



UNIVERSITAT DE
BARCELONA

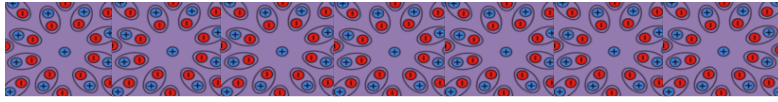
Estudio del impacto de la polarización electrónica en interacciones biomoleculares

Silvana de Souza Pinheiro

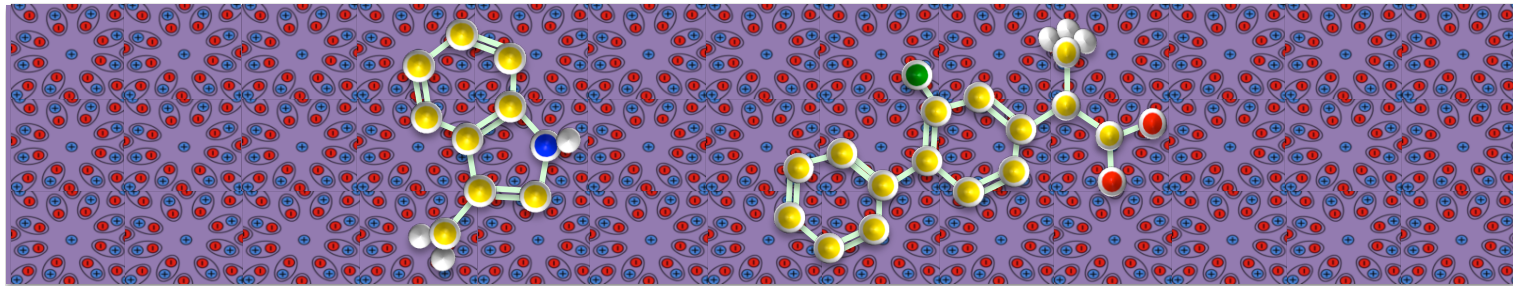
ADVERTIMENT. La consulta d'aquesta tesi queda condicionada a l'acceptació de les següents condicions d'ús: La difusió d'aquesta tesi per mitjà del servei TDX (www.tdx.cat) i a través del Dipòsit Digital de la UB (diposit.ub.edu) ha estat autoritzada pels titulars dels drets de propietat intel·lectual únicament per a usos privats emmarcats en activitats d'investigació i docència. No s'autoritza la seva reproducció amb finalitats de lucre ni la seva difusió i posada a disposició des d'un lloc aliè al servei TDX ni al Dipòsit Digital de la UB. No s'autoritza la presentació del seu contingut en una finestra o marc aliè a TDX o al Dipòsit Digital de la UB (framing). Aquesta reserva de drets afecta tant al resum de presentació de la tesi com als seus continguts. En la utilització o cita de parts de la tesi és obligat indicar el nom de la persona autora.

ADVERTENCIA. La consulta de esta tesis queda condicionada a la aceptación de las siguientes condiciones de uso: La difusión de esta tesis por medio del servicio TDR (www.tdx.cat) y a través del Repositorio Digital de la UB (diposit.ub.edu) ha sido autorizada por los titulares de los derechos de propiedad intelectual únicamente para usos privados enmarcados en actividades de investigación y docencia. No se autoriza su reproducción con finalidades de lucro ni su difusión y puesta a disposición desde un sitio ajeno al servicio TDR o al Repositorio Digital de la UB. No se autoriza la presentación de su contenido en una ventana o marco ajeno a TDR o al Repositorio Digital de la UB (framing). Esta reserva de derechos afecta tanto al resumen de presentación de la tesis como a sus contenidos. En la utilización o cita de partes de la tesis es obligado indicar el nombre de la persona autora.

WARNING. On having consulted this thesis you're accepting the following use conditions: Spreading this thesis by the TDX (www.tdx.cat) service and by the UB Digital Repository (diposit.ub.edu) has been authorized by the titular of the intellectual property rights only for private uses placed in investigation and teaching activities. Reproduction with lucrative aims is not authorized nor its spreading and availability from a site foreign to the TDX service or to the UB Digital Repository. Introducing its content in a window or frame foreign to the TDX service or to the UB Digital Repository is not authorized (framing). Those rights affect to the presentation summary of the thesis as well as to its contents. In the using or citation of parts of the thesis it's obliged to indicate the name of the author.

