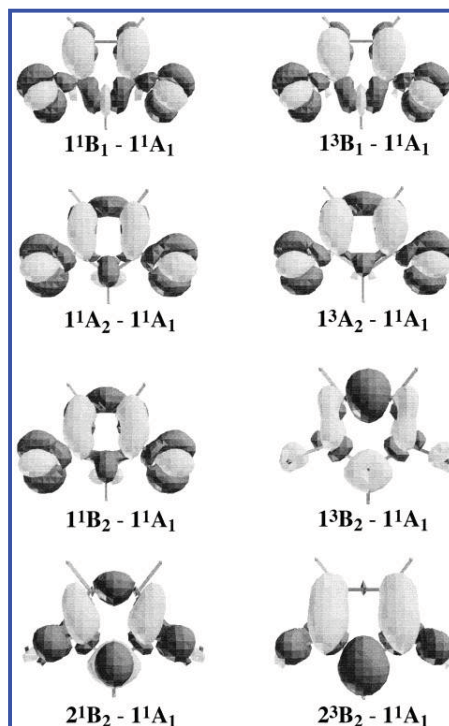


Figure 3. Density difference with respect to the ground state for the eight electronic excited states in their respective equilibrium geometry.

of the configurations involved is, therefore, a laborious task because of the different sets of natural orbitals that have to be detailed simultaneously. We have preferred to rationalize the computed molecular structures in terms of the electron density difference (EDD) between the ground and the excited state at the equilibrium geometry of the excited state. Thus, to analyze in more detail the relationship between the electronic nature and equilibrium geometry, the computed EDDs for the eight low-lying excited states of valence character are shown in Figure 3.

The darkest surfaces (plotted at an electronic density of 0.004 electrons/ au^3) represent regions where the density difference is negative—"where the electron comes from". By a rapid inspection of Figure 3, one can easily notice the involvement and the important role of the LUMO-like orbital in the description of the eight excited states. Accordingly, the increase of electron density in the two C—C regions of the excited state is paralleled by a decrease of the two equivalent C—C bond distances. To a lesser or greater extent, it holds true for all of the cases considered. For the $n\pi^*$ transitions the electron density is moved from the lone-pair MOs ($1p_1$ or $1p_2$), causing an elongation of the carbonyl bond length. A pronounced increase of the N—C and C=C bond distances occurs for the $1^1B_2(\pi\pi^*)$ and $1^3B_2(\pi\pi^*)$ states, respectively, in accordance with the electron density shift from those respective regions. Conversely, by inspection of the EDDs related to the $2^1B_2(\pi\pi^*)$ and $2^3B_2(\pi\pi^*)$ electronic states, one can understand the lengthening of the C=O bond distance taking place in these excited states.

TABLE 2: Computed PMCAS-CI and MS-CASPT2 Excitation Energies (eV), Related Oscillator Strengths (f), and PMCAS-CI Dipole Moments (μ , D) for the Vertical Excited States of Maleimide, Employing the C,N,O[4s3p1d]/H[2s1p] + 1s1p1d (Rydberg Functions) ANO-type Basis Set and Available Experimental Data

state ^a	PMCAS-CI	MS-CASPT2	f	μ	exptl
ground state (1^1A_1)				-1.78	
		Singlet Excited States			
$1^1B_1(lp_1 \rightarrow \pi^*)$	6.99	2.48	0.0004	-0.16	
$1^1A_2(lp_2 \rightarrow \pi^*)$	7.97	3.29	<i>b</i>	-0.08	3.33(0-), ^c 3.60 ^d
$1^1B_2(\pi\pi^*)$	6.61	4.44	0.0221	+0.50	4.34, ^e 4.46 ^d
$2^1B_2(\pi\pi^*)$	8.67	5.59	0.4526	-0.92	5.53, ^e 5.72 ^d
$3^1B_2(lp_1 \rightarrow 3s)$	9.86	5.98	0.0318	-3.23	
$2^1A_2(lp_1 \rightarrow \pi^*)$	11.46	6.00	<i>b</i>	-1.80	
$3^1A_2(lp_1\pi \rightarrow \pi^*\pi^*)$	11.17	6.35	<i>b</i>	-0.40	
$4^1A_2(lp_1 \rightarrow 3p_x)$	10.40	6.75	<i>b</i>	-1.94	
$4^1B_2(lp_1 \rightarrow 3p_x)$	10.36	6.78	0.0000	+0.89	
$2^1A_1(lp_1 \rightarrow 3p_x)$	9.47	6.84	0.0356	+1.83	
$3^1A_1(lp_2 \rightarrow 3s)$	9.49	6.96	0.0727	-3.54	6.99, ^e 6.94 ^f
$4^1A_1(\pi\pi^*)$	8.31	7.13	0.0335	+1.18	6.99, ^e 6.94 ^f
$2^1B_1(lp_2 \rightarrow \pi^*)$	12.36	7.30	0.0000	-0.84	
$3^1B_1(\pi_1 \rightarrow 3s)$	7.56	7.35	0.0001	-3.90	
$5^1A_2(lp_1 \rightarrow 3d_{xz})$	10.98	7.38	<i>b</i>	+2.40	
$5^1B_2(lp_1 \rightarrow 3d_{z^2})$	10.86	7.41	0.0001	+3.33	
$4^1B_1(lp_1 \rightarrow 3d_{xy})$	10.44	7.62	0.0003	+0.41	
$5^1A_1(lp_2 \rightarrow 3p_x)$	10.20	7.67	0.0134	+2.01	7.81 ^f
$5^1B_1(lp_2 \rightarrow 3p_x)$	10.99	7.75	0.0005	-1.40	
$6^1B_2(lp_1 \rightarrow 3d_{x^2-y^2})$	11.15	7.77	0.0083	+0.94	
$7^1B_2(lp_2 \rightarrow 3p_x)$	11.92	7.79	0.0005	+1.79	
$6^1A_1(lp_1 \rightarrow 3d_{xz})$	10.11	7.80	0.0016	-1.26	
$6^1B_1(\pi_2 \rightarrow 3s)$	8.41	7.81	0.0259	+1.84	
$7^1B_1(\pi_1 \rightarrow 3p_x)$	8.41	8.20	0.0014	+2.88	
$7^1A_1(\pi_1 \rightarrow 3p_x)$	7.62	8.22	0.0099	-1.06	
$8^1A_1(lp_2 \rightarrow 3d_{z^2})$	10.71	8.28	0.0218	+2.61	
$8^1B_2(lp_2 \rightarrow 3d_{xz})$	12.39	8.42	0.0045	-2.49	
$6^1A_2(lp_2 \rightarrow 3d_{xy})$	11.47	8.50	<i>b</i>	+0.38	
$9^1A_1(\pi\pi^*)$	9.72	8.53	0.0160	+1.04	8.13, ^e 8.12 ^f
$7^1A_2(\pi_1 \rightarrow 3p_x)$	8.32	8.54	<i>b</i>	+3.94	
$8^1B_1(\pi_2 \rightarrow 3p_x)$	9.03	8.56	0.0022	-1.24	
$9^1B_1(lp_2 \rightarrow 3d_{xz})$	11.66	8.61	0.0011	+2.01	
$10^1A_1(lp_2 \rightarrow 3d_{x^2-y^2})$	10.89	8.66	0.0077	+1.03	8.68 ^f
$11^1A_1(\pi_2 \rightarrow 3p_x)$	8.54	8.74	0.0272	-0.51	
$10^1B_1(\pi_1 \rightarrow 3d_{z^2})$	8.93	8.77	0.0029	+5.85	
$9^1B_2(\pi\pi^*)$	10.42	8.77	0.0151	-0.23	
$8^1A_2(\pi_2 \rightarrow 3p_x)$	9.03	8.93	<i>b</i>	+3.04	
$12^1A_1(\pi_1 \rightarrow 3d_{xz})$	8.44	9.00	0.0001	+4.68	
$10^1B_2(\pi_1 \rightarrow 3d_{xy})$	9.37	9.00	0.0292	+1.98	
$9^1A_2(\pi_1 \rightarrow 3d_{xz})$	8.81	9.14	<i>b</i>	+0.71	
$11^1B_1(\pi_1 \rightarrow 3d_{x^2-y^2})$	9.20	9.23	0.0008	+2.14	
$12^1B_1(\pi_2 \rightarrow 3d_{z^2})$	9.73	9.26	0.0101	-8.69	
$11^1B_2(\pi_2 \rightarrow 3d_{xy})$	10.14	9.42	0.0000	-3.28	
$13^1A_1(\pi_2 \rightarrow 3d_{xz})$	9.32	9.62	0.0000	-6.69	
$10^1A_2(\pi_2 \rightarrow 3d_{yz})$	9.81	9.69	<i>b</i>	-11.66	
$13^1B_1(\pi_2 \rightarrow 3d_{x^2-y^2})$	10.08	9.73	0.0054	-2.67	
		Triplet Excited States			
$1^3B_1(lp_1 \rightarrow \pi^*)$	6.34	2.31		+0.08	
$1^3A_2(lp_2 \rightarrow \pi^*)$	7.34	3.14		+0.15	
$1^3B_2(\pi\pi^*)$	4.12	3.49		-2.52	
$2^3B_2(\pi\pi^*)$	5.21	3.84		+1.16	

^a The character and ordering of the states are those obtained with the MS-CASPT2 method. ^b Forbidden. ^c From the absorption spectrum of maleimide in EPA mixed solvent at 77 K.⁹ ^d From the absorption spectrum of maleimide in water.¹³ ^e From the vapor absorption spectrum of *N*-methyl maleimide.⁹ ^f From the vapor absorption spectrum of *N*-methyl maleimide.¹²

The optimized geometries were employed to characterize the vertical electronic spectrum (absorption and emission) and nonvertical electronic transitions. These topics will be treated in turn.

3.2. Vertical Spectra of Maleimide. The computed vertical excitation energies at the PMCAS-CI and MS-CASPT2 levels of theory, together with oscillator strengths of the corresponding transitions, dipole moments, and selected experimental data are listed in Table 2. The nature of the states, valence or Rydberg, as computed from the corresponding PMCAS-CI wave function, has been characterized on the basis of the dominant orbital

configurations and the expectation values $\langle x^2 \rangle$, $\langle y^2 \rangle$, and $\langle z^2 \rangle$, which are appreciably larger for Rydberg states than are those of the ground state and valence excited states. The resulting assignment of the excited state is given within parentheses in the first column of Table 2, identifying the different excited states of maleimide. The detailed interpretation of the results is discussed below.

Valence Singlet Excited States. At the highest level of theory employed, four valence singlet excited states are placed below the lowest Rydberg transition. They correspond to the excited states presented in the previous section and can be related to

the main features of the recorded electronic spectrum up to 6 eV. The two lowest singlet excited states are of $n\pi^*$ type, the $1^1B_1(lp_1 \rightarrow \pi^*)$ and $1^1A_2(lp_2 \rightarrow \pi^*)$ states, and are calculated to lie vertically 2.48 and 3.29 eV above the ground state, respectively. The intensities of the transitions are predicted to be exceedingly weak. In fact, the $1^1A_1 \rightarrow 1^1A_2$ electronic transition is electric-dipole-forbidden. The computed excitation energy for the 1^1A_2 state is consistent with the spectroscopic data of maleimide in EPA with a band origin at 3.33 eV.⁹ The band maximum moves up to 3.60 eV in the maleimide absorption spectrum in water.¹³ From the magnitude of the dipole moment of the valence excited states with respect to that of the ground state, the occurrence of a blue shift in polar solvents can be predicted qualitatively for most of the valence transitions considered. Therefore, considering the uncertainty in the experimental determination of the band-origin peak, the error bars of the computation (0.1–0.2 eV), and the solvent influence, the agreement between theory and experiment can be regarded at first sight as satisfactory. Nevertheless, the situation is actually more complex. Vibronic interactions with the higher-energy allowed $\pi\pi^*$ state, through the activity of out-of-plane vibrations, are proposed as the source of the observed intensity in the lowest-energy band of maleimide, similarly as it is documented to occur for the forbidden $n\pi^*$ transition in the simplest carbonyl compound, formaldehyde,¹⁹ and homologous (also forbidden) transitions in related systems such as *p*-benzoquinone.²¹ It is worth recalling that in the observed spectrum of maleimide,¹³ the $n\pi^*$ feature looks like a shoulder of the lowest-energy $\pi\pi^*$ band. The energy difference of the computed band origin for the $1^1A_1 \rightarrow 1^1A_2$ transition and the vertical excitation is 0.59 eV (see below). It is consistent with the expected long Franck–Condon progression because of the pronounced geometric changes taking place in the excited state with respect to the ground state (see Figure 2). In the recorded spectrum, the band maximum could easily be somewhat deviated from the band origin and the vertical transition. Moreover, the overall contribution of the lowest excited state (1^1B_1) to the observed feature is difficult to predict at present, although it is expected to play a minor role. Therefore, in case that a higher accuracy would be required to fully describe on theoretical grounds the observed low-intensity, low-energy band of maleimide, vibronic interactions would have to be taken into account and the two lowest excited states would have to be considered. The $n\pi^*$ transitions have been previously obtained at the semiempirical level, but the lowest excited state, 1^1B_1 , instead of the second excited state, 1^1A_2 , was assigned to be responsible for the observed band in view of the forbidden character of the electronic transition involving the latter.⁹

There has been controversy along the years concerning the character, valence or Rydberg, of the main bands of maleimide. From the MS-CASPT2 results, it is clear that both the second- and third-energy observed bands can be related each to a single electronic transition: $1^1A_1 \rightarrow 1^1B_2$ and $1^1A_1 \rightarrow 2^1B_2$, respectively. Both are valence $\pi\pi^*$ transitions, the former with the vertical transition computed at 4.44 eV and the latter at 5.59 eV. Comparatively, the lowest $\pi\pi^*$ transition is weaker than the third-energy band, predicted to be the most intense feature of the absorption spectrum of maleimide in the gas phase. Nevertheless, the lowest Rydberg excited state is predicted in the higher-energy side of the third-energy band. The theoretical findings are consistent with the high-pressure effect investigation reported by Robin.¹² Both transitions show vibrational structure in the vapor absorption spectrum of *N*-methyl maleimide with band maximum at 4.34 and 5.53 eV.⁹ In solution, only the third-

TABLE 3: Computed CASSCF and CASPT2 Vertical Ionization Potentials (IP) Employing the C,N,O[4s3p1d]/H[2s1p] + 1s1p1d (Rydberg functions) ANO-type Basis Set^a

IP	state	CASSCF	CASPT2
IP ₁	$1^2B_2(lp_1\text{-hole})$	11.33	9.41
IP ₂	$1^2A_1(lp_2\text{-hole})$	12.35	10.32
IP ₃	$1^2B_1(\pi_1\text{-hole})$	9.95	10.61
IP ₄	$2^2B_1(\pi_2\text{-hole})$	11.19	11.08
IP ₅	$1^2A_2(\pi\text{-hole})$	12.75	13.43
IP ₆	$3^2B_1(\pi\text{-hole})$	15.37	15.30

^a The character and ordering of the states are those obtained with the CASPT2 method. Results in eV.

energy band of maleimide is structured. The respective band maxima are slightly blue-shifted to 4.46 and 5.72 eV.¹³ Despite the large Stokes shifts predicted from the optimized geometries of the excited states involved, the relative concordance of the recorded band maximum with the corresponding excitation energy computed vertically seems to point out that the description of the lowest $\pi\pi^*$ electronic transitions within the Franck–Condon scheme is appropriate.

Higher valence excited states are predicted at 6.00, 6.35, 7.13, 7.30, 8.53, and 8.77 eV (MS-CASPT2 results). They are related to the excited states $2^1A_2(lp_1 \rightarrow \pi^*)$, $3^1A_2(lp_1\pi \rightarrow \pi^*\pi^*)$, $4^1A_1(\pi\pi^*)$, $2^1B_1(lp_2 \rightarrow \pi^*)$, $9^1A_1(\pi\pi^*)$, and $9^1B_2(\pi\pi^*)$, respectively. The 2^1A_2 state involves a one-electron promotion from the lp_1 MO to the valence orbital $3a_2$. The 3^1A_2 state is of mixed character, implying simultaneous electron promotion from the lp_1 and occupied π orbitals. The related transitions are electric-dipole-forbidden and are expected to play a minor role in the description of the spectrum. Transition from the ground to the 2^1B_1 state, described mainly by the singly excited configuration $lp_2 \rightarrow \pi^*$ (b_1 -type), is also predicted with a weak intensity. The remaining valence excited states, 4^1A_1 , 9^1A_1 , and 9^1B_2 , are of $\pi\pi^*$ character. Those two of the same symmetry as the ground state show complex multiconfigurational wave functions. The 4^1A_1 state has a large weight (31%) of the configuration involving the two-electron promotion (HOMO \rightarrow LUMO)². It is the expected doubly excited state. The placement of a valence transition at about 7 eV, overlapped with a transition $n \rightarrow 3s$ (cf. Table 2), is consistent with previous suggestion.¹² A feature observed at 8.12 eV of vertical nature was deduced by Robin to be of valence $\pi\pi^*$ character in the gas spectrum of *N*-methyl maleimide.¹² It can be tentatively related here to the $1^1A_1 \rightarrow 9^1A_1$ transition computed at 8.53 eV, although the $1^1A_1 \rightarrow 8^1A_1$ Rydberg transition at 8.28 eV has a larger oscillator strength and is energetically closer to the detected peak. The 9^1B_2 state is the highest valence state computed (8.77 eV) and can be described primarily by the one-electron promotion from the $2b_1'$ MO to the valence MO $3a_2$ (weight 75%).

Rydberg Singlet Excited States. The calculations comprise the 3s, 3p, and 3d members of the Rydberg series converging to the first four ionization potentials (IPs). Table 3 lists the vertical IPs computed at the CASSCF and CASPT2 levels. A comparison of the CASSCF and CASPT2 results demonstrates the importance of dynamic correlation effects for an accurate prediction of the IPs, which are highly differential for the different states of the maleimide cation. Indeed, the lowest IP at the CASSCF level is computed to have a π -hole in the cation whereas the actual ground state of the maleimide cation is the $1^2B_2(lp_1\text{-hole})$ state. The energy difference between the CASPT2 and the CASSCF results gives the effect of dynamic correlation on the obtained IPs, clearly larger for the n -hole states than for π -hole states. It points out the difficulties that approximate methods may encounter in the treatment of maleimide, a very

complex molecule to be described on theoretical grounds, although at first sight it may appear to be simple because of its relatively small molecular size.

The first (IP_1) and second (IP_2) vertical ionization potentials of maleimide lie within 0.9 eV of each other (CASPT2 results). As a consequence, corresponding members of the Rydberg series converging to these two ionization potentials are expected to lie within 1 eV of each other. Consistently, the energy difference between the $3^1A_1(lp_2 \rightarrow 3s)$ and $3^1B_2(lp_1 \rightarrow 3s)$ is 0.98 eV at the MS-CASPT2 level. The $3^1B_1(\pi_1 \rightarrow 3s)$ state appears at 7.35 eV, 0.39 eV above the $3^1A_1(lp_2 \rightarrow 3s)$ state. The energy difference can be related to the difference $IP_3 - IP_2$ (0.29 eV). The placement of the $6^1B_1(\pi_2 \rightarrow 3s)$ state at 7.81 eV, 0.46 eV above the $3^1B_1(\pi_1 \rightarrow 3s)$ state, also keeps the trend ($IP_4 - IP_3 = 0.47$ eV). Because 1.5 eV approximates the usual 3s and 3d Rydberg-state separation, Rydberg states leading to the first four ionized limits are expected to overlap in energy even for the early series members. Here we discuss the MS-CASPT2 energies and assignments for the states described by excitations out of the lp_1 -, lp_2 -, π_1 -, and π_2 -like MOs and compare them with the experimental assignments performed in *N*-methyl maleimide in the gas phase.¹²

The calculated vertical transition energy from the ground to the lowest 3s Rydberg state, $3^1B_2(lp_1 \rightarrow 3s)$, is 5.98 eV. The oscillator strength is around 0.03. It is predicted at the high-energy region of the most intense band. In the gas-phase absorption spectrum of maleimide, the 3s Rydberg feature would probably look like a sharp peak overlapping the vibrational structure of the strongest $\pi\pi^*$ band. This prediction is consistent with the findings reported by Robin concerning the optical spectrum of *N*-methyl maleimide vapor. Robin investigated the high-pressure effect on the 5.53-eV band and found that the 0–0 and a few higher vibronic bands remained sharp under perturbation but several others displayed the broadening to higher frequencies characteristic of Rydberg transitions. Consequently, it was suggested that certain members of the $\pi\pi^*$ vibronic band are strongly mixed with the $n \rightarrow 3s$ Rydberg excitation.¹² The assignment from the current results is clear: the Rydberg state involved in the mixing is precisely the $3^1B_2(lp_1 \rightarrow 3s)$ state.

The $4^1A_2(lp_1 \rightarrow 3p_x)$ state is calculated to lie 6.75 eV above the ground state. Transitions to the $4^1B_2(lp_1 \rightarrow 3p_x)$ and $2^1A_1(lp_1 \rightarrow 3p_y)$ states appear slightly above, at 6.78 and 6.84 eV, respectively. The former is predicted to be exceedingly weak, and the latter has a computed oscillator strength similar to that obtained for the $1^1A_1 \rightarrow 3^1B_2(lp_1 \rightarrow 3s)$ transition.

The energy of the first component of the Rydberg series to the second ionization potential is computed to be 6.96 eV. The $1^1A_1 \rightarrow 3^1A_1(lp_2 \rightarrow 3s)$ transition has a significant oscillator strength, 0.07. It is observed at 6.94–6.99 eV in *N*-methyl maleimide.^{9,12} In the same energy region, an overlapping valence shell transition (to the 4^1A_1 state) is predicted, in agreement with the conclusions achieved from the analysis performed by Robin.¹²

The members of the $lp_1 \rightarrow 3d$ series have been computed in the energy range 7.38–7.80 eV. The transitions are predicted with weak intensity. Within the same energy interval, the $1^1A_1 \rightarrow 5^1A_1(lp_2 \rightarrow 3p_z)$ transition at 7.67 eV is clearly allowed and can be related to the 3p assignment in *N*-methyl maleimide at 7.81 eV.¹² In addition, on the basis of expected term values, the peak at 8.68 eV in the methylated compound was assigned as a transition to 3d. This is consistent with the computed results for maleimide (cf. Table 2), in particular with the $1^1A_1 \rightarrow 10^1A_1(lp_2 \rightarrow 3d_{z^2-y^2})$ electronic transition computed at 8.66 eV.

Because the symmetries of the transition moments are not known experimentally and many closely spaced states are calculated in the higher spectral region, coincidence between experimental and computed transition energy may be accidental. In this region, therefore, the experimental (*N*-methyl maleimide) and theoretical results (maleimide) can be best combined to deduce the spectral assignments (*N*-methyl maleimide), but neither can be said to confirm the other. We just realize that the assignments performed by Robin in *N*-methyl maleimide and our results on maleimide follow the same pattern. Larger deviation in the correlation between experimental data and theoretical results is obtained for the valence transition to the 9^1A_1 state. It is unclear whether the discrepancy is because of the influence of the methyl group, uncertainties in the experimental–theoretical correlation, limitations in the resolution of the observed spectrum, or limitations in our own calculations. However, both molecules have a similar scheme: a valence feature in maleimide (*N*-methyl maleimide) 8.53 (8.12 eV) is between two Rydberg peaks located at 7.67 (7.81) and 8.66 (8.68) eV.

Many closely spaced states are predicted in the energy region 9–9.7 eV. The average of the excitation energies corresponding to $\pi_1 \rightarrow 3d$ and $\pi_2 \rightarrow 3d$ is 9.03 and 9.54 eV, respectively. Thus, they are about half of an electronvolt apart from each other, as is the separation between the fourth and third ionization potentials computed vertically (0.47 eV).

Low-Lying Triplet Excited States. The singlet–triplet spectrum computed at the ground-state geometry places the lowest triplet state $1^3B_1(lp_1 \rightarrow \pi^*)$ at 2.31 eV. The second triplet state is also of $n\pi^*$ character and lies 3.14 eV above the ground state. The ordering and nature of the four lowest valence singlet and triplet states is similar. The $\pi\pi^*$ triplet states, 1^3B_2 and 2^3B_2 , are found at 3.49 and 3.84 eV, respectively. Therefore, a considerably larger singlet–triplet splitting is computed for the $\pi\pi^*$ states than for the $n\pi^*$ states. It can be easily rationalized in terms of the exchange integral involved in the energy difference between a triplet and singlet state, both described by the same orbitals, in a simple molecular orbital model. Because the penetration of a $\sigma-\pi$ excitation is smaller than that of a $\pi-\pi^*$, the exchange integral is larger in the latter and consequently the singlet–triplet splitting is more pronounced.

One of the key questions of spectroscopy of maleimide is the source of the observed luminescence, mainly phosphorescence. A more detailed discussion on this topic shall be considered in the next section, involving the nonvertical features of the spectrum. We just mention here that from the dipole moments of the triplet states a slight blue shift for the $n\pi^*$ states and a red shift for the lowest $\pi\pi^*$ state can be inferred to occur in polar solvents.

3.3. Emission and Nonvertical Transition Energies of Maleimide. With the optimized geometry of the excited states, the corresponding emission and nonvertical features have also been characterized. Knowledge of the energy hypersurfaces of the ground and low-lying excited states of maleimides is essential for the determination of spectroscopic properties, which are particularly relevant in an ample group of biologically active chromophores. The basic understanding of the spectroscopic behavior of these compounds requires not only the accurate computation of the vertical states but also the characterization of the emission and nonvertical transition energies.

The computed transitions involve the energy difference between the ground and the excited state at the respective optimized geometry and the vertical emission from the excited-state minimum vertically to the ground state. The 0–0 absorption, also labeled T_0 , is the first component of each absorption

TABLE 4: Vertical, Adiabatic, and Emission Electronic Transitions Computed for Maleimide^a

state	vertical absorption		0-0 transition		emission maximum	
	PMCAS-CI	MS-CASPT2	PMCAS-CI	MS-CASPT2	PMCAS-CI	MS-CASPT2
Singlet-Singlet Spectra						
$1^1B_1(lp_1 \rightarrow \pi^*)$	6.99	2.48	6.29	2.38	5.73	1.94
$1^1A_2(lp_2 \rightarrow \pi^*)$	7.97	3.29	7.20	2.70	6.71	2.43
$1^1B_2(\pi\pi^*)$	6.61	4.44	5.03	3.65	3.72	2.62
$2^1B_2(\pi\pi^*)$	8.67	5.59	7.78	4.99	7.05	4.55
Singlet-Triplet Spectra						
$1^3B_1(lp_1 \rightarrow \pi^*)$	6.34	2.31	5.98	2.08	5.48	1.68
$1^3A_2(lp_2 \rightarrow \pi^*)$	7.34	3.14	6.91	2.48	6.47	2.21
$1^3B_2(\pi\pi^*)$	4.12	3.49	3.33	2.77	2.61	2.20
$3^1B_2(\pi\pi^*)$	5.21	3.84	4.57	3.07	3.92	2.65

^a Geometry has been optimized at the CASSCF level (see Figure 2). The ANO-type C,N,O[4s3p1d]/H[2s1p] enlarged with 1s1p1d diffuse functions has been used in the spectral study. Results in eV.

band and corresponds to a transition from and to the lowest vibrational state. The computed value is, however, T_e . The energy difference between T_0 and T_e is expected to be small in maleimide (within 0.1 eV). On the other hand, the vertical emission (fluorescence or phosphorescence) is obtained from experiment as the maximum of the emission band. The basis set includes diffuse functions, and the active spaces are those used in the geometry optimization and for the vertical transitions (π -valence plus lone pair MOs as appropriate). Then a consistent comparison can be performed between the computed absorption and emission spectra. The study includes determination of the adiabatic transition and emission maximum for the four lowest valence excited states in the singlet and triplet manifold. Because the geometry of the excited states has been optimized within the C_{2v} symmetry, the computed band origins can be considered upper bounds of the actual gas-phase data. On the other hand, considering that any deviation from planarity would certainly increase the energy of the ground state, the emission maxima might be somewhat smaller than those reported here. Table 4 lists the vertical, adiabatic, and fluorescence and phosphorescence maximum for the $n\pi^*$ and $\pi\pi^*$ states. The vertical ordering of the states is maintained in the adiabatic and emission transitions. For the lowest singlet-singlet feature ($n\pi^*$ -type), the vertical and adiabatic transitions are computed within 0.1 eV, and the fluorescence maximum is predicted at 1.94 eV. The $1^1A_1 \rightarrow 1^1A_2$ with the band origin around 0.59 eV below the vertical feature seems to be responsible for the weak band overlapping the lowest $\pi\pi^*$ band. The computed origin at 2.70 eV can be compared to the datum reported for maleimide in EPA mixed solvent at 77 K, 3.33 eV.⁹ As stated earlier, this is a complex band and its intensity should arise by vibronic coupling. Nevertheless, the deviation of the theoretical results with respect to the experimental is somewhat large, even invoking a blue shift due to solvent effects. In this respect, one should recall that the component located at 3.33 eV was considered as the most likely to be the 0-0 band, probably because it is the lowest feature recorded. It is clear that higher resolution in this energy region is required to determine accurately the band origin (and also the most intense peak, if different).

In the vapor absorption spectrum of *N*-methyl maleimide at 360 K, the 0-0 vibrational band of the lowest $\pi\pi^*$ band could not be determined because the onset of this transition was found to be weak, overlapping the $n\pi^*$ transition. However, by inspection of the observed spectrum (see Figure 3 in Seliskar and McGlynn⁹), the band origin is estimated to be around 3.5 eV, consistent with the computed result at 3.65 eV. We have related the third-energy band to the $1^1A_1 \rightarrow 2^1B_2$ electronic transition. The origin is computed at 4.99 eV, that is, 0.6 eV

below the vertical transition. The absorption spectrum of *N*-methyl maleimide at 300 K (Figure 4 in Seliskar and McGlynn⁹), which is slightly unresolved because of instrumental limitations, has the most prominent peak at 5.53 eV. On the basis of the corresponding vibrational analysis, it was assigned by these authors as the band origin.⁹ The prominent feature was subsequently classified as adiabatic-vertical by Robin.¹² The fact is that the recorded spectrum begins at 5.4 eV,⁹ and the MS-CASPT2 result suggests that the origin should be around half an electronvolt below. Can the influence of the methyl group produce such a deviation? It is at present an open question. Once more, it is clear that a higher resolution of the vapor absorption spectrum of maleimides would be highly appreciated. It is also worth noting that the theoretical results, both vertical and adiabatic, are very reasonable compared to the absorption spectrum of maleimide in water,¹³ where the full spectrum from 3.1 to 6.8 eV is depicted. A shift of 0.1-0.2 eV is easily rationalized because of solvent effects, including possible maleimide-water hydrogen bonding, and limitations of the computation.

Maleimide exhibits appreciable luminescence in solution, and its *N*-alkyl derivatives are only very weakly emissive. The emission appears to be solely of phosphorescent nature. The triplet excited state has been tentatively assigned to $n\pi^*$ character on the basis of the observed lifetime (10 ms) of the phosphorescence of maleimide in EPA mixed solvent at 77 K. Not much further evidence was given to support the assignment.⁹ Because of the weak luminescence, it was difficult to locate accurately the 0-0 vibrational band. It was assumed to be at 2.80 eV, the first clear feature of the recorded phosphorescence spectrum, where a maximum around 2.50 eV can be also noted (see Figure 7 reported by Seliskar and McGlynn⁹).

The current results give an additional rationalization of the observed phosphorescence spectrum. The $1^1A_2(lp_2 \rightarrow \pi^*)$ and $1^3B_2(\pi\pi^*)$ states are computed to be, within the error bars of the method, nearly degenerate with adiabatic transitions at 2.70 and 2.77 eV. In fact, at the optimized geometry of the 1^1A_2 state, the 1^3B_2 state is found at 2.72 eV. The small energy gap between the two states and their different nature, $1(n\pi^*)$ and $3(\pi\pi^*)$, suggest that an efficient intersystem crossing can actually take place. The observed phosphorescence with an origin estimated at 2.80 eV and an emission maximum at 2.50 eV approximately is consistent with the computed results for the $1^3B_2 \rightarrow 1^1A_1$ transition, 2.77 eV (0-0) and 2.20 eV (emission maximum). The picture is further complicated by the occurrence of the emission maximum from the $1^3A_2(lp_2 \rightarrow \pi^*)$ state, which is also predicted at 2.21 eV. Thus, taking into account spin-orbit interactions, those states can further interact. It points out that phosphorescence comes mainly from the

$^3(\pi\pi^*)$ state with some mixing of $^3(n\pi^*)$ as suggested earlier by Scharf and Leismann.¹¹ Comparison of the dipole moment of the lowest triplet of $\pi\pi^*$ character and the ground state reveals that the excited state should suffer a larger stabilization in polar solvents; as a consequence, a red shift of the computed transition in vacuo can be expected. Using similar arguments, we can infer a slight blue shift for the triplet $n\pi^*$ transitions. Therefore, even if in the gas phase the lowest two triplet states are computed to be of $n\pi^*$ character, we suggest that the main source of the observed phosphorescence in solution comes most probably from a $\pi\pi^*$ triplet state.

4. Summary and Conclusions

The excited states of maleimide have been studied using ab initio quantum-chemical methods based on multiconfigurational wave functions. Forty-six singlet and four triplet excited states have been characterized. Employing the optimized geometry of the ground state at the CASSCF level, the vertical excitation energies have been computed with the MS-CASPT2 method. The calculation comprised both valence and Rydberg excited states in the singlet manifold. In particular, the 3s, 3p, and 3d members of the Rydberg series converging to the first four ionization potentials have been explicitly taken into account. Four valence singlet excited states have been identified below the lowest 3s Rydberg state. They have been related to the observed absorption spectrum of maleimide in solution and *N*-methyl maleimide in the gas phase. To estimate the band origin and emission maximum, geometry optimizations for the lowest four singlet and triplet excited states have also been performed at the CASSCF level.

We conclude that each of the three observed bands in the absorption spectrum of maleimide (up to 6 eV) is related to a single electronic excitation. The lowest-energy band has $n\pi^*$ character. The lone pair MO can be described as the symmetric linear combination of the lone pair of the oxygen atoms (a_1 symmetry in the C_{2v} point group), and the π^* is the LUMO. Transition to the 1A_2 state is therefore electric-dipole-forbidden, and the source of weak feature observed can be attributed to vibronic coupling. Below the 1A_2 state, the 1B_1 state is computed to be the lowest excited state. It is also of $n\pi^*$ character, where n represents the minus linear combination of the oxygen lone pairs (b_2 symmetry) and π^* is the LUMO. It seems that it has not been detected experimentally. The 1B_1 and 1A_2 states have been placed vertically at 2.48 and 3.29 eV, respectively. The second- and third-energy bands are of $\pi\pi^*$ nature. They have been related to the $^1A_1 \rightarrow ^1B_2$ (4.44 eV) and $^1A_1 \rightarrow ^2B_2$ (5.59 eV) electronic transitions, the latter predicted to be the most intense of the spectrum. The lowest Rydberg state is computed in the high-energy side of the most intense feature at 5.98 eV. The Rydberg states follow the trends expected from the computed ionization potentials and are interleaved among four dipole-allowed excited states of valence character placed vertically at 7.13, 7.30, 8.53, and 8.77 eV. The results complement the available data on the higher-energy region of the spectrum of *N*-methyl maleimide and show the same trends as those analyzed by Robin.

An attempt to rationalize the observed phosphorescence of maleimide has also been performed. In summary, we propose that the relaxation of the singlet state responsible for the lowest-energy band, the $^1A_2(lp_2 \rightarrow \pi^*)$ state, causes the intersection between the $^1A_2(lp_2 \rightarrow \pi^*)$ and $^3B_2(\pi\pi^*)$ potential hypersurfaces in polar solvents, that is, an optimal situation for an efficient intersystem crossing. Therefore, the observed phosphorescence comes mainly from the triplet $^3B_2(\pi\pi^*)$ state,

which is predicted to be the lowest triplet state in solution. The results are consistent with the recorded phosphorescence pattern of maleimide in EPA.

Acknowledgment. Financial support from the MCYT of Spain through the Project BQU2001-2926 and by the Generalitat Valenciana is gratefully acknowledged.

References and Notes

- (1) Miller, C. W.; Jönsson, E. S.; Hoyle, C. E.; Viswanathan, K.; Valente, E. J. *J. Phys. Chem. B* **2001**, *105*, 2707.
- (2) Antczak, C.; Baubois, B.; Monneret, C.; Florent, J. C. *Bioorg. Med. Chem.* **2001**, *9*, 2843.
- (3) Schopfer, L. M.; Salhany, J. M. *Anal. Biochem.* **1998**, *257*, 139.
- (4) Liu, X. H.; Wang, H. K.; Herron, J. N.; Prestwich, G. D. *Bioconjugate Chem.* **2000**, *11*, 755.
- (5) LaFemina, J. P.; Arjavalingam, G.; Hougham, G. *J. Chem. Phys.* **1989**, *90*, 5154.
- (6) Schaffner, K.; Gärtner, W. *The Spectrum* **1999**, *12*, 1.
- (7) González-Luque, R.; Garavelli, M.; Bernardi, F.; Merchán, M.; Robb, M. A.; Olivucci, M. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 9379.
- (8) Molina, V.; Merchán, M. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 4299.
- (9) Seliskar, C. J.; McGlynn, S. P. *J. Chem. Phys.* **1971**, *55*, 4337.
- (10) Seliskar, C. J.; McGlynn, S. P. *J. Chem. Phys.* **1972**, *56*, 1417.
- (11) Scharf, H.-D.; Leismann, H. Z. *Naturforsch.* **1973**, *28*, 662.
- (12) Robin, M. B. *Higher Excited States of Polyatomic Molecules*; Academic Press: New York, 1975; Vol. II.
- (13) Johnson, E. A. In *UV-Vis Atlas of Organic Compounds*; Perkamppus, H.-H., Ed.; VCH: Weinheim, Germany, 1992; p 175.
- (14) Roos, B. O.; Fülischer, M. P.; Malmqvist, P.-Å.; Merchán, M.; Serrano-Andrés, L. In *Theoretical Studies of Electronic Spectra of Organic Molecules*; Langhoff, S. R., Ed.; Kluwer Academic Publishers: Dordrecht, The Netherlands, 1995; p 357.
- (15) Roos, B. O.; Andersson, K.; Fülischer, M. P.; Malmqvist, P.-Å.; Serrano-Andrés, L.; Pierloot, K.; Merchán, M. In *Multiconfigurational Perturbation Theory: Applications in Electronic Spectroscopy*; Prigogine, I., Rice, S. A., Eds.; J. Wiley & Sons: New York, 1996; p 219.
- (16) Merchán, M.; Serrano-Andrés, L.; Fülischer, M. P.; Roos, B. O. In *Multiconfigurational Perturbation Theory Applied to Excited States of Organic Compounds*; Hirao, K., Ed.; World Scientific Publishing Company: Amsterdam, 1999; Vol. 4, p 161.
- (17) (a) Serrano-Andrés, L.; Merchán, M.; Nebot-Gil, I.; Roos, B. O.; Fülischer, M. *J. Am. Chem. Soc.* **1993**, *115*, 6184. (b) Roos, B. O.; Malmqvist, P.-Å.; Molina, V.; Serrano-Andrés, L.; Merchán, M. *J. Chem. Phys.* **2002**, *116*, 7526.
- (18) Serrano-Andrés, L.; Fülischer, M. P.; Roos, B. O.; Merchán, M. *J. Chem. Phys.* **1996**, *100*, 6484.
- (19) Merchán, M.; Roos, B. O. *Theor. Chim. Acta* **1995**, *92*, 227.
- (20) Merchán, M.; Roos, B. O.; McDiarmid, R.; Xing, X. *J. Chem. Phys.* **1996**, *104*, 1791.
- (21) (a) Pou-Amérgo, R.; Merchán, M.; Ortí, E. *J. Chem. Phys.* **1999**, *110*, 9536. (b) Pou-Amérgo, R.; Serrano-Andrés, L.; Merchán, M.; Ortí, E.; Forsberg, N. *J. Am. Chem. Soc.* **2000**, *122*, 6067.
- (22) Andersson, K.; Malmqvist, P.-Å.; Roos, B. O.; Sadlej, A. J.; Wolinski, K. *J. Phys. Chem.* **1990**, *94*, 5483.
- (23) Andersson, K.; Malmqvist, P.-Å.; Roos, B. O. *J. Chem. Phys.* **1992**, *96*, 1218.
- (24) Finley, J.; Malmqvist, P.-Å.; Roos, B. O.; Serrano-Andrés, L. *Chem. Phys. Lett.* **1998**, *288*, 299.
- (25) Widmark, P.-O.; Malmqvist, P.-Å.; Roos, B. O. *Theor. Chim. Acta* **1990**, *77*, 291.
- (26) For reviews of the CASSCF methods, see different contributions in: *Ab Initio Methods in Quantum Chem.-II*; Lawley, K. P., Ed.; J. Wiley & Sons Ltd.: New York, 1987.
- (27) Forsberg, N.; Malmqvist, P.-Å. *Chem. Phys. Lett.* **1997**, *274*, 196.
- (28) Crespo, R.; Merchán, M.; Michl, J. *J. Phys. Chem. A* **2000**, *104*, 8593.
- (29) Malmqvist, P.-Å. *Int. J. Quantum Chem.* **1986**, *30*, 479.
- (30) Malmqvist, P.-Å.; Roos, B. O. *Chem. Phys. Lett.* **1989**, *155*, 189.
- (31) Andersson, K.; Baryz, M.; Bernihardsson, A.; Blomberg, M. R. A.; Boussard, P.; Cooper, D. L.; Fleig, T.; Fülischer, M. P.; Hess, B.; Karlström, G.; Lindh, R.; Malmqvist, P.-Å.; Neogrády, P.; Olsen, J.; Roos, B. O.; Sadlej, A. J.; Schimmelpfennig, B.; Schütz, M.; Seijo, L.; Serrano-Andrés, L.; Siegbahn, P. E. M.; Ståhring, J.; Thorsteinsson, T.; Verezavov, V.; Wahlgren, U.; Widmark, P.-O. *MOLCAS*, version 5.0; Department of Theoretical Chemistry, Chemical Centre, University of Lund: Lund, Sweden, 2000.
- (32) Harsányi, L.; Vajda, E.; Hargittai, I. *J. Mol. Struct.* **1985**, *129*, 315.
- (33) Császár, P.; Császár, A.; Harsányi, L. *J. Mol. Struct.* **1986**, *136*, 323.