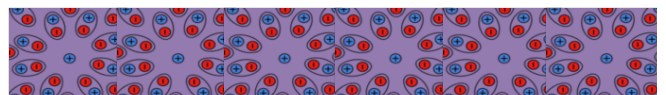


Estudio del Impacto de la Polarización Electrónica en Interacciones Biomoleculares



Silvana de Souza Pinheiro

*Tesis Doctoral
2017*





UNIVERSITAT DE
BARCELONA

UNIVERSITAT DE BARCELONA

FACULTAT DE FARMÀCIA I CIÈNCIES DE L'ALIMENTACIÓ
DEPARTAMENT DE FARMÀCIA, TECNOLOGIA FARMACÈUTICA I
FISICOQUÍMICA

**ESTUDIO DEL IMPACTO DE LA POLARIZACIÓN ELECTRÓNICA
EN INTERACCIONES BIOMOLECULARES**

SILVANA DE SOUZA PINHEIRO

BARCELONA, 2017



UNIVERSITAT DE
BARCELONA

UNIVERSITAT DE BARCELONA
FACULTAT DE FARMÀCIA I CIÈNCIES DE L'ALIMENTACIÓ
DEPARTAMENT DE FARMÀCIA, TECNOLOGIA FARMACÈUTICA I
FISICOQUÍMICA
PROGRAMA DE DOCTORAT EN RECERCA, DESENVOLUPAMENT I CONTROL DE
MEDICAMENTS

**ESTUDIO DEL IMPACTO DE LA POLARIZACIÓN ELECTRÓNICA
EN INTERACCIONES BIOMOLECULARES**

Memoria presentada por Silvana de Souza Pinheiro para optar al título de Doctora por la
Universidad de Barcelona.

Dr. Carles Eduard Curutchet Barat
Director

Dr. Ramon Pouplana i Solé
Director

Silvana de Souza Pinheiro
Doctoranda

Dr. Ramon Pouplana i Solé
Tutor

SILVANA DE SOUZA PINHEIRO

BARCELONA, 2017

*A vocês papai e mamãe, com
todo meu amor e gratidão.*

Agradecimientos

*Antes de tudo, quero agradecer a Deus, por ter abençoado todos os dias da minha vida, iluminar sempre o meu caminho e me dar força e muita coragem para não desistir diante de tantos obstáculos. Te agradeço meu Deus por ter me acompanhado e guiado ao longo de minha carreira, por ser minha fortaleza nos momentos de debilidade e por brindar-me com uma vida cheia de aprendizagens, experiências e aventuras. Agradeço também pelas pessoas que o Senhor colocou em meu caminho. Algumas delas me inspiram, me ajudam, me desafiam e me encorajam a ser cada dia melhor. Eu Te agradeço, por todas as coisas boas e más que me aconteceram. Cada uma delas, ao seu modo, me fizeram chegar onde eu cheguei, e me fizeram ser quem eu sou. Foi a minha jornada de tropeços, vitórias e derrotas, que me fez enxergar o verdadeiro significado e beleza da vida. “Tudo posso Naquele que me fortalece! **Filipenses 4:13**”*

En la fase final de esta “*superaventura*” de ir a otro lado del océano a estudiar, de dejar a la familia, trabajo, amigos, en fin, de abandonar una vida y aventurarme en una nueva meta por cumplir tanto a nivel personal como profesional. Al mirar hacia atrás y ver todo lo dejado, al mirar en el presente y ver todo lo aprendido, conquistado y disfrutado de esta experiencia y al mirar hacia el futuro y ver las posibilidades que hay en mi camino, es preciso agradecer a quienes han participado conmigo de esta “*superaventura*”. Así que, intentaré resumir en unas líneas la gratitud que siento a todas estas personas.