Publicación 2

Unified Model for the Ultrafast Decay of Pyrimidine Nucleobases

Merchán, M.; González-Luque, R.; Climent, T.; Serrano-Andrés, L.;
Rodríguez, E.; Reguero, M.; Pelaez, D.

J. Phys. Chem. B **2006**, 110, 26471.

Unified Model for the Ultrafast Decay of Pyrimidine Nucleobases

Manuela Merchán,[†] Remedios González-Luque,[†] Teresa Climent,[†] Luis Serrano-Andrés,^{*,‡} Elena Rodríguez,[‡] Mar Reguero,[‡] and Daniel Peláez[§]

Instituto de Ciencia Molecular, Universitat de València, Apartado 22085, ES-46071 Valencia, Spain,
Departament de Química Física i Inorgànica, Universitat Rovira i Virgili, Marcel·li Domingo s/n, ES-43007
Tarragona, Spain, and Departamento de Química Física, Facultad de Ciencias, Universidad de Málaga,
ES-29071 Málaga, Spain

Received: October 19, 2006

Ultrafast decay processes detected after absorption of UV radiation in gas-phase pyrimidine nucleobases uracil, thymine, and cytosine are ascribed to the barrierless character of the pathway along the low-lying $^1(\pi\pi^*)$ hypersurface connecting the Franck–Condon region with an out-of-plane distorted ethene-like conical intersection with the ground state. Longer lifetime decays and low quantum yield emission are on the other hand related to the presence of a $^1(\pi\pi^*)$ state planar minimum on the S_1 surface and the barriers to access other conical intersections. A unified model for the three systems is established on the basis of accurate multiconfigurational CASPT2 calculations, whereas the effect of the different levels of theory on the results is carefully analyzed.

1. Introduction

Photostability after UV light absorption is one of the basic properties of DNA/RNA nucleobases. It prevents photodamage and makes them optimal constituents of the genetic code.¹ Since quenched DNA fluorescence in nucleobase monomers at room temperature was first reported,² new advances have improved our knowledge on the dynamics of photoexcited DNA. Recent femtosecond pump–probe experiments in molecular beams detected multiexponential decay channels in the femtosecond (fs) and picosecond (ps) timescales for the isolated nucleobases.^{1,3–5} Lack of strong solvent effects and similar ultrafast decays obtained for nucleosides and nucleotides suggest that ultrashort lifetimes of nucleobases are intrinsic molecular properties, intimately related to nonadiabatic dynamic processes leading to radiationless relaxation toward the ground state.¹ In modern photochemistry, the efficiency of nonradiative decays between different electronic states taking place in internal conversion processes is associated with the presence of crossings of different potential energy hypersurfaces (PEH) in regions or seams of conical intersection (CI) points which behave as energy funnels where the probability for nonadiabatic, nonradiative jumps is high.⁶ Although only future reliable molecular dynamics methods based on sound electronic structure calculations can be expected to fully elucidate the subtleties of the decay process, a quantum-chemical description based on the calculation of minimum energy paths (MEP), geometry optimizations, and location of conical intersections is highly needed establish the main aspects of the photochemical process. It is worth mentioning that because of the complexity of the mechanisms, the interpretation of the experimental evidence is carried out nowadays closely following the theoretical findings, which are required, more than ever, to be accurate and conclusive.

The extremely low fluorescence quantum yields reported for the isolated nucleobases^{2,7–9} together with the detected ultrafast femtosecond energy relaxation suggest that the system finds a very efficient path to deactivate the bright spectroscopic $\pi\pi^*$ singlet state, initially populated by direct absorption, and reach the ground state via an accessible CI implying S_1 and S_0 . Different and recent ab initio quantum chemical studies on the PEHs of the pyrimidine nucleobase monomers^{10–16} have described CIs representing crossings of S_1 , both with $\pi\pi^*$ and $n\pi^*$ nature, and S_0 PEHs, and even CIs involving three states have been determined in cytosine¹⁷ and uracil.¹⁸ The photochemical problem cannot be understood by simply describing conical intersections between the different states at various levels of calculation. It is becoming, nowadays, evident that a large number of CIs can be found for multidimensional systems in regions of the PEHs with a complex electronic state structure where degeneracy is frequent.^{19,20} As important as getting the energy and location of a CI is to determine how accessible it might be along the path of energy relaxation, in this case from the initially populated $\pi\pi^*$ state, therefore, the key information is contained in the energy barriers found along the evolution of the different states. In a recent theoretical CASPT2 study on the 9H-tautomer of the adenine molecule²¹ we found that the MEP along the $\pi\pi^*$ HOMO \rightarrow LUMO $^1(\pi\pi^*$ HL) singlet excited state starting at the ground state (gs) Franck–Condon (FC) geometry evolved without finding any barrier toward a $(gs/\pi\pi^*)_{CI}$ conical intersection involving twisting of a N–C bond and a perpendicular C–H bond in the six-membered ring. The barrierless character of the path was also obtained in similar studies of adenine,^{22–25} even when, instead of MEPs, geometry optimizations or linear interpolations from the FC geometry and the CI structure were performed.^{22–24} In pyrimidine nucleobases a similar CI consisting of the twisting of the ethylenic C_5 – C_6 bond (see Figure 1) was first identified in studies on the guanine-cytosine pair²⁶ and uracil.¹⁵ A number of calculations have been performed independently on uracil (U),¹⁵ thymine (T),¹³ and cytosine (C)^{11,14} where this ethene-like CI was located, but there

* To whom correspondence should be addressed. E-mail: luis.serrano@uv.es.

[†] Universitat de València.

[‡] Universitat Rovira i Virgili.

[§] Universidad de Málaga.

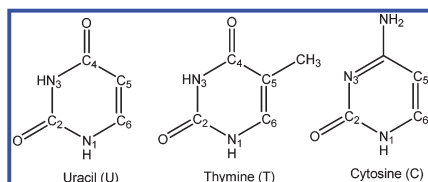


Figure 1. Structure and labeling of the pyrimidine nucleobases.

is no firm and common evidence of which are the profiles of the pathways for accessing the CI with a sound quantum chemical method and no indication of the role played in the photophysics by the minima and other $(gs/\pi\pi^*)_{CI}$ and $(gs/n\pi^*)_{CI}$ conical intersections found as $(S_0/S_1)_{CI}$ PEH crossings,^{10,12} whereas the origin of the weak fluorescence measured for the systems in aqueous media^{1,2,7–9,16} has to be merged into the scheme. Therefore, there is a need for a unified model explaining the fate of the low-lying singlet excited-state population in the three nucleobases.

In the present paper multiconfigurational CASPT2/CASSCF calculations on the low-lying singlet states of the pyrimidine nucleobases U, T, and C have been used to elaborate a framework to explain the most representative feature of their photochemistry, the ultrafast femtosecond decay responsible for their photostability and emission quenching. The main goal of this contribution is twofold. First, we want to provide a common framework for the understanding of the main photochemical event described in the nucleobases, that is, the ultrafast decay of the energy after absorption of UV light, a property which makes them extremely photostable. Despite a number of studies in the field, it seems that some basic features of the mechanism are not firmly established, such as the involvement of the lowest $^1(\pi\pi^* HL)$ HOMO \rightarrow LUMO state in the energy absorption and deactivation, both via internal conversion to the ground state and fluorescence, and the relationship of the ultrafast character of the decay with the *barrierless* profile toward the ethene-like conical intersection found along the relaxation process. Those concepts are here fully established and highlighted within a common scheme for the three pyrimidine nucleobases and we think that this is a message that will help clarify the scenario. The second and no less important aim is related to the reliability of theoretical calculations, specifically multiconfigurational approaches like the CASSCF and CASPT2 methods, which although difficult to use, are undoubtedly the main protagonists of the theoretical studies on the photochemistry of these and similar systems. In this paper we perform a detailed analysis of different levels of calculation and their relationship with the final results, in particular, the dependence on the level of dynamic electron correlation employed in the description of the pathways and the quality of the one-electron basis sets. We consider such an analysis timely in this specific field, taking into account that it is not uncommon to find unbalanced descriptions yielding results that may misguide the interpretations of the experimental evidence. This information is relevant not only for theoreticians but also for experimentally oriented researchers who have to interpret the theoretical findings and need hints to evaluate their reliability.

2. Methods and Computational Details

All calculations reported in the present contribution have been performed using the CASPT2/CASSCF protocol in which geometry optimizations, including minima, surface crossings, and minimum energy paths (MEPs), were carried out at the

multiconfigurational CASSCF level, whereas electronic energy calculations used the second-order multiconfigurational perturbation approach, CASPT2.^{27–30} Different active spaces have been employed to assess the converged character and reliability of the findings. The results used in the discussion to support the main conclusions are based on geometry optimizations of states minima, minimum energy paths, and searches for conical intersections between the ground and the lowest $\pi\pi^*$ excited state at the CASSCF level with an active space comprising all $\pi\pi^*$ electrons and orbitals (MO). The stability of the obtained results was checked, as explained in the text, by increasing the active spaces. In particular, the final results included minimum energy paths computed by adding three extra correlating π^* orbitals. The final calculations represented in the figures, at each of the CASSCF optimized structures, used the ANO-S-type one-electron basis set contracted to C, N, O [3s2p1d]/H [2s1p] (hereafter ANO-S)³¹ and an active space comprising 14 electrons distributed in 10 orbitals, including also lone pairs, and can therefore be labeled CASPT2/CASSCF(14,10)/ANO-S C, N, O [3s2p1d]/H [2s1p]. The standard zeroth-order Hamiltonian was employed in the CASPT2 calculations, which include an imaginary level-shift correction of 0.2 au to avoid the presence of intruder states.³² The accuracy of the computational strategies used in the present paper has been proved many times in different studies.^{10,21,28–30,33–37} Additional technical details on the active spaces and the algorithms used to compute MEPs and CIs can be found in the Supporting Information, together with the obtained geometries. All calculations used the MOL-CAS-6.0 set of programs.³⁸

3. Results and Discussion

A scheme summarizing the present findings and the photochemistry of pyrimidine DNA and RNA nucleobases is displayed in Figure 2. The most relevant photoinduced chemistry of isolated DNA and RNA nucleobases begins with the absorption of the energy to the bright $^1(\pi\pi^* HL)$ singlet state, computed vertically at 5.02 (U), 4.89 (T), and 4.41 eV (C), with related oscillator strengths near 0.2 (0.09 for C), in agreement with previous theoretical and gas-phase experimental data, 5.1 (U), 4.8 (T), and 4.6 eV (C).^{1,39,40} Direct absorption to close-lying $n\pi^*$ states is essentially forbidden, and these states will be populated only by internal conversion processes taking place along the main relaxation route of the $^1(\pi\pi^* HL)$ state. They are not expected to play major roles in the primary photochemical event. After photoexcitation, the relaxation path along $^1(\pi\pi^* HL)$ in the three systems leads in a barrierless form to the ethene-like conical intersection $(gs/\pi\pi^*)_{CI}$ in which the hydrogen atoms (or carbon in T) display a dihedral angle near 120° (see Figure 2), with the CI lying adiabatically (from the ground-state minimum) at 3.9 (U), 4.0 (T), and 3.6 eV (C). Most of the absorbed energy will decay nonradiatively to the ground state through the $(gs/\pi\pi^*)_{CI}$ funnel in an ultrafast relaxation process that we assign to the femtosecond component of the multiexponential decay measured in molecular beams at 130 (U), 105 (T), and 160 fs (C).⁴

Up to now only multiconfigurational approaches have been able to cope with the description of conical intersections and barrier heights in the excited states of nucleobases, although the accurate determination of the different pathways seems to depend on the amount of correlation energy employed. For instance, in U a CASSCF/6-31G(d,p) (valence π) geometry optimization of the $^1(\pi\pi^* HL)$ state starting at the FC geometry leads directly to a planar structure, whereas addition of three extra active correlating orbitals was required to reach the CI. A

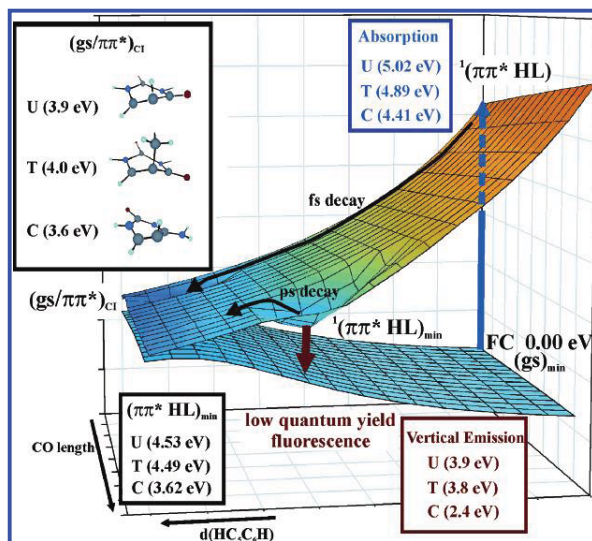


Figure 2. Global picture of the photochemistry of uracil (U), thymine (T), and cytosine (C) as suggested by the CASPT2 calculations.

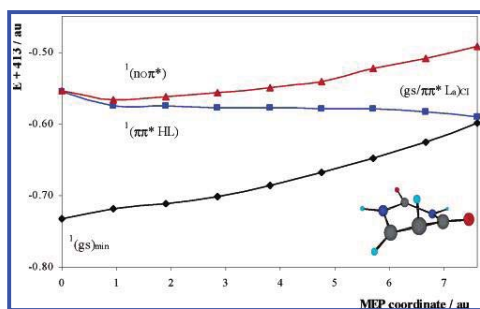


Figure 3. Evolution of the ground and lowest singlet excited states for uracil from the FC geometry along the $^1(\pi\pi^* \text{HL})$ MEP.

MEP performed for uracil clearly shows the barrierless profile of the pathway from FC to the ethene-like $(gs/\pi\pi^*)_{CI}$, related to the measured femtosecond decay, and representing the most efficient path for energy deactivation in nucleobases. Figure 3 displays the evolution of the ground and two lowest singlet excited states of the uracil molecule from the Franck-Condon (FC) geometry along the $^1(\pi\pi^* \text{HL})$ MEP computed at the CASSCF(10,11)/ANO-S level of theory, with an active space of 10 electrons distributed in 11 orbitals comprising all the valence π space of the system plus the addition of three extra correlating orbitals required to properly describe the steepest descent character of the MEP (the same conclusion was obtained with the 6-31G(d) basis set at the same level of calculation). At each optimized point of the MEP, CASPT2//CASSCF(14,10)/ANO-S calculations were performed for the states of interest. The profile of the spectroscopic $^1(\pi\pi^* \text{HL})$ state (adiabatically S_1) is proven to be barrierless toward the ethene-like conical intersection $(gs/\pi\pi^*)_{CI}$. As mentioned above, reduced active spaces excluding the correlating orbitals were unable to evolve toward the ethene-like CI in uracil, in contrast, leading to a planar minimum. In the same way as the addition

of extra correlating orbitals seems to improve the level of the CASSCF treatment, leading to results co-incident with those obtained with more correlated methods such as MRCI,¹⁵ the improvement of the one-electron basis sets also points in the same direction. For instance, CASSCF geometry optimizations for the lowest $^1(\pi\pi^* \text{HL})$ state of the thymine molecule using valence π active spaces and the 6-31G(d) or 6-31G(d,p) basis sets led directly from the FC region to the planar $^1(\pi\pi^* \text{HL})$ state minimum, a local minimum located in a totally different region of the hypersurface. Improvement to the ANO-S basis set, however, changed the outcome, and the optimization led to the ethene-like conical intersection instead. ANO-type basis sets are known to be more accurate and get better account of the correlation energy, whereas it is proved that they provide much better results than segmented basis sets with the same number of contracted functions.^{30,37,41} Both improvements, increasing the quality of the basis set or adding extra correlating orbitals to the active space, partially compensate the lack of dynamic correlation of the π -valence CASSCF method. To ensure the barrierless profile of the relaxation also in thymine, an equivalent MEP as in uracil, that is, from the FC geometry leading to the ethene-like CI, has been computed and it is displayed in Figure 4.

The problems in the description of the relaxation pathways are more severe for cytosine. Previously, we reported for C a MEP computed at the CASSCF level (valence π space) showing that no apparent energy barrier occurred on going from the FC region toward a planar S_1 $^1(\pi\pi^* \text{HL})_{min}$ minimum.³⁴ Energy optimizations adding extra correlating orbitals like in uracil and thymine or improving the basis set do not change the outcome. Does it mean that cytosine is different from the other nucleobases in that there is an energy barrier along the main path from FC to the ethene-like CI? Probably not, but the fact is that, unlike for the other systems, the two critical points, the ethene-like CI $(gs/\pi\pi^*)_{CI}$ and the planar minimum $^1(\pi\pi^* \text{HL})_{min}$, lie in cytosine almost isoenergetic, and therefore the description of the barriers heavily depends on subtle effects of the correlation level used in the optimization approach. Evidence pointing in that direction

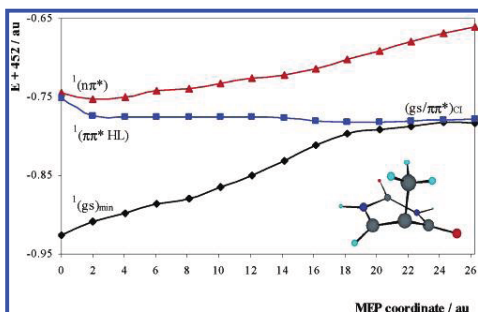


Figure 4. Evolution of the ground and lowest singlet excited states for thymine from the FC geometry along the $^1(\pi\pi^*)$ HL MEP.

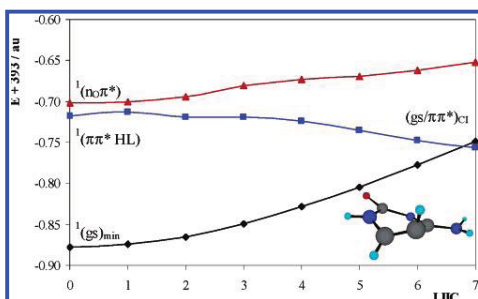


Figure 5. Evolution of the low-lying singlet $\pi\pi^*$ excited states in cytosine from the FC geometry to the geometry of the ethene-like conical intersection $(gs/\pi\pi^*)_{CI}$ along a linear interpolation in internal coordinates (LIIC).

becomes clear when a linear interpolation in internal coordinates (LIIC) is performed in C between the optimized FC and $(gs/\pi\pi^*)_{CI}$ geometries. Figure 5 displays the evolution of the low-lying singlet $\pi\pi^*$ excited states in cytosine from the FC geometry to the geometry of the ethene-like conical intersection $(gs/\pi\pi^*)_{CI}$ along a LIIC computed at the CASPT2//CASSCF(14,10)/ANO-S level of calculation. The initial FC and final $(gs/\pi\pi^*)_{CI}$ geometries were determined in separated calculations. The profile obtained along the $^1(\pi\pi^*)$ HL state in the LIIC procedure has a small barrier of 2.5 kcal/mol at the CASPT2 level found at the beginning of the interpolation, when the dihedral angle HC_3C_6H (see Figure 1) increases up to 16° . Typically, a barrier found along a linear interpolation, in particular when a tight set of points is employed, represents an upper bound for the actual barrier, even when the initial point here is the S_1 state at the FC geometry, a situation without excess energy. To check this aspect, we will perform interpolations in situations we have proved barrierless like uracil and thymine. In the present case, a barrier expected to be lower than 2.5 kcal/mol in cytosine implies in practice that the pathway can be considered essentially barrierless. Increasing the basis set to ANO-S C, N, O [4s3p1d]/H [2s1p]³¹ and ANO-L C, N, O [4s3p2d1d]/H [3s2p]⁴² leads to barriers of 1.0 and 1.1 kcal/mol, respectively, in the initial sections of the interpolation. Similar barriers, never larger than 2 kcal/mol, were obtained when enlarging the active spaces. The barriers were however much higher at the CASSCF level of calculation. For instance, the CASSCF results corresponding to those of Figure 3 display a barrier of 6.6 kcal/mol. This is an indication of how difficult it is to describe the path from FC to the ethene-like CI at the

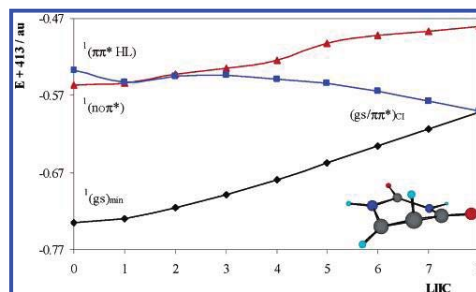


Figure 6. Evolution of the low-lying singlet excited states in uracil between the FC and the ethene-like conical intersection $(gs/\pi\pi^*)_{CI}$ structures along a linear interpolation in internal coordinates (LIIC).

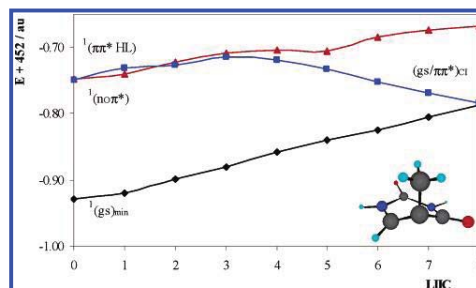


Figure 7. Evolution of the low-lying singlet excited states in thymine between the FC and the ethene-like conical intersection $(gs/\pi\pi^*)_{CI}$ structures along a linear interpolation in internal coordinates (LIIC).

CASSCF level because the path toward the planar minimum will be preferred, enforcing the need of using dynamic correlation for the optimization in cytosine. Therefore, it can be concluded that the described pathway is barrierless also in C and that it is the lack of dynamic electron correlation in the optimization which prevents a more accurate description. Further confirmation can be found if LIIC paths are computed for uracil and thymine at the same level as the MEPs (see Figures 3 and 4) was performed. Figures 6 and 7 display the evolution of the low-lying singlet $\pi\pi^*$ excited states in uracil and thymine, respectively, from the FC geometry to the geometry of the ethene-like conical intersection $(gs/\pi\pi^*)_{CI}$ along a linear interpolation in internal coordinates (LIIC), computed at the CASPT2//CASSCF(14,10)/ANO-S level of calculation. The profile obtained along the $^1(\pi\pi^*)$ HL state in the LIIC procedure renders an energy barrier of near 5 kcal/mol in uracil at a dihedral angle HC_3C_6H close to 30° , and near 12 kcal/mol in thymine at an angle close to 16° . Such energy barriers can be compared to the actual barrierless character of the MEPs in both systems. As in cytosine the barrier derived from the LIIC is much lower than that obtained for uracil and thymine; the barrierless character of the relaxation pathway in C is almost guaranteed.

We cannot discard, especially in C, that the pathways to the planar $^1(\pi\pi^*)_{min}$ minima play a competitive role in energy relaxation. As a matter of fact, as suggested earlier,¹⁰ the longer near 1 ps component decays reported in the molecular beam experiments⁴ could be related to the state deactivation and energy barriers from the minima toward different CIs, not only the ethylenic one but also those involving C=O stretching, previously described.^{10,34} An accurate determination of those

barriers for the three molecules at the highest possible level of calculation will have to be performed to be able to analyze the role of those paths for the secondary decays. Adiabatically, planar $^1(\pi\pi^* \text{HL})_{\text{min}}$ minima for the three systems are computed here at 4.53 (U), 4.49 (T), and 3.62 eV (C). Vertical emission are on the other hand calculated at 3.9 (U), 3.8 (T), and 2.4 eV (C). Those minima can be considered the source of the low quantum yield ($\sim\phi_F = 10^{-4}$) fluorescence observed in solution.^{1,2,7-9,16,43} Fluorescence band origins in aqueous media have been measured near 4.6 (U) and 4.5 eV (T),¹⁶ whereas their respective emission band maxima were reported near 3.9 eV both for U and T.¹⁶ The computed results clearly relate to the experimental evidence. Even if minima corresponding to $n_0\pi^*$ transitions could lie lower in energy, emission takes place essentially from the planar $^1(\pi\pi^* \text{HL})_{\text{min}}$ minima in these systems. Computed radiative lifetimes for U and T, obtained by means of the Strickler–Berg relationship,^{44,45} are $\tau_{\text{rad}} \sim 9$ ns, a value in agreement with the experimental estimate 7–8 ns,¹⁶ which points to a $\pi\pi^*$ state as responsible of the emission. The agreement is not so good for cytosine, and the computed results, both adiabatic and vertical emission, are somewhat lower than the measured fluorescence band maxima in water, ~ 3.8 eV.¹² The assignment of the observed emission band to the $^1(\pi\pi^* \text{HL})_{\text{min}}$ minimum in C has also been proven in previous theoretical studies,^{10,11} whereas the computed radiative lifetime of 33 ns matches also the experimental estimate, 40 ns.¹⁰ The obtained structures at the minima are quite different in C than in the other two nucleobases. The C₂=O (see Figure 1) bond length increases from 1.198 Å in the ground state structure in all three systems to 1.203 (U), 1.202 (T), and 1.353 Å (C) at $^1(\pi\pi^* \text{HL})_{\text{min}}$. In the same relaxation process, C₄=O, on the other hand, increases from 1.199 to 1.249 Å in U and from 1.201 to 1.243 Å in T. In C, the $^1(\pi\pi^* \text{HL})_{\text{min}}$ minimum, lying isoenergetic with the ethene-like CI, has a structure with a more elongated C=O bond, which reflects a close interaction with the nearby $n_0\pi^*$ excited state.^{10,34} For U and T the $^1(\pi\pi^* \text{HL})_{\text{min}}$ minima lie adiabatically 0.5–0.6 eV above the ethene-like CI and maintain a purely $\pi\pi^*$ character. Other $n\pi^*$ minima on the hypersurfaces, even when they may lie lower in energy, at least in the gas phase where they are less perturbed (a blue shift of 0.5 eV is estimated in water⁴⁶), will surely play a minor role in the emission properties. As a main conclusion, however, it must be remembered that also the contribution of the planar $^1(\pi\pi^* \text{HL})_{\text{min}}$ minima, which should be considered a local minimum in the S₁ hypersurface, is not relevant for the main photophysical event, the ultrafast deactivation of the systems, because they lie in a region of the hypersurface far from the main deactivation path involving out-of-plane distortions toward the ethene-like CI. The fact that geometry optimizations or MEPs at the highest level of calculation directly lead to the ethene-like CI does not imply that the planar minima do not exist, but simply that the steepest-descendent MEP describing the main photochemical path from the FC structure does not include such minima. The minor population reaching the planar minima by direct absorption or upon excess energy can be considered responsible for the residual emission observed in solution and also for longer lifetime decays ascribed to the barriers connecting the planar $^1(\pi\pi^* \text{HL})_{\text{min}}$ minima with other low-lying CIs^{10,12} and also to the presence of close-lying singlet $n\pi^*$ and triplet states.

4. Summary and Conclusions

High-level CASPT2//CASSCF quantum chemical methods and ANO-type basis sets have been employed to study the low-lying singlet states of the three pyrimidine DNA and RNA

nucleobases uracil, thymine, and cytosine. The goal of the present research has been to present a unified view to explain the main photochemical events in the three systems after light irradiation, such as the measured ultrafast sub-picosecond decay and the origin of the low quantum yield fluorescence, and to analyze the difficulties that quantum chemical methods encounter when mapping potential energy hypersurfaces in these systems. Whereas two low-lying singlet $\pi\pi^*$ states were used to describe the most basic photochemistry of adenine,²¹ the leading photophysical events in the pyrimidine nucleobases can be explained on the basis of the $^1(\pi\pi^* \text{HOMO} \rightarrow \text{LUMO})$ surface. The ultrafast energy deactivation measured in different environments for the isolated systems can be related to the barrierless character of the path toward the ethene-like out-of-plane $^1(\text{gs}/\pi\pi^*)_{\text{CI}}$ conical intersection. The presence or not of a barrier along this path and the accurate description of the profiles in the different reaction pathways have been shown to depend on the level of calculation employed, in particular, on the amount of correlation energy included in the computational procedure. The absence of dynamic electron correlation energy in a multiconfigurational procedure like CASSCF may easily lead to finding spurious energy barriers along the different paths and missing the main relaxation pathway, as has been shown for uracil and thymine. Addition of extra correlating orbitals into the active space or improvement of the one-electron basis sets, using for instance ANO-type functions, leads to more balanced treatments. In the present case, e.g., for uracil and thymine, MEPs and geometry optimizations show the barrierless character of the path from the FC region to the ethene-like CI only when the higher level computational strategies are employed. It is true that geometry optimizations with a method like CIS, which in practice does not include electronic correlation, is able in these molecules to reach directly the ethene-like CI.^{13,14,26,47} However, because of the well-known weaknesses of the CIS method,^{28,30,41} such an agreement can only be considered fortuitous and cannot be expected to be general in other cases. On the other hand, measured longer lifetime decays and low quantum yield emission in the three nucleobases are here suggested to be connected to the presence of a $^1(\pi\pi^*)$ state planar minimum on the S₁ surface and the barriers to access other conical intersections. That minimum is proved here to be the source of the weak emission recorded for the systems, although such planar structure and other close-lying excited states of $n\pi^*$ type are not expected to participate in the key photochemical event taking place along the main decay pathway toward the ethene-like CI.

Acknowledgment. The research has been supported by projects CTQ2004-01739 and CTQ2005-08459-CO2-02 of the Spanish MEC/FEDER and GV06-192 of the *Generalitat Valenciana*.

Supporting Information Available: Computational details, Cartesian coordinates of the optimized structures, and complete ref 38 (PDF). This material is available free of charge via the Internet at <http://pubs.acs.org>.

References and Notes

- (1) Crespo-Hernández, C. E.; Cohen, B.; Hare, P. M.; Kohler, B. *Chem. Rev.* **2004**, *104*, 1977.
- (2) Daniels, M.; Hauswirth, W. W. *Science* **1971**, *171*, 675.
- (3) Samoylova, E.; Lippert, H.; Ullrich, S.; Hertel, I. V.; Radloff, W.; Schultz, T. *J. Am. Chem. Soc.* **2004**, *127*, 1782.
- (4) Canuel, C.; Mons, M.; Piuze, F.; Tardivel, B.; Dimicoli, I.; Elhanine, M. *J. Chem. Phys.* **2005**, *122*, 074316.

- (5) Kuimova, M. K.; Dyer, J.; George, M. W.; Grills, D. C.; Kelly, J. M.; Matousek, P.; Parker, A. W.; Sun, X. Z.; Towrie, M.; Whelan, A. M. *Chem. Commun.* **2005**, 1182.
- (6) Olivucci, M., Ed. *Computational Photochemistry*; Elsevier: Amsterdam, 2005.
- (7) Hauswirth, W.; Daniels, M. *Chem. Phys. Lett.* **1971**, *10*, 140.
- (8) Callis, P. R. *Chem. Phys. Lett.* **1979**, *61*, 563.
- (9) Callis, P. R. *Ann. Rev. Phys. Chem.* **1983**, *34*, 329.
- (10) Merchán, M.; Serrano-Andrés, L. *J. Am. Chem. Soc.* **2003**, *125*, 8108.
- (11) Tomic, K.; Tatchen, J.; Marian, C. M. *J. Phys. Chem. A* **2005**, *109*, 8410.
- (12) Blancafort, L.; Cohen, B.; Hare, P. M.; Kohler, B.; Robb, M. A. *J. Phys. Chem. A* **2005**, *109*, 4431.
- (13) Zgierski, M. Z.; Patchkovskii, S.; Fujiwara, T.; Lim, E. C. *J. Phys. Chem. A* **2005**, *109*, 9384.
- (14) Zgierski, M. Z.; Patchkovskii, S.; Lim, E. C. *J. Chem. Phys.* **2005**, *123*, 081101.
- (15) Matsika, S. *J. Phys. Chem. A* **2004**, *108*, 7584.
- (16) Gustavsson, T.; Banyasz, A.; Lazzarotto, E.; Markovitsi, D.; Scalamani, G.; Frisch, M. J.; Barone, V.; Impropa, R. *J. Am. Chem. Soc.* **2006**, *128*, 607.
- (17) Blancafort, L.; Robb, M. A. *J. Phys. Chem. A* **2004**, *108*, 10609.
- (18) Matsika, S. *J. Phys. Chem. A* **2005**, *109*, 7538.
- (19) Truhlar, D. G.; Mead, C. A. *Phys. Rev. A* **2003**, *68*, 032501.
- (20) Dreuw, A.; Worth, G. A.; Cederbaum, L. S.; Head-Gordon, M. *J. Phys. Chem. B* **2004**, *108*, 19049.
- (21) Serrano-Andrés, L.; Merchán, M.; Borin, A. C. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 8691.
- (22) Marian, C. M. *J. Chem. Phys.* **2005**, *122*, 104314.
- (23) Perun, S.; Sobolewski, A. L.; Domcke, W. *J. Am. Chem. Soc.* **2005**, *127*, 6257.
- (24) Chen, H.; Li, S. H. *J. Phys. Chem. A* **2005**, *109*, 8443.
- (25) Blancafort, L. *J. Am. Chem. Soc.* **2006**, *128*, 210.
- (26) Sobolewski, A. L.; Domcke, W. *Phys. Chem. Chem. Phys.* **2004**, *6*, 2763.
- (27) Andersson, K.; Malmqvist, P. Å.; Roos, B. O. *J. Chem. Phys.* **1992**, *96*, 1218.
- (28) Roos, B. O.; Fülcher, M. P.; Malmqvist, P. Å.; Serrano-Andrés, L.; Pierloot, K.; Merchán, M. *Adv. Chem. Phys.* **1996**, *93*, 219.
- (29) Merchán, M.; Serrano-Andrés, L.; Fülcher, M. P.; Roos, B. O. In *Recent Advances in Multireference Methods*; Hirao, K., Ed.; World Scientific Publishing: Singapore, 1999.
- (30) Merchán, M.; Serrano-Andrés, L. In *Computational Photochemistry*; Olivucci, M., Ed.; Elsevier Publishing: Amsterdam, 2005.
- (31) Pierloot, K.; Dumez, B.; Widmark, P.-O.; Roos, B. O. *Theor. Chim. Acta* **1995**, *90*, 87.
- (32) Forsberg, J.; Malmqvist, P. Å. *Chem. Phys. Lett.* **1997**, *274*, 196.
- (33) Serrano-Andrés, L.; Fülcher, M. P. *J. Am. Chem. Soc.* **1998**, *120*, 10912.
- (34) Merchán, M.; Serrano-Andrés, L.; Robb, M. A.; Blancafort, L. *J. Am. Chem. Soc.* **2005**, *127*, 1820.
- (35) González-Luque, R.; Garavelli, M.; Bernardi, F.; Merchán, M.; Robb, M. A.; Olivucci, M. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 9379.
- (36) Serrano-Andrés, L.; Merchán, M.; Lindh, R. *J. Chem. Phys.* **2005**, *122*, 104107.
- (37) Serrano-Andrés, L.; Merchán, M. In *Encyclopedia of Computational Chemistry*; Schleyer, P. v. R., Schreiner, P. R., Schaefer, H. F., III, Jørgensen, W. L., Thiel, W., Glen, R. C., Eds.; Wiley: Chichester, 2004.
- (38) Andersson, K. et al. *MOLCAS, version 6.0*; Department of Theoretical Chemistry, Chemical Centre, University of Lund, P.O.B. 124, S-221 00 Lund, Sweden, 2004.
- (39) Fülcher, M. P.; Roos, B. O. *J. Am. Chem. Soc.* **1995**, *117*, 2089.
- (40) Lorentzon, J.; Fülcher, M. P.; Roos, B. O. *J. Am. Chem. Soc.* **1995**, *117*, 9265.
- (41) Serrano-Andrés, L.; Merchán, M. *J. Mol. Struct. Theochem* **2005**, *729*, 99.
- (42) Widmark, P. O.; Malmqvist, P. A.; Roos, B. O. *Theor. Chim. Acta* **1990**, *77*, 291.
- (43) Vigny, P.; Duquesne, M. *Photochem. Photobiol.* **1974**, *20*, 15.
- (44) Strickler, S. J.; Berg, R. A. *J. Chem. Phys.* **1962**, *37*, 814.
- (45) Rubio-Pons, O.; Serrano-Andrés, L.; Merchán, M. *J. Phys. Chem. A* **2001**, *105*, 9664.
- (46) Zazza, C.; Amadei, A.; Sanna, N.; Grandi, A.; Chillemi, G.; Di, Nola, A.; D'Abramo, M.; Aschi, M. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1385.
- (47) Shukla, M. K.; Mishra, P. C. *Chem. Phys.* **1999**, *240*, 319.

Publicación 3

Theoretical Insight into the Spectroscopy and Photochemistry of Isoalloxazine, the Flavin Core Ring

Climent, T.; González-Luque, R.; Merchán, M.; Serrano-Andrés, L.

J. Phys.Chem. **2006**, *110*, 13584.

Theoretical Insight into the Spectroscopy and Photochemistry of Isoalloxazine, the Flavin Core Ring

Teresa Climent, Remedios González-Luque, Manuela Merchán, and Luis Serrano-Andrés*

Instituto de Ciencia Molecular, Universitat de València, Apartado 22085, ES-46071 Valencia, Spain

Received: September 5, 2006; In Final Form: October 14, 2006

The electronic singlet–singlet and singlet–triplet electronic transitions of the isoalloxazine ring of the flavin core are studied using second-order perturbation theory within the framework of the CASPT2//CASSCF protocol. The main features of the absorption spectrum are computed at 3.09, 4.28, 4.69, 5.00, and 5.37 eV. The lowest singlet (S_1) and triplet (T_1) excited states are found to be both of π character with a singlet–triplet splitting of 0.57 eV. On the basis of the analysis of the computed spin–orbit couplings and the potential energy hypersurfaces built for the relevant excited states, the intrinsic mechanism for photoinduced population of T_1 is discussed. Upon light absorption, evolution of the lowest singlet excited state along the relaxation pathway leads ultimately to the population of the lowest triplet state, which is mediated by a singlet–triplet crossing with a state of $n\pi^*$ type. Subsequently a radiationless decay toward T_1 through a conical intersection takes place. The intersystem crossing mechanism and the internal conversion processes documented here provide a plausible route to access the lowest triplet state, which has a key role in the photochemistry of the flavin core ring and is mainly responsible for the reactivity of the system.

1. Introduction

Substances that contain the *isoalloxazine* ring are called *flavins* and the flavin-containing proteins are known as *flavoproteins*.¹ Phototropin is a plasma membrane associated protein that constitutes a relevant and well-known example of flavoprotein and regulates phototropism response of higher plants by using two isoforms (phot1 and phot2), as well as chloroplast photorelocation, and cell stomatal opening.^{2,3} Phototropin acts as a blue-light sensitive photoreceptor having two LOV (light, oxygen, and voltage) domains, LOV1 and LOV2, and uses as chromophore a noncovalently bound single flavin mononucleotide (FMN) (see Figure 1) per domain.⁴

It is accepted among the scientific community that the primary step of the FMN photocycle after light absorption involves population of the lowest singlet excited state (S_1) and subsequently a rapid decay to the lowest triplet excited state (T_1) takes place through an intersystem crossing (ISC) mechanism.^{5,6,7,8,9} The fate of the photoinduced electron transfer to triplet flavins, which behave like acceptors, has been indeed extensively analyzed by using different experimental techniques.^{10,11} Nevertheless, a detailed account for the singlet–triplet nonradiative process within the framework of nonadiabatic photochemistry has not been yet provided. In particular, the key issue seems to be whether relaxation to the triplet state involves a direct coupling between the lowest singlet and triplet states or proceeds via a third excited state. Surprisingly, despite the relevance of flavoproteins, just a few theoretical studies on the excited states of related systems have been previously carried out, either by employing density functional theory (DFT) schemes or by using semiempirical methodologies.^{6,12–14} To the best of our knowledge, there are no high-level *ab initio* studies offering a detailed mechanism on the photochemistry of isoalloxazine (benzo[*g*]pteridine-2,4(3*H*,10*H*)-dione), the flavin

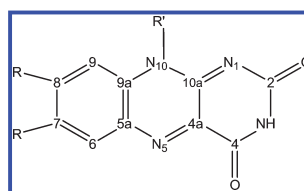


Figure 1. Structure and labeling of the isoalloxazine ring and related flavins.

core ring. Analysis of the photochemical behavior of the isoalloxazine ring on the basis of accurate (predictive) quantum-chemical computations has been the underlying reason to undertake the present research. As a first step toward the understanding of the isoalloxazine photochemistry on theoretical grounds, the main vertical and adiabatic electronic transitions of the molecule are initially considered. The electronic spectra of lumiflavin, 8-methylisoalloxazine, and riboflavin show the same characteristics and can be described as typical of isoalloxazines.¹⁵ In the energy region up to 5.5 eV, the absorption spectra of these systems are composed mainly of three bands. On the other hand, no significant differences have been found in the polarization spectra of lumiflavin, riboflavin, and FMN with respect to the parent molecule, the oxidized form of the isoalloxazine ring,¹⁶ which can therefore be expected to be responsible for the flavin photochemistry.

We report here comprehensive theoretical research on the valence excited states of the isoalloxazine molecule making possible confident assignments. Furthermore, the results provided can serve as calibration for less computationally demanding methods. The excited states of the system have been characterized by using multiconfigurational second-order perturbation theory through the CASPT2 method.¹⁷ The successful performance of the CASPT2 approach in the interpretation of the electronic spectra of a large number of organic compounds

* To whom correspondence should be addressed. E-mail: Luis.Serrano@uv.es.

is well established.^{18–26} Particularly related to the present work are the previous studies performed on chromophores of different biomolecules, such as the protonated Schiff base of retinal (PSB),²⁷ the chromophore of the Photoactive Yellow Protein (PYP),²⁸ and the DNA nucleobases cytosine²⁹ and adenine.³⁰ The present work includes determination of vertical transition energies, band origins, and emission maxima related to the low-lying valence singlet and triplet excited states. In addition, on the basis of the analysis performed for the relaxation pathway along the S_1 hypersurface, it is concluded that the coupling between $S_1(\pi\pi^*)$ and $T_1(\pi\pi^*)$ does not occur directly but is mediated by a high-lying electronic triplet state, ${}^3(n_N-\pi^*)$, which involves the one-electron promotion from the minus linear combination of the lone-pair molecular orbitals (MOs), implying the nitrogen atoms, to the π system.

Computational details are described in the following section. Next, the equilibrium geometry computed for the ground and low-lying excited states, as well as the CASPT2 vertical and adiabatic singlet \rightarrow singlet and singlet \rightarrow triplet absorption spectra, together with other spectroscopic features, are presented and discussed. Finally, the current picture on the photochemistry of isoalloxazine obtained from global research is addressed. Our conclusions are summarized in the last section.

2. Computational Details

Geometry optimization of the ground and low-lying valence excited states were carried out at the CASSCF level.³¹ The π -valence active space of the isoalloxazine ring involves 18 electrons and 16 molecular orbitals (MOs), hereafter denoted by CASSCF(18,16). For the computation of the excited states of $\pi\pi^*$ character, and after careful calibration calculations, it could be decreased to 14 π -valence electrons distributed among 13 π -valence MOs, resulting in the CASSCF(14,13) wave function. In order to study the $n\pi^*$ states, the active space was enlarged with the corresponding lone-pair MOs, keeping inactive the π MO with the highest occupation number in the preceding CASSCF(14,13) treatment and considering as secondary the π MO with the lowest occupation number, finally yielding the CASSCF(16,13) wave function. It is a reasonable choice that largely fulfills the required flexibility to describe the $n\pi^*$ states. The basis set 6-31G(d) was used throughout. Inclusion of polarization functions on the hydrogen atoms by using the 6-31G(dp) basis set had a minor effect on the computed electronic transition energies.

Exploratory calculations in different regions of the hypersurfaces starting from well distorted geometries led to basically planar singular points in all the cases for ground and excited states; therefore, according to our research, the molecule keeps planarity during the photochemical reaction. The vertical transition energies were then computed within the constraints of C_s symmetry, making the computation more technically tractable. The MOs for the vertical excited states were obtained from average CASSCF computations, where the averaging includes all states of interest within a given symmetry: eight roots for ${}^1A'$ states, two roots for ${}^1{}^3A''$ states involving the lone-pair MOs (treating those located on the oxygen or nitrogen atoms in separate calculations but using the same scheme for the active space), and one root for the lowest ${}^3A'$ state. To take into account the remaining electron correlation effects, the CASSCF wave functions were employed as reference in a single-state second-order perturbation CASPT2 treatment.¹⁷ In addition, the effect of weakly interacting intruder states was minimized by using the so-called imaginary level-shift technique.³² After the usual calibration, a shift of 0.3 au was selected

to compile the final results. The vertical energy of each excited state is referred to the ground-state energy computed with the same active space. The carbon, nitrogen, and oxygen 1s electrons were kept frozen in the form determined by the ground-state closed-shell Hartree–Fock wave function and were not correlated at the second-order level.

The CASSCF state interaction (CASSI) method^{33,34} was used to calculate the transition dipole moments (TDMs) from the CASSCF wave functions. The corresponding CASSCF TDM and the energy difference obtained in the CASPT2 computation were used in the formula for the oscillator strength, and similarly to estimate the radiative lifetimes by using the Strickler–Berg relationship.^{35,36}

From the Franck–Condon (FC) geometry, a minimum energy path (MEP) for the lowest singlet excited electronic state (S_1) of the system was computed by using a new recently developed algorithm for constrained optimizations.³⁷ The evolution of the states was mapped by computing at each MEP coordinate the CASPT2 energies, for both the singlet and triplet excited states of interest. The spin–orbit coupling (SOC) has been estimated as described in detail in a related study on the triplet-state formation along the ultrafast decay of excited cytosine.³⁸ All computations in the present paper use the MOLCAS 6.0 quantum chemistry package.^{39,40}

3. Results and Discussion

As a preliminary step toward the theoretical understanding of the electronic spectra of isoalloxazine, the optimized geometries computed for the ground state and the low-lying singlet and triplet excited states of the molecular model considered are first presented. The properties of the excited states calculated vertically are next analyzed. After that, the nonvertical transition energies are discussed. Finally, the photoinduced population of the lowest triplet state along the relaxation path of S_1 is addressed. The present findings will be compared with the available experimental data and previous theoretical findings.

3.1. Equilibrium Structures. The CASSCF optimized geometric parameters for the ground state (gs)_{min} and the lowest singlet and triplet excited states of $\pi\pi^*$ character, ${}^1(\pi\pi^*)$ _{min} and ${}^3(\pi\pi^*)$ _{min}, respectively, are depicted in Figure 2. In addition, the results for the triplet state ${}^3(n_N-\pi^*)$, particularly relevant for the rationalization of photochemical behavior of the molecule, are also included. The Cartesian coordinates for the optimized structures can be found in the Supporting Information.

Most of the computed bond distances for the ground state deviate by less than 0.020 Å from the available crystallographic data.⁴¹ Since isoalloxazine tautomerizes immediately to alloxazine by a N_{10} to N_1 proton shift, the simplest suitable derivative to compare with corresponds to 10-methylisoalloxazine.⁴¹ The agreement between the CASSCF results obtained for the ground state and the X-ray data is satisfactory. Therefore, it can be inferred that the methyl substituent plays a minor role in determining the equilibrium structure of the fused-ring molecular system. A similar situation is expected for the spectroscopic properties.

The major geometric changes taking place for the ${}^1(\pi\pi^*)$ _{min} and ${}^3(\pi\pi^*)$ _{min} structures, with respect to the ground state (gs)_{min}, involve an increase of the N_1-C_{10a} and $C_{4a}-N_5$ bond distances and of the $C_{5a}-C_{9a}$ bond length to a lesser extent, whereas a decrease of the $C_{4a}-C_{10a}$ bond distance is also noted. In terms of the natural orbitals (NOs) of the CASSCF wave function, the excited states ${}^1(\pi\pi^*)$ and ${}^3(\pi\pi^*)$ are described mainly by a one-electron promotion from the NO topologically equivalent to the highest occupied MO (HOMO-like) to the lowest

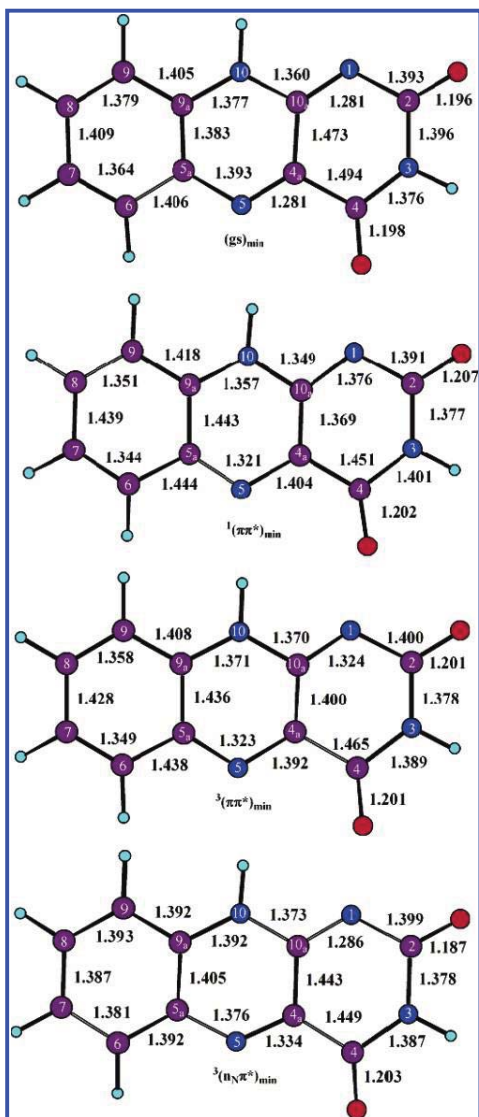


Figure 2. Optimized CASSCF/6-31G(d) structures for the ground and low-lying excited states.

unoccupied MO (LUMO-like). Therefore, the pattern found for the variations in geometry upon excitation correlates nicely to the relative electron-density shifts concomitant to the HOMO/LUMO-like electron redistribution. As can be seen in the Supporting Information, in the HOMO-like NO the electron density is mainly localized in the nitrogen atom N₁ and in the C_{5a}–C_{9a} bonding region. For the LUMO-like NO, the electron density increases considerably in the region of the C_{4a}–C_{10a} bond and the nitrogen atom N₅. The occurrence of the reported geometric changes can be seen just as the response of the

electronically excited molecule to such electron reorganization. Similarly, the geometry computed for the ³($n_N\pi^*$) state is consistent with migration of electron density from the lone pairs located on the nitrogen atoms to the π system, basically toward the LUMO-like NO.

The optimized geometries were employed to characterize the vertical electronic spectra (absorption and emission) and non-vertical electronic transitions.

3.2. Vertical Spectra. The computed vertical excitation energies at the CASSCF and CASPT2 levels of theory, together with oscillator strengths of the related electronic transitions, dipole moments for the states, and selected experimental data are listed in Table 1. The resulting assignments for the excited states are given within parentheses in the first column of Table 1, identifying the different excited states obtained for the isoalloxazine ring, where n_{N-} and n_{N+} represent the minus and plus linear combinations of the in-plane lone pairs belonging to the N₁ and N₅ atoms (see Figure 1 for atom labeling). The same applies for the two oxygen atoms: n_{O-} and n_{O+} . For the sake of brevity, the lowest singlet ²1A'($\pi\rightarrow\pi^*$) and triplet ³1A'($\pi\rightarrow\pi^*$) excited states will be hereafter called S₁ and T₁, respectively, and the ³1A''($n_N\pi^*$) state will be denoted by T_N. As discussed below, this state might be especially involved in the deactivation of the S₁ state and subsequent population of the lowest triplet T₁ state.

At the highest level of theory employed, five valence electronic transitions involving excited states of $\pi\pi^*$ character are computed with significant oscillator strength and can be related to the main features of the recorded electronic spectrum up to 5.5 eV. Only one transition is found in the visible range at 3.09 eV, corresponding to the vertical absorption from the ground state (S₀) to the lowest singlet excited state. The S₀ → S₁ transition has an associated oscillator strength of about 0.2. As stated above, in terms of the NOs of the CASSCF wave function, S₁ involves electron reorganization from the HOMO-like to the LUMO-like NOs. As the S₁ dipole moment (9.7 D) is much larger than that of the ground state (2.4 D), the occurrence of a red shift in polar solvents can be expected for the lowest energy transition. This is consistent with the absorption spectra reported by Sun et al.¹⁵ for methylisalloxazine derivatives in ethanol at 77 K, where a maximum peaking at 2.85 eV is observed for 8-methylisalloxazine and at 2.76 eV for lumiflavin, 3-methyllumiflavin, and riboflavin. The lowest singlet excited state of $n\pi^*$ type (T_N) is calculated to lie vertically at 3.34 eV, that is, close to the lowest transition. As expected, the intensity of the transition is predicted to be exceedingly weak (oscillator strength 0.007). Nevertheless, because of the close proximity of the $n\pi^*$ and $\pi\pi^*$ states, they could further interact by vibronic coupling and the resulting feature could be related to the shoulder observed in the high region of the lowest energy visible band recorded in ethanol.¹⁵

The second energy band recorded experimentally can be related to the electronic transition ¹1A' → ³1A' of $\pi\pi^*$ nature computed in vacuo at 4.28 eV and with a slightly smaller oscillator strength than that obtained for the lowest transition. The vertical result is somewhat blue shifted with respect to the corresponding band maximum observed at 3.62–3.76 eV in solution (depending on the methyl derivative).¹⁵ Absorption and fluorescence studies⁴² made on different isoalloxazines have concluded that an increase of the solvent polarity has only a little influence on the lowest energy absorption maximum, whereas the maximum in the second energy band exhibits a pronounced bathochromic shift. Accordingly, the dipole moment calculated for the ³1A' state is indeed larger than that for the

TABLE 1: Experimental and Computed CASSCF And CASPT2 Excitation Energies, Related Oscillator Strengths (f), and CASSCF Dipole Moments (μ) for the Vertical Excited States of Isoalloxazine

state ^a	excitation energy (eV)		f	μ (D)	experimental ^b
	CASSCF	CASPT2			
ground state					
1 ¹ A'				2.4	
singlet excited states					
2 ¹ A'($\pi \rightarrow \pi^*$)	4.15	3.09	0.239	9.7	2.85
1 ¹ A''(n _N - $\rightarrow \pi^*$)	4.79	3.34	0.007	5.4	
2 ¹ A''(n _O - $\rightarrow \pi^*$)	5.15	3.75	0.001	1.9	
3 ¹ A'($\pi \rightarrow \pi^*$)	5.47	4.28	0.158	10.1	3.76
3 ¹ A''(n _{O+} - $\rightarrow \pi^*$)	5.62	4.43	0.000	2.6	
4 ¹ A''(n _{N+} - $\rightarrow \pi^*$)	5.83	4.66	0.001	5.0	
4 ¹ A'($\pi \rightarrow \pi^*$)	6.27	4.69	0.104	7.8	
5 ¹ A'($\pi \rightarrow \pi^*$)	5.88	4.90	0.035	7.9	
6 ¹ A'($\pi \rightarrow \pi^*$)	6.78	4.92	0.058	1.8	
7 ¹ A'($\pi \rightarrow \pi^*$)	7.16	5.00	0.337	7.1	4.68–4.77
8 ¹ A'($\pi \rightarrow \pi^*$)	7.48	5.37	0.641	5.5	
triplet excited states					
1 ³ A'($\pi \rightarrow \pi^*$)	3.14	2.52		6.0	
1 ³ A''(n _N - $\rightarrow \pi^*$)	4.30	2.97		6.3	
2 ³ A''(n _O - $\rightarrow \pi^*$)	4.99	3.70		2.2	
3 ³ A''(n _{O+} - $\rightarrow \pi^*$)	5.47	4.39		2.8	
4 ³ A''(n _{N+} - $\rightarrow \pi^*$)	5.57	4.53		5.1	

^a The character and ordering of the states are those obtained with the CASPT2 method. ^b Band maxima (eV) from the absorption spectrum of 8-methylisoalloxazine in ethanol at 77 K.¹⁵

2¹A' state and consequently a larger red shift should occur in polar solvents for the second energy band maximum. Since two $n\pi^*$ states involving the lone pairs of the carbonyl group are energetically close, an interaction with the 2¹A'($\pi \rightarrow \pi^*$) state may also occur, leading to a complex band envelope where the maximum might not necessarily be coincident with the vertical feature here computed, which is more a rule than an exception as far as the spectroscopy of the carbonyl group is concerned.²⁴ Because of the large difference between the ground and excited state geometries, the spectroscopic features related to CO bonds are characterized by long vibrational progressions, and as a result the maximum of the band does not match the vertical transition, as is commonly assumed within the FC principle.^{22,43} Taking into account the error bars of the computation (0.1–0.2 eV), the absence of the methyl group in the calculation, and the possible solvent influence, one could easily conclude that the agreement between theory and experiment can be regarded as satisfactory. Nevertheless, experience has shown us that when the deviation of vertical CASPT2 result from the observed maximum is significant, as much as 1/2 eV as is the case here, it points to the nonvertical nature of the recorded band. This issue could only be resolved by carrying out computationally a vibrational resolution of the band, a formidable task at present for a system of the molecular size of isoalloxazine, but that might hopefully be undertaken soon with the expected advances to be seen from both software and hardware standpoints.

The third energy band of isoalloxazines in ethanol solution has a band maximum in the 4.7–4.8 eV energy range.¹⁵ From 4.7 to 5.4 eV five electronic transitions are computed that might contribute to the observed spectra. Most of the states have dipole moments higher than the ground state, and consequently the computed transitions are expected to be somewhat blue shifted with respect to the recorded data in ethanol, a constant behavior found in the present system for the transitions bearing the largest oscillator strengths. Thus, we conclude that the third energy band is the result of several $\pi\pi^*$ electronic transitions, in particular those taking place from the ground state to the 4¹A' (4.69 eV), 7¹A' (5.00 eV), and 8¹A' (5.37 eV) excited states, which have significant oscillator strengths (cf. Table 1) and therefore constitute the main contributors to the band. The

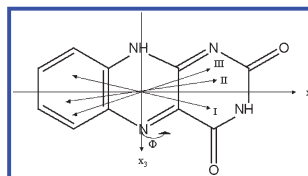


Figure 3. Polarization directions for the low-energy electronic transitions (I, II, and III) in the absorption spectrum of isoalloxazine. Using a previous criterion,¹⁶ the angle Φ has been calculated with respect to the x_3 short axis.

CASSCF wave functions of the 4¹A' and 7¹A' excited states have a dominant participation of the singly excited HOMO \rightarrow LUMO+1 configuration. The two $\pi\pi^*$ states 5¹A' and 6¹A' can be described mainly as the result of the HOMO-2 \rightarrow LUMO one-electron configuration with contributions of the two-electron promotions from the HOMO to the LUMO in the former and the HOMO-3 \rightarrow LUMO in the latter.

As can be seen in Table 1, a number of $n\pi^*$ states are interleaved among the main valence transitions. The 1¹A''(n_N- $\rightarrow \pi^*$), 2¹A''(n_O- $\rightarrow \pi^*$), 3¹A''(n_{O+}- $\rightarrow \pi^*$), and 4¹A''(n_{N+}- $\rightarrow \pi^*$) states are placed vertically at 3.34, 3.75, 4.43, and 4.66 eV, respectively. The computed oscillator strengths associated with those transitions are negligible and therefore do not have special relevance with regard to the absorption spectrum.

Further support for the proposed assignments can be obtained by analyzing the polarization directions computed for the main transitions, which are depicted in Figure 3, with respect to the available experimental data.

The results $\Phi_I = 75^\circ$, $\Phi_{II} = 100^\circ$, and $\Phi_{III} = 120^\circ$, where Φ_I , Φ_{II} , and Φ_{III} are the angles of the transition dipole moment of each band with respect to the short molecular axis, are consistent with those determined experimentally.^{15,16,44} In particular, the relative angle between the computed TDM obtained between the lowest and second energy bands is close to that deduced by Sun et al.,¹⁵ $\Phi_{I,II} = 20^\circ \pm 5^\circ$, from fluorescence polarization data. As compared to the data reported by Johansson et al.,¹⁶ $\Phi_I = 58^\circ \pm 4^\circ$, $\Phi_{II} = 97^\circ \pm 3^\circ$, and $\Phi_{III} = 119^\circ \pm 2^\circ$, a large deviation for Φ_I is noted, although both

TABLE 2: Computed Vertical Transition, Band Origin, and Vertical Emission Energies and Radiative Lifetimes (τ_{rad}) for the Low-Lying Singlet Excited States of Isoalloxazine

state	vertical transition (eV)		band origin (eV)		vertical emission (eV)		τ_{rad}
	CASSCF	CASPT2	CASSCF	CASPT2	CASSCF	CASPT2	
singlet states							
$\pi\pi^*$	4.15	3.09	3.20	2.69	2.21	2.04	15 ns
$n_N\pi^*$	4.79	3.34	3.82	3.33	2.65	2.61	467 ns
$n_O\pi^*$	5.15	3.75	3.52	3.16	2.19	2.33	6458 ns
triplet states							
$\pi\pi^*$	3.14	2.52	2.41	2.03	1.73	1.75	116 ms
$n_N\pi^*$	4.30	2.97	3.50	2.73	3.00	2.37	
$n_O\pi^*$	4.99	3.70	3.40	3.10	2.20	2.34	

theory and experiment agree in the long-axis polarization character of the band. It is worth mentioning, however, that the theoretical prediction matches the value given by Eaton et al.⁴⁴ ($\Phi_1 = 75^\circ$). The present results would hopefully help to elucidate the discrepancy between the mentioned sets of experimental data.^{16,44}

The ordering and nature of the five lowest valence triplet states is the same as that obtained for the singlet excited states. The singlet–triplet spectrum computed at the ground-state geometry of isoalloxazine places the lowest triplet state $1^3A'(\pi \rightarrow \pi^*)$ at 2.52 eV. Next, at 2.97 eV, the $1^3A''(n_N \rightarrow \pi^*)$ state is located. As can be seen in Table 1, the remaining $n\pi^*$ triplet states are found within the energy range 3.70–4.53 eV. As expected, the largest singlet–triplet splitting is computed for the low-lying $\pi\pi^*$ states (0.57 eV). Assuming that both singlet and triplet states are described by the same MOs, in a simple molecular orbital model the exchange integral is related to the singlet–triplet energy difference. Normally, since that integral is relatively larger for $\pi\pi^*$ than for $\sigma\pi^*$ states, the singlet–triplet splitting becomes more pronounced in states of $\pi\pi^*$ character. Thus, except for the pair $1^3A''(n_N \rightarrow \pi^*)$, which has a singlet–triplet energy difference of 0.37 eV, the other singlet–triplet gaps are at most 0.1 eV. Since the states discussed are described mainly by singly excited configurations, that elementary reasoning holds true despite the use of large and complex CASSCF wave functions.

As far as we are aware, no theoretical studies have been performed on the excited states of isoalloxazine itself. However, time-dependent DFT results on related compounds have offered a similar three-band view of the singlet–singlet spectrum attributed to ($\pi \rightarrow \pi^*$) transitions,⁶ whereas the relative placement of the lowest ($n \rightarrow \pi^*$) transition seems to depend on the details of the computational procedure.^{6,11}

3.3. Emission and Nonvertical Transition Energies. Knowledge of the energy hypersurfaces of the ground and low-lying excited states of isoalloxazine is essential for the determination of spectroscopic properties, which are particularly relevant in an ample group of biologically active chromophores. The basic understanding of the spectroscopic behavior of these compounds requires not only the accurate computation of the vertical states but also the characterization of the emission and nonvertical transition energies.

The computed transitions involve the energy difference between the ground state and the excited state at the respective optimized geometries and the vertical emission from the excited-state minimum vertically to the ground state. The lowest component of each absorption band is the 0–0 absorption, also labeled T_0 , and involves the lowest vibrational states. The value computed here is, however, T_e , obtained as the difference of the adiabatic electronic energy values. On the other hand, the computed vertical emission (fluorescence or phosphorescence) can be considered as a lower bound for the observed maximum

of the emission band. Since the active spaces are the same as those employed in the geometry optimizations and for the computation of vertical transitions (a π -valence plus lone-pairs MO as appropriate), a consistent comparison can be then performed between the computed absorption and emission spectra. The study includes determination of both vertical and adiabatic transitions related to absorption and emission bands for the three lowest valence excited $\pi\pi^*$ and $n\pi^*$ states in the singlet and triplet manifolds as listed in Table 2.

For the lowest singlet–singlet feature ($\pi\pi^*$ type), the vertical and the adiabatic transitions are computed within 0.4 eV, with a fluorescence band origin predicted at 2.69 eV and a vertical emission at 2.04 eV. The obtained results can be directly related to the reported spectra of 8-methylisoalloxazine in ethanol,¹⁵ displaying low-lying absorption and emission band origins near 2.7 eV and fluorescence band maxima between 2.5 and 2.3 eV. The computed radiative lifetime, 15 ns, is also in agreement with the experimental datum for the fluorescence lifetime of the methylisoalloxazine derivatives (15.1 ns for 3,10-dimethylisoalloxazine in acetonitrile solution⁴²).

On the basis of relatively long lifetimes and the tendency for negative phosphorescence polarization in the 0–0 region with respect to the lowest $\pi\pi^*$ transitions, the phosphorescent triplet state of isoalloxazines has been assigned as a $^3(\pi, \pi^*)$ type.¹⁵ Our computed results are in agreement with that suggestion, predicting that the lowest triplet state has (π, π^*) character with the phosphorescence band origin at 2.03 eV and vertical emission at 1.75 eV, in relatively good agreement with the reported values for 8-methylisoalloxazine in ethanol,¹⁵ near 2.2 and 2.0 eV, respectively. A detailed account for the intrinsic mechanism of the T_1 population is subsequently considered.

3.4. Triplet-State Formation. Upon near-UV light irradiation, the isoalloxazine system is expected to be promoted to the lowest excited state of the same symmetry as the ground state (singlet). As presented above, the vertical transition to S_1 at the FC geometry was computed to be 3.09 eV with an oscillator strength of around 0.2. To analyze the evolution of S_1 , once that the initial absorption took place, a minimum energy path (MEP) along the S_1 reaction coordinate was computed at the CASSCF level. At the stationary geometries so obtained along the path, a number of singlet and triplet excited states were subsequently calculated at the CASPT2 level (see Supporting Information), yielding in this manner a static picture of the energetic variations that suffer the different states along the S_1 MEP. A representation of the main photochemical processes occurring is schematically shown in Figure 4.

Along the relaxation pathway on S_1 , the first event is the occurrence of a near degeneracy between S_1 and T_N , close in the FC region, leading to a singlet–triplet crossing denoted by (S_1/T_N)_{STC}. In the vicinity of such a crossing the ISC mechanism is favorable in the two required aspects, i.e., the close energetic proximity of both states and the relatively large SOCs. As can

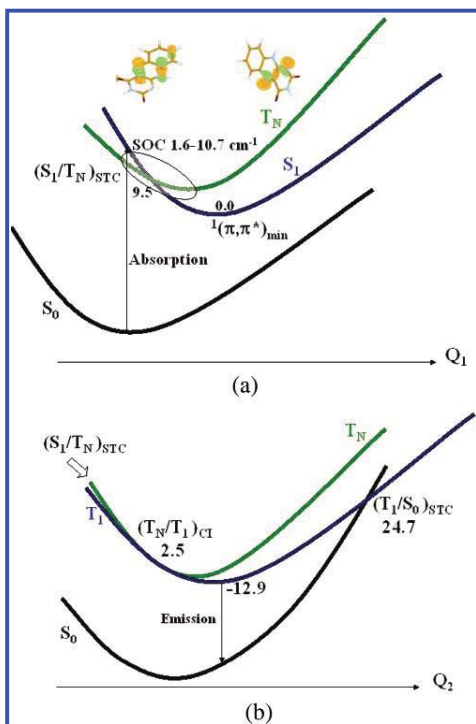


Figure 4. Schematic representation of the photochemical events after light absorption once isoalloxazine has been promoted vertically to S_1 . (a) Population of T_N in the vicinity of Franck-Condon region by intersystem crossing mechanism. (b) Internal conversion toward T_1 mediated by the $(T_N/T_1)_{CI}$ leading ultimately to the $^3(\pi,\pi^*)_{min}$ structure. Energies are in $\text{kcal}\cdot\text{mol}^{-1}$. Q_1 and Q_2 are reaction coordinates.

be seen in Figure 4a, the computed electronic SOC around this region ranges from 2 to 11 cm^{-1} , reflecting the spin-orbit allowed character of the $^1(\pi,\pi^*)/{}^3(n,\pi^*)$ transition as enunciated by the El-Sayed selection rules for intersystem crossing.⁴⁵ Further along the path the equilibrium structure of the lowest singlet excited state, $^1(\pi,\pi^*)_{min}$, is ultimately reached. It is worth mentioning that T_1 is placed below S_1 all the way along the S_1 relaxation path. Thus, one can envisage that upon absorption of a photon, the isoalloxazine ring is promoted to its lowest singlet excited state, S_1 , which partially decays in times shorter than the experimental time resolution to an intermediate related to T_N . The population transfer occurs before the S_1 minimum is encountered, which is consistent with the fact that the fluorescence quantum yields of the flavin compounds are lower than the intersystem crossing quantum yield and triplet state formation.⁴⁶ Polar solvents seem to increase ISC and decrease fluorescence quantum yields in isoalloxazines.^{15,42} According to their nature, $S_1(\pi,\pi^*)$ and $T_N(n,\pi^*)$ would be in polar media stabilized and destabilized, respectively, with respect to the in vacuo values, leading to a decrease in the vertical energy gap and a singlet-triplet crossing, which might occur closer to the FC region, favoring the trends observed experimentally.

As an inset of Figure 4a, the structures of isoalloxazine showing the π and N_5 lone-pair orbitals involved in the spin flip are shown. A comparative analysis of the Mulliken

population per center at the geometries around the FC zone, with respect to the $^1(\pi,\pi^*)_{min}$ and $^3(n,\pi^*)_{min}$ structures, allows us to conclude that N_5 is the center directly implied in the spin flip. The fact that the carbon atom C_{4a} has a minor relevance in this issue is somewhat counterintuitive since the molecule distortion in the T_N state with respect to S_1 occurs in the axis involving the $N_1-C_{10a}-C_{4a}-N_5$ atoms and comprises the shortening of the N_1-C_{10a} and $C_{4a}-N_5$ bond distances, as well as an increase of the $C_{10a}-C_{4a}$ bond length.

Starting from $(S_1/T_N)_{STC}$, evolution of the $^3(n,\pi^*)$ state has been studied by following a similar computational strategy. The CASSCF MEP along the $^3(n,\pi^*)$ state was first calculated and, at the corresponding T_N -MEP coordinates, the singlet and triplet states were computed at the CASPT2 level. The results from this part of the research are summarized in Figure 4b. Once T_N is populated, a rapid T_N to T_1 internal conversion takes place, mediated by the $(T_N/T_1)_{CI}$ conical intersection, which is computed to be at 2.5 kcal/mol (0.11 eV) above the $^1(\pi,\pi^*)_{min}$, the reference structure chosen for the relative energies depicted in Figure 4. The T_N MEP leads ultimately to the equilibrium structure of the T_1 state, $^3(\pi,\pi^*)_{min}$, placed 12.9 kcal/mol (0.56 eV) below $^1(\pi,\pi^*)_{min}$. From relaxed T_1 , three different photochemical events could take place: emission (phosphorescence), reactivity of the long-lived T_1 isoalloxazine with different protein residues, and radiationless decay to the ground state through a $(T_1/S_0)_{STC}$. In order to explore the latter possibility, the $(T_1/S_0)_{STC}$ structure was actually computed. With the system located initially at the $^3(\pi,\pi^*)_{min}$ point of the hypersurface, the energy barrier to overcome in order to access the $(T_1/S_0)_{STC}$ is 37.6 kcal/mol at the CASPT2 level. Therefore, to deactivate T_1 isoalloxazine to the ground state nonradiatively via the $(T_1/S_0)_{STC}$ crossing ultimately depends on the excess of vibrational energy in the system at this point; otherwise the molecule will just emit. As compiled in Table 2, the radiative lifetime for T_1 is on the order of 116 ms, consistent with the nature of the state. On the other hand, it is known that, after T_1 formation, the flavin core ring, the chromophore of phototropin, reacts with the sulfur atom of the cysteine-39 to form an adduct. To get further insight into that process, and as a natural continuation of the present study, an ongoing project is addressed toward modeling the reactivity of isoalloxazine in T_1 with sulfur-containing compounds.

4. Summary and Conclusions

The electronic excited states of isoalloxazine were studied using high-level ab initio quantum-chemical methods based on multiconfigurational wave functions. In particular, the CASPT2 multiconfigurational second-order perturbation theory, in conjunction with equilibrium geometries determined at the CASSCF level, i.e., the CASPT2//CASSCF protocol, was employed. The vertical excitation energies were computed at the CASPT2 level using the CASSCF optimized geometry for the ground state. The study comprised excited states in the singlet manifold up to 5.5 eV. Eleven singlet and five triplet vertical excited states were characterized. The overall picture of the obtained spectra as regards the nature of the states, polarization directions of the transitions, and relative intensities of the main absorption bands shows satisfactory agreement with the available experimental data. The main vertical features of the absorption spectrum computed in vacuo at 3.09, 4.28, 4.69, 5.00, and 5.37 eV are, however, somewhat blue shifted with respect to the experimental data in solution. To estimate the band origin and emission maximum, geometry optimizations for the lowest singlet and triplet excited states were also carried at the CASSCF level. The lowest vertical singlet-singlet (at 3.09 eV) and

singlet–triplet (at 2.52 eV) transitions are of $\pi\pi^*$ character. The corresponding band origins are placed at 2.69 and 2.03 eV, respectively. We have analyzed in detail the possibility of populating the lowest triplet state of isoalloxazine along the relaxation path of the lowest singlet state. For this purpose, spin–orbit couplings and potential energy hypersurfaces for the relevant excited states were computed. The theoretical analysis reveals that the intrinsic mechanism for photoinduced population of T_1 has two steps. First, in a wide region close to the vertical transition of S_1 , a singlet–triplet crossing by an intersystem crossing mechanism involving a triplet state of $n\pi^*$ type is favored. Second, from this point of the hypersurface, further along the relaxation pathway of the triplet $n\pi^*$ state, a conical intersection implying the two lowest triplet states is found, leading directly to relaxed T_1 , mainly responsible for the inherent reactivity of the flavin core ring.

We can finally conclude that triplet-state formation, responsible for triggering the photochemical cycle of flavin-related compounds, is revealed from the present results in vacuo as an intrinsic property of the isoalloxazine molecule. Although the protein environment does indeed increase the intersystem crossing rate⁴⁶ and is clearly necessary for adduct formation, the chromophore itself becomes the main actor in the initial part of the photocycle, showing an implicit ability to populate efficiently the lowest triplet state, a property that has been most probably evolved as the optimal biochemical-function response of flavoproteins to sunlight and environmental exposure.

Acknowledgment. The research reported has been supported by the Spanish MEC-FEDER, Project No. CTQ2004-01739, and the Generalitat Valenciana, Project No. GV06-192.

Supporting Information Available: Cartesian coordinates of the most relevant structures employed, calculated CASPT2 MEPs along the $S_1(\pi\pi^*)$, $T_N(n\pi\pi^*)$, and $T_1(\pi\pi^*)$ hypersurfaces, shapes of the HOMO and LUMO natural orbitals, and complete ref 39. This material is available free of charge via the Internet at <http://pubs.acs.org>.

References and Notes

- Voet, D.; Voet, J. *Biochemistry*; John Wiley and Sons, Inc.: New York, 1990.
- Briggs, W. R.; Beck, C. F.; Cashmore, A. R.; Christie, J. M.; Hughes, J.; Jarillo, J. A.; Kagawa, T.; Kanegae, H.; Liscum, E.; Nagatanni, A.; Okada, K.; Salomon, M.; Rüdiger, R.; Sakai, T.; Zakano, M.; Wada, M.; Watson, J. C. *Plant Cell* **2001**, *13*, 993.
- Dürr, H.; Bouas-Laurent, H. *Photochromism Molecules and Systems*; Elsevier Science: New York, 2003.
- Christie, J. M.; Salomon, M.; Nozue, K.; Wada, M.; Briggs, W. R. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 8779.
- Swartz, T. E.; Corchnoy, S. B.; Christie, J. M.; Lewis, J. W.; Szundi, I.; Briggs, W. R.; Bogomolni, R. A. *J. Biol. Chem.* **2001**, *276*, 36493.
- Neiss, C.; Saalfrank, P. *Photochem. Photobiol.* **2003**, *77*, 101.
- Kennis, J. T. M.; Crosson, S.; Gauden, M.; Stokkum, I. H. M.; Moffat, K.; van Grondelle, R. *Biochemistry* **2003**, *42*, 3385.
- Kottke, T.; Heberle, J.; Hehn, D.; Dick, B.; Hegemann, P. *Biophys. J.* **2003**, *84*, 1192.
- Scheleicher, E.; Kowalczyk, R. M.; Kay, C. W. M.; Hegemann, P.; Bacher, A.; Fischer, M.; Bittl, R.; Richter, G.; Weber, S. *J. Am. Chem. Soc.* **2004**, *126*, 11067.
- Crovetto, L.; Braslavsky, E. *J. Phys. Chem. A* **2006**, *110*, 7307 and references cited therein.
- Kowalczyk, R. M.; Schleicher, E.; Bittl, R.; Weber, S. *J. Am. Chem. Soc.* **2004**, *126*, 11393.
- Song, P.-S. *J. Phys. Chem.* **1968**, *72*, 536.
- Wouters, J.; Durant, F.; Champagne, R.; André, J. M. *Int. J. Quantum Chem.* **1997**, *64*, 721.
- Sikorska, E.; Khmelinskii, I. V.; Koput, J.; Bourdelande, J. L.; Sikorski, M. *J. Mol. Struct.* **2004**, *697*, 137.
- Sun, M.; Moore, T. A.; Song, P.-S. *J. Am. Chem. Soc.* **1972**, *94*, 5.
- Johansson, L. B.-A.; Davidsson, A.; Lindblom, G.; Naqvi, K. R. *Biochemistry* **1979**, *18*, 4249.
- Andersson, K.; Malmqvist, P.-Å.; Roos, B. O. *J. Chem. Phys.* **1992**, *96*, 1218.
- Roos, B. O.; Fülischer, M. P.; Malmqvist, P.-Å.; Merchán, M.; Serrano-Andrés, L. *Theoretical Studies of Electronic Spectra of Organic Molecules*; Langhoff, S. R., Ed.; Kluwer Academic Publishers: Dordrecht, The Netherlands, 1995; p 357.
- Roos, B. O.; Andersson, K.; Fülischer, M. P.; Malmqvist, P.-Å.; Serrano-Andrés, L.; Pierloot, K.; Merchán, M. *Adv. Chem. Phys.* **1996**, *93*, 219.
- Merchán, M.; Serrano-Andrés, L.; Fülischer, M. P.; Roos, B. O. *Multiconfigurational Perturbation Theory Applied to Excited States of Organic Compounds*; Hirao, K., Ed.; World Scientific Publishing Company: Amsterdam, 1999; Vol. 4, p 161.
- Merchán, M.; Serrano-Andrés, L. In *Computational Chemistry*; Olivucci, M., Ed.; Elsevier: Amsterdam, 2005.
- Serrano-Andrés, L.; Merchán, M. In *Encyclopedia of Computational Chemistry*; Schleyer, P. v. R.; Schreiner, P. R.; Schaefer, H. F., III, Jorgensen, W. L.; Thiel, W.; Glen, R. C., Eds.; Wiley: Chichester, U.K. 2004.
- (a) Serrano-Andrés, L.; Merchán, M.; Nebot-Gil, I.; Roos, B. O.; Fülischer, M. *J. Am. Chem. Soc.* **1993**, *115*, 6184. (b) Roos, B. O.; Malmqvist, P.-Å.; Molina, V.; Serrano-Andrés, L.; Merchán, M. *J. Chem. Phys.* **2002**, *116*, 7526. (c) Serrano-Andrés, L.; Fülischer, M. P.; Roos, B. O.; Merchán, M. *J. Chem. Phys.* **1996**, *100*, 6484.
- (a) Merchán, M.; Roos, B. O. *Theor. Chim. Acta* **1995**, *92*, 227. (b) Merchán, M.; Roos, B. O.; McDiarmid, R.; Xing, X. *J. Chem. Phys.* **1996**, *104*, 1791.
- (a) Pou-Américo, R.; Merchán, M.; Ortí, E. *J. Chem. Phys.* **1999**, *110*, 9536. (b) Pou-Américo, R.; Serrano-Andrés, L.; Merchán, M.; Ortí, E.; Forsberg, N. *J. Am. Chem. Soc.* **2000**, *122*, 6067.
- Climent, T.; González-Luque, R.; Merchán, M. *J. Phys. Chem. A* **2003**, *107*, 6995.
- González-Luque, R.; Garavelli, M.; Bernardi, F.; Merchán, M.; Robb, M. A.; Olivucci, M. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 9379.
- Molina, V.; Merchán, M. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 4299.
- Merchán, M.; Serrano-Andrés, L. *J. Am. Chem. Soc.* **2003**, *125*, 8108.
- Serrano-Andrés, L.; Merchán, M.; Borin, A. C. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 8691.
- For reviews of the CASSCF methods, see different contributions in: *Ab Initio Methods in Quantum Chemistry—II*; Lawley, K. P., Ed.; Wiley: New York, 1987.
- Forsberg, N.; Malmqvist, P.-Å. *Chem. Phys. Lett.* **1997**, *274*, 196.
- Malmqvist, P.-Å. *Int. J. Quantum Chem.* **1986**, *30*, 479.
- Malmqvist, P.-Å.; Roos, B. O. *Chem. Phys. Lett.* **1989**, *155*, 189.
- Strickler, S. J.; Berg, R. A. *J. Chem. Phys.* **1962**, *37*, 814.
- Rubio-Pons, O.; Serrano-Andrés, L.; Merchán, M. *J. Phys. Chem. A* **2001**, *105*, 9664.
- De Vico, L.; Olivucci, M.; Lindh, R. *J. Chem. Theory Comput.* **2005**, *1*, 1029.
- Merchán, M.; Serrano-Andrés, L.; Robb, M. A.; Blancafort, L. *J. Am. Chem. Soc.* **2005**, *127*, 1820.
- Andersson, K.; et al. *MOLCAS*, version 6.0; Department of Theoretical Chemistry, Chemical Centre, University of Lund: Lund, Sweden, 2004.
- Veryazov, V.; Widmark, P.-O.; Serrano-Andrés, L.; Lindh, R.; Roos, B. O. *Int. J. Quantum Chem.* **2004**, *100*, 626.
- Wang, M.; Fritchie, J., Jr. *Acta Crystallogr.* **1973**, *B29*, 2040.
- Visser, A. J. W. G.; Müller, F. *Helv. Chim. Acta* **1979**, *62*, 593.
- Serrano-Andrés, L.; Forsberg, N.; Malmqvist, P.-Å. *J. Chem. Phys.* **1998**, *108*, 7202.
- Eaton, W. A.; Hofrichter, J.; McKinen, M. W.; Andersen, D. R.; Ludwig, M. L. *Biochemistry* **1975**, *14*, 2146.
- Turro, N. J. *Modern Molecular Photochemistry*; University Science Books: Sausalito, CA, 1991.
- Shüttrigkeit, T. A.; Kompa, C. K.; Salomon, M.; Rüdiger, W.; Michel-Beyerle, M. E. *Chem. Phys.* **2003**, *294*, 501.

Publicación 4

On the intrinsic population of the lowest triplet state of uracil

Climent, T.; González-Luque, R.; Merchán, M.; Serrano-Andrés, L.

Chem. Phys. Lett. **2007**, 441, 327.