Agradezco a mi director de tesis Dr. Carles Curutchet, muchas gracias *Curu* por guiarme en el reto de realizar una tesis en un tema súper complejo, que a principio lo veía todo muy raro y confieso que ganas de salir corriendo no me han faltado, pero luego con su manera tranquila y calma, me lo hizo entender el complejo mundo de la *energy transfer*. Gracias por recibirme de manera muy cariñosa en tu *petit* grupo de las “*Curutchetas*”. Eres de verdad un súper jefe pero también una gran persona. Gracias por haber estado siempre disponible, ya sea para arreglar los infinitos papeleos, o para discutir/discordar de cualquier tema. Gracias no solamente por ayudarme en el desarrollo de esta bonita tesis, sino también en mi formación como investigadora. Por depositar su confianza en mí, ofreciéndome libertad y también apoyo constante, lo que me ha permitido, no solamente aprender muchísimo del mundo cuántico, sino disfrutar enormemente durante el aprendizaje en estos cuatro años. De corazón será un orgullo que haya sido mi director de tesis y además ser la primera “*Dr. Curutcheta*.”

Agradezco al Dr. Javier Luque, muchas gracias *Javi* por su disponibilidad y por su gran colaboración en esta importante etapa de mi vida. Por ser un jefe ejemplar y sobretodo una gran persona. Le agradezco que me haya abierto las puertas de tu grupo de investigación *Computational Biology & Drug Design*, dándome la oportunidad de tener una visión más amplia del mundo de la investigación y sobretodo por ayudarme cuando lo he necesitado. Al Dr. Ramon Pouplana, muchas gracias por tu colaboración en esta tesis, siendo con sus sugerencias, con tus historias graciosas, o con *les seves ricas calçotades*.

También agradezco a todas aquellas personas con las que, en mayor o menor medida y durante diferentes periodos de tiempo, he tenido contacto en estos cuatro años. Son muchos los compañeros con los que he compartido inolvidables momentos: A los primeros compañeros de trabajo que me han recibido tan bien e hicieron de la fase de adaptación menos difícil. Gracias por enseñarme de todo: desde degustar un buen vino y una rica comida (Ornella), *parlar una mica el catalá* (Marina) hasta los interminables *scripts* (Jordinho). Constantí, por hacerme reír con su acento brasileño *poderosooo*. A Toni y Carol y Axel por sus inúmeras ayudas con mis ordenadores. A Salo y Jordi, muchas gracias por ayudarme cuando lo he necesitado y sobretodo por haberme salvado de los abuelos malos. Les agradezco que me hayan abierto las puertas de su súper dúplex en *el carrer d'En Cortines 6*, mismo cambiando la decoración, no me han echado. A la *mara* de Roc, Marina, por su disponibilidad y súper *dicas* de ahorro. *Moltes gràcies* por recibirme tan bien y ayudarme siempre con mi castellano, mismo siendo una verdadera catalana (*¡Visca el Barça y Visca Catalunya*). Han sido muchas aventuras juntas, de verdad he aprendido y disfrutado muchísimo contigo *nená*. También a los demás “*Curuchetas*” Sonia, Téó y Andreana por su participación en muchos momentos de esta “*superaventura*”. A los primeros “chicos de Xavi”: Dani, Montse, Yvonne, Kevin y Sergi. También a la segunda generación: Míriam, Serena, Guillermo y Galdeano y luego a la tercer generación: Maciej, Moira y Melpomeni. Agradezco a Ignacio por haber empezado este tema y haber sido un colaborador en esta tesis. También a JP Gelpí, François y Chipot. A Campa por su disponibilidad y por presentarnos el *delicios oli d'oliva Siurana*. A Xavi por sus sugerencias en este trabajo. Gracias también a Maria Luisa (Marisa) por su disponibilidad y hacer mi vida más fácil con los inúmeros papeleos. A las “*profes*” del departamento de fisicoquímica de la facultad de Farmacia, muchas gracias por los momentos ricos de las comidas diarias.

Aos meus Amigos, que sempre souberam desculpar minha ausência e sempre tiveram uma palavra de encorajamento e ânimo. Estou absolutamente convencida de que, se algo que eu possa me orgulhar nesta vida é de grandes amigos que tenho, o que me faz sentir uma pessoa de muita sorte. Quero agradecer de montão ao meu amigo Nelsão, por estar sempre disponível, desde o início dessa vida de teórica, com as primeiras dinâmicas e homologias. Muito obrigada Nelson por ter acreditado e sonhado comigo esta conquista, me direcionando e ajudando em uma decisão muito difícil que foi fazer o doutorado no exterior. Agradeço também a minha grande amiga Sol, que desde o mestrado, nunca mais me deixou e sempre esteve comigo altas horas (a sua filha queridinha). As minhas queridas amigas Thaíris, Emília e Cléo por me acompanharem sempre. Aos meus inspiradores amigos garrotes (Renatinha, Anderson e Davis), por me incluírem nesse mundo da pesquisa e ajudarem a fazer do meu currículo *lattes* um verdadeiro passaporte, e graças a isso eu tive a oportunidade de conhecer de norte a sul do meu país. Aos meus amigos Nilmissíma, Alê e Pedro, por cuidarem tão bem, do bem mais precioso que tenho, meus pais. A Cleice e Felipe por estarem sempre torcendo por mim. Ao meu grande amigo Davi por suas orações e lindas

mensagens de ânimo. E a todos os meus lindos *Zamigos*. Os dedico esse título com toda minha gratidão.

Muito obrigada as minhas amigas/os “*brasinholas*”: *Pani* Nubis a primeira que conheci de maneira muito especial, que com todo seu glamour e *popozão*, já imaginava que era brasileira. A minha linda *mami* Pepeu, por todo seu cuidado, carinho e dedicação comigo e Will. Ao meu *papi* Manuel por me ensinar tudo sobre as castas das uvas e me apresentar o delicioso vinho português. Tami, obrigada por ser tão doce e sempre disposta. Joan por los lindos días de playas y las muchas comidas em tu casa, sois muy especiales. Japinha, muito obrigada pelos lindos passeios e por apresentar os melhores restaurantes de Barcelona, eres um grande amigo. Dadá minha flor, obrigada pelo seu carinho. Railda e Lourí, por todo o carinho e amizade. Aos profs. Claudio e Jerônimo por estarem sempre dispostos a ajudar.

Tudo isso não seria possível sem as orações, amparo, ajuda e amor incondicional da minha amada família. Agradeço imensamente aos meus pais Marlete e Manoel, vocês são minha vida, foi muito difícil ficar muito tempo sem o abraço, carinho e amor de vocês, mas mesmo do outro lado do oceano eu sempre me sentir muito amada. Obrigada por confiarem em mim e sonharem junto comigo cada conquista, os amo infinitamente. Aos meus irmãos Avenilson, Silvano, Simone, Silvane e Júnior, agradeço todo o apoio, carinho e compreensão. Aos meus *fofoxos* sobrinhos: Andreiva, Kauã, Kauê, Marllon e Lara, obrigado pelos momentos únicos e pelos milhões de vídeos e fotos engraçadas. Agradeço também aos meus primíssimos e primíssimas, por todo o carinho que recebo de cada um de vocês, de verdade me sinto muito querida pela minha família. Aos meus queridos tios Ailton, por seu grande amor e tio Irivaldo (*in memoriam*), que representam minha grande e amada família, herança da minha inesquecível avó/mãe Osvaldina (*in memoriam*). Este título também é de vocês.

Y para cerrar esta “*superaventura*”, mis mas sinceras gratitudes a Will, por haber estado incondicionalmente siempre a mi lado. Gracias por todo los momento lindos y inolvidables a tu lado. Gracias por estar animándome siempre a continuar y por todos los esfuerzos que ha hecho, por hacerme creer cada día que podía hacerlo, por toda su ayuda, dedicación y amor. Por todo eso y mucho más este trabajo también le pertenece.

Meus agradecimentos ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pela bolsa de doutorado pleno no exterior e por todo o financiamento desta tese. Al Consorci de Serveis Universitaris de Catalunya (CSUC), por proporcionar acceso a recursos computacionales, Agència de Gestió d’Ajuts Universitaris i de Recerca (Generalitat de Catalunya) y al Ministerio de Economía y Competitividad.

A todos, mi mayor reconocimiento y gratitude!