On the intrinsic population of the lowest triplet state of uracil

Teresa Climent, Remedios González-Luque, Manuela Merchán, Luis Serrano-Andrés *

Instituto de Ciencia Molecular, Universitat de València, Apartado 22085, ES-46071 Valencia, Spain

Received 23 February 2007; in final form 3 April 2007

Available online 17 May 2007

Abstract

From CASPT2//CASSCF quantum-chemical computations it is determined that the lowest triplet state of uracil can be efficiently populated from the initially activated singlet manifold through respective singlet–triplet crossings of the singlet state with the low-lying $^3\pi\pi^*$ state at 4.6 eV and with the lowest $^3\pi\pi^*$ state at 4.2 eV located along the minimum energy path of the low-lying $^1\pi\pi^*$ state. Large spin–orbit coupling elements predict, in particular for the former case, efficient intersystem crossing processes. The wavelength dependence measured for the triplet quantum yield can be explained by the location of the singlet–triplet crossing regions.

© 2007 Elsevier B.V. All rights reserved.

1. Introduction

We have recently presented a unified model for the rapid internal conversion of the singlet excited pyrimidine (Pyr) nucleobases uracil (U), thymine (T), and cytosine (C) [1], where the ultrafast decay component observed in these systems, both in the gas phase and in solution, can be interpreted in terms of the barrierless character of the minimum energy path (MEP) associated to the lowest singlet state of $\pi\pi^*$ character toward a conical intersection (CI) with the ground state, as it has also been reported for adenine [2–4]. The description of the conical intersection for uracil has been also performed at different levels of calculation [5–8]. The involved state can be described mainly by the one-electron promotion from the highest occupied molecular orbital (HOMO) to the lowest unoccupied molecular orbital (LUMO), hereafter denoted as $^1(\pi\pi^* \text{ HL})$. Despite the short lifetimes associated to this state, which is the main contributor to the photophysics of the system, formation of photodimers $\text{Pyr} \rightleftharpoons \text{Pyr}$ has been observed for the monomers in solution, as well as in solid state, for oligonucleotides, and DNA [9]. Since the 60s, the determination of the mechanism of the photoinduced formation of cyclobutane dimers has been the subject of numerous studies [9–14].

One of the most classic models that has been proposed for the photodimerization of Pyr nucleobases in solution invokes photoexcitation of a molecule to a singlet state followed by population of a triplet state by an intersystem crossing mechanism (ISC) and subsequent reaction of the triplet Pyr with a second molecule in the ground state [9,13–15]. In particular, the presence of the triplet state of uracil in solution has been observed by using different experimental techniques, such as flash photolysis [16], and indirect evidence for the involvement of a triplet state has also been provided by using specific triplet quenchers, as well as external photosensitizers [9,10,16–18]. An interesting feature concerning the intersystem crossing quantum yield ϕ_{ISC} is the fact that its value is strongly dependent on the excitation wavelength in the lowest-energy absorption band. In the case of uracil, the value increases from 1.4×10^{-3} at 280 nm (4.43 eV) to 1.6×10^{-2} at 230 nm (5.39 eV) [9,10]. Inspired by those findings, the purpose of this Letter is to analyze on theoretical grounds whether they can be rationalized on the basis of the singlet–triplet crossings (STCs) occurring along the major deactivation path responsible of the rapid internal process detected. Thus, we address here two main questions: Can the intrinsic population of the lowest triplet state in uracil take place in vacuo? And, what is the molecular basis for the wavelength dependence on ϕ_{ISC} ? For this purpose, high level ab initio methods are applied within the framework of the current theoretical paradigm

* Corresponding author. Fax: +34 963544427.

E-mail address: Luis.Serrano@uv.es (L. Serrano-Andrés).

for nonadiabatic photochemistry [19]. Since the calculations have been performed in vacuo, without the explicit consideration of solvent effects, the answer here provided can be regarded as a characteristic molecular property of the uracil system, which might be expected to be somewhat disturbed by the specific environment exposed to in solution, in solid, or in vivo.

2. Methods and computational details

The stationary points computed (minima, MEP points, CIs, and STCs) were obtained initially at the CASSCF level, except for the minimum of the $^1(\pi\pi^*)$ state, which was obtained at the CASPT2 level (see Supplementary data). For the geometry optimization of the ground state and the $\pi\pi^*$ states the active space comprised the full π system, ten electrons distributed in eight active MOs, hereafter denoted by π -CASSCF(10/8). Geometry determinations of the $^{1,3}(n\pi^*)$ states, where n_O labels generically a lone-pair MO located on oxygen atom O_4 of uracil (see Fig. 1), were carried out by using a CASSCF(12/9) wave function, that is, including the corresponding lone-pair MO and electrons in the previous active space. The points forming the MEP were optimized using a CASSCF(10/11) active space, which adds three extra correlating orbitals to the π active space. For the final results at the obtained geometries CASPT2(14/10) point calculations were used. The basis set 6-31G(d,p) has been employed throughout, with control calculations to check the stability of the results with the enlargement of the basis sets. The methodology employed, CASPT2//CASSCF has repeatedly proved its accuracy in excited state chemistry [1–4,20–24]. More detailed technical aspects of the calculations can be found in our previous papers [1–3,22,25] and in the Supplementary data. All the calculations reported used the quantum-chemical methods implemented in the MOLCAS 6.0 package [26].

3. Results and discussion

As stated above the main goal of this contribution is to explore the mechanism for efficient population of the lowest triplet state of uracil, a state protagonist of a rich photochemistry.

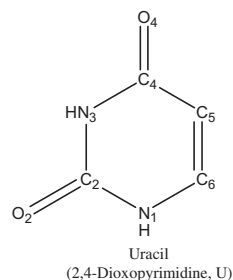


Fig. 1. Uracyl structure and labeling.

It is undoubtedly that upon absorption most of the population reaches initially singlet excited states. In particular, as listed in Table 1, at the Franck–Condon (FC) region it is the transition to the S_1 $^1(\pi\pi^*)$ HL excited state at 5.18 eV which has the largest oscillator strength up to 6 eV, 0.1955. As it was reported previously [1], and it is shown also here in Fig. 2, the MEP computed from FC on the S_1 $^1(\pi\pi^*)$ HL hypersurface leads in a barrierless way toward a conical intersection with the ground state ($gs/{}^1\pi\pi^*$)_{CI}. Unlike simple geometry optimizations, the use of the MEP technique guarantees the absence of energy barriers along the lowest-energy path. As displayed in Fig. 2, the structure of the CI at the end of the MEP can be characterized as ethene-like, involving combined stretching and twisting of the ethylenic bond [1]. Most of the population on S_1 will decay nonradiatively along this path toward the ground state, explaining the ultrafast deactivation reported for uracil in the fs range [27,28]. As described earlier [1] the low quantum yield ($\sim\phi_F = 10^{-4}$) fluorescence detected for uracil with band origin at 4.6 eV in water [8] can be better related to the presence of a high-lying planar $^1(\pi\pi^*)$ HL minimum computed adiabatically from the ground state at 4.48 eV, whereas a nonfluorescent $^1(n_O\pi^*)$ minimum is found at 4.03 eV (see Table 1) with a minor contribution to the emissive properties. Similar vertical and adiabatic energy values have been found at other levels of theory [5].

The efficient population of the triplet manifold should take place essentially along the main decay process on S_1 $^1(\pi\pi^*)$ HL. As can be seen in Fig. 2, along the $^1(\pi\pi^*)$ HL state MEP, and apart from the occurrence of a CI with the weakly interacting $^1(n_O\pi^*)$ singlet state, it is clear that two singlet–triplet crossings (STC) take place: first, with the $^3(n_O\pi^*)$ triplet state and, further along the relaxation path, with the $^3(\pi\pi^*)$ T_1 triplet state, in the vicinity of the CI with the ground state. Those two STC will be responsible of the depopulation of S_1 $^1(\pi\pi^*)$ HL toward the low-lying $^3(\pi\pi^*)$ T_1 triplet state. Fig. 3 displays a scheme describing the population of T_1 based on our CASPT2 calculations and the two mechanisms proposed to explain the overall intersystem crossing process. Near the third point of the $^1(\pi\pi^*)$ HL MEP there is a STC ($^3(n_O\pi^*)/{}^1\pi\pi^*$)_{STC} with the lowest

Table 1
Computed spectroscopic properties for the low-lying singlet and triplet excited states of uracil at the CASPT2//CASSCF(14/10)/6-31G(d,p) level

State	Vertical transition (eV)		Band origin (T_e , eV)		τ_{rad}
	CASSCF	CASPT2 ^a	CASSCF	CASPT2	
$^1(n_O\pi^*)$	5.18	4.93 (0.0006)	4.07	4.03	3051 ns
$^1(\pi\pi^*)$	6.82	5.18 (0.1955)	6.30 ^b	4.48 ^b	7 ns
$^1(\pi\pi^*)$	7.29	6.18 (0.0733)			
$^3(\pi\pi^*)$	3.98	3.80	3.16	3.15	135 ms
$^3(n_O\pi^*)$	4.87	4.71	3.81	3.91	
$^3(\pi\pi^*)$	5.76	5.33			

^a Oscillator strength within parentheses.

^b Minimum optimized at the CASPT2 level.

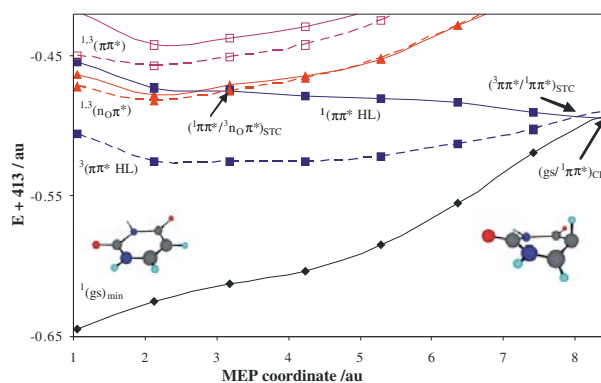


Fig. 2. Evolution of the ground and lowest singlet excited states for uracil from the FC geometry along the $^1(\pi, \pi^* \text{ HL})$ MEP computed at the CASPT2//CASSCF(14,10)/6-31G(d,p) level.

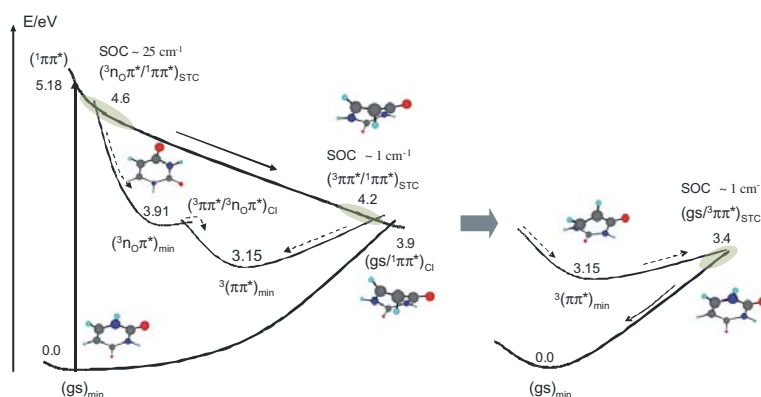


Fig. 3. Scheme, based on CASPT2 results, of the photochemistry of uracil focused on the population of the lowest-energy triplet state.

$^3(n_O \pi^*)$ triplet state, almost 0.6 eV below the vertical absorption energy to the $^1(\pi \pi^* \text{ HL})$ state at the FC region. As expected by the different $n\pi^*$ and $\pi\pi^*$ nature of the states, in the crossing region there is a large spin-orbit coupling (SOC), $\sim 25 \text{ cm}^{-1}$, between the $^3(n_O \pi^*)$ and $^1(\pi \pi^* \text{ HL})$ states, supporting an efficient population transfer from the initial singlet excited state toward the triplet state. From the third point of the $^1(\pi \pi^* \text{ HL})$ MEP (see Fig. 2) the MEP on the $^3(n_O \pi^*)$ triplet state hypersurface (see Supplementary data) leads to the minimum of the state, $^3(n_O \pi^*)_{\text{min}}$, placed at 3.91 eV adiabatically from the ground state. Just slightly above the minimum, at 0.01 eV, we have located a conical intersection connecting the triplet states $^3(n_O \pi^*)$ and $^3(\pi \pi^*)$ and favoring the population switch toward the lowest $^3(\pi \pi^*)$ state. From $(^3n_O \pi^*/^3\pi \pi^*)_{\text{Cl}}$ it is possible to follow the T_1 hypersurface and reach the $^3(\pi \pi^*)$ state minimum. As it was not possible to directly follow a MEP toward the

minimum, we performed a linear interpolation in internal coordinates (LIIC) connecting $(^3n_O \pi^*/^3\pi \pi^*)_{\text{Cl}}$ and $^3(\pi \pi^*)_{\text{min}}$ (see Supplementary data). A small barrier of near 0.02 eV (0.50 kcal/mol) was found at the initial part of the LIIC which, although probably prevents to directly follow a MEP, it is small enough to be of minor importance in the evolution of the $^3(\pi \pi^*)$ state toward its own minimum. Similar mechanisms for the population of the lowest $^3(\pi \pi^*)$ triplet state have been described previously in other biological chromophores such as isoalloxazine [29] and psoralen [30].

The mechanism described above is not the only possibility to populate the lowest T_3 $^1(\pi \pi^*)$ state from the initially excited $^1(\pi \pi^* \text{ HL})$ state. As displayed in Figs. 2 and 3, close to the ethene-like conical intersection between the $^1(\pi \pi^* \text{ HL})$ and the ground state, $(\text{gs}/^1\pi \pi^*)_{\text{Cl}}$, there is a STC connecting the lowest singlet and triplet $\pi \pi^*$ excited states, in particular near 4.2 eV adiabatically from the ground state

minimum. The same distortions that lead the molecule from the FC region to $(gs/{}^1\pi\pi^*)_{CI}$ along the MEP, stretching and twisting of the double bond $C_5=C_6$, are found in the structure of the STC (${}^3\pi\pi^*/{}^1\pi\pi^*$)_{STC}. Unlike in the STC involving the $n\pi^*$ state, the electronic SOC computed in this case is low, 1 cm^{-1} , and from this viewpoint the intersystem crossing could be expected somewhat less efficient. Both mechanisms can however contribute to the overall population of the lowest triplet state, which, as mentioned in the introduction, occurs with low intersystem crossing yields, undoubtedly due to the ultrafast character of the energy dissipation along the singlet manifold [1,27,28]. 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 [8]. Despite those effects, both singlet and triplet $n\pi\pi^*$ -type states are estimated to lie in the solvent below the ${}^1(\pi\pi^*$ HL) state at the FC geometry [8], guaranteeing the existence of the STC crossing upon decay along the ${}^1(\pi\pi^*$ HL) state. The reported wavelength dependence on the intersystem crossing quantum yield in uracil can be understood by the location of the STC crossing regions. At higher excitation energies like 5.39 eV the measured ϕ_{ISC} [9,10], 1.6×10^{-2} , is larger than at 4.43 eV, 1.4×10^{-3} , probably because in the former case the two described ISC mechanisms, with STC crossing regions near 4.6 ($n\pi\pi^*$ -driven) and 4.2 eV ($\pi\pi^*$ -driven), can be active, whereas in the latter case the ${}^1(\pi\pi^*$ HL) state is populated below the higher STC crossing region, and only the $\pi\pi^*$ -driven ISC can contribute to the overall process.

In either case that the lowest triplet state may be populated, and from both the conical intersection with the ${}^3(n\pi\pi^*)$ state or via the singlet–triplet crossing with the ${}^1(\pi\pi^*$ HL) state (see Fig. 3), the MEP on ${}^3(\pi\pi^*)$ leads to the state minimum, ${}^3(\pi\pi^*)_{min}$, in which the molecule displays a distorted structure with a ring deformation including the dihedral angle $C_2N_1C_6C_5$ as 45° and an increased bond length C_5C_6 of 1.503 Å. The high reactivity attributed to this triplet state originates from its biradical character on C_5 and C_6 . The minimum is placed at 3.15 eV adiabatically (see Table 1) from the ground state optimized minimum, a value that can be compared with the 3.3 eV estimated for the location of the triplet state for uracil mononucleotide in aqueous solution and room temperature [18] and it is consistent with previous theoretical determinations at around 3.2 [5] and 3.1 eV [31]. As a final aspect of the evolution along the triplet manifold in uracil 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 ${}^3(\pi\pi^*)_{min}$ (see Supplementary data). The crossing is placed at near 3.4 eV from the ground state minimum, which means that there is a barrier of 0.21 eV (4.7 kcal/mol) to reach $(gs/{}^3\pi\pi^*)_{STC}$ from ${}^3(\pi\pi^*)_{min}$. The molecule recovers at the STC the planarity, and the computed electronic SOC is somewhat low, $\sim 1\text{ cm}^{-1}$, predicting for the triplet state a long lifetime and a slow relaxation, becoming therefore prone to react.

4. Summary and conclusions

CASPT2//CASSCF quantum-chemical calculations have been performed on the low-lying electronic singlet and triplet states hypersurfaces of the uracil molecule in order to determine the most efficient pathways for population of the reactive lowest triplet state. By computing minimum energy paths and locating singlet–triplet crossings, conical intersections, and state minima we have identified two different mechanisms in which the population can be transferred from the initially populated ${}^1(\pi\pi^*$ HL) state to the triplet manifold through respective singlet–triplet crossings (STC): At 4.6 eV there is the STC (${}^3n\pi\pi^*/{}^1\pi\pi^*$)_{STC}, where the spin–orbit coupling (SOC) has been computed high, $\sim 25\text{ cm}^{-1}$, and at 4.2 eV the (${}^3\pi\pi^*/{}^1\pi\pi^*$)_{STC}, where the SOC is obtained much lower, $\sim 1\text{ cm}^{-1}$. In the former case the population has to switch from the ${}^3(n\pi\pi^*)$ state to the ${}^3(\pi\pi^*)$ state through a conical intersection (${}^3n\pi\pi^*/{}^3\pi\pi^*$)_{CI}. The observed wavelength dependence of the intersystem crossing quantum yield in uracil, one order of magnitude larger at excitation energies near 5.4 eV than that near 4.4 eV, is here interpreted by the contribution of both or just the lowest-energy ISC mechanisms, respectively.

Acknowledgments

The research has been supported by projects CTQ2004-01739 of the Spanish MEC/FEDER and GV06-192 of the *Generalitat Valenciana*.

Appendix A. Supplementary data

Technical details of the calculation of MEP points, conical intersections, singlet–triplet crossings, spin–orbit coupling elements, and radiative lifetimes. Figures describing significant minimum energy paths and linear interpolations on the excited states hypersurface. Cartesian coordinates of the optimized geometries. Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.cplett.2007.05.040.

References

- [1] M. Merchán, R. González-Luque, T. Climent, L. Serrano-Andrés, E. Rodríguez, M. Reguero, D. Peláez, *J. Phys. Chem. B* 110 (2006) 26471.
- [2] L. Serrano-Andrés, M. Merchán, A.C. Borin, *Proc. Nat. Acad. Sci. USA* 103 (2006) 8691.
- [3] L. Serrano-Andrés, M. Merchán, A.C. Borin, *Chem. Eur. J.* 12 (2006) 6559.
- [4] L. Blancafort, *J. Am. Chem. Soc.* 128 (2006) 210.
- [5] C.M. Marian, F. Schneider, M. Kleinschmidt, J. Tatchen, *Eur. Phys. J. D* 20 (2002) 357.
- [6] S. Matsika, *J. Phys. Chem. A* 108 (2004) 7584.
- [7] M.Z. Zgierski, S. Patchkovskii, T. Fujiwara, E.C. Lim, *J. Phys. Chem. A* 109 (2005) 9384.
- [8] T. Gustavsson et al., *J. Am. Chem. Soc.* 128 (2006) 607.

- [9] J. Cadet, P. Vigny, in: H. Morrison (Ed.), *Bioorganic Photochemistry*, Wiley, New York, 1990.
- [10] I.H. Brown, H.E. Johns, *Photochem. Photobiol.* 8 (1968) 273.
- [11] R. Beukers, J. Ijstra, W. Berends, *Rev. Trav. Chim. Pays. Bas.* 79 (1960) 101.
- [12] V.I. Danilov, O.N. Slyusarchuk, J.L. Alderfer, J.J.P. Stewart, P.R. Callis, *Photochem. Photobiol.* 59 (1994) 125.
- [13] J. Leszczynski (Ed.), *Computational Molecular Biology*, Elsevier, Amsterdam, 1999.
- [14] L.A. Ericsson (Ed.), *Theoretical Biochemistry. Processes and Properties in Biological Systems*, Elsevier, Amsterdam, 2001.
- [15] E. Sztumpf-Kulikowska, D. Shugar, J.W. Boag, *Photochem. Photobiol.* 6 (1967) 41.
- [16] D.W. Whillans, H.E. Johns, *Photochem. Photobiol.* 6 (1969) 323.
- [17] I.G. Gut, P.D. Wood, R.W. Redmond, *J. Am. Chem. Soc.* 118 (1996) 2366.
- [18] P.D. Wood, R.W. Redmond, *J. Am. Chem. Soc.* 118 (1996) 4256.
- [19] M. Olivucci (Ed.), *Computational Photochemistry*, Elsevier, Amsterdam, 2005.
- [20] K. Andersson, P.-Å. Malmqvist, B.O. Roos, *J. Chem. Phys.* 96 (1992) 1218.
- [21] B.O. Roos, M.P. Fülcher, P.-Å. Malmqvist, L. Serrano-Andrés, K. Pierloot, M. Merchán, *Adv. Chem. Phys.* 93 (1996) 219.
- [22] M. Merchán, L. Serrano-Andrés, M.A. Robb, L. Blancafort, *J. Am. Chem. Soc.* 127 (2005) 1820.
- [23] L. Serrano-Andrés, M. Merchán, *J. Molec. Struct. Theochem.* 729 (2005) 99.
- [24] L. Serrano-Andrés, N. Forsberg, P.-Å. Malmqvist, *J. Chem. Phys.* 108 (1998) 7202.
- [25] O. Rubio-Pons, L. Serrano-Andrés, M. Merchán, *J. Phys. Chem. A* 105 (2001) 9664.
- [26] K. Andersson, et al., *MOLCAS*, version 6.0. Department of Theoretical Chemistry, Chemical Centre, University of Lund, P.O. Box 124, S-221 00 Lund, Sweden, 2004.
- [27] C.E. Crespo-Hernández, B. Cohen, P.M. Hare, B. Kohler, *Chem. Rev.* 104 (2004) 1977.
- [28] C. Canuel, M. Mons, F. Piuze, B. Tardivel, I. Dimicoli, M. Elhanine, *J. Chem. Phys.* 122 (2005) 074316.
- [29] T. Climent, R. González-Luque, M. Merchán, L. Serrano-Andrés, *J. Phys. Chem. A* 110 (2006) 13584.
- [30] J.J. Serrano-Pérez, M. Merchán, L. Serrano-Andrés, *Chem. Phys. Lett.* 434 (2007) 107.
- [31] M.T. Nguyen, R. Zhang, P.-C. Nam, A. Ceulemans, *J. Phys. Chem. A* 108 (2004) 6554.

Publicación 5

Deciphering Intrinsic Deactivation/Isomerization Routes in a Phytochrome Chromophore Model

Altoe, P.; Climent T.; De Fusco, G.; Stenta, M.; Bottoni, A.;
Serrano-Andrés, L.; Merchán, M.; Orlandi, G.; Garavelli, M.

J. Phys. Chem. B **2009**, *113*, 15067.

Deciphering Intrinsic Deactivation/Isomerization Routes in a Phytochrome Chromophore Model

Piero Altoè,[†] Teresa Climent,[‡] Giulia C. De Fusco,[§] Marco Stenta,[†] Andrea Bottoni,[†] Luis Serrano-Andrés,[‡] Manuela Merchán,^{*,‡} Giorgio Orlandi,[†] and Marco Garavelli^{*,†}

Dipartimento di Chimica "G. Ciamician", Università di Bologna, Via Selmi 2, Bologna I-40126, Italy, Instituto de Ciencia Molecular, Universitat de València, ES-46071 Valencia, Spain, and Department of Chemistry, Thomas Young Centre, Imperial College London, South Kensington Campus, London SW7 2AZ, U.K.

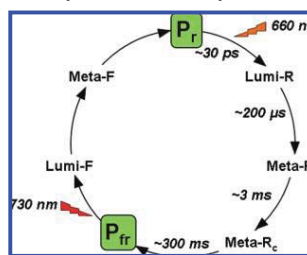
Received: April 1, 2009

High level ab initio correlated (CASPT2) computations have been used to elucidate the details of the photoinduced molecular motion and decay mechanisms of a realistic phytochrome chromophore model in vacuo and to explore the reasons underneath its photophysical/photochemical properties. Competitive deactivation routes emerge that unveil the primary photochemical event and the intrinsic photoisomerization ability of this system. The emerged in vacuo based static (i.e., nondynamical) reactivity model accounts for the formation of different excited state intermediates and suggests a qualitative rationale for the short (picosecond) excited state lifetime and ultrafast decay of the emission, its small quantum yield, and the multiexponential decay observed in both solvent and phytochromes. It is thus tentatively suggested that this is a more general deactivation scheme for photoexcited phytochrome chromophores that is independent of the surrounding environment. Spectroscopic properties have also been simulated in both isolated conditions and the protein that satisfactorily match experimental data. For this purpose, preliminary hybrid QM/MM computations at the correlated (CASPT2) level have been used in the protein and are reported here for the first time.

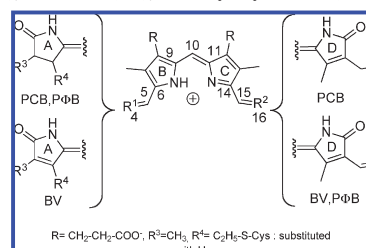
1. Introduction

Phytochromes (phys) are a large and heterogeneous family of red and far-red photoreceptors present in eukaryotes and prokaryotes that can switch between two photoconvertible forms, P_r and P_{fr} (see Scheme 1).¹ This light-induced transition initiates transcriptional signaling cascades that regulate plant growth and development and is responsible for fundamental functions that spread from the control of phototaxis and pigmentation in bacteria to seed germination and shade avoidance in plants.¹ The three chromophores found in phys (phytochromobilin (PΦB), biliverdin (BV), and phycocyanobilin (PCB)) are very similar and share a linear fully conjugated tetrapyrrolic system with differences on the first (A) and last (D) rings (see Scheme 2). Their photophysics and photochemistry have been widely investigated experimentally both in the protein and in solution.^{2–6} Changes in P_r/P_{fr} spectral properties are usually assigned to an initial $Z \rightarrow E$ photoisomerization of the C15=C16 double bond, which leads to the Lumi-R intermediate within few tens of picoseconds,^{7,8} and a subsequent isomerization of the adjacent single bond C14–C15, although the latter has been recently questioned on the basis of new evidence that either favors another rotation at the A–B bridge^{9,10} or even rules out any other isomerization but the photochemical one.^{11,12} The primary photoinduced event has been shown to follow the same complex multiexponential kinetics both in the protein and in solvent, with a fast (about 2–3 ps) and a slow (about 30–40 ps) component, suggesting a similar deactivation mechanism regardless the

SCHEME 1: Phytochromes Photocycle



SCHEME 2: Phys Chromophores: Phytochromobilin (PΦB), Biliverdin (BV), and Phycocyanobilin (PCB)^a



^a The tetrapyrrolic model system employed has been simplified by replacing R, R³, and R⁴ with H.

environment. Interestingly, it has been recently shown (via combined vis/vis and vis/IR pump/probe experiments)⁷ that only

* To whom correspondence should be addressed. For M.M.: e-mail, Manuela.Merchan@uv.es. For M.G.: e-mail, marco.garavelli@unibo.it.

[†] Università di Bologna.

[‡] Universitat de València.

[§] Imperial College London.

the slow (>30 ps) component is the photoreactive one leading to C15=C16 photoisomerization (and Lumi-R) with a rather small quantum yield (~9%) that nicely matches the one found for P_{fr} formation (~8%).¹³ Notably, it has been suggested⁷ that the isomerization occurs within the lifetime of S₁ (for which an even shorter decay component (0.7 ps) was detected), although a multistep process that is initiated in the excited state and, eventually, is finalized on S₀ cannot be ruled out. Indication of a C15=C16 Z → E photoisomerization that is complete on the excited state surface is also supported by very recent femtosecond stimulated Raman spectroscopies.¹⁴ These results suggest initial formation (within 0.6 ps) of a predistorted excited state intermediate (I*), followed by formation (in ~3 ps) of a later excited state Lumi-R-like photoproduct (Lumi-R*), where C15=C16 isomerization has been mostly finalized.

Despite the several ultrafast time-resolved spectroscopic studies, the primary events holding the key to photon energy storage and photocycle initiation remain obscure. Computational photochemistry^{15,16} would provide the way for a molecular scale understanding of those fast events, as it offers the tools for analyzing the details of the photoinduced molecular motion. Because of the complexity and size of the system, few computational studies have appeared^{17–19} and details of the photoinduced molecular motion and decay mechanism are still uncertain. In order to elucidate these processes, energy profiles for excited state relaxation paths have been investigated for a bare cationic PΦB model in vacuo via high level ab initio correlated (CASPT2) methods. Competitive deactivation routes emerge that unveil the primary photochemical event and the intrinsic photoisomerization ability of phys chromophores and suggest a qualitative rationale for the multiexponential decay and short (picosecond) excited state lifetime observed.⁶ Spectroscopic properties have been also simulated in both isolated conditions and the protein (via a QM and QM/MM approach, respectively) that satisfactorily match (within a 4 kcal/mol error) experimental data.

2. Methodology

Molecular Model Selection. To reduce computational complexity and cost, a simplified tetrapyrrolic chromophore has been selected as the PΦB model: the thioether, the methyl groups on the ring A, and the propionic carboxyl groups are replaced by hydrogen atoms (see Scheme 2). A ZZZasa chromophore geometry is assumed in this work (see Scheme 2) as suggested by several spectroscopic experiments in phys.²⁰ Anyway, it must be pointed out that the initial conformation for the chromophore in the protein is still widely under discussion, as good arguments of vibrational spectroscopy for both a ZZZasa and a ZZZssa structure have been presented.^{1,4} Furthermore, the chromophore is likely to adopt a cyclic (or helical) porphyrin-like ZZZsss geometry when free in solution.^{1,21} Despite the specific conformation selected in this work, the photoisomerization processes documented below are considered to be unaffected by the geometry adopted, as these isomerizations are all one-bond-flip events that are highly independent of the stereochemistry of other bonds, and the presented results are suggested to have a more general flavor.

Computational Methods: Selection and Validation. To select and validate the computational approach adopted for geometry optimization and energy evaluation, a full arsenal of different computational methods (such as HF, MP2, DFT, CIS, RASSCF, CASSCF, CASPT2, and RICC2) have been used and the results compared. Briefly (details of this analysis are reported in the subsections below and Supporting Info (SI)), ground (S₀)

and excited state (S₁) geometry optimizations are carried out using Gaussian03²² at the MP2 and CIS levels, respectively, whereas all energies are reevaluated at the CASPT2²³ level with MOLCAS 6.0²⁴ to account for correlation effects and to compute accurate energy profiles and spectroscopic properties. A 6-31G(d) basis set is employed throughout as a compromise between computational cost and reliability of the results; because of the size of the PΦB model employed, any larger (although more reliable) basis set would be too demanding.

Ground State Optimizations. Ground state geometry optimization has been performed according to different methods, namely, HF, MP2, DFT/B3LYP, and CASSCF. While application of HF/MP2/DFT methods is straightforward, the choice of the active space is critical for CASSCF computations. Unfortunately, in a system as large as PΦB, it is not possible to include the full π system (28 electrons in 25 π orbitals), and it must be drastically reduced. An 8 electrons/8 orbitals active π space has been initially selected and used for CASSCF(8,8) geometry optimizations. This comprises the last three π occupied HF orbitals and the first three π* virtual HF orbitals (that are all located on the central B and C rings) and the two π/π* orbitals describing the C15=C16 double bond (i.e., the biologically active one involved in the photoisomerization). Unfortunately, it is apparent that this choice is not perfectly balanced (i.e., the active space is not uniformly/symmetrically distributed on the tetrapyrrolic system): it biases C15=C16 photoisomerization, and it may create some problems for the correct geometrical description of the system during CASSCF geometry optimizations as some regions are better described than others, which may result in unrealistic asymmetries of the optimized structures. Thus, to avoid this problem, also a 12 electrons/11 orbitals active π space has been used for RASSCF(12,11) (where all single and double excitations within the active space have been considered) geometry optimization, which is perfectly balanced and uniformly distributed on the tetrapyrrolic system: four electrons in the four π/π* orbitals describing the C4=C5 and C15=C16 double bonds, plus eight electrons in the seven π/π* orbitals of the central (B_{ring}-C₁₀H-C_{ring}) part of PΦB.

While a delocalized resonance-hybrid appears as the unique stable ground state structure at the HF, MP2, and DFT/B3LYP levels (with the three methods substantially returning a similar structure where the positive charge is equally shared between the B and C rings), both CASSCF(8,8) and RASSCF(12,11) do produce a two-minima ground state topology, with the two stable geometries corresponding to the two localized resonant Lewis structures (and the positive charge that is localized either on one (B) or on the other (C) of the two central rings), regardless the active space used in MC-SCF computations (see Table S1 in SI for geometrical parameters). These results shed some doubts on the reliability of MC-SCF geometries, perhaps because of the reduced active space we are forced to use or because dynamic correlation becomes crucial (and must be accounted for) to describe delocalization effects. This is typical for systems (such as PΦB) where there is a competition between localized and delocalized resonant structures. In this case, it is well-known that inclusion of dynamic correlation normally favors the more equalized structure. For instance, ground-state free base porphyrin has a D_{2h} geometry at the MP2 level, whereas lower level computations lead to a C_{2v} structure.^{25,26}

To further validate this delocalized picture, single-point ground-state CASPT2 computations have been also performed on top of the MP2 and CASSCF optimized structures to have a homogeneous and uniform set of accurate correlated energies (see Table S2 in SI for absolute energy values). These

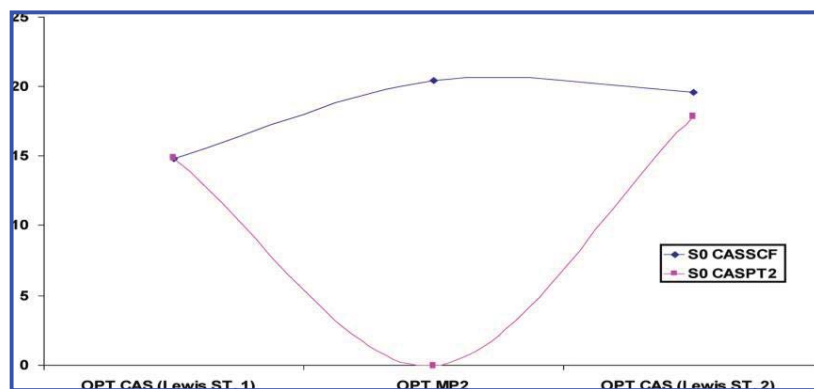


Figure 1. CASSCF and CASPT2 energy profiles (kcal/mol) connecting CASSCF and MP2 ground state minima.

calculations substantially confirm the single minimum MP2 topology (which is also apparent at the HF and DFT level): the optimized resonance-hybrid MP2 geometry appears at least 15 kcal/mol lower in energy than the two localized CASSCF minima (see Figure 1), thus clearly ruling out CASSCF ground state structures and calling for the delocalized resonance-hybrid as the real minimum in the ground state.

Excited State Optimizations. The problems experienced by CASSCF for S_0 geometry optimizations do suggest that CASSCF is perhaps not a suitable choice for excited state optimizations as well. Unfortunately, methods capable of performing excited state geometry optimization at the highly correlated level are scarce²⁷ and they certainly become technically impracticable for a system of the molecular size as the chromophore under study, especially for determination of conical intersections. As it has been documented for DNA nucleobases, upon geometry optimization the CIS method yields directly to the S_1/S_0 CI of DNA/RNA nucleobases. This has been subsequently confirmed by using higher-level methodologies.²⁸ Thus, inspired by (i) those findings, (ii) the fact that the HF approach (that is its ground state counterpart) does also appear to work properly on the ground state (HF accounts for the correct resonance-hybrid structure, delivering a geometry that is very close to the MP2 one; see discussion above), and (iii) the almost pure single (HOMO \rightarrow LUMO) excitation character of the S_1 state, the CIS approach has been selected for preliminary geometry determinations on the lowest excited state S_1 of the P Φ B model. This choice is further supported by CASPT2 calibration computations on optimized P Φ B structures, showing that S_1 CIS optimized geometries do work better (i.e., they always return a lower CASPT2 energy) than the corresponding CASSCF optimized structures. Thus, quite surprisingly and unexpectedly, it is concluded that CASSCF does not represent a reliable method for geometry optimization in this case, neither in the ground nor in the excited state, while CIS is capable of mimicking the work of high-level geometry determinations, even though it does not account for electron correlation. In summary, different structures belonging to the S_1 hypersurface have been characterized at the CIS level, just as a practical computational strategy and after a careful analysis of other alternative computational approaches, including the much more expensive CASSCF approach.

Ground and Excited State Energy Estimates. On top of the optimized ground/excited state structures, single point CASPT2

computations have been employed to account for correlation energy and compute accurate energy differences (see Tables S3 and S4 in SI for absolute energy values). The successful performance of CASPT2 in computing experimentally accurate energies (with errors usually below 4 kcal/mol) and spectroscopic properties is well-established and is the reason for using this method in evaluating the ground and excited state energies of the investigated system.

The choice of the active space is critical for CASPT2. Unfortunately, in a system as large as P Φ B, it is not possible to include the full π system, which is composed of 28 electrons in 25 orbitals, and a calibration is needed in order to select the proper reduced active space. For this purpose, the optimized ground-state MP2/6-31G(d) geometry computed for the tetrapyrrolic model designed was used. The selection of the active space has been performed by analyzing the convergence of the vertical spectroscopic properties (transition energies and related oscillator strengths) upon a systematic increasing of the number of active orbitals and active electrons (see Table S5 in SI). A 11 π/π^* orbitals, 10 electrons active space, joined to a three-roots state average procedure, leads to CASSCF/CASPT2(11,10) results that can be considered in practice to be converged within the error bars of the method (see SI for details). The standard zeroth-order Hamiltonian was employed in the CASPT2 calculations, which include an imaginary level-shift correction of 0.2 au in order to avoid the presence of intruder states. The CASSCF state interaction (CASSI) method²⁹ was used (as implemented in MOLCAS 6.0²⁴) to calculate the transition dipole moments from the reference zeroth-order CASSCF wave function. The latter moments were used, together with CASPT2 corrected energies, to compute oscillator strengths (f). These data unambiguously reveal that the single HOMO \rightarrow LUMO excitation state S_1 is the bright state, in agreement with all the other methods employed.

Finally, multireference (i.e., CASPT2) and single reference (i.e., TD-RICC2) correlated energies have been compared at all the points of the investigated channels. This is done to access the validity of a multireference perturbation vs a single reference-CC2 approach and disclose the reliability of the selected protocol for energy estimates. Because of the relevant charge transfer character along the photoisomerization paths (see the results below), TD-RICC2 appears unsuitable for the

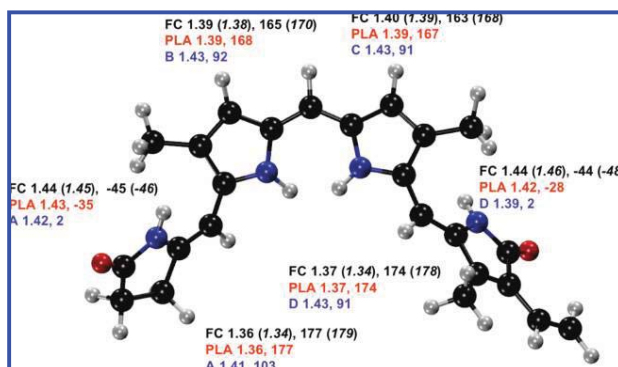


Figure 2. Relevant geometrical parameters (bonds in Å and twisting angles in degrees) computed for the MP2 and HF (numbers in italics) optimized ground state minimum (FC) and the CIS optimized excited state minima (PLA, A, B, C, D). The molecular geometry reported corresponds to the FC one.

description of the photoisomerization processes and CASPT2 turns out to be the most reliable choice (see SI for further details).

QM/MM Computations. Hybrid QM/MM computations in the protein (the resolved phys structure 1ZTU¹¹ has been used, as it was the only one available when we started this investigation) have been performed employing our recently developed COBRAMM interface³⁰ that combines AMBER8³¹ and Gaussian²² and accounts for chromophore/protein electrostatic interactions at the highest (QM) level by employing an electrostatic embedding scheme. Details of this hybrid approach are presented elsewhere. Briefly, the method is based on a hydrogen link-atom scheme.³² QM and MM layers interact in the following ways: (i) all QM atoms feel the electrostatic potential of all MM point charges (i.e., an electrostatic embedding scheme has been adopted); (ii) all the bonding (i.e., stretching, bending, and torsion) and the van der Waals QM/MM cross-terms are described by the standard MM potential; (iii) electrostatic interactions involving MM atoms are all accounted for classically.

The QM region is represented by the chromophore atoms (note that the chromophore is BV rather than PΦB in the available crystallographic structure) except its two propionic groups that are included in the MM region comprising also the crystallographic waters and the rest of the protein. The same QM approach as described above for PΦB in vacuo is used for BV, while the ff99 force field³¹ is employed for the surrounding MM region. The chromophore atoms, the two water molecules HOH12 and HOH18, and the side chain of the LYS24, which is covalently linked to the QM part, are left free to move, while all the rest is kept frozen at its crystallographic position during all calculations. Using the standard Amber package tools the protein has been solvated with a water cup centered on BV and with a 20 Å radius. The hydrogen atoms, lacking in the crystallographic structure, have been added using the H++³³ program. The protonation state of the residues close to the chromophore (HIS260, ASP207, HIS290) and the orientation of HOH12 and HOH18 have been corrected in order to preserve the hydrogen bond network previously documented.¹¹ In order to minimize spurious steric interactions between nearby protein residues and BV, residues within 10 Å around the chromophore have been preoptimized at the MM (ff99) level.

To account for correlation energy, CASPT2 computations are performed on top of the optimized points in the bath of the

protein (AMBER) point charges by adopting the same procedure described above for PΦB in vacuo.

Further details on the method and protein setup are reported in the SI.

3. Results and Discussion

As depicted in Figure 2, the ground state minimum at the Franck–Condon region (called FC thereafter) (which is slightly twisted to minimize steric interactions due to the rings substituents) appears as a stable resonance-hybrid where the positive charge is equally shared by B and C rings. Its structure is roughly symmetric with respect to the central carbon atom C10. These properties are conserved in the relaxed, roughly planar excited state structure (S_1 minimum, called PLA thereafter), where a rearrangement in the skeletal bonds occurs (a maximum difference of 0.04 Å is seen for HF \rightarrow CIS C14–C15 bond shortening; see Figure 2).

The S_1 state is mainly described by the singly excited configuration HOMO \rightarrow LUMO, with both orbitals delocalized in the two central (B, C) rings and exocyclic (C4–C5, C15–C16) bonds, leaving unaffected the external (A, D) rings (see Figure S1 in the SI). Thus, the electronic transition has no effect on the dipole moment of the molecule because the two orbitals share the same spatial region and no intramolecular charge transfer takes place at FC and PLA. The calculated vertical $S_0 \rightarrow S_1$ absorption energy is in reasonable agreement with the experimental values recorded both in solution³⁴ and in phys³⁵ (see Table 1). It is noted that the conformation of the chromophore in solution may be different (ZZZsss) from the one considered here (ZZZasa, as suggested in phys²⁰). However, since the excitation is located at the center of the chromophore that is structurally preserved in the two conformers, the effects on transition energies are expected to be minor and fall well within the error of the employed computational approach. The S_1 minimum (PLA) produced by the initial excited state dynamics out of the Franck–Condon region may be assigned reasonably to the emitting state: here, it is where the system gets funneled following absorption and stands for a while before internal vibrational energy redistribution occurs and deactivation channels leading to ground state recovery are populated. In this sense, it is worth mentioning that the $S_1 \rightarrow S_0$ vertical energy

TABLE 1: Absorption and Emission Features Computed in Vacuo and in the Protein (λ in nm)^a

abs/em	P Φ B			protein (BV)	
	CASPT2 (vacuo)	exptl		CASPT2 (QM/MM)	Exp.
λ_{abs}	642	590	666	709	700 ^d
λ_{em}	673	680–690	685	742	e

^a Experimental values are also reported. ^b Reference 34: The dihydrobiliverdin chromophore in CH_2Cl solution slightly differs from P Φ B, as it contains a methoxycarbonylmethyl substituent at C3 and an ethyl group, rather than a vinyl, at C18. This may partly account for the blue-shifted value. ^c Spectra recorded in a P Φ B containing phytochrome; see ref 35. ^d Reference 36. ^e Not detected in BV containing phys.

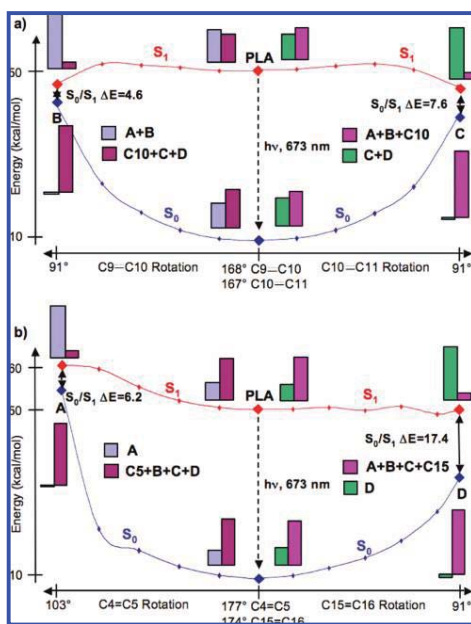


Figure 3. CIS-relaxed scans along the four torsional modes. The starting (PLA) and twisted ending points (part a, paths B and C; part b, paths A and D) are CIS fully optimized minima. Energy profiles are recalculated at the CASPT2 level (i.e., CASPT2 energy corrections are performed on top of all the points along the four paths). The amount of positive charge carried by each rotating fragment is displayed in bar diagrams (Mulliken charges as from the zeroth-order CASSCF wave function are considered) and underscores the charge transfer character of the electronic state.

gap computed at PLA matches the emission wavelength observed for the chromophore in the different environments (see Table 1).

PLA appears as a transient located on a very flat region of the S_1 potential energy surface. In particular, four possible one-bond-flip isomerization paths have been mapped to escape PLA, each one involving the twisting of one specific exocyclic bond: C4–C5 (path A), C9–C10 (path B), C10–C11 (path C), and C15–C16 (path D); see Figure 3. Three of them are competitive, almost barrierless, isomerization channels (paths B, C, and D with activation energies of 1.9, 1.5, and 0.7 kcal/mol, respec-

tively), while a fourth path (A) is unfavored (ending at a point on S_1 that is 15.7 kcal/mol higher than PLA). Interestingly, the S_1 – S_0 energy gaps at the twisted (ending) points of paths B and C are rather small (only 5 and 8 kcal/mol, respectively). On the other hand, and rather surprisingly, the S_0 – S_1 energy gap at the twisted minimum along the biologically active path D (involving the isomerization of the C15=C16 bond) is much larger (17.4 kcal/mol). Charge analysis reveals that whereas no significant intramolecular charge transfer takes place at FC and PLA for S_1 (HOMO \rightarrow LUMO) and the excitation stays mainly localized at the center of the chromophore, the positive charge progressively moves and gets localized into the shortest twisted half, and a pure twisted intramolecular charge transfer (TICT) state appears at the twisted optimized points (see Figure 3). These results clearly rule out computational approaches that do not satisfactorily account for photoinduced charge transfers (such as TD-DFT with standard functionals).³⁷ A TD-DFT/b3lyp computational study has recently appeared¹⁹ reporting results for the photoinduced molecular motion of a P Φ B model that are seemingly consistent with the present work (path D is less favored than path B/C), although for different reasons; a barrier appears along path D, while paths B/C are substantially barrierless. Additionally, CIs and efficient deactivation channels have not been located. It is doubtful whether TD-DFT/b3lyp can give reliable results along the photoisomerization paths, specifically when highly twisted structures are explored, because of the above-mentioned charge transfer character characterizing the twisted region. Thus, while such an approach may deliver some insight into the photoinduced motion at planar (or nearly planar) structures, attention must be paid not to overinterpret reaction barriers that occur along the isomerization paths at highly twisted geometries.

Considering the small CASPT2 S_1 – S_0 energy gap at the twisted minima B and C, it is suggested that these points are (or are close to) conical intersection (CI) funnels, whereas the energy gap at D seems far too large (at least at this level, vide infra) to support the existence of a CI at this point. Thus, according to the CASPT2//CIS results, a deactivation scheme emerges for isolated P Φ B with multiple competitive internal conversion (IC) routes: two photoisomerization paths (B and C) leading to ultrafast decay through directly accessed twisted CIs and a third one (D) populating a twisted excited state minimum and, possibly, a nearby CI related to a slower IC (see below). Interestingly, this isolated molecule model, displaying easily accessible twisted minima and/or CIs, can qualitatively account for most of the properties recorded in both solvents and the proteins, namely, the ultrafast (picosecond) decay of the emission, its rather small quantum yield (QY, $\sim 10^{-3}$), the short S_1 lifetime, its multiexponential decay, the faster (2–3 ps)⁶ lifetime component that is here related to paths B and C, and the slow (>30 ps) component that has been shown to be related to C15=C16 photoisomerization⁷ (i.e., path D). Since the observed multiexponential decay and transient spectroscopic features are the same regardless of the environment (i.e., solvent or protein), this reactive scheme is tentatively suggested to be a more general model for the photobehavior of phys chromophores.

It must be pointed out that the viability of the slow (but still ultrafast, i.e., tens of picoseconds long)^{7,8} component via path D depends on the existence and accessibility of a CI from the twisted minimum on S_1 , as the S_1 – S_0 energy gap documented here at the CASPT2//CIS level is inconsistent with a ~ 30 ps decay. We believe that a CI exists and is close in energy and geometry to the twisted D minimum and that the surprisingly

features are reproduced, and the agreement with the available experimental data is satisfactory. This supports the adequacy of the computational approach and model systems employed and the reliability of the drawn photoreactivity scheme, at least on a qualitative ground.

Acknowledgment. Financial support is acknowledged from MURST PRIN 2005, Projects CompRenDe (Bologna University), CTQ2007-61260, and CSD2007-0010 Consolider-Ingenio in Molecular Nanoscience of the Spanish MEC/FEDER. M.G. and P.A. thank E4-Computer Engineering Spa for granted computational time and technical assistance.

Supporting Information Available: Analysis of methods, MO, QM/MM details and protein setup, geometries, and energies. This material is available free of charge via the Internet at <http://pubs.acs.org>.

References and Notes

- (1) Rockwell, N. C.; Su, Y.-S.; Lagarias, J. C. *Annu. Rev. Plant Biol.* **2006**, *57*, 837.
- (2) Schaffner, K.; Braslavsky, S. E.; Holzwarth, A. R. Protein Environment and Photophysics and Photochemistry of the Prosthetic Chromophores of Biliproteins. In *Frontiers in Supramolecular Organic Chemistry and Photochemistry*; Schneider, H.-J., Dürr, H., Eds.; VCH Verlagsgesellschaft: Weinheim, Germany, 1991.
- (3) Holzwarth, A. R.; Venuti, E.; Braslavsky, S. E.; Schaffner, K. *Biochim. Biophys. Acta* **1992**, *1140*, 59.
- (4) Gärtner, W.; Braslavsky, S. E. The Phytochromes: Spectroscopy and Function. In *Photoreceptors and Light Signaling*; Batschauer, A., Ed.; Royal Society of Chemistry: Cambridge, U.K., 2004; Vol. 3, pp 136–180.
- (5) Falk, H. *The Chemistry of Linear Oligopyrroles and Bile Pigments*; Springer-Verlag: New York, 1989.
- (6) Rentsch, S.; Hermann, G.; Bischoff, M.; Strehlow, D.; Rentsch, M. *Photochem. Photobiol.* **1997**, *66*, 585.
- (7) Schumann, C.; Groß, R.; Michael, N.; Lamparter, T.; Diller, R. *ChemPhysChem* **2007**, *8*, 1657.
- (8) Müller, M. G.; Lindner, I.; Martin, I.; Gärtner, W.; Holzwarth, A. R. *Biophys. J.* **2008**, *94*, 4370.
- (9) Mroginiski, M. A.; Murgida, D. H.; Hildebrandt, P. *Acc. Chem. Res.* **2007**, *40*, 258.
- (10) Thor, J. J. v.; Ronayne, K. L.; Towrie, M. *J. Am. Chem. Soc.* **2007**, *129*, 126.
- (11) Wagner, J. R.; Brunzelle, J. S.; Forest, K. T.; Vierstra, R. D. *Nature* **2005**, *438*, 325.
- (12) Inomata, K.; Hammam, M. A. S.; Kinoshita, H.; Murata, Y.; Khawn, H.; Noack, S.; Michael, N.; Lamparter, T. *J. Biol. Chem.* **2005**, *280*, 24491.
- (13) Hughes, J.; Lamparter, T.; Mittmann, F.; Hartmann, E.; Gärtner, W.; Wilde, A.; Borner, T. *Nature* **1997**, *386*, 663.
- (14) Dasgupta, J.; Frontiera, R. R.; Taylor, K. C.; Lagarias, J. C.; Mathies, R. A. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 1784.
- (15) Olivucci, M. *Computational Photochemistry*; Elsevier: Amsterdam, 2005; Vol. 16.
- (16) Domcke, W.; Yarkony, D.; Köppel, H. *Conical Intersections: Electronic Structure, Dynamics & Spectroscopy*; World Scientific: Singapore, 2004.
- (17) Durbeej, B.; Eriksson, L. A. *Phys. Chem. Chem. Phys.* **2006**, *8*, 4053.
- (18) Borg, O. A.; Durbeej, B. *Phys. Chem. Chem. Phys.* **2008**, *10*, 2528.
- (19) Durbeej, B. *Phys. Chem. Chem. Phys.* **2009**, *11*, 1354.
- (20) Kneip, C.; Hildebrandt, P.; Schlamann, W.; Braslavsky, S. E.; Mark, F.; Schaffner, K. *Biochemistry* **1999**, *38*, 15185.
- (21) Knipp, B.; Müller, M.; Metzler-Nolte, N.; Balaban, T. S.; Braslavsky, S. E.; Schaffner, K. *Helv. Chim. Acta* **1998**, *81*, 881.
- (22) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03*, revision C.02; Gaussian, Inc.: Wallingford, CT, 2004.
- (23) Andersson, K.; Malmqvist, P. A.; Roos, B. O. *J. Chem. Phys.* **1992**, *96*, 1218.
- (24) Andersson, K.; Barysz, M.; Bernhardtsson, A.; Blomberg, M. R. A.; Carissan, Y.; Cooper, D. L.; Cossi, M.; Fülfscher, M. P.; Gagliardi, L.; de Graaf, C.; Hess, B.; Hagberg, G.; Karlström, G.; Lindh, R.; Malmqvist, P.-Å.; Nakajima, T.; Neogrády, P.; Olsen, J.; Raab, J.; Roos, B. O.; Ryde, U.; Schimmelpfennig, B.; Schütz, M.; Seijo, L.; Serrano-Andrés, L.; Siegbahn, P. E. M.; Stålring, J.; Thorsteinsson, T.; Vyzayov, V.; Widmark, P. O. *MOLCAS 6.0*; Department of Theoretical Chemistry, Chemical Centre, University of Lund: Lund, Sweden, 2004.
- (25) Almlöf, J.; Fischer, T. H.; Gassman, P. G.; Ghosh, A.; Häser, M. *J. Phys. Chem.* **1993**, *97*, 10964.
- (26) Merchán, M.; Orf, E.; Roos, B. O. *Chem. Phys. Lett.* **1994**, *221*, 136.
- (27) Merchán, M.; Serrano-Andrés, L. *THEOCHEM* **2005**, *729*, 99.
- (28) Merchán, M.; Gonzalez-Luque, R.; Climent, T.; Serrano-Andrés, L.; Rodríguez, E.; Reguero, M.; Pelaez, D. *J. Phys. Chem. B* **2006**, *110*, 26471.
- (29) Malmqvist, P.-Å.; Roos, B. O. *Chem. Phys. Lett.* **1989**, *155*, 189.
- (30) Altoè, P.; Stenta, M.; Bottoni, A.; Garavelli, M. *Theor. Chem. Acc.* **2007**, *118*, 219.
- (31) Case, D. A.; Cheatham, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. *J. Comput. Chem.* **2005**, *26*, 1668.
- (32) Singh, U. C.; Kollman, P. A. *J. Comput. Chem.* **1986**, *7*, 718.
- (33) Gordon, J. C.; Myers, J. B.; Folta, T.; Shoja, V.; Heath, L. S.; Onufriev, A. *Nucleic Acids Res.* **2005**, *33*, 368.
- (34) Ditto, M.; Brunner, H.; Lippitsch, M. E. *Chem. Phys. Lett.* **1991**, *185*, 61.
- (35) Büchler, R.; Hermann, G.; Lap, D. V.; Rentsch, S. *Chem. Phys. Lett.* **1995**, *233*, 514.
- (36) Wagner, J. R.; Zhang, J.; von Stetten, D.; Gunther, M.; Murgida, D. H.; Mroginiski, M. A.; Walker, J. M.; Forest, K. T.; Hildebrandt, P.; Vierstra, R. D. *J. Biol. Chem.* **2008**, *283*, 12212.
- (37) Dreuw, A.; Head-Gordon, M. *J. Am. Chem. Soc.* **2004**, *126*, 4007.
- (38) Sineshchekov, V.; Ogorodnikova, O.; Thiele, A.; Gatz, C. *J. Photochem. Photobiol., B* **2000**, *59*, 139.
- (39) Sineshchekov, V. A.; Ogorodnikova, O. B.; Devlin, P. F.; Whitelam, G. C. *J. Photochem. Photobiol., B* **1998**, *42*, 133.
- (40) Sineshchekov, V. A.; Ogorodnikova, O. B.; Weller, J. L. *J. Photochem. Photobiol., B* **1999**, *49*, 204.

JP904669X

Publicación 6

The role of pyrimidine nucleobase excimers in DNA photophysics and photoreactivity

González-Ramírez, I.; Climent, T.; Serrano-Pérez, J.; González-Luque, R.; Merchán, M.; Serrano-Andrés, L.

Pure Appl. Chem. **2009**, *81*, 1695.

The role of pyrimidine nucleobase excimers in DNA photophysics and photoreactivity*

Israel González-Ramírez, Teresa Climent, Juan José Serrano-Pérez, Remedios González-Luque, Manuela Merchán, and Luis Serrano-Andrés[‡]

Instituto de Ciencia Molecular, Universitat de València, Apartado de Correos 22085, ES-46071 Valencia, Spain

Abstract: Quantum chemical studies using the accurate CASPT2//CASSCF procedure show that π -stacked interactions in biochromophores such as pyrimidine (Pyr) DNA/RNA nucleobases pairs yield excimer-like situations which behave as precursors of processes like charge transfer (CT) or photoreactivity and are the source of the emissive properties in DNA. Examples are the CT between adjacent DNA nucleobases in a strand of oligonucleotides and the photodimerization taking place in cytosine (C) pairs leading to cyclobutanecytosine (CBC) mutants. These processes take place through nonadiabatic photochemical mechanisms whose evolution is determined by the presence and accessibility of conical intersections (CIs) and other surface crossings between different electronic states.

Keywords: bioexcimers; charge transfer; DNA lesions; DNA photochemistry; quantum chemistry.

INTRODUCTION

It is by now clearly established that π -stacked interactions between piled DNA and RNA nucleobases play a basic role in the stability, dynamics, and reactivity of the genetic material [1]. We have recently reported [2] that such types of conformations are favorable sources for the formation of excimers or exciplexes, that is, molecules which are homo- or heterodimers, respectively, associated in an excited electronic state and dissociative in the ground state [3]. The existence of excimers/exciplexes (excimer will be used as a general term hereafter) is common in different fields, including biochemistry and photobiology [1,3–5]. In particular, we have recently studied the binding energies in different states of the neutral and ionic cytosine (C) homodimer, and the existence of a bound excited state was proved by means of quantum chemical calculations in various relative conformations of the homodimer [6]. Binding and interaction change dramatically with the relative orientation of the two stacked C molecules. For instance, at the B-DNA-like type of orientations the binding energy for the low-lying excited singlet state, S_1 , is just 0.11 eV, a value that increases to 0.58 eV in the so-called ¹(LE) conformation, a sandwich-like structure obtained by symmetrically approaching the molecules along the intermonomer distance, R , and to 1.10 eV in the conformation we coined $(C^*C)_{exc}$, the most stable structure for the excimer obtained by geometry optimization [6]. The stabilization is favored in the two latter orientations because of the maximum overlap obtained between the π -clouds in the homodimer. As

*Paper based on a presentation at the XXIInd IUPAC Symposium on Photochemistry, 28 July–1 August 2008, Gothenburg, Sweden. Other presentations are published in this issue, pp. 1615–1705.

[‡]Corresponding author: E-mail: Luis.Serrano@uv.es

will be shown below, the presence of so low-lying conformation in the excited state will have important consequences in the photoinduced reactivity of the nucleobases.

Before going deeper into the consequences of the excimer formation in DNA environments, and how they behave as precursors in the formation of DNA-base photoproducts [6,7], participating in the charge-transfer (CT) processes in DNA, we shall firstly establish their existence as an intrinsic property of all nucleobases, something that will be illustrated here for the homodimers of C, uracil (U), and thymine (T). Later, the importance of excimers in the rationalization of DNA emission properties and dynamics, in the formation of cyclobutane pyrimidine (Pyr) dimers, often leading to mutations, and in the CT process taking place along the DNA strands shall be discussed.

METHODOLOGY

Our methodology involves the multiconfigurational CASSCF and CASPT2 quantum chemical methods and highly accurate ANO-S-type one-electron basis sets contracted to C,N,O [3s2p1d]/H [2s1p], a strategy which has proved its accuracy [8–11]. An active space of 12 electrons in 12 orbitals was employed in the description of the CASPT2 potential energy curves (PECs) for the ground and low-lying singlet excited states of the different homodimers along the intermolecular distance between the two π -stacked systems using a C_s ground-state geometrical arrangement and C_1 symmetry to obtain the wavefunction. The inclusion of the basis set superposition error (BSSE) correction, estimated by the counterpoise correction [2], was required to get the proper description of the system. Geometry optimizations, minimum energy paths (MEPs), and determination of hypersurface crossings in the dimers were carried out initially at the CASSCF level using 16 electrons distributed among 14 molecular orbitals (MOs), that is, all the π -system except for the all-in-phase deep π -orbital localized on each monomer, which was treated as inactive. No symmetry constraints were imposed on the wavefunction. Additional computational details can be found elsewhere [2,6,7,12]. All calculations employed the MOLCAS 6.0 program quantum chemistry package [13,14].

RESULTS AND DISCUSSION

Excimer formation in π -stacked Pyr nucleobases

Our initial goal is to prove that all three Pyr nucleobases—C, T, and U—yield excimers when stacked through their π -structure. Figures 1 and 2 display the CASPT2 PECs for the ground and low-lying and triplet singlet excited states of the U homodimer along the intermolecular distance between the two π -stacked systems. A face-to-face sandwich-type arrangement has been selected because, taking into account the maximum overlap between the π -clouds, it represents the most favorable conformation to yield an excimer, which, on the other hand, is also present in other structures, for instance, the B-DNA form. As noted in Fig. 1, the lack of BSSE causes an overestimation of the states binding energies, a situation that is corrected in Fig. 2. Whereas the ground state is dissociative, the lowest singlet excited state becomes clearly bound (0.48 eV) at the corrected BSSE-CASPT2(12,12) level, with a S_1 minimum placed at $R(C_5-C_5) = 2.910$ Å. In the PECs, we have chosen to monitor the intermolecular distance $R(C_5-C_5)$, which is particularly relevant in the formation of cyclobutane photodimers. In the figures, energies are referred to two ground-state separated U molecules. In the asymptotic limit, S_1 and S_2 become degenerate. They are related to the equivalent situations $U + U^*$ and $U^* + U$, where U and U^* represent the ground-state U and its lowest singlet excited state, respectively.

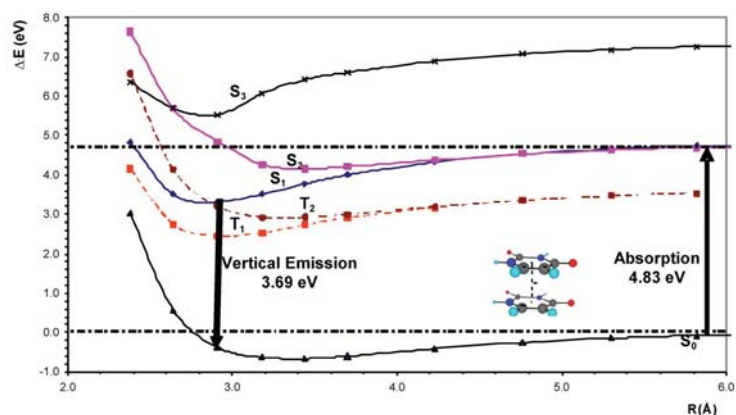


Fig. 1 CASPT2(12,12)/ANO-S C,N,O[3s2p1d]/H[2s1p] PECs built with respect to the intermolecular distance considering the center of mass of two face-to-face π -stacked U molecules.

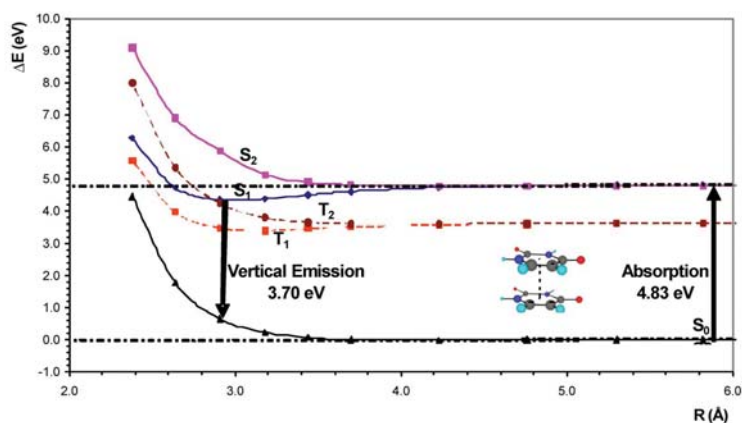


Fig. 2 Corrected BSSE-CASPT2(12,12)/ANO-S C,N,O[3s2p1d]/H[2s1p] PECs built with respect to the intermolecular distance considering the center of mass of two face-to-face π -stacked U molecules.

Similar behavior has been found for all three Pyr nucleobases. The binding energies (BSSE corrected) are compiled in Table 1.

The homodimers of the three natural Pyr nucleobases display similar behavior with respect to their binding properties, although the formation of excimers in T is hampered by the steric problems consequence of the presence of the methyl group. In particular (see Table 1) for the comparable profile as C and U, the binding energy for the T homodimer lowers to 0.36 eV and the corresponding mini-

mum takes place at a larger distance, 3.703 Å. In any case, the important conclusion obtained is that all three homodimers give rise to excimers, and, as mentioned above, the interaction occurs at many other relative orientations of the two nucleobases. Considering, therefore, the inherent flexibility of the DNA strand, there will be a large number of relative orientations in which the formation of excimers will be highly favored, even increasing the binding energy, as it was found for C (1.10 eV) upon optimization of the geometry for the S_1 state of the dimer [6].

Table 1 BSSE-corrected and -uncorrected binding energies (E_b), intermonomer distance (R_{\min}), and VE for the face-to-face π -stacked homodimers of C, T, and U.

Homodimer	E_b /eV	BSSE- E_b /eV	R_{\min} /Å	VE/eV
Cytosine	1.55	0.58	3.076	3.40
Thymine	1.05	0.36	3.703	4.38
Uracil	1.51	0.48	2.910	3.70

Red-shifted emission in DNA: Excimer radiative deactivation

Eisinger et al. [15] determined that in chains of oligonucleotides a source of emission was found at energies red-shifted from the weak features of the monomers, which were reported around 4.0 eV in water [16]. More detailed studies followed. For instance, the fluorescence maximum observed in aqueous solution for the dimer d(C)₂ and the 15-mer d(C)₁₅ ($\lambda_{\max} = 385$ nm; 3.22 eV) [17] is considerably red-shifted as compared to that of the monomer ($\lambda_{\max} = 313$ nm; 3.96 eV) [18]. Nowadays, the available experimental data for the fluorescence band maxima in dinucleotides, polynucleotides, and DNA range 3.2–3.4 eV [17].

It is, therefore, clear that π -stacked nucleobases give rise to a new source of fluorescence related to the association of the monomers. Our description of excimers in the homodimers of the Pyr nucleobases at different orientations of the moieties and the CASPT2 computed vertical emissions (VEs) supports the excimer origin of the red-shifted fluorescence observed in Pyr oligonucleotides [15,17], whereas preliminary studies indicate that excimers are also present for purine nucleobase dimers. Our best estimates have differences near 0.2 eV with respect to the recorded emission maxima, however, 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.

The evidence of the excimer origin of the emissive properties of DNA highlights the importance of π -stacking and excimer formation in the modulation of DNA relaxation dynamics, which will combine emission with nonradiative decays through accessible CIs connecting the ground and low-lying singlet excited state—such as those found for the monomers at 3.6 (C), 3.9 (T), 4.0 (U), 4.0 (adenine, A), and 4.3 eV (guanine, G) [7,11,19–21]—favoring hole and electron transfer along the strand, and formation of lesions as covalently bound Pyr dimers [1,6].

Charge transport along the DNA strand: Micro-hopping mechanism

Nucleobase cations and anions can be formed via direct ionization of the DNA strand as primary radical product, by means of the exposure of living matter to ionizing radiation and, also indirectly, as secondary radical products of the nucleobase interaction with reactive species present in DNA and its vicinity, that is, endogenous compounds that have been originated from irradiation of other DNA components or water: sugar and phosphate radicals, $\cdot\text{OH}$, the hydrated electron, H^\cdot , and H_3O^+ [22,23]. The energy required for the formation of charged nucleobases, related to the corresponding ionization potential (IP) and electron affinity (EA), has been evaluated both experimentally and theoretically. Once

the charged DNA bases are formed, the produced hole or electron can migrate up to long distances through the π -stacked structure of the double helix [24,25]. This phenomenon is known in general as charge transport or transfer (CT), and it is usually denoted hole transfer (HT) and excess electron transfer (EET) when it takes place under oxidative and reductive stress, respectively. Because of its biological implications and its technological and medical applications, CT in DNA has been during the last decades an area of increasing interest [26–29].

Whereas it is currently established by using different spectroscopic, biochemical, electrochemical, and photophysical studies that DNA provides a favorable medium for CT [30–33], the detailed mechanism on how it takes place is not yet known. The available models have been particularly focused on HT and, among the different hypotheses suggested, three main mechanisms have been put forward: (1) coherent single-step tunneling process between the initial and final charged nucleobases [34–36] (2) random-walk multistep hopping process between the initial and final sites mediated by G [37–39] and, in special cases, by A [40], (3) and polaron-like hopping process [41–44]. We have recently applied the highest available levels of theory to produce accurate theoretical (in vacuo) values for IPs and EAs [45,46], and our goal here is to summarize our preliminary efforts to establish a comprehensive model to rationalize at molecular level the CT process. Focusing on the HT type of mechanism, and in order to get further insight into the mechanism of CT, we have analyzed at the CASPT2 level the lowest doublet states of different oxidized nucleobase homo- and heterodimers in different orientations, from the highly overlapped π -stacked face-to-face arrangement (see Figs. 1 and 2) to the B-DNA form (see Figs. 3 and 4) and their evolution upon different displacements [47–49]. As an illustration we show here the case of the cationic A–C heterodimer ($C_3A_5^+$) in the B-DNA form [12] along the rise coordinate, that is, the displacement along the DNA helix axis [48].

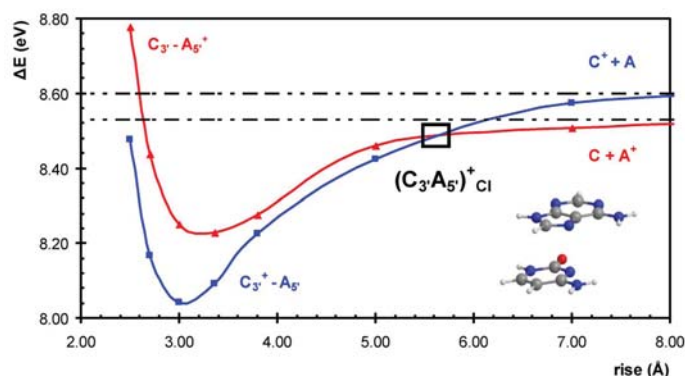


Fig. 3 BSSE-corrected CASPT2/ANO-S C,N,O [321]/H [21] PECs, built with respect to the intermonomer distance (rise) for the two low-lying doublet states of the oxidized C–A heterodimer ($C_3A_5^+$). The CI between the ground and lowest excited state of the dimer is represented by $(C_3A_5^+)_{CI}$ and marked with a black square.

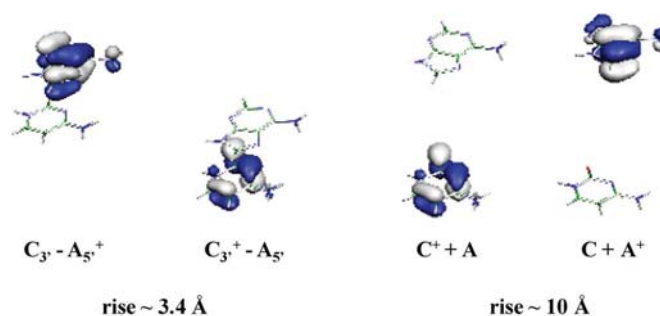


Fig. 4 Singly occupied natural orbitals (SONOs) of the ground and lowest excited state of the cationic B-DNA form of the A-C heterodimer (C_3A_5)⁺ at a geometry close to the average distance in DNA (3.4 Å) (left) and at large distances (right).

By inspection of the PECs for the C_3A_5 ⁺ heterodimer, it is evident that an oxidized C is strongly stabilized by the presence of A. As it can be seen in Fig. 3, in the heterodimer analyzed here, (C_3A_5)⁺, the cationic C and A are bound by the adjacent neutral nucleobase with decreases in their IPs of 0.56 eV at *rise* = 3.00 Å and 0.30 eV at *rise* = 3.36 Å, respectively. Since the Pyr nucleobase is more stabilized than the purine one and such stabilization is higher than the difference between the IPs of the isolated nucleic bases, the two low-lying states of the system, $C_3^+A_5$ and $C_3A_5^+$, which are characterized by the location of the hole in the C or A, respectively, cross-over in a point that takes place at intermolecular distances around 5.5–6.0 Å and corresponds to the conical intersection (CI) that mediates the HT between both nucleobases, (C_3A_5)⁺_{CI}. This result implies that the order of the IPs of C and A is inverted when these nucleic bases appear in adjacent positions of a DNA chain, and therefore the electron is removed from the Pyr nucleobase when the A-C heterodimer is oxidized. The remaining hole of the oxidation can be filled again by the promotion of one electron from the A partner by means of a thermally activated increase of the intermolecular distance between both nucleobases through the (C_3A_5)⁺_{CI}. This behavior is illustrated by the composition of the singly occupied MOs displayed in Fig. 4. Similar conclusions can be obtained by studying other nucleobase dimers upon other displacements, isolated or within the oligomer strand and under different environmental conditions, something which is currently under analysis by using QM/MM procedures. In all cases, the cornerstone of the mechanism is the presence of PEC CI crossings enabling ultrafast hole (or electron) transfer to take place. It is also true that the degree of transfer between the two states will be modulated by the strength of the electronic coupling between them, which is, together with the energy degeneration, the main condition for efficient charge or energy transfer. Comparison between the strength of the energy coupling between the different nucleobase dimers will be the goal of future research.

As a conclusion, the formation of stable excimer-like structures, revealed to be an intrinsic property of stacked nucleobases, also provides the framework to assemble the process of charge migration within the same basic unique scheme or unified theory. In this context, we propose a *micro-hopping* mechanism for CT along a single strand of nucleobase molecules formulated as a sequence of steps. Starting from an arrangement of π -stacked monomers at typical intermolecular distances of 3.1–3.4 Å, and considering the inherent flexibility of the DNA helix, in each step the two adjacent nucleobases may change their relative orientations by, for instance, increasing and decreasing their rise or twisting distance, by means of a thermally activated process, leading the system to reach accessible CIs between the lowest doublet states of the system and to transfer the excess electron (or hole charge) between the two molecules in an ultrafast manner. Therefore, charge migrates from an initial charged nucleobase to

the adjacent moiety in one step and between the initial and final charged DNA-sites in a sequence of such steps, thus making possible the CT process along the strand. That the process is more or less favorable will be surely related to the intrinsic properties of the nucleotide. It is therefore understandable that G sites, displaying the highest IPs, behave as hole charge trapping locations where the CT process is hindered, because they need a strongly distorted interacting conformation in order to cross with another state and transfer the hole to the neighbor moiety. Further research has to be performed in order to fully validate this hypothesis, extending the study to all combination of nucleobases and environments and considering the role of the Watson–Crick nucleobase pairs.

Photoreactivity in DNA: Formation of mutated Pyr dimers

One of the main motivations for studying the excited states of nucleic acids relies on the observation that UV illumination causes lesions and mutations due to photochemical modifications, the most common involving cycloaddition reactions of Pyrs T and C. Although the production of T–T cyclobutane dimers is most frequent, those involving C lead to mutation. T dimerization has recently been determined [50] to be an ultrafast reaction along the singlet manifold, whereas time-resolved studies of T dimer formation [51] show that direct excitation of $(dT)_{20}$ leads to cyclobutane T dimers ($T \leftrightarrow T$) in less than 200 ns with a remarkably absence of any triplet absorption from the transient spectra of the oligonucleotide. These evidences are in contrast to previous suggestions about the role of triplet states in DNA chemistry, in particular on the formation of Pyr complexes, $\text{Pyr} \leftrightarrow \text{Pyr}$ [52], and are also at odds with the well-known dimerization at the bipyrimidine sites under triplet photosensitization conditions [53]. Effective population of the triplet manifold is feasible, and it has been documented in detail for C [54], U [55], and T [56] along the ultrafast internal conversion channel. It is therefore not surprising that in addition to the presence of singlet excimer states which will behave as precursors of the photodimerization process, we have found corresponding triplet excimer states. For instance, in the C homodimer, the binding energy for the lowest triplet state computed at the CASPT2 level (plus BSSE) is 0.22 eV, with a predicted VE (phosphorescence) of 3.23 eV and a 0–0 triplet–singlet transition of 3.44 eV. Consequently, it is concluded that the triplet excimer is also bound, although the binding energy is reduced about 60 % with respect to singlet excimer. The possibility of excimer formation arises from the Watson–Crick structure in which hydrogen-bonded pairs A–T and G–C are situated inside a double helix, the backbone formed by two sugar-phosphate chains. One turn of the helix involves 10 base pairs and is 34 Å high. Thus, the interplanar distance between neighboring base pairs is about 3.4 Å, a value that is often found in excimer-type organic crystals [3].

Because of the increased stability of the lowest excited state, geometries around the face-to-face sandwich-type structure can be considered as the best candidates as precursors of photodimers. It seems that the ideal twist angle between successive base pairs makes the geometry of the B-DNA (and A-DNA) nonreactive. According to recent experimental evidence, the static Pyr–Pyr conformations and not conformational motions after photoexcitation determines the formation of $(\text{Pyr} \leftrightarrow \text{Pyr})$ photoproducts. Within the model proposed by Schreier et al. [50], 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 $(\text{Pyr} \leftrightarrow \text{Pyr})$ formation. As shown by these authors, dimerization occurs only for T residues that are already in a reactive arrangement at the instant of excitation because the rate of formation by favorably oriented T pairs is much faster than the rate of orientation change. A similar situation can therefore be assumed in C oligomers. We have taken the face-to-face arrangements, both for the singlet and triplet excimers, as the starting point for the study of the dimerization reaction occurring along the singlet and triplet manifold, respectively, in order to clarify at the molecular level the controversial and poorly understood mechanism of photodimer formation.

Figure 5 displays a scheme, based on the CASPT2 results, on the photodimerization of two C molecules taking place along both the singlet and the triplet manifold. The photoreactions taking place belong to the class of the [2+2] photocycloadditions. Starting from the triplet, MEP computations from

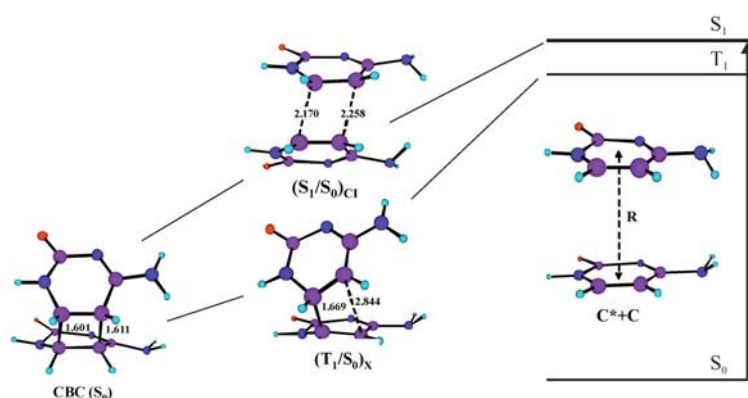


Fig. 5 Scheme based on CASPT2 calculations for the photoinduced formation of a CBC from a π -stacked C excimer ($C^* + C$) and along the singlet and triplet manifold, involving a CI (S_1/S_0)_{CI} and a singlet-triplet intermediate (T_1/S_0)_X, respectively.

the face-to-face arrangement lowest-energy minima of T_1 lead directly in a barrierless fashion to a triplet step-wise intermediate, which is characterized by the formation of a single covalent intermonomer bond computed to be 1.669 Å, whereas the other intermonomer CC distance remains elongated, about 2.8 Å. Remarkably, at this optimal structure the ground state of the system becomes degenerate. In other words, the triplet state is coincident with a triplet-singlet crossing, (T_1/S_0)_X, a region of the hypersurface where decay to the ground state becomes particularly favored. Such singlet-triplet degeneracy can be understood on the basis of the biradical character of the triplet-singlet states, having the unpaired electrons located on the bound C atoms. The intermediate triplet state lies 2.70 eV above two ground-state C molecules, an energy difference that can be considered a lower bound for the triplet energy of C in DNA.

It can be envisaged that exogenous photosensitizers could populate the relaxed triplet state of the monomer, which will subsequently evolve toward the intermediate and then toward the formation of the mutated dimers. Thus, the required energy can be related to the threshold observed experimentally for a given compound to become a potential DNA photodamager via ($C \leftrightarrow C$) or ($T \leftrightarrow T$) formation. The computed result for the singlet-triplet crossing structure of C, 2.70 eV, is fully consistent with the triplet energy of T in DNA deduced experimentally, 2.80 eV [53]. The intermediate represents, thus, a channel for photodimer formation from the triplet state of π -stacked C (and presumably also for T) in DNA and provides the basic understanding of potential photogenotoxicity via triplet-triplet sensitization. The efficiency of the overall process along the triplet manifold will be affected by the magnitude of the spin-orbit coupling (SOC), which is directly related to the efficiency of the intersystem crossing (ISC) process, and that would strongly rely on the actual environment of the biopolymer.