Resumen

Esta tesis doctoral tiene como objetivo profundizar en el impacto que tiene la polarización electrónica en interacciones en sistemas biológicos. La inclusión explícita de estos efectos en campos de fuerzas es uno de los grandes retos de las técnicas de modelización molecular típicamente usadas para el estudio de biomoléculas, donde tales efectos solo se incluyen de forma implícita. La investigación desarrollada se centra en dos casos de interacciones moleculares, tanto a nivel de estado fundamental (interacciones catión- π -catión) como de estados electrónicos excitados (acoplamientos excitónicos), donde se espera que el impacto de la polarización electrónica sea particularmente importante. La tesis demuestra, en este sentido, que los efectos de polarización no aditivos son muy significativos en interacciones catión- π -catión a tres cuerpos, además de demostrar que estas interacciones poco estudiadas son comunes en proteínas y pueden contribuir en la estabilización de su estructura terciaria. Los resultados obtenidos muestran pues que los campos de fuerzas aditivos habitualmente usados en simulaciones no pueden describir correctamente estos contactos, y presenta un conjunto de energías de interacción de referencia que pueden ser usadas en futuros desarrollos de campos de fuerzas polarizables. Respecto a interacciones a nivel de estados excitados, la tesis revela que el entorno heterogéneo polarizable modula de forma significativa procesos de transferencia de energía electrónica que involucran al ligando en complejos proteína-ligando, y que tal impacto depende del modo de unión correspondiente. Esto indica que la interpretación de datos de fluorescencia relacionada la eficiencia de la transferencia únicamente con la distancia intercromofórica, como asume la teoría de Förster típicamente usada en este contexto, supone una aproximación importante. En este sentido, la tesis demuestra el potencial de simulaciones basadas en estructura a fin de interpretar los cambios de fluorescencia asociados a eventos de unión de ligandos a proteínas y ayudar de esta forma a caracterizar nuevos sitios de unión.

Summary

This thesis aims at advancing our comprehension about the impact of electronic polarization in interactions in biological systems. The explicit inclusion of these effects in force fields is one of the big challenges of molecular modeling techniques typically used for the study of biomolecules, where such effects are included only in an implicit way. The research performed is centered in two cases of molecular interactions, both at the level of ground state (cation- π -cation interactions) and excited states (excitonic couplings), where the impact of electronic polarization is expected to be particularly important. The thesis demonstrates, in this context, that polarization non-additive effects are very significant in three-body cation- π -cation interactions, and demonstrates as well that these kind of interactions are common in proteins and can contribute to the stabilization of protein tertiary structures. The results obtained thus show that additive force fields usually adopted in simulations cannot describe correctly such interactions, and provide a set of reference interaction energies that can be used in future developments of polarizable force fields. Regarding interactions at the level of excited states, the thesis points to a significant modulation of electronic energy transfer processes that involve ligands in ligand-protein complexes due to the heterogeneous polarizable environment, such impact depending on the corresponding binding mode. This indicates that the interpretation of fluorescence data that relates changes in transfer efficiency only with interchromophoric distance, as it is assumed in Förster theory, represents an important approximation. In this sense, the thesis demonstrates the potential of structure-based simulations to interpret fluorescence changes associated to ligand binding events and help in the characterization of novel binding sites.