Regarding the mechanism of C photodimer formation along the singlet manifold, and in line with the similar excited-state [2+2] cycloaddition reactions of two ethene molecules [57], a CI is the funnel of decay toward ground-state cyclobutane C dimer (CBC) from the singlet excimer. Computations from different unrelaxed excimer lead to such CI structure. The process in the triplet manifold is also expected to follow a steepest descent path as it occurs along the singlet hypersurface. Thus, the current view supports the hypothesis that the dimerization photoreaction of two C molecules occurs barrierless, both on the singlet and triplet hypersurfaces. It would depend on the experimental conditions whether

the singlet or triplet mechanism becomes activated, fully operative, or even competitive with each other. The different mechanisms proposed in the literature involve singlet and triplet states of the monomers and vertical stacking to account for dimerization in solution and solid state, respectively. All those are here supported on the basis of CASPT2 results [6,58,59]. The efficiency of the photodimerization would markedly depend on the experimental conditions (solvent, aggregation conditions, pH, degree of hydration), the sequence of nucleotides, the type (A-, B- C-like) of DNA conformation [50,52]. Also important for the relative efficiency of the dimer formation is the presence of relaxed excimer structures at energies lower than the CI, as it occurs in C in contrast to T, a feature that lowers the formation yield of CBC. In fact, if dimer formation occurs with reasonable yields between monomeric solute molecules in solution, the dimer must have a triplet precursor, because singlet lifetimes simply are not long enough to permit excited bimolecular reactions to occur [60]. Nevertheless, as Eisinger and Shulman have emphasized [60], the same reaction which proceeds via triplet state in solution may have a singlet-state precursor when the biochromophores are held together, as is the case in frozen solutions or in a biopolymer. Theory predicts that the photoinduced reactions both on the singlet and triplet hypersurfaces are essentially barrierless, and singlet and triplet excimers play an active role in the photophysics outcome and in the photochemical properties of C-containing biopolymers. The present results also offer a nice rationale to the known fact that Pyr dimers are formed under triplet photosensitization conditions [52].

ACKNOWLEDGMENTS

Financial support is acknowledged from projects CTQ2007-61260 and CSD2007-0010 Consolider-Ingénio in Molecular Nanoscience of the Spanish MEC/FEDER. I.G.R. acknowledges a Ph.D. grant from the Spanish MEC (FPI).

REFERENCES

1. C. E. Crespo-Hernández, B. Cohen, P. M. Hare, B. Kohler. *Chem. Rev.* **104**, 1977 (2004).
2. G. Olaso-González, D. Roca-Sanjuán, L. Serrano-Andrés, M. Merchán. *J. Chem. Phys.* **125**, 231102 (2006).
3. J. B. Birks. *Rep. Prog. Phys.* **38**, 903 (1975).
4. W. Klöpffer. "Intramolecular excimers", in *Organic Molecular Photophysics*, J. B. Birks (Ed.), p. 357, Wiley-Interscience, London (1973).
5. G. Olaso-González, M. Merchán, L. Serrano-Andrés. *J. Phys. Chem. B* **110**, 24734 (2006).
6. D. Roca-Sanjuán, G. Olaso-González, I. González-Ramírez, L. Serrano-Andrés, M. Merchán. *J. Am. Chem. Soc.* **130**, 10768 (2008).
7. L. Serrano-Andrés, M. Merchán. "Photostability and photoreactivity in biomolecules: Quantum chemistry of nucleic acid base monomers and dimers", in *Radiation Induced Molecular Phenomena*, J. Leszczynski, M. Shukla (Eds.), Springer, Berlin (2008).
8. K. Andersson, P.-Å. Malmqvist, B. O. Roos. *J. Chem. Phys.* **96**, 1218 (1992).
9. L. Serrano-Andrés, M. Merchán, I. Nebot-Gil, R. Lindh, B. O. Roos. *J. Chem. Phys.* **98**, 3151 (1993).
10. B. O. Roos, K. Andersson, M. P. Fülscher, P.-Å. Malmqvist, L. Serrano-Andrés, K. Pierloot, M. Merchán. *Adv. Chem. Phys.* **93**, 219 (1996).
11. L. Serrano-Andrés, M. Merchán, A. C. Borin. *Proc. Natl. Acad. Sci. USA* **103**, 8691 (2006).
12. D. Roca-Sanjuán, L. Serrano-Andrés, M. Merchán. *Chem. Phys.* **349**, 188 (2008).
13. G. Karlström, R. Lindh, P.-Å. Malmqvist, B. O. Roos, U. Ryde, V. Veryazov, P.-O. Widmark, M. Cossi, B. Schimmelpfennig, P. Neogrady, L. Seijo. *Comp. Mater. Sci.* **28**, 222 (2003).
14. V. Veryazov, P.-O. Widmark, L. Serrano-Andrés, R. Lindh, B. O. Roos. *Int. J. Quantum Chem.* **100**, 626 (2004).
15. J. Eisinger, M. Guéron, R. G. Shulman, T. Yamane. *Proc. Natl. Acad. Sci. USA* **55**, 1015 (1966).

16. P. R. Callis. *Annu. Rev. Phys. Chem.* **34**, 329 (1983).
17. R. Plessow, A. Brockhinke, W. Eimer, K. Kohse-Höinghaus. *J. Phys. Chem. B* **104**, 3695 (2000).
18. P. R. Callis. *Annu. Rev. Phys. Chem.* **34**, 329 (1983).
19. M. Merchán, R. González-Luque, T. Climent, L. Serrano-Andrés, E. Rodríguez, M. Reguero, D. Peláez. *J. Phys. Chem. B* **110**, 26471 (2006).
20. L. Serrano-Andrés, M. Merchán, A. C. Borin. *Chem.—Eur. J.* **12**, 6559 (2006).
21. L. Serrano-Andrés, M. Merchán, A. C. Borin. *J. Am. Chem. Soc.* **130**, 2473 (2008).
22. M. Aida, M. Kaneko, M. Dupuis. In *Computational Molecular Biology*, J. Leszczynski (Ed.), Elsevier, Amsterdam (1999).
23. S. D. Wetmore, L. A. Eriksson, R. J. Boyd. In *Theoretical Biochemistry: Processes and Properties of Biological Systems*, L. A. Eriksson (Ed.), Elsevier, Amsterdam (2001).
24. M. E. Núñez, D. B. Hall, J. K. Barton. *Chem. Biol.* **6**, 85 (1999).
25. D. Ly, L. Sani, G. B. Schuster. *J. Am. Chem. Soc.* **121**, 9400 (1999).
26. S. Delaney, J. K. Barton. *J. Org. Chem.* **68**, 6475 (2003).
27. E. M. Boon, J. K. Barton. *Curr. Opin. Struct. Biol.* **12**, 320 (2002).
28. C. R. Treadway, M. G. Hill, J. K. Barton. *Chem. Phys.* **281**, 409 (2002).
29. E. B. Starikov. *Int. J. Quant. Chem.* **77**, 859 (2000).
30. M. E. Núñez, J. K. Barton. *Curr. Opin. Chem. Biol.* **4**, 199 (2000).
31. G. B. Schuster. *Acc. Chem. Res.* **33**, 253 (2000).
32. B. Giese. *Annu. Rev. Biochem.* **71**, 51 (2002).
33. F. D. Lewis, R. L. Letsinger, M. R. Wasielewski. *Acc. Chem. Res.* **34**, 159 (2001).
34. R. A. Marcus, N. Sutin. *Biochim. Biophys. Acta* **811**, 265 (1985).
35. M. Ratner. *Nature* **397**, 480 (1999).
36. J. Jortner, M. Bixon, T. Langenbacher, M. E. Michel-Beyerle. *Proc. Natl. Acad. Sci. USA* **95**, 12759 (1998).
37. B. Giese. *Acc. Chem. Res.* **33**, 631 (2000).
38. Y. A. Berlin, A. L. Burin, M. A. Ratner. *J. Am. Chem. Soc.* **123**, 260 (2001).
39. J. Joseph, G. B. Schuster. *J. Am. Chem. Soc.* **128**, 6070 (2006).
40. B. Giese, J. Amaudrut, A.-K. Köhler, M. Spormann, S. Wessely. *Nature* **412**, 318 (2001).
41. P. T. Henderson, D. Jones, G. Hampikian, Y. Kan, G. B. Schuster. *Proc. Natl. Acad. Sci. USA* **96**, 8353 (1999).
42. E. M. Conwell, S. V. Rakhmanova. *Proc. Natl. Acad. Sci. USA* **97**, 4556 (2000).
43. E. M. Conwell, S. M. Bloch. *J. Phys. Chem. B* **110**, 5801 (2006).
44. E. M. Conwell, S. M. Bloch, P. M. McLaughlin, D. M. Basko. *J. Am. Chem. Soc.* **129**, 9175 (2007).
45. D. Roca-Sanjuán, M. Rubio, M. Merchán, L. Serrano-Andrés. *J. Chem. Phys.* **125**, 084302 (2006).
46. D. Roca-Sanjuán, M. Merchán, L. Serrano-Andrés, M. Rubio. *J. Chem. Phys.* **129**, 095104 (2008).
47. R. Chandrasekaran, S. Arnott. In *Landolt-Börnstein Numerical Data and Functional Relationships in Science and Technology, Group VIII/1b, Nucleic Acids*, W. Saenger (Ed.), p. 31, Springer-Verlag, Berlin (1989).
48. S. Arnott. In *Oxford Handbook of Nucleic Acid Structure*, S. Neidle (Ed.), p. 1, Oxford University Press, Oxford (1999).
49. X.-J. Lu, W. K. Olson. *Nucleic Acids Res.* **31**, 5108 (2003).
50. W. J. Schreier, T. E. Schrader, F. O. Koller, P. Gilch, C. E. Crespo-Hernández, V. N. Swaminathan, T. Carell, W. Zinth, B. Köhler. *Science* **315**, 625 (2007).
51. S. Marguet, D. Markovitsi. *J. Am. Chem. Soc.* **127**, 5780 (2005).
52. J. Cadet, P. Vigny. "The photochemistry of nucleic acids", in *Bioorganic Photochemistry*, H. Morrison (Ed.), John Wiley, New York (1990).

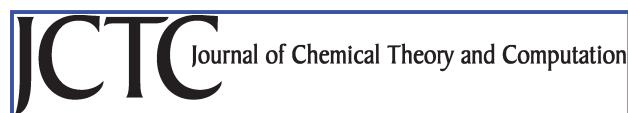
53. F. Bosca, V. Lhiaubet-Vallet, M. C. Cuquerella, J. V. Castell, M. A. Miranda. *J. Am. Chem. Soc.* **128**, 6318 (2006).
54. M. Merchán, L. Serrano-Andrés, M. A. Robb, L. Blancafort. *J. Am. Chem. Soc.* **127**, 1820 (2005).
55. T. Climent, R. González-Luque, M. Merchán, L. Serrano-Andrés. *Chem. Phys. Lett.* **441**, 327 (2007).
56. J. J. Serrano-Pérez, R. González-Luque, M. Merchán, L. Serrano-Andrés. *J. Phys. Chem. B* **111**, 11880 (2007).
57. F. Bernardi, M. Olivucci, M. A. Robb. *Acc. Chem. Res.* **23**, 405 (1990).
58. M. Boggio-Pasqua, G. Groenhof, L. V. Schäfer, H. Grubmüller, M. A. Robb. *J. Am. Chem. Soc.* **129**, 10996 (2007).
59. L. Blancafort, A. Migani. *J. Am. Chem. Soc.* **129**, 14540 (2007).
60. J. Eisinger, R. G. Shulman. *Science* **161**, 1311 (1968).

Publicación 7

Singlet-Triplet States Interaction Regions in DNA/RNA Nucleobase Hypersurfaces

González-Luque, R.; Climent, T.; González-Ramírez, I.; Merchán,
M.; Serrano Andrés, L.

J. Chem. Theor. Comp. **2010**, 6, 2103.



Singlet–Triplet States Interaction Regions in DNA/RNA Nucleobase Hypersurfaces

Remedios González-Luque, Teresa Climent, Israel González-Ramírez, Manuela Merchán, and Luis Serrano-Andrés*

Instituto de Ciencia Molecular, Universitat de València, Apartado 22085, ES-46071 Valencia, Spain

Received March 26, 2010

Abstract: The present study provides new insight into the intrinsic mechanisms for the population of the triplet manifold in DNA nucleobases by determining, at the multiconfigurational CASSCF/CASPT2 level, the singlet–triplet states crossing regions and the main decay paths for their lowest singlet and triplet states after near-UV irradiation. The studied singlet–triplet interacting regions are accessible along the minimum energy path of the initially populated singlet bright $^1\pi\pi^*$ state. In particular, all five natural DNA/RNA nucleobases have, at the end of the main minimum energy path and near a conical intersection of the ground and $^1\pi\pi^*$ states, a low-energy, easily accessible, singlet–triplet crossing region directly connecting the lowest singlet and triplet $\pi\pi^*$ excited states. Adenine, thymine, and uracil display additional higher-energy crossing regions related to the presence of low-lying singlet and a triplet $n\pi^*$ state. These funnels are absent in guanine and cytosine, which have the bright $^1\pi\pi^*$ state lower in energy and less accessible $n\pi^*$ states. Knowledge of the location and accessibility of these regions, in which the singlet–triplet interaction is related to large spin–orbit coupling elements, may help to understand experimental evidence such as the wavelength dependence measured for the triplet formation quantum yield in nucleobases and the prevalence of adenine and thymine components in the phosphorescence spectra of DNA.

1. Introduction

Phosphorescence spectra of DNA at low temperatures have been established as consisting of two basic components which originate mainly from thymine and, to a lesser extent, from adenine.^{1–3} Although triplet state formation and phosphorescence data of individual nucleobases and different derivatives in several media and conditions have been reported and reviewed,^{4–7} including recent studies employing external photosensitizers,^{8–10} the specifics of the intrinsic population mechanism of the triplet manifold in each of the nucleobases has not been understood so far. The different fates of their triplet states, explaining, for instance, the prevalence of two of the bases in the phosphorescence spectra of DNA, the absence of triplet guanine signals, or the triplet state involvement in the fast relaxation processes of nucleobases,

in particular for thymine,¹¹ have still to be elucidated. Triplet states of molecular systems are frequent intermediates in important photoinduced reactions. Both their usual biradical character and relatively long lifetimes make them reactive species prone to interacting with other compounds.¹² Triplet states of DNA/RNA purine and pyrimidine nucleobases are not an exception, and they have been determined to participate in UV-promoted photoreactions as the formation of phototherapeutic nucleobase–pharmakon adducts¹³ or the photodimerization of pyrimidine nucleobases, considered to be the most frequent genetic lesion taking place after UV-light irradiation.^{7,14–16} Since most of the recent attention has been focused on the rapid dynamics of the initially populated singlet states of DNA/RNA nucleobases,^{17–22} their intersystem crossing (ISC) mechanisms and triplet states' decay processes are only now starting to be analyzed.^{23–26} The present study aims to present a unified scheme, based on quantum chemical grounds, for the description of the main

* Corresponding author fax: (+34) 96-3544427, e-mail: Luis.Serrano@uv.es.

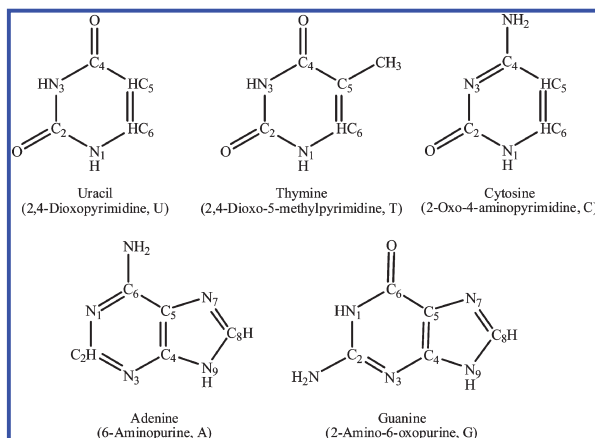


Figure 1. Structure, labeling, common name, IUPAC name, and acronym used for the five natural DNA/RNA nucleobases.

decay pathways for the singlet and triplet states of the five natural DNA/RNA nucleobases, thymine (T), uracil (U), cytosine (C), adenine (A), and guanine (G) (see Figure 1), locating the singlet–triplet crossing regions and computing the related spin–orbit coupling terms in order to provide insight into the intrinsic mechanisms of triplet state population in these molecules and to help rationalize the observed experimental data.

The triplet state population may proceed via endogenous or exogenous photosensitization from other triplet species or by efficient intersystem crossing (ISC) from the initially excited singlet state. There is an essential consensus that efficient radiationless transitions among states of the same multiplicity leading to internal conversion (IC) take place in the close vicinity of conical intersection (CI) regions and that the probability for the decay and the IC rates relate to the size of the nonadiabatic coupling matrix elements between the states.^{27–29} The situation is more complex for the computation of ISC rates. In this case, the efficiency of the interaction between states of different multiplicities, for instance, singlet–triplet, seems to be reasonably well described by the Fermi Golden Rule, which relates the strength of the interaction to the extent of the vibronic spin–orbit coupling (SOC) factors and the Franck–Condon (FC) weighted density of states.²⁷ Recent studies of Marian and co-workers^{26,30,31} have proved that the efficiency of an ISC process relies on a subtle balance of effects, including an enlarged density of vibrational states and a proper overlap of vibrational wave functions which, in turn, enhance the vibronic SOC effects. The decrease of the energy gap between singlet and triplet states, and in particular the presence of singlet–triplet degeneracies, crossing regions, especially when related to the existence of low-energy out of plane vibrational modes, is a good indication of a high density of states, and it is therefore conceivable that singlet–triplet crossings play an important role for increased ISC population transfer rates.¹² This relevance is well-known in the emerging field of multistate reactivity,^{27,32,33} in which the presence of singlet–triplet crossings and the occurrence

of corresponding ISC processes in the vicinity of the ground-state transition state regions become crucial for the enhancement of the reaction rates.³⁴ As they compete with generally faster internal conversion processes, intersystem crossings or spin crossovers can also be expected to be more efficient in energy trapping regions, for instance, near singlet states minima or sloped singlet–singlet conical intersections.^{35,36} Full reaction dynamics calculations including in a balanced and accurate way nonadiabatic and spin–orbit coupling effects for polyatomic systems like those considered here have not been performed yet. Until those studies are available, calculations of ISC rates in which the vibronic spin–orbit and overlap coupling effects are considered give the best information about the efficiency of the ISC process.^{26,30} Our goal in the present research is to determine the presence and accessibility of the singlet–triplet degeneracy regions in natural nucleobases along the main singlet decay pathways and provide hints of their relevance for ISC by computing also electronic SOC terms.

The strategy employed here starts by obtaining the minimum energy paths (MEPs) leading from the primary step of the photochemical process after UV light absorption in DNA nucleobases, being basically the population of the spectroscopically bright singlet excited state, here always the so-called $^1(\pi\pi^* L_a)$ state, toward the singlet–triplet degeneracy regions and finally the lowest-energy and reactive triplet excited state $^3(\pi\pi^* L_a)$, and calculating electronic SOC terms between relevant states.

Recent quantum-chemical *ab initio* CASPT2 studies have provided a unified model for the rapid internal conversion (IC) of the singlet excited DNA/RNA nucleobases manifold^{18,20,29,37–46} that allowed a proper rationalization of the experimental findings.^{17,47} The observed ultrafast decay component in all natural nucleobases, both in the gas phase and in solution, can be interpreted in terms of the barrierless character of the minimum energy path (MEP) associated with the lowest singlet state of the $\pi\pi^*$ type, $^1(\pi\pi^* L_a)$, toward a conical intersection (CI) with the ground state, $(gs/\pi\pi^*)_{CI}$. Secondary decay paths involving the lowest $^1n\pi^*$ state and

even a higher $^1\pi\pi^*$ state have been also identified.^{20,23,29,38–41} Within the context of the photochemical reaction path approach⁴⁸ and the current theoretical paradigm for non-adiabatic photochemistry,^{28,29} it is possible to analyze how the lowest triplet state can be reached efficiently by finding the singlet–triplet crossing (STC) regions more easily accessed from the FC MEP on the $^1(\pi\pi^* L_a)$ state, which represents the major deactivation path responsible of the rapid IC process detected in the molecule. Further studies combining the calculation of ISC rates and wave packet evolution will have to determine how efficient actually are our proposed channels. The obtained results suggest that enhancements in the population yield of the lowest triplet state of the natural DNA/RNA nucleobases can be related to the presence in three of them, T, U, and A, of more ISC channels along the singlet state MEP, in particular those related with low-lying singlet and triplet $n\pi^*$ states that act as intermediate population switchers, unlike what occurs in C and G. The obtained scheme may help to understand how the intrinsic population of the lowest triplet state can take place in vacuo for all the nucleobases, why T and A triplet states seem to prevail on the DNA phosphorescence spectrum and can be expected to have a larger quantum yield of formation (ϕ_{ISC}) than the other nucleobases, and what the molecular basis is for the detected wavelength dependence of ϕ_{ISC} .⁷ Since the calculations have been performed in vacuo, without the explicit consideration of solvent effects, the answer provided here can be regarded as a characteristic molecular property of the nucleobases, which might be expected to be somewhat disturbed by the specific environment in solution, in a solid, in vitro, or in vivo. The presence of the same ultrafast decays has been, however, identified in strands of oligonucleotides in solution,⁴⁹ probably related with the channels of the monomers in relatively unstacked nucleobases.⁵⁰

II. Methods and Computational Details

The present calculations include CASSCF geometry optimizations, MEPs, CIs, and STC searches, followed by multiconfigurational perturbation theory, CASPT2, calculations at the optimized geometries. SOC terms and transition dipole moments (TDM) have also been computed. Radiative lifetimes have been estimated by using the Strickler–Berg relationship,⁵¹ as explained elsewhere,⁵² although their applicability is restricted to cases where radiative deactivation predominates. Their magnitude, otherwise, is only indicative of the prospective emissive characteristics of the state related with the TDM values. For the sake of consistency with previous calculations on the singlet states of the systems, the same one-electron basis sets and active spaces were employed. For the pyrimidine T, U, and C and purine A and G nucleobases, basis sets of the ANO-S type contracted to C,N,O[3s2p1d]/H[2s1p] and 6-31G(d,p) were used, respectively. The final results can be described as CASPT2-(14,10) for T, U, and C, involving an active space of 14 electrons distributed in 10 orbitals, with all valence $\pi\pi^*$ and lone-pair orbitals, and CASPT2(14,12) for A and G, which include all $\pi\pi^*$ orbitals except those related to the deepest canonical orbital plus two lone-pair orbitals. Other active spaces were employed in the optimization procedures,

Table 1. Computed Properties for the Low-Lying Singlet and Triplet Excited States of Adenine

State	vertical transition (eV)		band origin (T_e , eV)		τ_{rad}^b
	CASSCF	CASPT2 ^a	CASSCF	CASPT2	
$^1(n\pi^*)$	5.95	4.96 (0.004)	4.88	4.52	334 ns
$^1(L_b \pi\pi^*)$	5.56	5.16 (0.004)	4.92	4.83	251 ns
$^1(L_a \pi\pi^*)^c$	7.03	5.35 (0.175)			
$^3(L_a \pi\pi^*)$	3.77	4.00	3.52	3.36 ^d	359 ms
$^3(n\pi^*)$	5.38	4.91	4.84	4.41	
$^3(\pi\pi^*)$	5.07	4.95			

^a Oscillator strengths within parentheses. ^b Computed using the Strickler–Berg approximation. See SI. ^c Geometry optimization leads directly to a conical intersection with the ground state, (gs/ $\pi\pi^*$)CI, at 4.0 eV. See refs 29 and 38. ^d Phosphorescence band origin and maximum in solution/glasses: 3.43 and 3.05 eV, respectively. See refs 58 and 59.

following a strategy which was proved successful previously. More detailed technical aspects of the calculations can be found in our previous papers^{23–25,37–39} and in the Supporting Information (SI). All the reported calculations used the quantum-chemical methods implemented in the MOLCAS 7 package.^{53,54}

III. Results and Discussion

The research effort in our group has been focused in recent years on the main singlet decay channels involving DNA/RNA nucleobases as well as several derivatives.^{18,20,37–39} In addition, studies were reported on the lowest triplet population mechanisms of the pyrimidine nucleobases thymine,^{24,26} uracil,^{25,26} and cytosine.²³ Other recent theoretical studies on the vertical and adiabatic energies of the nucleobases' triplet states have also been reported.^{55,56} In the present paper, we outline a unified scheme describing prospective population paths of the triplet manifold in all five natural DNA/RNA nucleobases T, U, C, A, and G, in order to obtain an overall model able to explain the common and the distinct behavior of the systems. Fully new results on the triplet states of the purine nucleobases A and G shall be presented, whereas our previous studies on T and U and new complementary calculations on C will be used and commented upon. The following subsections describe the results for each of the nucleobases. The most relevant conclusions are summarized in the last section.

A. Population of the Triplet Manifold in Adenine.

Table 1 compiles vertical transitions, band origins, oscillator strengths, and radiative lifetimes computed for the transitions to the singlet and triplet states of adenine at the CASSCF and CASPT2 levels of theory. Unless indicated, CASPT2 results will be used in the discussion. Both at the FC region and adiabatically, the lowest-lying singlet excited state is of the $n\pi^*$ ($n\pi\pi^*$) type, whereas the one carrying the largest intensity for the related transition, and therefore getting initially most of the population at low energies almost up to 6.0 eV, is the $^1(\pi\pi^*)$ HOMO (H) \rightarrow LUMO (L) (hereafter L_a) singlet excited state at 5.35 eV. The ultrafast nonradiative decay undergone by adenine in the femtosecond range^{17,47} can be rationalized by the barrierless character of the path on this state leading from the FC region toward a CI seam with the ground state, (gs/ $^1\pi\pi^*$)CI,^{29,38,41,43,57} and it is shown also here in Figure 2. Unlike simple geometry optimizations,

these two STC regions, the computed electronic SOC terms are $2 (^3n\pi^*/^1\pi\pi^*)$ and $7 \text{ cm}^{-1} (^3\pi\pi^*/^1\pi\pi^*)$. These values can be considered in agreement with the qualitative El-Sayed rules, which pointed to large SOC terms for states of different natures and small otherwise.⁶² El-Sayed rules were developed for molecules near the FC region, where most of the (organic) molecules considered were planar, and their identity, $\pi\pi^*$, $n\pi^*$, etc., could be qualitatively described as such. Far from the FC region, in particular, close to a strongly distorted and puckered geometry like the $^2\text{E CI}$, the same rules are not so easy to apply. For instance, the $\pi\pi^*$ state at this region, due to the out-of-plane distortion, has a close diradical character with two electrons in orbitals that are almost perpendicular to each other, the same as the $n\pi^*$ state in the FC region. This effect is particularly true for the low-energy singlet–triplet crossing region, which will be shown to be common in all nucleobases. The presence of the STCs combined with large electronic SOC terms are necessary, but not sufficient, conditions to guarantee efficient ISC processes, but they are good indications of relevant regions in which the population transfer toward the triplet states may take place, provided that the wave packet remains there for a long enough time for the ISC process to take place. The high-energy ($\sim 5.2 \text{ eV}$) $^1\pi\pi^* \rightarrow ^3n\pi^*$ STC area, not far from the FC absorption region, fulfills those conditions. On the other hand, recent reaction dynamics calculations suggest⁴⁵ that the region of the $(\text{gs}/^1\pi\pi^*)_{\text{CI}}$ (reached in femtoseconds), where also the STC takes place, represents an area in which the system stays trapped for some time (due to the structure of the CI) until the population switch toward the ground state takes place, which could also explain the slower picosecond channel observed in nucleobases.¹⁷ Figure 3 includes a scheme describing the population of T_1 based on our CASPT2 calculations.

From each one of the STC regions, we have computed corresponding MEPs along the populated triplet states, $^3(n\pi^*)$ and $^3(\pi\pi^*)$, for the suggested high- and low-energy ISC channels, respectively (they can be found in the SI). Soon, along the MEP on $^3(n\pi^*)$, a crossing with the lowest-lying $^3(\pi\pi^*)$ state takes place. The corresponding CI, $(^3n\pi^*/^3\pi\pi^*)_{\text{CI}}$, represents another funnel for efficient energy transfer within the triplet manifold. Additionally, as the singlet $^1(n\pi^*)$ state lies very close to the triplet counterpart and their PEHs run almost parallel, an STC ($^1n\pi^*/^3\pi\pi^*$) also occurs at that region. Considering that the computed SOC term in this case rises to 15 cm^{-1} , the corresponding ISC process toward the $^3(\pi\pi^*)$ state should be considered very favorable. A subsequent MEP from the $(^3n\pi^*/^3\pi\pi^*)_{\text{CI}}$ along the $(^3\pi\pi^*)$ PEH leads to the lowest triplet state minimum (see SI). Regarding the STC described at 4.0 eV , the MEP computed from the $(^3\pi\pi^*/^1\pi\pi^*)_{\text{STC}}$ along the $(^3\pi\pi^*)$ state leads directly to the minimum of the triplet state (see SI). The involvement of a dark singlet $n\pi^*$ state on adenine relaxation dynamics was previously suggested by other authors to explain slow decay features.^{17,63,64}

After the lowest triplet state is populated by any of the previous ISC processes, the system is finally expected to evolve toward the triplet state minimum, $^3(\pi\pi^*)_{\text{min}}$ (see Figure 3), which is characterized by a structure with almost planar rings but with the terminal hydrogen C_8H lifted near

40° and with an increased bond length C_2N_3 of 1.389 \AA (compared to 1.311 \AA in the ground state), in agreement with previous estimations.⁵⁵ The reactivity that could be attributed to this triplet state originates from its biradical character on C_2 and N_3 . The minimum is placed at 3.36 eV adiabatically (see Table 1) from the ground state optimized minimum, a value consistent with the measured phosphorescence band origin in solution at 3.43 ,⁵⁸ and other theoretical results.^{43,55} We have also located the singlet–triplet crossing connecting the $^3(\pi\pi^*)$ and the ground state and mapped the MEP leading from such an STC toward $^3(\pi\pi^*)_{\text{min}}$ (see SI). The crossing is placed near 4.0 eV from the ground state minimum, which means that there is a barrier of near 0.6 eV (14.0 kcal/mol) to reach $(\text{gs}/^3\pi\pi^*)_{\text{STC}}$ from $^3(\pi\pi^*)_{\text{min}}$. The distortion of the five-membered ring is larger at the STC point, and the computed electronic SOC is somewhat low, $\sim 2 \text{ cm}^{-1}$, suggesting for the triplet state a long lifetime and a slow relaxation, becoming therefore prone to reacting or transferring its energy by photosensitization mechanisms.^{8–10}

In summary, we have identified in adenine (see Figure 3) three possible intrinsic ISC channels toward the lowest triplet state which can be easily accessed from the main barrierless MEP for singlet decay dynamics, two of them mediated by $n\pi^*$ states. In all three cases, the magnitude of the computed SOC terms between the relevant states is high enough to suggest an efficient population of the triplet manifold in adenine upon UV irradiation. This type of $^1\pi\pi^*/^3\pi\pi^*$ ISC mechanism via intermediate $n\pi^*$ states can be suggested here as favorable, even far from the FC region, as it has been recently reported also for other biological chromophores such as isoalloxazine⁶⁵ and psoralen.⁶⁶ Both mechanisms described here can in any case contribute to the overall population of the lowest triplet state. In principle, in different environments, such as in polar solvents, it is expected that the $n\pi^*$ -type excited state will become destabilized with respect to $\pi\pi^*$ -type excited states.⁶⁷ Despite those effects, both singlet and triplet $n\pi^*$ -type states are estimated to lie in the solvent below the $^1(\pi\pi^* L_a)$ state at the FC geometry,⁶⁸ guaranteeing the existence of the STC crossing upon decay along the $^1(\pi\pi^* L_a)$ state. Intersystem crossing quantum yields have been measured by means of nanosecond laser photolysis in adenine to be 0.23×10^{-2} higher than in guanine.⁷ Likewise, phosphorescence quantum yields of 4.5×10^{-2} for adenine in frozen solutions at 77 K have been reported, slightly higher than for guanine, 3.6×10^{-2} and 2×10^{-2} ,^{69,70} and lower than thymine.⁷ For the purine nucleobases, the ISC yield has been measured to be lower in the nucleotide.⁷ Also in adenine,⁷¹ although less clearly documented as in pyrimidine nucleobases, a wavelength dependence of the intersystem crossing quantum yield in nucleobases has been reported, as it can be expected by the contribution of the three (at excitation energies higher than 5.0 eV) or just the lowest-energy (at energies close to 4.0 eV) ISC mechanisms. This point requires further experimental confirmation.

B. Population of the Triplet Manifold in Guanine. The same strategy as for adenine has been followed in the calculations of guanine. Table 2 lists the main spectroscopic properties of the lowest-lying singlet and triplet states of the

Table 2. Computed Properties for the Low-Lying Singlet and Triplet Excited States of Guanine

state	vertical transition (eV)		band origin (T_e , eV)		τ_{rad}^b
	CASSCF	CASPT2 ^a	CASSCF	CASPT2	
¹ (L _a ππ*) ^c	6.36	4.93 (0.158)			
¹ (n _o π*)	5.70	5.54 (0.002)	4.04	4.56	6800 ns
¹ (L _b ππ*)	7.04	5.72 (0.145)	6.07	5.69	5 ns
³ (L _a ππ*)	3.97	4.11	3.13	3.15	3562 ms
³ (ππ*)	5.08	4.76			
³ (ππ*)	5.41	5.14			
³ (n _o π*)	5.82	5.30	4.64	4.17	

^a Oscillator strengths within parentheses. ^b Computed using the Strickler–Berg approximation. See SI. ^c Geometry optimization leads directly to a conical intersection with the ground state, (gs/ππ*)CI, at 4.3 eV. See ref.³⁹

molecule. As compared to adenine, a couple of important aspects of the electronic structure of guanine have to be highlighted. First of all is the low energy displayed by the ¹(ππ* L_a) HOMO → LUMO state, placed at 4.93 eV at the FC region as the lowest-energy feature. The value of the related oscillator strength, 0.158, indicates that this is the bright singlet state basically populated in the low-energy absorption spectrum, and that the relevant photophysics of the system will take place along the MEP on such a state. The second aspect is related to the high-energy of the low-lying nπ* states, which are placed near 0.6 (singlet) and 0.4 (triplet) eV above the ¹(ππ* L_a) state (even higher in solution). As is clear from Table 2, and also from Figure 4, the gap between the initially populated singlet state and the nπ* states is much larger than in adenine. At the FC region, it is therefore expected that an ISC process relating the ¹(ππ* L_a) and ³nπ* states is less favorable than for adenine.

Figure 4 displays the MEP from the FC structure and along the ¹(ππ* L_a) state. At the beginning of the MEP, the singlet state only crosses with the second triplet ³(ππ*) state. The computed electronic SOC terms are small (<0.1 cm⁻¹), and only strongly coupled vibronic terms would enhance in this region the ISC rate. Near point 9 of the MEP, the singlet state crosses with the lowest triplet state, as it occurred in adenine. The STC region, placed adiabatically at 4.3 eV, is not far from the CI between the singlet and the ground state. The corresponding SOC terms are much larger here, 8 cm⁻¹, and therefore a more efficient ISC process leading directly to the population of the lowest ³(ππ*) state can be therefore expected, or at least proposed. As compared with adenine, however, the overall population of the triplet manifold cannot be expected to be favorable. Even when the ³(n_oπ*) excited state minimum lies lower in energy than the (gs/ππ*)CI, and therefore a crossing with the ¹(ππ*) state takes place at some other region, the key point is that such a crossing cannot be easily accessed from the photochemically relevant MEP, that is, the main decay path for singlet deactivation. As a matter of fact, we have computed the STC crossing structure (³n_oπ*/¹ππ*)_{STC}, which lies almost degenerate with the computed (¹n_oπ*/¹ππ*)_{CI} (see ref 39), at 4.6 eV, but far from the main MEP region, because it represents the stretching and twisting of the C₆N₁ bond. Even when such a structure, in which the SOC is large enough, 8 cm⁻¹, can be accessed with excess energy, it cannot be considered as favorable as those reached via the MEP-related channels.

For the sake of completeness, we have connected the mentioned STC points with the minimum of the lowest ³(ππ*) state by computing the corresponding MEPs: (i) from the computed (³ππ*/¹ππ*)_{STC} and (³n_oπ*/¹ππ*)_{STC} structures along the ³(ππ*) and ³(n_oπ*) states, leading to their respective minima, (ii) from the computed (³n_oπ*/³ππ*)_{CI} to the ³(ππ*) minimum, and (iii) from the singlet–triplet (¹gs/³ππ*)_{STC} toward the final ³(ππ*) minimum. All them are possible paths leading to the population of the lowest triplet state, although we emphasize that, unlike adenine, only the lowest-lying 4.3 eV ISC mechanism related to the (³ππ*/¹ππ*)_{STC} should be initially considered efficient, because it is the only one taking place in the proximity of the main ¹(ππ* L_a) MEP (see Figure 5 for a scheme of the triplet photophysics in guanine). Finally, the ³(ππ*) minimum has been connected through a corresponding MEP with the STC with the ground state, (gs/³ππ*)_{STC}. Although the SOC terms at this point are higher than in adenine, the barrier from the minimum, placed at 3.15 eV, is too large (0.85 eV) to expect an efficient decay to the ground state. All computed MEPs can be found in the SI. At the ³(ππ*) minimum, the molecule displays a slightly puckered envelope structure on the six-membered ring,³⁹ with the C₂N₃ bond having a biradical character and enlarged up to 1.438 Å, as compared to 1.286 Å at the FC ground-state geometry.

It has to be finally mentioned that guanine is the only natural nucleobase in which no phosphorescence data or triplet state formation has been reported for the parent compound, although intersystem crossing⁷ and phosphorescence quantum yields of 0.042 and 0.095 have been reported for the nucleoside and nucleotide in ethanol,⁶ 5 to 7 times larger than the fluorescence quantum yields. It has to be remembered also that the natural keto form of 9H-guanine is not the most stable in the gas phase and that other close tautomers can contribute to the measurements for the isolated system,^{39,72} not in an oligomer sample.

C. Population of the Triplet Manifold in Pyrimidine Nucleobases: Thymine, Uracil, and Cytosine. For the sake of brevity, we will discuss the triplet manifold population of the pyrimidine nucleobases together within the same framework. The computational strategies followed have been those described above for adenine and guanine. As uracil has a state structure and triplet photophysics very similar to that of thymine, we will refer to our previous results²⁵ and concentrate on the latter. Thymine has, at the FC region, a low-lying ¹(n_oπ*) state (basically related to the O₄ atom), placed 0.2 eV below the spectroscopic ¹(ππ* L_a) HOMO → LUMO state, this one lying at 4.89 eV with a related oscillator strength of 0.167 (see Table 3). The photophysical mechanisms proposed for the population of the lowest triplet state will be very similar to those already explained for adenine, as Table 3 and Figures 6 and 7 can confirm. Once more, the key point is that three different STC regions can be easily accessed through the main decay path of the energy, as it is the FC ¹(ππ* L_a) MEP, being prospective channels for ISC toward the lowest-lying triplet state.

Soon after the beginning of the MEP (see Figure 6), the ¹(ππ* L_a) state crosses with both singlet and triplet nπ* states. Apart from a possible IC toward the singlet ¹(n_oπ*)

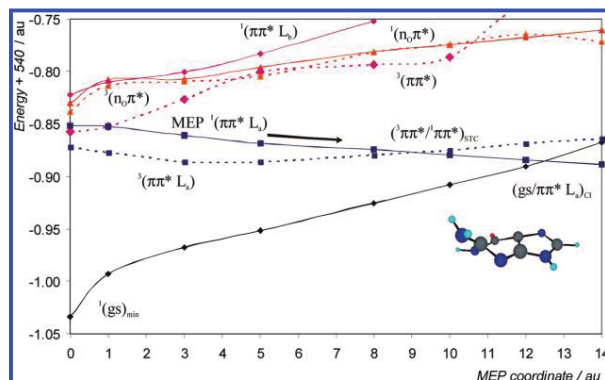


Figure 4. Evolution of the ground and lowest singlet and triplet excited states for guanine from the FC geometry along the ${}^1(\pi\pi^* L_a)$ MEP.

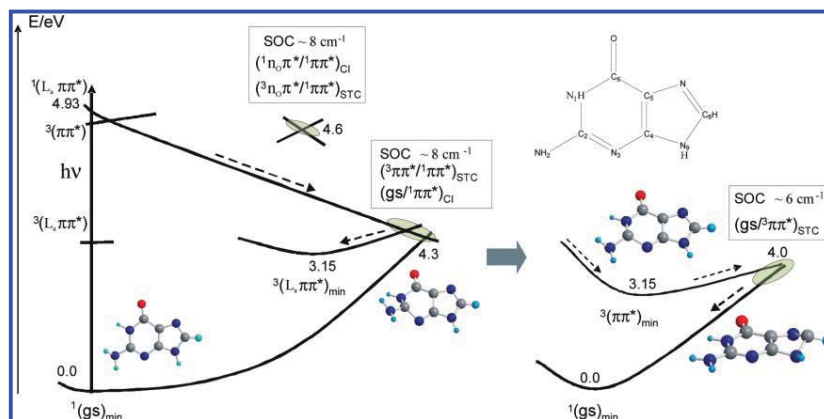


Figure 5. Scheme, based on CASPT2 results, of the photochemistry of guanine focused on the population of the lowest-energy triplet state. Unless otherwise stated, ${}^1\pi\pi^*$ represents the ${}^1L_a \pi\pi^*$ state.

Table 3. Computed Properties for the Low-Lying Singlet and Triplet Excited States of Thymine^a

state	vertical transition (eV)		band origin (T_e , eV)		τ_{rad}^c
	CASSCF	CASPT2 ^b	CASSCF	CASPT2	
${}^1(n\sigma\pi^*)^d$	5.41	4.77 (0.004)	4.23	4.05	2501 ns
${}^1(\pi\pi^* L_a)$	6.52	4.89 (0.167)	6.07	4.49	9 ns
${}^1(\pi\pi^*)$	7.36	5.94 (0.114)			
${}^3(\pi\pi^* L_a)$	3.95	3.59	2.99	2.87	17 ms
${}^3(n\sigma\pi^*)^d$	5.21	4.75	3.84	3.93	
${}^3(\pi\pi^*)$	5.86	5.14			

^a See also ref. 24. ^b Oscillator strengths within parentheses. ^c Computed using the Strickler-Berg approximation. See SI. ^d Involving basically O_4 (in ortho position with methyl group).

state through a corresponding CI, this region may be responsible for the first ISC process taking place in thymine at high energies (4.8 eV), in which the ${}^3(n\sigma\pi^*)$ state could be populated from the initially activated singlet $\pi\pi^*$ state. The SOC terms, computed as 8 cm^{-1} , point to the efficiency of the process. Another MEP computed from this crossing and along the ${}^3(n\sigma\pi^*)$ PEH leads the system toward the

minimum of this state, in whose neighborhood we have found the conical intersection with the lowest triplet state, $({}^3n\sigma\pi^*/{}^3\pi\pi^*)_{\text{CI}}$, near 3.9 eV. As the singlet and triplet $n\pi^*$ PEHs are always very close along the MEP, near the CI we have also found the $({}^1n\sigma\pi^*/{}^3\pi\pi^*)_{\text{STC}}$. In case some population reaches the ${}^1(n\sigma\pi^*)$ state via the higher-energy crossing with ${}^1(\pi\pi^* L_a)$ —and a decay path through this dark intermediate has been recently reported⁶⁸—the energy switch toward the lowest triplet state would be extremely favorable, because the computed SOC term increases in the $({}^1n\sigma\pi^*/{}^3\pi\pi^*)_{\text{STC}}$ region to 61 cm^{-1} . It is possible to confirm our suggestions about the effectiveness of this type of mechanism thanks to the recent study by Etinski et al.,²⁶ which has established the efficiency of the $({}^1n\sigma\pi^*/{}^3\pi\pi^*)_{\text{STC}}$ ISC channel by computing vibrational FC factors and ISC rates. Either by triplet-triplet IC or by singlet-triplet ISC, the final population process of the lowest ${}^3(\pi\pi^*)$ state should be considered to be extremely favorable. As in the other nucleobases, a low-energy STC region lies close to the end of the FC ${}^1(\pi\pi^*$

Table 4. Computed Properties for the Low-Lying Singlet and Triplet Excited States of Cytosine

state	vertical transition (eV)		band origin (T _e , eV)		τ _{rad} ^b
	CASSCF	CASPT2 ^a	CASSCF	CASPT2	
¹ (L _a ππ*) ^c	5.22	4.41 (0.069)	4.14	3.62	30 ns
¹ (n _o π*)	5.23	4.95 (0.001)	3.68	3.72	1200 ns
¹ (n _N π*) ^d	5.59	5.06 (0.003)			
¹ (L _b ππ*)	6.17	5.89 (0.106)			
³ (L _a ππ*)	3.64	3.53	2.85	2.98	437 ms
³ (ππ*)	4.87	4.45			
³ (n _o π*)	5.13	4.63	3.49	3.66	
³ (n _N π*)	5.31	4.94			

^a Oscillator strengths within parentheses. ^b Computed using the Strickler–Berg approximation. See SI. ^c The MEP to the minimum and the CI, (gs/ππ*)_{CI}, at 3.6 eV, are competitive. See ref 37. ^d Geometry optimization leads directly to a CI with the ground state, (gs/n_Nπ*)_{CI}. See ref 76.

measured for pyrimidine (~0.6 s) than for purine (~2.0 s) nucleobases in ethanol glasses.⁶ Similar conclusions can be derived for uracil, which has a state structure and properties very similar to those of thymine.^{25,26}

Regarding cytosine, the values in Table 4 help to understand (and predict to some extent) the behavior of its triplet photophysics. As in guanine, cytosine has a lowest-lying singlet ¹(ππ* L_a) state, whose initial interaction with the nπ* states, placed higher in energy, will not be strong either at the FC region or along the ¹(ππ* L_a) decay pathway (see Figure 8). The singlet relaxation in C is somewhat more complex than in the other nucleobases. The presence of a low-lying planar minimum for the ¹(ππ* L_a) state at 3.62 eV, nearly isoenergetic with the ethene-like (gs/ππ*)_{CI}, generates several competitive decay paths, as has been analyzed before.^{23,37,75} The possibilities for displaying different ISC processes are therefore larger, but always at low, not at high, energies like, for instance, in thymine, uracil, or adenine. In particular, we show in Figure 8 a linear interpolation in internal coordinates (LIIC) path from the FC region toward the ethene-like CI with the ground state. The barrier along the ¹(ππ* L_a) state, computed 2.5 kcal mol⁻¹ as a higher bound, is very small, and in practice the path can be considered barrierless. As in the other nucleobases, an STC between the lowest ππ* states takes place close to the CI, at 3.6 eV, yielding a SOC term value of 6 cm⁻¹. In a previous study,²³ we analyzed ISC processes taking place at other low-energy regions, obtaining also large SOC values and expectedly favorable situations for the lowest triplet population.

As a result of the excited state structure in C, obtained at the CASPT2 level, the photophysical scheme for the population of the lowest triplet state of the molecule can be summarized in Figure 9. Unlike in the other two pyrimidine nucleobases, where three basic channels for the possible triplet manifold population were found, one at high energies (close to FC and nπ* mediated) and another at low energies (caused by the common ethene-like CI type of decay present in all nucleobases), in C, only low-energy channels seem to be accessible. This feature could probably help to explain the absence of cytosine (guanine too) components in DNA phosphorescence at low temperatures,^{1–3} and also the generally lower phosphorescence quantum yields obtained for cytosine and its derivatives as compared to other

nucleobases.⁶ The same trends are obtained for ISC yields from flash photolysis experiments in nucleotides, although not in nucleobases.⁷ Higher yields of nπ* formation have been suggested for cytosine than thymine,⁶⁸ but theoretical evidence indicates that the higher-lying nπ* states of cytosine will be less accessible from the main relaxation pathways than in thymine due to the large potential energy barriers found in the former.⁷⁶

IV. Summary and Conclusions

Calculation of PEHs for the low-lying singlet and triplet states of natural DNA/RNA nucleobases adenine, guanine, thymine, uracil, and cytosine at the *ab initio* multiconfigurational CASPT2//CASSCF quantum-chemical level have been carried out in order to help to establish general mechanisms for the population of the triplet manifold of the molecules. The proposed framework is an attempt to rationalize the reported triplet states properties of DNA components, in particular the measurement of larger quantum yields of phosphorescence than of fluorescence in the individual systems,^{4,7} the observed wavelength dependence of the triplet state formation,^{7,73} or the prevalence of adenine and thymine components in the phosphorescence signals of DNA at low temperatures.^{1–3} It can be considered that an efficient ISC channel is easily accessible from the regions close to the main decay pathway of the initially populated singlet state. We have analyzed the accessibility of the ISC channels for the population of the lowest triplet state along such a pathway, a strategy that requires computation of minimum energy paths on the different states and determination of singlet–triplet crossings and conical intersections. This is, however, only a necessary but not sufficient condition to establish the efficiency of an ISC process. Computation of vibronic contributions to the ISC rates and reaction dynamic calculations establishing the temporal evolution of the system are encouraged in a close future in order to unambiguously determine if the proposed accessible singlet–triplet crossing regions fulfill all the requirements: close singlet–triplet energies, a high density of vibronic states, large vibronic contributions to the spin–orbit coupling terms, and regions where the population gets trapped for long enough of a time to allow the ISC process to take place in competition with the internal conversion decay, for instance, close to the FC region, to a singlet state minimum, or near a sloped conical intersection. Recent ISC rate calculations on thymine and uracil²⁶ confirm the main role of some of our proposed ISC mechanisms in these systems.

Our results indicate that three STC regions can be easily accessed from the singlet main decay pathway in adenine, thymine, and uracil, two of them located at high energies and mediated by the presence of lowest-lying singlet and triplet nπ* states, and a third one at low energies close to the end of the main MEP on the ¹(ππ*) singlet excited state and the ethene-like (pyrimidines) or methanamine-like (purines) conical intersection of this state with the ground state. These three regions are proposed as prospective ISC channels. At least those related to the ¹nπ*–³ππ* STC seem to be confirmed as such by recent calculations on ISC rates on pyrimidine nucleobases.²⁶ Additionally, the wavelength

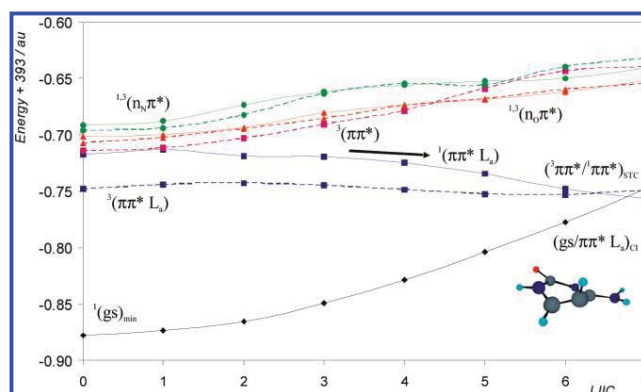


Figure 8. Evolution of the ground and lowest singlet and triplet excited states of cytosine from the FC geometry to the $(gs/\pi\pi^*)_{Cl}$ along an LIIC path competitive with the ${}^1(\pi\pi^*)_{La}$ MEP.

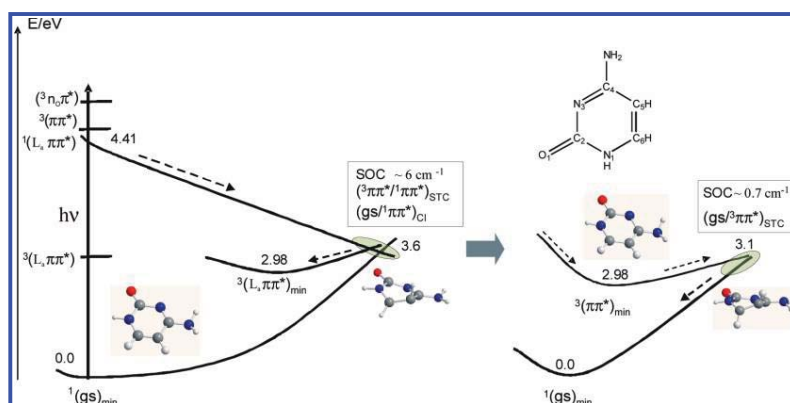


Figure 9. Scheme, based on CASPT2 results, of the photochemistry of cytosine focused on the population of the lowest-energy triplet state.

dependence of the triplet formation quantum yield reported in these three molecules is suggested to be related to the activation of the three (both at high and low excitation energies) or only one (at low energies) ISC channels. On the other hand, guanine and cytosine, having a much lower spectroscopic ${}^1(\pi\pi^*)$ singlet excited state below the $n\pi^*$ -type states, are not expected to display the $n\pi^*$ -mediated ISC mechanisms in regions close to the main MEP and may have only efficient ISC funnels at low energies, close to the singlet CI, a feature common to all nucleobases. The present results explain the fact that guanine and cytosine contribute much less to the phosphorescence of DNA, as it has been established.^{1–3} It is noteworthy to indicate that the phosphorescence spectrum of RNA was also reported,⁷⁷ and it was shown, first, to be determined mainly by the individual properties of the ribonucleotides' π -electron systems, and second, to be composed by triplet signals of adenosine groups and centers of an unknown nature with structureless long-wavelength phosphorescence different from that in DNA. The present results would indicate that adenine and, in this

case, uracil nucleobases should be preferably considered as sources of phosphorescence in RNA, as adenine and thymine are in DNA. It is clear that the present results for the isolated systems cannot be directly extrapolated to polymeric DNA/RNA. As already explained before, the described properties should be, however, considered intrinsic features of the nucleobases that, even if they may change in condensed phases or, in general, in the biological environment for the single monomers, are expected to maintain their basic characteristics, as occurs for the singlet states properties and it seems also for triplet states.⁷⁷

Acknowledgment. This research was supported by projects CTQ2007-61260, CTQ2010-14892, and CSD2007-0010 Consolider-Ingenio in Molecular Nanoscience of the Spanish MEC/FEDER and UV-EQUIP09-5764 of the Universitat de València.

Supporting Information Available: Additional computational details, reaction paths, and Cartesian coordinates

of the singular points. This material is available free of charge via the Internet at <http://pubs.acs.org>.

References

- (1) Imakubo, K. *J. Phys. Soc. Jpn.* **1968**, *24*, 1124.
- (2) Szerenyi, P.; Dearman, H. H. *Chem. Phys. Lett.* **1972**, *15*, 81.
- (3) Arce, R.; Rodríguez, G. *J. Photochem.* **1986**, *33*, 89.
- (4) Gueron, M.; Eisinger, J.; Lamola, A. A. Excited States of Nucleic Acids. In *Basic Principles in Nucleic Acid Chemistry*; Tso, P. O. P., Ed.; Academic Press: New York, 1974; Vol. 1, pp 311–398.
- (5) Daniels, M. In *Photochemistry and Photobiology of Nucleic Acids*; Wang, S. Y., Ed.; Academic Press: New York, 1976; Vol. 1, pp 23–108.
- (6) Görner, H. *J. Photochem. Photobiol. B: Biol.* **1990**, *5*, 359.
- (7) Cadet, J.; Vigny, P. In *Bioorganic Photochemistry*; Morrison, H., Ed.; John Wiley & Sons: New York, 1990; Vol. 1, pp 1–272.
- (8) Gut, I. G.; Wood, P. D.; Redmond, R. W. *J. Am. Chem. Soc.* **1996**, *118*, 2366.
- (9) Wood, P. D.; Redmond, R. W. *J. Am. Chem. Soc.* **1996**, *118*, 4256.
- (10) Bosca, F.; Lhiaubet-Vallet, V.; Cuquerella, M. C.; Castell, J. V.; Miranda, M. A. *J. Am. Chem. Soc.* **2006**, *128*, 6318.
- (11) Kang, H.; Lee, K. T.; Jung, B.; Ko, Y. J.; Kim, S. K. *J. Am. Chem. Soc.* **2002**, *124*, 12958–12959.
- (12) Klessinger, M. In *Theoretical Organic Chemistry - Theoretical and Computational Chemistry*; Párkányi, C., Ed.; Elsevier: Amsterdam, 1998; p 581.
- (13) Serrano-Pérez, J. J.; Merchán, M.; Serrano-Andrés, L. *J. Phys. Chem. B.* **2008**, *112*, 14002.
- (14) Brown, I. H.; Johns, H. E. *Photochem. Photobiol.* **1968**, *8*, 273.
- (15) Schreier, W. J.; Schrader, T. E.; Koller, F. O.; Gilch, P.; Crespo-Hernández, C. E.; Swaminathan, V. N.; Carell, T.; Zinth, W.; Kohler, B. *Science* **2007**, *315*, 625.
- (16) Roca-Sanjuán, D.; Olaso-González, G.; González-Ramírez, I.; Serrano-Andrés, L.; Merchán, M. *J. Am. Chem. Soc.* **2008**, *130*, 10768.
- (17) Crespo-Hernández, C. E.; Cohen, B.; Hare, P. M.; Kohler, B. *Chem. Rev.* **2004**, *104*, 1977–2019.
- (18) Serrano-Andrés, L.; Merchán, M. In *Radiation Induced Molecular Phenomena in Nucleic Acid: A Comprehensive Theoretical and Experimental Analysis*; Shukla, M. K., Leszczynski, J., Eds.; Springer: The Netherlands, 2008; pp 435–472.
- (19) Middleton, C. T.; De La Harpe, K.; Su, C.; Law, Y. K.; Crespo-Hernández, C. E.; Kohler, B. *Annu. Rev. Phys. Chem.* **2009**, *60*, 217.
- (20) Serrano-Andrés, L.; Merchán, M. *J. Photochem. Photobiol. C: Photochem. Rev.* **2009**, *10*, 21.
- (21) Conti, I.; Garavelli, M.; Orlandi, G. *J. Am. Chem. Soc.* **2009**, *131*, 16108.
- (22) Conti, I.; Altoè, P.; Stenta, M.; Garavelli, M.; Orlandi, G. *Phys. Chem. Chem. Phys.* **2010**, DOI: 10.1039/b926608a (accessed May 4, 2010).
- (23) Merchán, M.; Serrano-Andrés, L.; Robb, M. A.; Blancafort, L. *J. Am. Chem. Soc.* **2005**, *127*, 1820.
- (24) Serrano-Pérez, J. J.; González-Luque, R.; Merchán, M.; Serrano-Andrés, L. *J. Phys. Chem. B.* **2007**, *111*, 11880.
- (25) Climent, T.; González-Luque, R.; Merchán, M.; Serrano-Andrés, L. *Chem. Phys. Lett.* **2007**, *441*, 327.
- (26) Etinski, M.; Fleig, T.; Marian, C. M. *J. Phys. Chem. A* **2009**, *113*, 11809.
- (27) Klessinger, M.; Michl, J. *Excited States and Photochemistry of Organic Molecules*; VCH Publishers, Inc.: New York, 1995.
- (28) *Computational Photochemistry*; Olivucci, Ed.; Elsevier: Amsterdam, 2005.
- (29) Serrano-Andrés, L.; Merchán, M.; Borin, A. C. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 8691.
- (30) Tatchen, J.; Gilka, N.; Marian, C. M. *Phys. Chem. Chem. Phys.* **2007**, *9*, 5209.
- (31) Salzmann, S.; Tatchen, J.; Marian, C. M. *J. Photochem. Photobiol. A* **2008**, *198*, 221.
- (32) Carpenter, B. K. *Chem. Soc. Rev.* **2006**, *35*, 736.
- (33) Harvey, J. N.; Poli, R.; Smith, K. M. *Coord. Chem. Rev.* **2003**, *238–239*, 347.
- (34) González-Navarrete, P.; Coto, P. B.; Polo, V.; Andrés, J. *Phys. Chem. Chem. Phys.* **2009**, *11*, 7189.
- (35) Atchity, G. J.; Xantheas, S. S.; Ruedenberg, K. *J. Chem. Phys.* **1991**, *95*, 1862.
- (36) Ben-Nun, M.; Molnar, F.; Schulten, K.; Martínez, T. J. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 1769.
- (37) Merchán, M.; González-Luque, R.; Climent, T.; Serrano-Andrés, L.; Rodríguez, E.; Reguero, M.; Peláez, D. *J. Phys. Chem. B* **2006**, *110*, 26471.
- (38) Serrano-Andrés, L.; Merchán, M.; Borin, A. C. *Chem.—Eur. J.* **2006**, *12*, 6559.
- (39) Serrano-Andrés, L.; Merchán, M.; Borin, A. C. *J. Am. Chem. Soc.* **2008**, *130*, 2473.
- (40) Blancafort, L. *J. Am. Chem. Soc.* **2006**, *128*, 210.
- (41) Perun, S.; Sobolewski, A. L.; Domcke, W. *J. Am. Chem. Soc.* **2005**, *127*, 6257.
- (42) Perun, S.; Sobolewski, A. L.; Domcke, W. *J. Phys. Chem. A* **2006**, *110*, 13238.
- (43) Marian, C. M. *J. Chem. Phys.* **2005**, *122*, 104314.
- (44) Chen, H.; Li, S. H. *J. Phys. Chem. A.* **2005**, *109*, 8443.
- (45) Barbatti, M.; Lischka, H. *J. Am. Chem. Soc.* **2008**, *130*, 6831.
- (46) Hudock, H. R.; Martinez, T. J. *ChemPhysChem* **2008**, *9*, 2486.
- (47) Canuel, C.; Mons, M.; Pluzzi, F.; Tardivel, B.; Dimicoli, I.; Elhanine, M. *J. Chem. Phys.* **2005**, *122*, 074316.
- (48) Bernardi, F.; Olivucci, M.; Robb, M. A. *Pure Appl. Chem.* **1995**, *67*, 17.
- (49) Takaya, T.; Su, C.; De La Harpe, K.; Crespo-Hernández, C. E.; Kohler, B. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 10285.
- (50) Olaso-González, G.; Merchán, M.; Serrano-Andrés, L. *J. Am. Chem. Soc.* **2009**, *131*, 4368.
- (51) Strickler, S. J.; Berg, R. A. *J. Chem. Phys.* **1962**, *37*, 814.
- (52) Rubio-Pons, O.; Serrano-Andrés, L.; Merchán, M. *J. Phys. Chem. A* **2001**, *105*, 9664.

- (53) Veryazov, V.; Widmark, P.-O.; Serrano-Andrés, L.; Lindh, R.; Roos, B. O. *Int. J. Quantum Chem.* **2004**, *100*, 626.
- (54) Aquilante, F.; De Vico, L.; Ferré, N.; Ghigo, G.; Malmqvist, P.-Å.; Pedersen, T.; Pitonak, M.; Reiher, M.; Roos, B. O.; Serrano-Andrés, L.; Urban, M.; Veryazov, V.; Lindh, R. *J. Comput. Chem.* **2010**, *31*, 224.
- (55) Noguera, M.; Blancafort, L.; Sodupe, M.; Bertran, J. *Mol. Phys.* **2006**, *104*, 925.
- (56) Fleig, T.; Knecht, S.; Hättig, C. *J. Phys. Chem. A* **2007**, *111*, 5482.
- (57) Zgierski, M. Z.; Patchkovskii, S.; Lim, E. C. *Can. J. Chem.* **2007**, *85*, 124.
- (58) Cohen, B. J.; Goodman, L. *J. Am. Chem. Soc.* **1965**, *87*, 5487.
- (59) Lavík, J.; Jelínek, O.; Plášek, J. *Photochem. Photobiol.* **1979**, *29*, 491.
- (60) Sobolewski, A. L.; Domcke, W. *Phys. Chem. Chem. Phys.* **2004**, *6*, 2763.
- (61) Fülischer, M. P.; Serrano-Andrés, L.; Roos, B. O. *J. Am. Chem. Soc.* **1997**, *119*, 6168.
- (62) Lower, S. K.; El-Sayed, M. A. *Chem. Rev.* **1966**, *66*, 199.
- (63) Ullrich, S.; Schultz, T.; Zgierski, M. Z.; Stolow, A. *J. Am. Chem. Soc.* **2004**, *126*, 2262.
- (64) Ullrich, S.; Schultz, T.; Zgierski, M. Z.; Stolow, A. *Phys. Chem. Chem. Phys.* **2004**, *6*, 2796.
- (65) Climent, T.; González-Luque, R.; Merchán, M.; Serrano-Andrés, L. *J. Phys. Chem. A* **2006**, *110*, 13584.
- (66) Serrano-Pérez, J. J.; Merchán, M.; Serrano-Andrés, L. *Chem. Phys. Lett.* **2007**, *434*, 107.
- (67) Gustavsson, T.; Bányász, A.; Lazzarotto, E.; Markovitsi, D.; Scalamani, G.; Frisch, M. J.; Barone, V.; Improtà, R. *J. Am. Chem. Soc.* **2006**, *128*, 607.
- (68) Hare, P. M.; Crespo-Hernández, C. E.; Kohler, B. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 435.
- (69) Barenboim, G. M.; Domanskii, A. N. *Luminescence of biopolymers and cells*; Plenum Press: New York, 1969; p 85.
- (70) Snipes, W. *Electron spin resonance and the effects of radiation on biological systems*; National Academy of Sciences: Washington, DC, 1966, p55.
- (71) Gupron, M.; Shulman, R. G.; Eisinger, J. *Proc. Natl. Acad. Sci. U.S.A.* **1966**, *56*, 814.
- (72) Marian, C. M. *J. Phys. Chem. A* **2007**, *111*, 1545.
- (73) Nikogosyan, D. N.; Letokhov, L. S. *Riv. Nuov. Cim.* **1983**, *6*, 1.
- (74) Nguyen, M. T.; Zhang, R.; Nam, P.-C.; Ceulemans, A. *J. Phys. Chem. A* **2004**, *108*, 6554.
- (75) Blancafort, L. *Photochem. Photobiol.* **2007**, *83*, 603.
- (76) Merchán, M.; Serrano-Andrés, L. *J. Am. Chem. Soc.* **2003**, *125*, 8108.
- (77) Kudrya, V. Y.; Yashchuk, V. M.; Levchenko, S. M.; Melnik, V. I.; Zaika, L. A.; Govorun, D. M. *Mol. Cryst. Liq. Cryst.* **2008**, *497*, 425.

CT100164M

Publicación 8

Cyclobutane Pyrimidine Photodimerization of DNA/RNA Nucleobases in the Triplet State

Climent, T.; González-Ramírez, I.; González-Luque, R.; Merchán,
M.; Serrano-Andrés, L.

J. Phys. Chem. Lett. **2010**, 1, 2072.

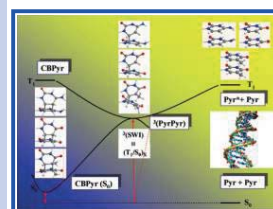
Cyclobutane Pyrimidine Photodimerization of DNA/RNA Nucleobases in the Triplet State

Teresa Climent, Israel González-Ramírez, Remedios González-Luque, Manuela Merchán, and Luis Serrano-Andrés*

Instituto de Ciencia Molecular, Universitat de València Apartado 22085, ES-46071 Valencia, Spain

ABSTRACT The photoinduced formation of cyclobutane pyrimidine dimers in the triplet excited state of the DNA/RNA pyrimidine nucleobases pairs has been studied at the CASPT2 level of theory. A stepwise mechanism through the triplet state of the homodimer is proposed for the pairs of nucleobases cytosine, thymine, and uracil involving a singlet–triplet crossing intermediary structure of biradical character representing the most favorable triplet state conformation of the nucleobases as found in the DNA environment. The efficiency of the mechanism will be modulated by two factors: the effectiveness of the triplet–triplet energy transfer process from a donor photosensitizer molecule, which relates to the relative position of the intermediate in the three acceptor systems, determined here to be lower in energy in the thymine and uracil dimers than in the cytosine pairs, and that of the intersystem crossing process toward the ground state of the photoproduct.

SECTION Dynamics, Clusters, Excited States



One of the most notorious examples of evidence of the photosensitivity of the genetic material to the action of ultraviolet (UV) light is the photoinduced formation of cyclobutane pyrimidine dimers (CBPyr or Pyr < > Pyr, see Figure 1) by pairs of DNA/RNA intrastrand adjacent pyrimidine nucleobases.¹ Those adducts constitute a major source of photoinduced DNA/RNA lesions, leading even to photomutagenesis and photocarcinogenesis, particularly in cellular DNA.² The process takes place in biological environments, solvents, and in the solid phase.¹ Femtosecond spectroscopy has proved that thymine (T) dimerization is an ultrafast photoreaction in which cyclobutane thymine dimers (CBT) are fully formed ~ 1 ps after UV illumination.³ Theoretical determinations have confirmed in thymine and cytosine (C) dimers a mechanism for a corresponding ultrafast non-adiabatic photoreaction mediated by the presence of a conical intersection (CI), an energy-degenerate structure between the low-lying singlet excited (S_1) and the ground state (S_0).^{4–7} In those quantum-chemical CASPT2 studies, the [2 + 2] photocycloaddition reaction leading to the formation of CBT and cyclobutane cytosine (CBC) dimers was characterized in the singlet manifold.^{4–7} Barrierless relaxation paths from favorable conformations of the nucleobases were shown to lead from an initially irradiated singlet state to a shearing-type CI structure, in which the nucleobases ethylenic C_5-C_6 and $C_5'-C_6'$ bonds laid parallel (parallelogram-type) and elongated, connecting the S_1 and S_0 states and allowing an efficient internal conversion process. Intrastrand nucleobase sequence and relative orientations were also proven to be essential for an efficient photoreaction to take place. This is particularly true for those conformations maximizing the

overlap between the π structures of stacked nucleobases, which formed favorable excimer arrangements.^{4–7} They were shown to yield the most stable structures leading to the photoreactive arrangements, in agreement with the higher yields obtained for photoproducts with cis-type parallel face-to-face conformations for the base pairs.^{1,4} We attributed the low yield measured for the CBC formation as compared to CBT to the stable excimer conformations, found in CC combinations lower in energy than the CI structure and opposite to what occurs in TT pairs.⁵

Apart from the nonadiabatic photoinduced reaction in the singlet manifold taking place after irradiation in the middle or far-UV, formation of the photoproducts mediated by the nucleobases triplet state is also conceivable from the initially excited singlet state of the monomer nucleobase to the ground state of the dimer photoproduct through successive intersystem crossing (ISC) processes.⁸ As it has been documented in detail experimentally⁹ and computationally for cytosine,¹⁰ uracil,^{11,12} and thymine,^{12,13} the lowest triplet state of the nucleobase can be populated along the ultrafast singlet-mediated internal conversion by different ISC mechanisms. Regarding dimers or oligomers, the experimental conclusions are not clear. Whereas recent time-resolved studies of CBT formation after irradiation of (dT)₂₀ oligonucleotides were unsuccessful at detecting the presence of the triplet state within a 200 ns time frame,¹⁴ there has also been

Received Date: May 10, 2010

Accepted Date: June 15, 2010

Published on Web Date: June 21, 2010

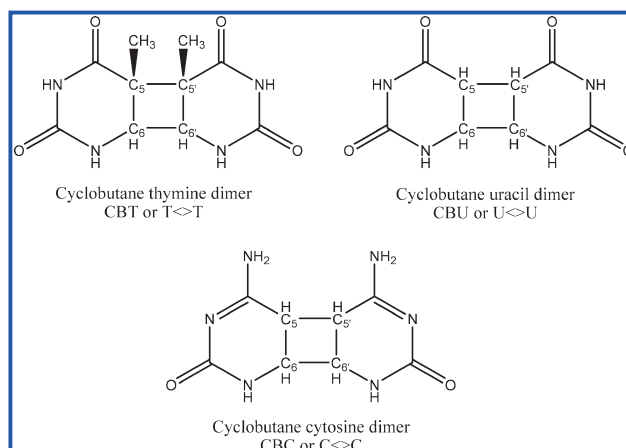


Figure 1. Structures and labeling of the DNA/RNA CBPyr homodimer photoproducts.

evidence of the ultrafast participation of the triplet state in the formation of the TT dimer.³ Even though the role of the triplet state of DNA oligomers in the dimerization process after direct singlet irradiation has still to be clarified, the participation of the triplet states in DNA photochemistry is uncontroversial. CBPyr adducts are notoriously formed upon nucleobase photosensitization by triplet–triplet energy transfer (TET) from favorable donors, such as different ketone or phthalimidine derivatives.^{1,2,15–18} The role of the nucleobases triplet states therefore seems to be linked to the ability of stacked DNA/RNA nucleobases to behave as acceptors under favorable triplet photosensitization conditions. As the efficiency of TET processes is strongly related to the relative triplet excited state energies of the donor and acceptor chromophores, the characterization of the nucleobases triplet intermediate responsible for the photosensitization action shall provide relevant insight into the photoreactive process. Great effort has been devoted to the determination of the nucleobases triplet (T_1) energy. Recent experimental studies on the TET formation of CBT dimers restrict the threshold observed for a given photosensitizer to become a potential DNA photodamager to substances with triplet state energies higher than 2.6–2.8 eV.^{18,19} Such limit was considered to represent the triplet energy of thymine in DNA, known to be much lower than that of the isolated nucleobase. In order to gain insight into the mechanisms of CBPyr adducts formation on the triplet state for the three nucleobase homodimers, CC, UU, and TT, this contribution reports calculations whose results are summarized in Table 1 and Figure 2, which also contain some of the data obtained in our previous study on CBC.⁴ Multiconfigurational CASPT2(14/10) and CASPT2(12/12) calculations for the monomers and the dimers, respectively, and ANO-S basis sets contracted to C,N,O[5s2p1d]/H[2s1p] were used, including CASSCF geometry optimizations of the singlet and triplet states minima, and minimum energy path (MEP) and minimum energy crossing point (MECP, singlet–triplet)

Table 1. Experimental and Theoretical CASPT2 Energies (eV) for the Lowest-Lying Triplet Excited State (T_1) of the DNA/RNA Pyrimidine Nucleobases^a

	monomer		dimer	monomer
	E_{VA}^b	T_e^c	$T_e(T_1/S_0)^d$	Exp ^e
cytosine	3.66	2.98	2.70	3.50
uracil	3.80	3.15	2.47	3.65
thymine	3.59	2.87	2.36	3.60

^aAll energies referred to the singlet ground state of two isolated nucleobases. See Figure 2. ^bVertical singlet–triplet excitation energy (E_{VA}) at the monomer ground-state (S_0) Franck–Condon minimum. ^cAdiabatic S_0 – T_1 monomer minimum to minimum energy (T_e): electronic excitation band origin. ^dAdiabatic S_0 – T_1 homodimer minimum to minimum energy, corresponding to the lowest-energy singlet–triplet crossing (T_1/S_0)_x and ³(SWI) intermediate. See text and Figure 2. ^eExperimental monomer singlet–triplet band maxima, approximately corresponding to E_{VA} . Gas-phase electron energy-loss spectroscopy (refs 12, 21, and 22).

determinations in the potential energy hypersurfaces (PEH) of a system of two nucleobases. Energies were corrected for the effect of the basis set superposition error using the counterpoise procedure. The MOLCAS-7 quantum-chemistry code was employed.²⁰ See further details in the Supporting Information (SI) file.

Results in Table 1 determine the location of the lowest-energy triplet state in DNA/RNA pyrimidine nucleobases, both for the isolated monomer and in the context of a stacked dimer. In the monomer, the three pyrimidine nucleobases have close-by experimental triplet state signals at 3.5–3.6 eV.^{12,21,22} Computation predicts, both for singlet–triplet vertical and adiabatic (band origin) transitions, that uracil lies higher in energy, followed by cytosine and thymine. Adiabatic energies cannot, however, properly represent triplet energies of the nucleobases in the biological environment, where many other effects have to be considered. We anticipate that the

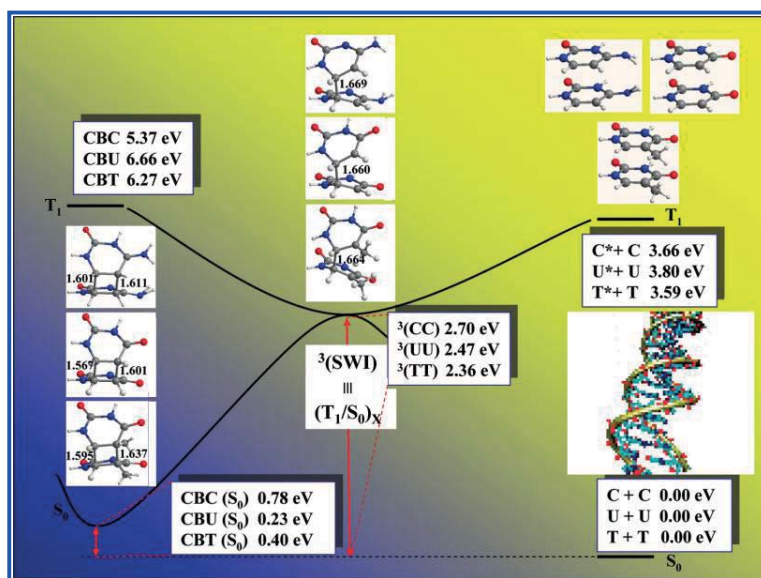


Figure 2. Scheme based on CASPT2 results of the triplet-mediated formation of CBPyr dimers (left) in isolated DNA/RNA nucleobase pairs. Center: singlet–triplet intermediate, $^3(\text{SWI})$; right: separated nucleobases stacked pairs. Bond lengths (Å) indicated when relevant: $\text{C}_6\text{--C}_6'$ in $^3(\text{SWI})$; $\text{C}_6\text{--C}_6'$ and $\text{C}_5\text{--C}_5'$ in the adducts.

major effect to take into account is the excimer stabilization undergone by the system in its excited state because of the interaction of pairs of stacked nucleobases. This would be followed, in turn, by the formation of a biradical stepwise intermediate $^3(\text{SWI})$, in which a covalent bond is actually established between the $\text{C}_6\text{--C}_6'$ carbons of the adjacent nucleobases (see Figure 2). MEP calculations prove (see Figures S5 to S7 in the SI) that such structures are reached in a barrierless way from the minimum of the triplet excimer arrangements formed in the parallel stacking of the nucleobases. The $^3(\text{SWI})$ intermediate has a biradical character with the unpaired electrons and the spin density on the two other ethylenic carbon atoms not forming the bond,⁴ i.e., C_5 and C_5' . The intermediate is an actual minimum in the T_1 PEH, and also corresponds to a singlet–triplet crossing (T_1/S_0)_X structure that is expected to mediate the formation of the CBPyr photoproduct. As we have demonstrated previously,⁴ in order to obtain a degenerated biradical singlet–triplet crossing situation in a [2 + 2] cycloaddition between ethylenic bonds, an asymmetric charge distribution at the carbon centers is required. This takes place in the stacked nucleobases, where the C_6 -type atoms are more electron deficient than the C_5 -type atoms because the electron affinity of the C_6 neighbor nitrogen is higher than that of the C_5 carbon atom of the pyrimidine ring. The substituted $\text{C}_6\text{--C}_6'$ carbon atoms compensate for the lack of electron density by getting closer to the other monomer and forming a new bond.

It is known that ultraviolet A (UVA) radiation (3.10–3.87 eV), which comprises a range of energies inaccessible for direct singlet state population but adequate to promote TET processes from certain photosensitizers, preferentially induces the production of cyclobutane dimers at TT sites without any detectable formation of Pyr(6–4)Pyr photoproducts.²³ Once the triplet excimer in the nucleobase dimer has accepted population from the donor triplet photosensitizer, the system will evolve favorably toward the formation of the biradical triplet intermediate, representing the most stable triplet state conformation in the dimer. We propose the $^3(\text{SWI})$ intermediate structure as the best candidate to trigger the cyclobutane photoadduct formation after ISC and evolution in the ground state PEH of the CBPyr adduct region, processes that have been suggested to have quasi unit efficiency.¹⁴ The actual efficiency of the decay process along the triplet manifold will also rely on the enhancement of the spin–orbit coupling terms, estimated here to be just a few inverse centimeters at the in vacuo $^3(\text{SWI})$ structures. In this respect, the solvent (or, in general, the environment) is expected to play a crucial role in the ISC process.²⁴ The location of the $^3(\text{SWI})$ intermediate has been found to be the lowest in energy in the thymine homodimers (2.36 eV), followed by uracil (2.47 eV) and cytosine (2.70 eV), and it can be also related to the minimum energy required to populate the triplet state in the DNA/RNA environment. These values are slightly lower than, but in relatively good agreement with, the experimental

estimation of 2.6–2.8 eV as a minimum bound for the occurrence of TET CBT formation in solvated environments.^{18,19} The present calculations therefore predict the production of thymine (CBT) and uracil (CBU) cyclobutane photoadducts in the triplet state taking place at similar energies and conditions. Higher energies would be, however, required to give rise to the cytosine adduct (CBC), which, as it occurred in the singlet manifold, might be also less favored by the formation of very stable cytosine excimers.⁴ Such energy ordering has also been deduced in aqueous solution and rigid media experiments.²⁵ Since the efficiency of the triplet state formation has been determined to be larger in isolated thymine than in cytosine,^{10,13} it is possible to rationalize the preference of the TT sites, and to a lesser extent CT, TC, and CC, to generate the cyclobutane photoproduct. CBT and CBU adducts have been found to dimerize faster and more efficiently than CBC in the presence of acetone and acetophenone.¹ Despite being obtained in lower yields, CC sequences are more relevant in the study of the damaging effects in DNA, because, unlike TT sequences, they are sites of relatively frequent CC-to-TT tandem mutations.²

The ³(SWI) structure connects, on one hand, with the S₀ state minimum of the CBPyr adduct, but also with the ground state of the separated monomers. From the adduct geometry, the intermediate can be reached in a barrierless way along the T₁ state PEH (see the corresponding MEPs in the SI file, Figures S5 and S6). The same type of behavior was found for the S₁ state with respect to the S₀/S₁ CI in CBC,⁴ and it is related with the observed photoreversibility of the adduct formation process after ultraviolet C (UVC) irradiation.¹ At the CBPyr structure the triplet state lays vertically at 4.59 (CBC), 6.43 (CBU), and 6.15 eV (CBT), whereas their respective S₁ states are placed (vertically) at 4.95, 6.28, and 5.35 eV, just in the UVC (middle to far-UV) range. It is worth mentioning that the nature of the T₁ state is different in CBC ($\pi\pi^*$), where the spin density is distributed on the C=N bond atoms in one of the monomers, and in the CBU and CBT pairs, in which the state, of $n\pi^*$ character, lies at considerably higher energy, and its spin density localizes on the C=O bond atoms neighbor to C₅. In these cases, the nature of the state changes adiabatically along the MEP, becoming of $\pi\pi^*$ in all cases at the ³(SWI) structure. It is worth mentioning that the intrinsic mechanism of CBPyr formation does not seem to largely change in the biological environment, as shown in our CASPT2/MM study on the reversibility of the cytosine dimer (CBC) formation in a biological environment (dC₁₈xdG₁₈ was simulated).²⁶ Although further studies are required in all cases, the computational evidence is strong enough to illustrate that the basics of the mechanism are properly described in the isolated dimer.

In summary, once the presence of precursor singlet excimer structures favoring ultrafast internal conversion through an S₀/S₁ CI was previously proposed^{3–7} as the major source of the formation of CBPyr dimers in the biopolymer nucleobases or in frozen solutions,¹ the present scheme tries to rationalize the mechanism for triplet–triplet photosensitized photoadducts from appropriate donors, an energy transfer process expected to be more favored in solution.¹ The overall lower yields of formation in cytosine dimers versus those in thymine (and uracil) can be related first to

the relative location on the former of stable singlet excimer structures lower in energy than the singlet CI triggering the internal conversion. Additionally, the triplet intermediate mediating the photoadduct formation by means of energy transfer processes is in cytosine higher (and less accessible) than those in the other pyrimidine nucleobases. Determining, along the lines followed in the present research, the nature and relative location of the triplet intermediates in different nucleobase duplexes will help to rationalize the phototoxicity of different photosensitizers, driven by the energy of their reactive triplet states, and to design health care phototherapeutic nucleobase-based drugs addressed to either enhance or decrease pyrimidine dimer formation by using, for instance, appropriate nucleobase derivatives.

SUPPORTING INFORMATION AVAILABLE Computational details, MEPs and scans on the PEHs, and Cartesian coordinates for the optimized conformations. This material is available free of charge via the Internet at <http://pubs.acs.org>.

AUTHOR INFORMATION

Corresponding Author:

*To whom correspondence should be addressed. E-mail: Luis.Serrano@uv.es.

ACKNOWLEDGMENT The research reported in this report has been supported by the Spanish MICINN/FEDER projects CTQ2007-61260, CTQ2010-14892, and CSD2007-0010 Consolider-Ingénio in Molecular Nanoscience and by the Generalitat Valenciana.

REFERENCES

- (1) Cadet, J.; Vigny, P. *The Photochemistry of Nucleic Acids*. In *Bioorganic Photochemistry*; Morrison, H., Ed.; John Wiley & Sons: New York, 1990; Vol. 1, pp 1–272.
- (2) Douki, T.; Cadet, J. Individual Determination of the Yield of the Main UV-Induced Dimeric Pyrimidine Photoproducts in DNA Suggests a High Mutagenicity of CC Photolesions. *Biochemistry* **2001**, *40*, 2495–2501.
- (3) Schreier, W. J.; Schrader, T. E.; Soller, F. O.; Gilch, P.; Crespo-Hernández, C. E.; Swaminathan, V. N.; Carell, T.; Zinth, W.; Kohler, B. Thymine Dimerization in DNA is an Ultrafast Photoreaction. *Science* **2007**, *315*, 625–629.
- (4) Roca-Sanjuán, D.; Olaso-González, G.; González-Ramírez, I.; Serrano-Andrés, L.; Merchán, M. Molecular Basis of DNA Photodimerization: Intrinsic Production of Cyclobutane Cytosine Dimers. *J. Am. Chem. Soc.* **2008**, *130*, 10768–10779.
- (5) Serrano-Pérez, J. J.; González-Ramírez, I.; Coto, P. B.; Merchán, M.; Serrano-Andrés, L. Theoretical Insight into the Intrinsic Ultrafast Formation of Cyclobutane Pyrimidine Dimers in UV-Irradiated DNA: Thymine versus Cytosine. *J. Phys. Chem. B* **2008**, *112*, 14096–14098.
- (6) Boggio-Pasqua, M.; Groenhof, G.; Schäfer, L. V.; Grubmüller, H.; Robb, M. A. Ultrafast Deactivation Channel for Thymine Dimerization. *J. Am. Chem. Soc.* **2007**, *129*, 10996–10997.
- (7) Blancafort, L.; Migani, A. Modeling Thymine Photodimerizations in DNA: Mechanism and Correlation Diagrams. *J. Am. Chem. Soc.* **2007**, *129*, 14540–14541.
- (8) Kwok, W. M.; Ma, C.; Phillips, D. L. A Doorway State Leads to Photostability or Triplet Photodamage in Thymine DNA. *J. Am. Chem. Soc.* **2008**, *130*, 5131–5139.

- (9) Hare, P. M.; Crespo-Hernandez, C. E.; Kohler, B. Internal Conversion to the Electronic Ground State Occurs via Two Distinct Pathways for Pyrimidine Bases in Aqueous Solution. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 435–440.
- (10) Merchán, M.; Serrano-Andrés, L.; Robb, M. A.; Blancafort, L. Triplet-State Formation Along the Ultrafast Decay of Excited Singlet Cytosine. *J. Am. Chem. Soc.* **2005**, *127*, 1820–1825.
- (11) Climent, T.; González-Luque, R.; Merchán, M.; Serrano-Andrés, L. On the Intrinsic Population of the Lowest Triplet State of Uracil. *Chem. Phys. Lett.* **2007**, *441*, 327–331.
- (12) Etinski, M.; Fleig, T.; Marian, C. M. Intersystem Crossing and Characterization of Dark States in the Pyrimidine Nucleobases Uracil, Thymine, and 1-Methylthymine. *J. Phys. Chem. A* **2009**, *113*, 11809–11816.
- (13) Serrano-Pérez, J. J.; González-Luque, R.; Merchán, M.; Serrano-Andrés, L. On the Intrinsic Population of the Lowest Triplet State of Thymine. *J. Phys. Chem. B* **2007**, *111*, 11880–11883.
- (14) Marguet, S.; Markovitsi, D. Time-Resolved Study of Thymine Dimer Formation. *J. Am. Chem. Soc.* **2005**, *127*, 5780–5781.
- (15) Gut, I. G.; Wood, P. D.; Redmond, R. W. Interaction of Triplet Photosensitizers with Nucleotides and DNA in Aqueous Solution at Room Temperature. *J. Am. Chem. Soc.* **1996**, *118*, 2366–2373.
- (16) Cadet, J.; Sage, E.; Douki, T. Ultraviolet Radiation-Mediated Damage to Cellular DNA. *Mutat. Res.* **2005**, *571*, 3–17.
- (17) Nguyen, M. T.; Zhang, R.; Nam, P.-C.; Ceulemans, A. Singlet-Triplet Energy Gaps of Gas-Phase RNA and DNA Bases. A Quantum Chemical Study. *J. Phys. Chem. A* **2004**, *108*, 6554–6561.
- (18) Bosca, F.; Lhiaubet-Vallet, V.; Cuquerella, M. C.; Castell, J. V.; Miranda, M. A. The Triplet Energy of Thymine in DNA. *J. Am. Chem. Soc.* **2006**, *128*, 6318–6319.
- (19) Albin, A.; Monti, S. Photophysics and Photochemistry of Fluoroquinolones. *Chem. Soc. Rev.* **2003**, *32*, 238–250.
- (20) Aquilante, F.; De Vico, L.; Ferré, N.; Ghigo, G.; Malmqvist, P.-Å.; Pedersen, T.; Pitonak, M.; Reiher, M.; Roos, B. O.; Serrano-Andrés, L.; Urban, M.; Veryazov, V.; Lindh, R. MOL-CAS 7: The Next Generation. *J. Comput. Chem.* **2010**, *31*, 224–247.
- (21) Abouaf, R.; Pommier, J.; Dunet, H.; Quan, P.; Nam, P.-C.; Nguyen, M. T. The Triplet State of Cytosine and its Derivatives: Electron Impact and Quantum Chemical Study. *J. Chem. Phys.* **2004**, *121*, 11668–11674.
- (22) Abouaf, R.; Pommier, J.; Dunet, H. Electronic and Vibrational Excitation in Gas Phase Thymine and 5-Bromouracil by Electron Impact. *Chem. Phys. Lett.* **2003**, *381*, 486–494.
- (23) Mouret, S.; Baudouin, C.; Charveron, M.; Favier, A.; Cadet, J.; Douki, T. Cyclobutane Pyrimidine Dimers are Predominant DNA Lesions in Whole Human Skin Exposed to UVA Radiation. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 13765–13770.
- (24) Hare, P. M.; Crespo-Hernández, C. E.; Kohler, B. Solvent-Dependent Photophysics of 1-Cyclohexyluracil: Ultrafast Branching in the Initial Bright State Leads Nonradiatively to the Electronic Ground State and a Long-Lived $^1n\pi^*$ State. *J. Phys. Chem. B* **2006**, *110*, 18641–18650.
- (25) Wood, P. D.; Redmond, R. W. Triplet State Interactions between Nucleic Acid Bases in Solution at Room Temperature: Intermolecular Energy and Electron Transfer. *J. Am. Chem. Soc.* **1996**, *118*, 4256–4263.
- (26) Roca-Sanjuán, D.; Olaso-González, G.; Rubio, M.; Coto, P. B.; Merchán, M.; Ferré, N.; Ludwig, V.; Serrano-Andrés, L. DNA Nucleobase Properties and Photoreactivity: Modeling Environmental Effects. *Pure Appl. Chem.* **2009**, *81*, 743–754.

Publicación 9

On the photoproduction of DNA/RNA cyclobutane pyrimidine dimers

González-Ramírez, I.; Roca-Sanjuán, D.; Climent, T.; Serrano-Pérez, J.J.;
Merchán, M.; Serrano- Andrés, L.

Theoretical Chemistry Accounts. **2010**, 128, 705.

On the photoproduction of DNA/RNA cyclobutane pyrimidine dimers

Israel González-Ramírez · Daniel Roca-Sanjuán ·
Teresa Climent · Juan José Serrano-Pérez ·
Manuela Merchán · Luis Serrano-Andrés

Received: 31 August 2010 / Accepted: 3 November 2010 / Published online: 23 November 2010
© Springer-Verlag 2010

Abstract The UV photoreactivity of different pyrimidine DNA/RNA nucleobases along the singlet manifold leading to the formation of cyclobutane pyrimidine dimers has been studied by using the CASPT2 level of theory. The initially irradiated singlet state promotes the formation of excimers between pairs of properly oriented nucleobases through the overlap between the π structures of two stacked nucleobases. The system evolves then to the formation of cyclobutane pyrimidine dimers via a shearing-type conical intersection activating a [2 + 2] photocycloaddition mechanism. The relative location of stable excimer conformations or alternative decay channels with respect to the reactive degeneracy region explains the differences in the photoproduction efficiency observed in the experiments for different nucleobases sequences. A comparative analysis of the main structural parameters and energetic profiles in the singlet manifold is carried out for thymine, uracil, cytosine, and 5-methylcytosine homodimers. Thymine and uracil dimers display the most favorable paths, in contrast

to cytosine. Methylation of the nucleobases seems to increase the probability for dimerization.

Keywords CASPT2 · Photochemistry · Conical Intersection · DNA/RNA · Cyclobutane Pyrimidine Dimers

1 Introduction

The UV radiation directly absorbed by the nucleic acids can produce a large number of lesions [1, 2], the most common corresponding to dimer or adduct formations involving adjacent pyrimidine bases of the DNA/RNA strand [3–7]. Two types of dimerized products are frequently found: the cyclobutane pyrimidine dimers (CPDs), in which the adduct is formed via a [2 + 2] photocycloaddition involving the C=C double bonds of the pyrimidines, and the pyrimidine (6-4) pyrimidone dimers, usually referred to as (6-4) photoproducts, (6-4)PPs, in which the cycloaddition takes place between C=C and C=O double bonds of the adjacent molecules. CPD lesions are more frequent than (6-4)PPs [3, 4]. Under UV-C (254-nm) radiation conditions, between two and ten CPDs per million bases are formed [5]. Two more orders of magnitude of lower-energy UV-B radiation are required to obtain the same result. Even though, in theory, various CPD diastereoisomers are possible, only the *cis-syn* conformer is found in the double helix. The *trans-syn* configuration can be also obtained, with lower yields, in single or double strands of denaturalized DNA, where the tertiary and secondary structure are lost to some extent with respect to the standard double helix [3]. Although the pyrimidine bases can react to form CPDs in many combinations, we will focus here on the homodimers of the canonical DNA/RNA bases thymine (T), uracil (U), cytosine (C) (and their adducts

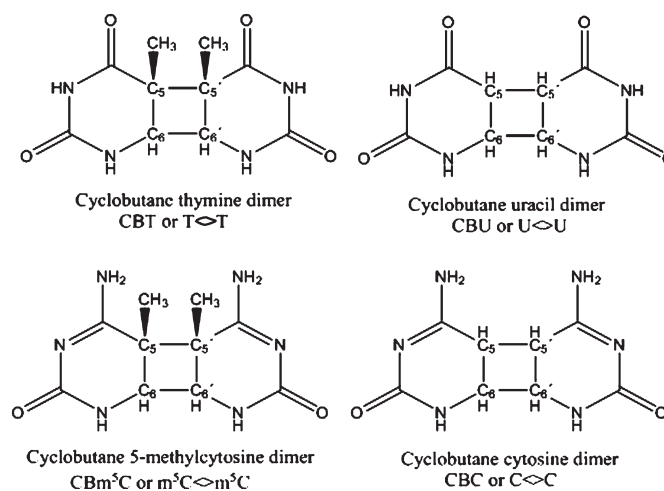
Published as part of the special issue celebrating theoretical and computational chemistry in Spain.

I. González-Ramírez · T. Climent · M. Merchán ·
L. Serrano-Andrés
Instituto de Ciencia Molecular, Universitat de València,
P. B. Box 22085, 46071 Valencia, Spain

D. Roca-Sanjuán (✉)
Department of Quantum Chemistry, Uppsala University,
Box 518, 75120 Uppsala, Sweden
e-mail: Daniel.Roca@kvac.uu.se

J. J. Serrano-Pérez
Department of Chemistry, Computational and Structural
Research Group, Imperial College London,
SW7 2AZ London, UK

Fig. 1 Structures and labeling of the DNA/RNA cyclobutane pyrimidine homodimer photoproducts



CBT, CBU, and CBC), and on the derivative 5-methylcytosine (m⁵C and CBm⁵C), as displayed in Fig. 1.

Despite all CPDs can be considered DNA/RNA lesions, not all of them are actually mutagenic hotspots, something that depends on the rate of the enzymatic repair mechanisms or of the transitions from one pyrimidine to another, for instance by deamination plus ketonization in C and m⁵C to U and T, respectively. Therefore, despite CBT dimers are more frequently found among the photodimers detected in DNA in different conditions [6], CC and m⁵Cm⁵C sites are potentially more damaging, because they might efficiently give rise to an actual mutation. Additionally, and in general, CPDs seem to undergo more rapid deamination than individual nucleobases [8–11]. The important role of the noncanonical m⁵C derivative [12], which is found in significant amounts in the DNA of many eukaryotic organisms (5% in human and calf thymus DNA and 31% in wheat DNA [13, 14]), is recognized since a decade ago. Whereas earlier works were unsuccessful in detecting significant amounts of m⁵C-containing CPDs [15, 16], it was found later that cyclobutane dimers of m⁵C are formed when irradiated with either UV-C or UV-B [12, 17, 18] and that methylation increases the photoproduction with respect to the canonical nucleobase [17]. Tommasi et al. [18] analyzed the CPDs formation in different combinations of pyrimidine nucleobases irradiated with UV-C, UV-B, and sunlight. The methylated DNA base m⁵C was the preferred target for CPD production when the natural sunlight was used.

Femtosecond spectroscopy has proved that T-dimerization is an ultrafast photoreaction in which CBTs are

fully formed ~1 ps after UV illumination [19]. From a theoretical standpoint, a few studies have confirmed for CBC [20] and CBT [21–24] dimers an ultrafast nonadiabatic photoreaction involving a barrierless path along the low-lying singlet excited (S₁) state. The concerted mechanism for the [2 + 2] photocycloaddition of two C- or T-molecules is mediated by the presence of a conical intersection (CI), an energy-degenerate structure between the low-lying singlet excited (S₁) and the ground state (S₀). The shearing-type CI structure—in which the nucleobases ethylenic C₅–C₆ and C_{5'}–C_{6'} bonds laid parallel (parallelogram type) and elongated—, connects the S₁ and S₀ states and allows an efficient internal conversion process [20]. Intrastrand nucleobase sequence and relative orientations were also proved to be essential for an efficient photoreaction to take place. In previous works on CC, TT, and UU pairs [20, 23, 25], we showed that those conformations maximizing the overlap between the π structures of stacked nucleobases formed favorable excimer arrangements, being the most stable structures leading to the photoreactive arrangements, in agreement with the higher yields obtained for photoproducts with *cis*-type parallel face-to-face conformations for the base pairs. Additionally, we determined that the formation of CPDs can be also obtained in the triplet manifold through a biradical intermediate involving a singlet–triplet crossing (S₀/T₁)_X relating the ground (S₀) and low-lying triplet state (T₁) [20, 26, 27], and explaining the high yields of CPDs detected in solution in presence of external photogenotoxic substances acting as triplet–triplet photosensitizers [3, 28, 29].

Obtaining an accurate mapping of the relative energies of the excimer and CI structures in the different dimers is crucial to understand the formation mechanisms and the observed photoreaction yields. Earlier, we attributed the low yield measured for the CBC formation when compared to CBT to the competitive presence of stable excimer conformations and the CI responsible for nucleobase monomer deactivation (CI_{mon}) at energies similar or lower than the [2 + 2] photocycloaddition sheared-type CI (CI_{dim}) in CC. In contrast, for TT CI_{dim} is the most stable structure [23, 26], favoring the reactive process. The goal of the present contribution is to complement the theoretical study on the CPDs formation in cytosine and thymine homodimers with the analysis of the photodimerization in the other pyrimidine nucleobase, uracil, and the non-canonical C₅-methylated cytosine base, providing a wider overview of the photodimerization process in pyrimidines. It is intended to give a rationale on the predominant presence of some specific photodimers over others in terms of the differences found among the potential energy hypersurfaces of the ground state and the lowest excited state related to the [2 + 2] photocycloaddition. Concerning U and m⁵C, and except for a previous DFT study on m⁵C [30], to our knowledge no reliable study at the required multiconfigurational level has been reported. Therefore, the present work is the first ab initio theoretical determination of the mechanism of photodimerization in uracil and the noncanonical m⁵C nucleobase, allowing an overall comparison of the CPDs formation mechanisms for the most important nucleobases homodimers.

2 Methods and computational details

In order to provide a comparative study of the cyclobutane dimer formation in pyrimidines, the same methodology as that previously employed in the cytosine and thymine homodimers has been used in the present work for the theoretical determination of the mechanism of photodimerization of two uracil and two 5-methylcytosine dimers [20, 23, 26]. The ANO-S basis set contracted to C,N,O[3s2p1d]/H[2s1p] was used throughout. Multiconfigurational CASPT2(14/10) and CASPT2(12/12) calculations [31–34] were performed, respectively, for the monomers and the dimers. This included CASSCF geometry optimizations of the singlet states minima, minimum energy paths (MEPs), and minimum energy crossing points (MECPs) determinations in the potential energy hypersurfaces (PEHs) of a system of two nucleobases, corrected at the CASPT2 level using point-wise calculations [20, 35]. Energies were also corrected on the effect of the basis set superposition error (BSSE) using the counterpoise (CP) procedure [20, 36]. As shown previously [20, 25, 37], the

inclusion of BSSE is crucial to accurately describe the binding energies and compare the different mechanisms. The MOLCAS quantum-chemistry code was employed throughout [38]. In order to minimize weakly interacting intruder states, the imaginary level-shift technique with a parameter 0.2 au has been employed [39]. In order to mimic the interaction of pyrimidines in the biologically relevant *cis-syn* diastereoisomer, geometry optimizations were initially performed within the constraints of the C_s symmetry. Since the CASSCF structures of the dimer on the MECPs do not represent a crossing at the CASPT2 level, MECPs were finally obtained with the CASPT2 methodology. Further technical details can be found in previous publications [20, 23].

3 Results and discussion

As mentioned in the introduction, previous experimental and theoretical studies [19–23, 40, 41] determined a general mechanism for the ultrafast photoproduction of CPDs in nucleobases oligomers strands. It is suggested that after initial radiative population of delocalized exciton states on the nucleobases multimers, the system evolves in an ultrafast manner to either a localized excited state of the nucleobase monomer or a fluorescent excimer/excplex state, depending on the smaller or larger degree of stacking, respectively [37]. Relaxation along the monomer path should be ultrafast ($\tau < 2$ ps), as known in the isolated systems [34, 42, 43], whereas from the long-lived excimer/excplex state the system is expected to decay to the ground state slowly ($\tau > 10$ ps). The slower relaxation paths for the stacked nucleobases, expanding from 10 to 200 ps, have been found dominant in the decay dynamics of dinucleotides after excitation at 267 nm (4.96 eV) [19], and it can be assigned to the formation of more or less stable excimer/excplex structures. There are additional accessible evolution paths for the pairs of bases, in particular for pyrimidine nucleobases, in which the excimer/excplex behave as precursor for the formation of cycloadducts like CPDs or (6-4)PPs dimerized structures. As shown, there is always a (S_0/S_1)_{CI} degeneracy region responsible for the photoprocess to take place. In order to elucidate the basics of the relaxation mechanisms in DNA photochemistry, it is required to provide a common framework in which all basic structures—CI of the monomer, excimers, and CI of the photoreaction—are determined at the same level of theory. The comparison between the mechanisms of photodimerization in the considered nucleobases will help to explain the distinct efficiency for photodimers production found in the experiments. Table 1 contains the relative energies, with respect to two isolated nucleobases in the ground state and selected

Table 1 Selected structural distances (R/Å) and energies (ΔE /eV) of the relevant structures along the singlet manifold in the photoproduction mechanism of cyclobutane pyrimidine dimers

Monomer/homodimer	R (C ₅ –C _{5'})	R (C ₆ –C _{6'})	ΔE
Mon (S₁)			
C* + C	–	–	4.41 ^a
m ⁵ C* + m ⁵ C	–	–	4.31 ^b
U* + U	–	–	5.02 ^a
T* + T	–	–	4.89 ^a
CI_{mon} (S₁/S₀)_{CI}			
C	–	–	3.60 ^a
m ⁵ C	–	–	3.64 ^b
U	–	–	3.90 ^a
T	–	–	4.00 ^a
Exc (S₁)			
C*C	3.427	3.219	3.31 ^c
m ⁵ C*m ⁵ C	3.594	3.346	3.46
U*U	2.503	2.315	3.68
T*T	2.650	2.380	3.64 ^d
CI_{dim} (S₁/S₀)_{CI}			
CC	2.258	2.170	3.51 ^c
m ⁵ Cm ⁵ C	2.491	2.056	3.56
UU	2.218	2.170	3.47
TT	2.350	2.220	3.26 ^d
(CPD)_{S0}			
CBC	1.611	1.601	0.78 ^c
CBm ⁵ C	1.648	1.593	0.71
CBU	1.601	1.568	0.23
CBT	1.637	1.595	0.40 ^d
(CPD)_{S1}			
CBC	1.611	1.601	5.35 ^c
CBm ⁵ C	1.648	1.593	5.63
CBU	1.601	1.568	5.87
CBT	1.637	1.595	5.88 ^d

All energies are referred to the singlet ground state of two isolated nucleobases (see Figs. 2 and 3)

^a Ref. [34]

^b This work

^c Ref. [20]

^d Ref. [23]

geometry parameters for the five relevant structures in the photodimerization mechanism in the singlet manifold: the monomeric UV absorption (Mon), the monomer CI (CI_{mon}), the excimer minimum (Exc), the dimer CI (CI_{dim}), and the photodimer (CPD) in the four homodimers studied here: thymine, uracil, cytosine, and 5-methylcytosine. Figure 2 summarizes the basics of the mechanism.

UV solar radiation extends from the edge of the visible light at 400 nm (3.1 eV) to the far UV range (200 nm, 6.2 eV) and beyond. As mentioned previously, the yield of

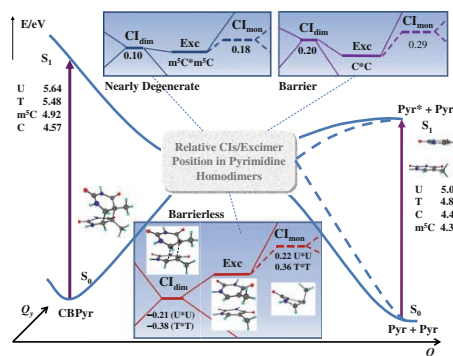


Fig. 2 Proposed scheme, based on actual CASPT2 results, for the decay path of the lowest singlet excited state S₁ of the U, T, C, and m⁵C dimers involving the relaxed excimer and the conical intersection (S₁/S₀)_{CI} leading to ground state cyclobutane pyrimidines (CBPyr or CPDs). Values inside the boxes correspond to the energies of the conical intersection of the dimer (CI_{dim}) and the monomer (CI_{mon}) structures with respect to the lowest-energy excimer (Exc). TT structures shown as an illustration

photoproduction of CPDs is wavelength dependent and increases with the energy of the absorbed radiation, because the available excess energy required to surmount energy barriers is larger for high-energy irradiations. In any case, the photoreactive process takes place ultimately in the S₁ excited state. The vertical excitation energy computed for the S₁ (HOMO → LUMO) state ranges from 4.3 eV in m⁵C to 5.0 in U for the four studied nucleobases. Independently from the procedure that the S₁ is reached—directly or from decay from higher-lying singlet states—the nucleobases strand will find many arrangements in which two of the monomers will overlap their π structures, yielding very favorable conformations ready to evolve toward stable excimer/excimer (excimer here since we are dealing with homodimers) minima in the excited state. Even when for weakly stacked pairs the excitation may localize in the monomer and decay to the ground state of the nucleobase through the CI_{mon} (S₀/S₁)_{CI} (see Table 1), many arrangements will be favorable for the formation of excimers. Even in their most common biological conformation, B-DNA, it is considered that nucleobases form weakly interacting or static excimers [37]. These two types of situations, decay in the monomer through localization with access to the monomer CI and formation of excimers, can be considered responsible for the ultrafast (>2 ps) and fast (>10 ps) decays observed in femtosecond transient absorption experiments, especially in purine strands [19].

Although different excimer arrangements of distinct stability are possible, we have obtained the most stable

excimer in the four homodimers studied here as a face-to-face quasi-parallel conformation that maximizes the π overlap (see Fig. 3 as an example). At the minimum (see Table 1), the intramolecular distances between the analogous C=C double bond atoms of the adjacent monomers (~ 2.3 – 2.6 Å) are almost one Å shorter for T*T and U*U when compared to the other homodimers. T*T and U*U seem to display a stronger interaction (probably because of the two carbonyl groups), showing bonding energies with respect to the isolated monomers near 1.3 eV. In contrast, C*C and m⁵C*m⁵C have larger intermonomer distances (~ 3.2 – 3.6 Å) and lower binding energies (~ 0.9 – 1.1 eV). The four homodimers have also a CI degeneracy region connecting the excited S₁ and the ground S₀ state in which the system displays very similar structures, with the two monomers arranged in a quasi-parallel shearing-type conformation (see Fig. 3). This CI_{dim} (S₀/S₁)_{CI} structure is responsible for the nonadiabatic [2 + 2] cyclophotoaddition reaction leading from the pair of stacked nucleobases to the final CPD photoproduct in the ground state.

The relative position of the most stable excimer conformation and the CI_{dim} (S₀/S₁)_{CI} region is different in the four homodimers studied. For TT and UU pairs, CI_{dim} is the lowest-energy structure of all studied S₁ hypersurface, lying near 0.4 and 0.2 eV, respectively, below the stable excimer. It is not surprising that all the computed minimum energy paths (MEPs) in TT leads in a barrierless way from different excimer arrangements to the CI_{dim} structure [21–23], as it is the case here with UU. Additionally, the CI_{dim} structure is near 0.7 (TT) and 0.4 eV (UU) lower in energy than the CI of the respective monomers, T and U. Therefore, there is no decay process that can efficiently compete with the access to CI_{dim} and subsequent nonadiabatic

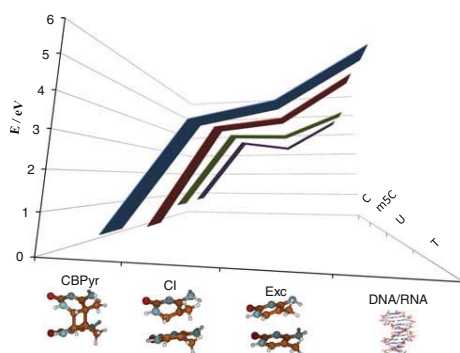


Fig. 3 Comparison of the photoreactive pathways leading to CPDs (or CBPyr) formation along the singlet manifold for the C, m⁵C, U, and T base pairs. The m⁵Cm⁵C structures obtained in this work are shown

transfer of energy to reach CBT or CBU, except for those systems evolving toward the splitting of the monomers. The described PEH profile is perfectly compatible with the observed high yields of production of CBT and CBU when compared with other adducts [3].

Regarding CC and m⁵Cm⁵C, the CI_{dim} structure has been computed 0.2 and 0.1 eV, respectively, higher in energy than their relaxed excimer minima. That means that there exist conformations that will become competitive with the nonadiabatic reaction, decreasing the rate and yield of the photoprocess in the two systems, especially in CC. The same argument can be used when comparing the dimer CI with the monomer CI, which is energetically just 0.1 eV above in both cases, opening a new decay route which may compete efficiently with the formation of CBC and CBm⁵C. Overall, the lower yields observed for CC tandems in contrast to TT [44] can be therefore understood by the presence in C-based dimers of several competitive structures—stable excimers and monomer decay routes—close in energy to the reactive CI. Regarding the comparison of CC and m⁵Cm⁵C, the former displays excimer structures somewhat more stable with respect to the CI_{dim} (0.2 eV) than in the case of the methyl derivative, in which CI_{dim} and excimer become almost degenerate. Even when the difference is small in both cases, this profile may explain the slightly higher efficiency found in the production of m⁵C dimers [12, 17]. Figure 3 depicts a scheme of the relative energy levels for the studied systems.

With respect to the geometry of the dimer CI in the four systems studied (rhomboid or parallelogram type, typical of the [2 + 2] cycloaddition [20]), and although the structural discrepancies are small, they might be relevant (see Fig. 3). Whereas the C₆–C₆' distance remains similar in the various cases, C₅–C₅' shows much more noticeable differences, reflecting a major or minor distortion with respect to the ideal rhomboid geometry. Hence, the m⁵C dimer presents the most distorted structure, with the largest C₅–C₅' distance, 2.491 Å, among the studied bases. This elongation is probably a consequence of the steric effect caused by the presence of both the methyl and amino groups. Thymine, also with a methyl group, comes next with a value of 2.350 Å, followed by the less distorted cytosine and uracil systems, with distances of 2.258 and 2.218 Å, respectively. As it could be expected, methylation increases the C₅–C₅' bond length in all computed structures, excimers, CIs, and adducts.

Once the CI_{dim} crossing is reached, dimers can evolve, in an ultrafast process, to form cyclobutane pyrimidines, whose structure is shown for the m⁵C in Fig. 3. The singlet pathway represented in this figure and Fig. 2 allows us to better understand the photodimer production mechanism that operates after the initial excitation taking place in DNA/RNA strands. Also, photodimers can absorb UV light

to finally obtain the separated base monomers through a photoreversibility process as shown in Fig. 2. In a previous study, performed using the CASPT2/MM methodology for the $dG_{18}xdC_{18}$ system, i.e., a 18-base-pair-long double helix of poly(C)-poly(G) surrounded by water molecules [45], we determined the barrierless path connecting the CBC S_1 state with the dimer CI, which may later lead to the separation of the monomers. Although the required excitation energy in the dimer to initiate the photoreversion process (4.57–5.64 eV, within the UV-C range) is higher than the energy needed for the monomer excitation, it represents a competitive photoreaction that may reduce the yield of the photoadduct. The formation of CC photodimers was studied by Tommasi et al. [18] under UV-B and UV-C irradiations, showing an increase in the CPD production when UV-B was used, in clear agreement with the results obtained here.

4 Conclusions

The formation of cyclobutane pyrimidine dimers following UV absorption of pairs of stacked DNA/RNA nucleobases has been studied for the homodimers of thymine, uracil, cytosine, and 5-methylcytosine at the theoretical, ab initio CASPT2 multiconfigurational level. Determination of states minima, minimum reaction paths, and conical intersections in the low-lying singlet states of the dimers and monomers has led us to establish a unified mechanism for the adduct formation, which proceeds via a nonadiabatic [2 + 2] photocycloaddition reaction in the singlet manifold. A sheared-like conical intersection connecting the S_1 and S_0 states of the dimer, $CI_{dim}(S_0/S_1)_{CI}$, is the funnel controlling the reactive process. The relative position of the dimer CI with respect to stable excimer structures or the monomer decay CI region determines the efficiency of the photoreactivity. Thus, the dimer CI is the lowest-energy structure in thymine and uracil homodimers, lacking other direct competitive decay processes and favoring the higher efficiency observed in the formation of the CBT and CBU adducts. In contrast, the dimer CI computed in cytosine and 5-methylcytosine is close in energy to the most stable excimer conformation, displaying a face-to-face quasi-parallel structure, and also to the corresponding monomer CIs. Those features can easily compete with the nonadiabatic dimerization reaction decreasing its rates and yields. In fact, the stable excimer structure is lower in energy than the dimer CI in the latter systems, in which a 0.2- and 0.1-eV barrier, respectively, is found to reach the degeneracy region. Those profiles can explain the higher yields found in TT sites when compared with CC. Also, it is shown that methylation, in T and m^5C with respect to U and C, respectively, destabilizes the excimer

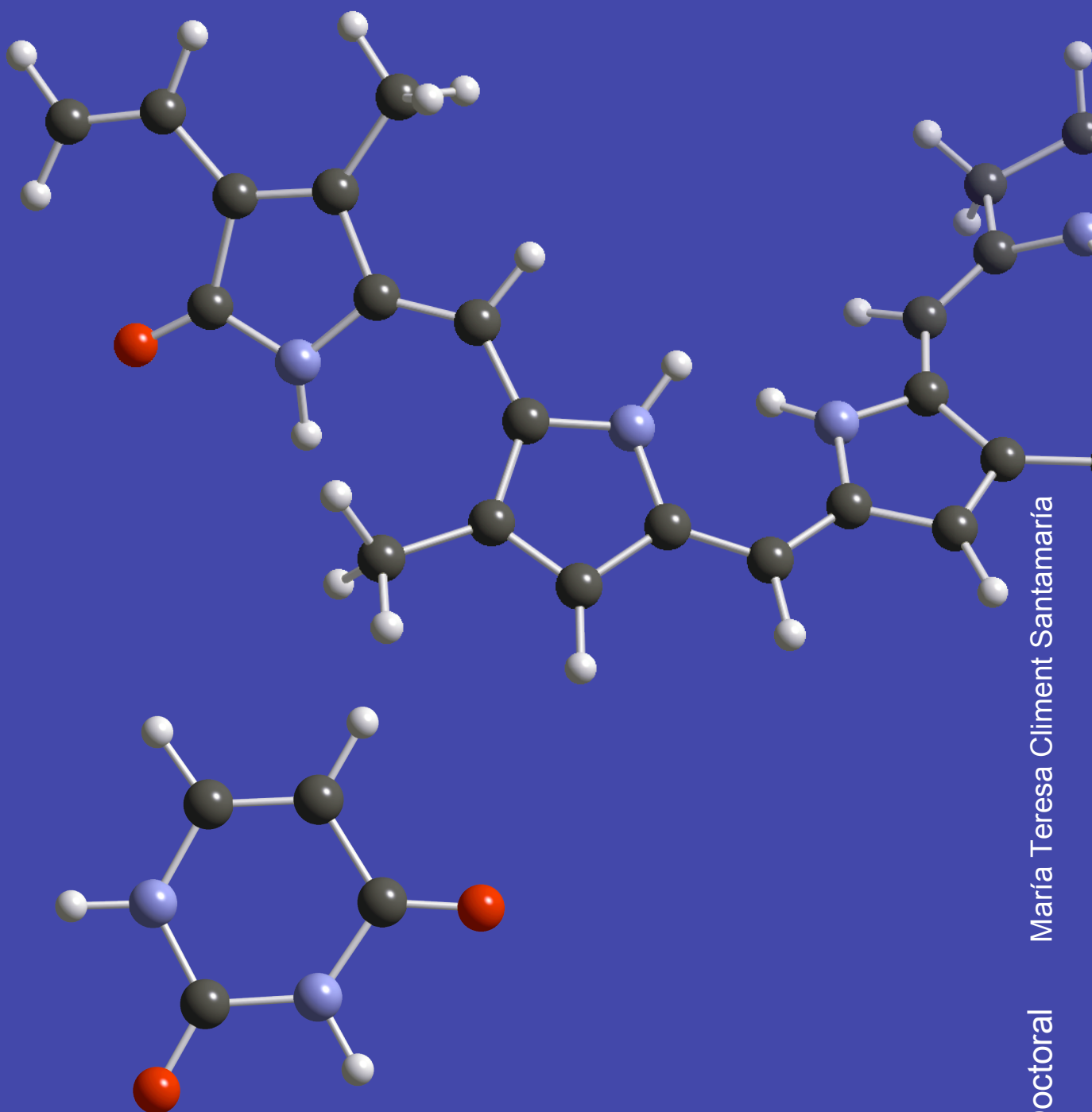
structures with respect to the CI. Thus, in the case of cytosine, the methylation favors the probability for dimerization by decreasing energy barriers leading to the reaction. It must be highlighted that a high level of theory is required to obtain accurate profiles. CASPT2, including exhaustively the correlation energy, in contrast to simply CASSCF (or lower level approaches like TDDFT) is required to remove undesired differential correlation effects [33, 46], whereas the inclusion of the BSSE effect is essential to obtain accurate and comparable binding energies among the different systems.

Acknowledgments In memory of Luis Serrano-Andrés, he was an excellent scientist and friend. Research supported by projects CTQ2007-61260, CTQ2010-14892, and CSD2007-0010 Consolider-Ingenio in Molecular Nanoscience of the Spanish MEC/FEDER and the Generalitat Valenciana. It has also received funding from the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013)/ERC grant agreements n° 255363 and n° 251955.

References

1. Olivera BM (1978) *Proc Natl Acad Sci USA* 75:238–242
2. Setlow RB (1974) *Proc Natl Acad Sci USA* 71:3363–3366
3. Cadet J, Vigny P (1990) In: Morrison H (ed) *Bioorganic photochemistry*. Wiley, New York, p 1
4. Friedberg EC, Walker GC, Siede W, Wood RD, Schultz RA (2006) Ellenberger T (ed) *DNA Repair and mutagenesis*. ASM Press, Washington, DC
5. Björn LO, McKenzie RL (2008) In: Björn LO (ed) *Photobiology. The science of life and light*. Springer, New York, p 503
6. Douki T, Cadet J (2001) *Biochemistry* 40:2495–2501
7. Schuch AP, Menck CFM (2010) *J Photochem Photobiol B Biol* 99:111–116
8. You YH, Li C, Pfeifer GP (1999) *J Mol Biol* 293:493–503
9. You YH, Pfeifer GP (2001) *J Mol Biol* 305:389–399
10. Ikehata H, Masuda T, Sakata H, Ono T (2003) *Environ Mol Mutagen* 41:280–292
11. Lee DH, Pfeifer GP (2003) *J Biol Chem* 278:10314–10321
12. Shetlar MD, Basus VJ, Falick AM, Mujeeb A (2004) *Photochem Photobiol Sci* 3:968–979
13. Friso S, Choi SW, Dolnikowski GG, Selhub J (2002) *Anal Chem* 74:4526–4531
14. Adams RLP, Burdon RH (1984) *Molecular biology of DNA methylation*. Springer, New York, pp 6–7 (Table 1.2)
15. Ehrlich M, Dove MF (1983) *Photobiochem Photobiophys* 6:121–126
16. Ehrlich M, Dove MF, Huang LH (1986) *Photobiochem Photobiophys* 11:73–79
17. Mitchell DL (2000) *Photochem Photobiol* 71:162–165
18. Tommasi S, Denissenko MF, Pfeifer GP (1997) *Cancer Res* 57:4727–4730
19. Schreier WJ, Schrader TE, Soller FO, Gilch P, Crespo-Hernández CE, Swaminathan VN, Carell T, Zinth W, Kohler B (2007) *Science* 315:625–629
20. Roca-Sanjuán D, Olaso-González G, González-Ramírez I, Serrano-Andrés L, Merchán M (2008) *J Am Chem Soc* 130:10768–10779
21. Boggio-Pasqua M, Groenhof G, Schäfer LV, Grubmüller H, Robb MA (2007) *J Am Chem Soc* 129:10996–10997

22. Blancafort L, Migani A (2007) *J Am Chem Soc* 129:14540–14541
23. Serrano-Pérez JJ, González-Ramírez I, Coto PB, Merchán M, Serrano-Andrés L (2008) *J Phys Chem B* 112:14096–14098
24. Zhang WY, Yuan SA, Li AY, Dou YS, Zhao JS, Fang WH (2010) *J Phys Chem C* 114:5594–5601
25. González-Ramírez I, Climent T, Serrano-Pérez JJ, González-Luque R, Merchán M, Serrano-Andrés L (2009) *Pure Appl Chem* 81:1695–1705
26. Climent T, González-Ramírez I, González-Luque R, Merchán M, Serrano-Andrés L (2010) *J Phys Chem Lett* 1:2072–2076
27. González-Luque R, Climent T, González-Ramírez I, Merchán M, Serrano-Andrés L (2010) *J Chem Theor Comp* 6:2103–2114
28. Bosca F, Lhiaubet-Vallet V, Cuquerella MC, Castell JV, Miranda MA (2006) *J Am Chem Soc* 128:6318–6319
29. Kwok WM, Ma C, Phillips DL (2008) *J Am Chem Soc* 130:5131–5139
30. Xiaoyi L, Eriksson LA (2005) *Chem Phys Lett* 401:99–103
31. Andersson K, Malmqvist PÅ, Roos BO (1992) *J Chem Phys* 96:1218–1226
32. Roos BO, Andersson K, Fülischer MP, Malmqvist PÅ, Serrano-Andrés L, Pierloot K, Merchán M (1996) *Adv Chem Phys* 93:219
33. Merchán M, Serrano-Andrés L (2005) In: Olivucci M (ed) *Computational photochemistry*. Elsevier, Amsterdam, p 35
34. Serrano-Andrés L, Merchán M (2009) *J Photochem Photobiol C Photochem Rev* 10:21–32
35. Serrano-Andrés L, Merchán M, Lindh R (2005) *J Chem Phys* 122:104107–104110
36. Boys SF, Bernardi F (2002) *Mol Phys* 100:65
37. Olaso-González G, Merchán M, Serrano-Andrés L (2009) *J Am Chem Soc* 131:4368–4377
38. Aquilante F, De Vico L, Ferré N, Ghigo G, Malmqvist PÅ, Pedersen T, Pitonak M, Reiher M, Roos BO, Serrano-Andrés L, Urban M, Velyazov V, Lindh R (2010) *J Comp Chem* 31:224–247
39. Forsberg N, Malmqvist PÅ (1997) *Chem Phys Lett* 274:196–204
40. Crespo-Hernández CE, De La Harpe K, Kohler B (2008) *J Am Chem Soc* 130:10844–10845
41. Takaya T, Su C, De La Harpe K, Crespo-Hernández CE, Kohler B (2008) *Proc Natl Acad Sci USA* 105:10285–10290
42. Crespo-Hernández CE, Cohen B, Hare PM, Kohler B (2004) *Chem Rev* 104:1977–2020
43. Serrano-Andrés L, Merchán M (2008) Shukla MK, Leszczynski J (eds) *Photostability and photoreactivity in biomolecules: quantum chemistry of nucleic acid base monomers and dimers*. In: *Radiation induced molecular phenomena in nucleic acids: a comprehensive theoretical and experimental analysis*. Springer, The Netherlands, p 435
44. Douki T, Cadet J (2001) *Biochemistry* 40:2495–2501
45. Roca-Sanjuán D, Olaso-González G, Rubio M, Coto PB, Merchán M, Ferré N, Ludwig V, Serrano-Andrés L (2009) *Pure Appl Chem* 81:743–754
46. Serrano-Andrés L, Merchán M (2005) *J Mol Struct Theochem* 729:99–108



Tesis Doctoral María Teresa Climent Santamaría

2012