Índice general

Índice

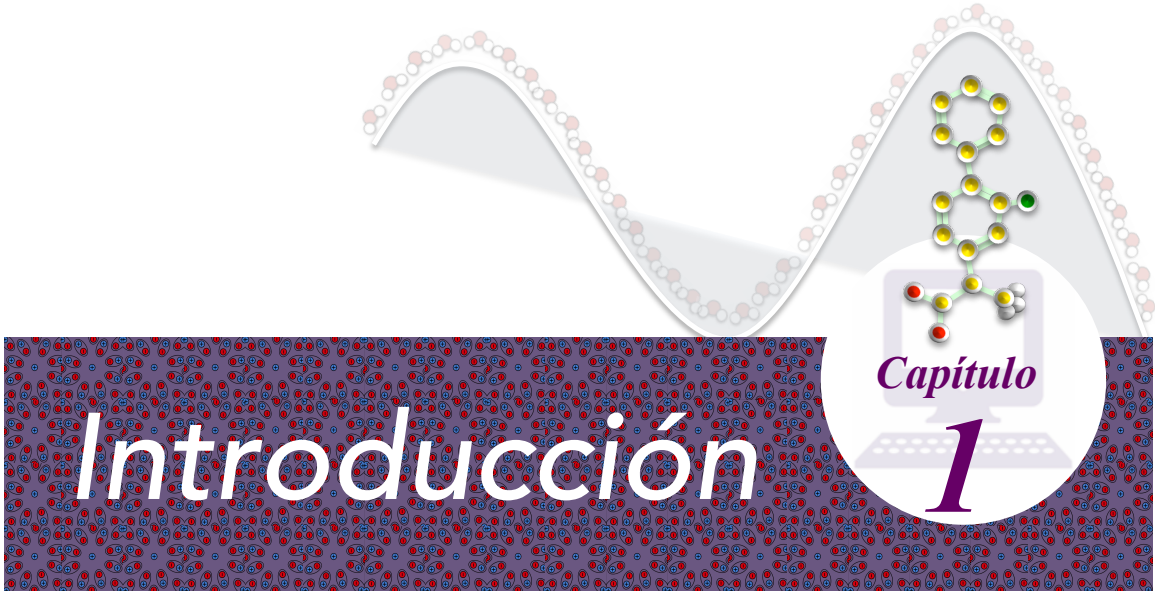
Resumen.....	13
Summary.....	14
Abreviaturas	19
1. INTRODUCCIÓN	25
1.1 Descripción de la polarización en la modelización de biosistemas.....	29
1.1.1 Campos de fuerzas clásicos.....	29
1.1.1.1 Campos de fuerzas aditivos	31
1.1.1.2 Campos de fuerzas polarizables	32
1.1.2 Métodos híbridos QM/MM	36
1.1.2.1 Modelos QM/MM basados en cargas puntuales	37
1.1.2.2 Modelos QM/MM polarizables	38
1.1.2.3 Extensión a TD-DFT	40
1.1.3 Modelos continuos de disolvente	40
1.1.3.1 QM/PCM.....	41
1.1.3.2 MM/GBSA	42
1.2 Impacto de la polarización en interacciones a nivel de estado electrónico fundamental.....	44
1.2.1 Interacciones intermoleculares y descomposición en términos físicos	44
1.2.2 Interacciones catión- π -catión.....	44
1.2.3 Cálculo de energías de interacción.....	46
1.3 Impacto de la polarización en interacciones entre estados excitados	49
1.3.1 Transferencia de energía electrónica.....	49
1.3.2 Teoría de Förster	50
1.3.3 Acoplamiento electrónico.....	52
2. OBJETIVOS	57
2.1 Impacto de las fuerzas de polarización electrónica en interacciones catión-π-catión.....	57
2.2 Impacto de la polarización en procesos de transferencia de energía.....	58

3. RESULTADOS	61
3.1 Artículo 1: Structural and energetic study of cation- π -cation interactions in proteins.....	65
3.2 Artículo 2: Can Förster theory describe stereoselective energy transfer dynamics in a protein-ligand complex?	83
4. DISCUSIÓN	101
4.1 Impacto de la polarización en interacciones a nivel de estado electrónico fundamental	103
4.1.1 Análisis estructural de los motivos catión-π-catión en proteínas	103
4.1.2 Análisis energético de los motivos catión-π-catión en proteínas	104
4.2 Impacto de la polarización en interacciones entre estados excitados	107
4.2.1 Sistema modelo FBP-TrpMe	109
4.2.2 Complejo FBP-HSA	111
5. CONCLUSIONES	119
5.1 Impacto de las fuerzas de polarización electrónica en interacciones catión-π-catión	119
5.2 Impacto de la polarización del entorno en interacciones entre estados excitados	120
Bibliografía	125
Apéndices	133
Apéndice I. Información suplementaria: Structural and energetic study of cation- π -cation interactions in proteins.....	135
Apéndice II. Información suplementaria: Can Förster theory describe stereoselective energy transfer dynamics in a protein-ligand complex?	155

Abreviaturas

AMBER	– / <i>Assisted Model Building with Energy Refinement</i>
B3LYP	– / <i>Becke three-parameter Lee-Yang-Parr exchange-correlation functional</i>
BSSE	Error de Superposición de Base / <i>Basis-Set Superposition Error</i>
CAM-B3LYP	– / <i>Becke three-parameter Lee-Yang-Parr exchange-correlation functional based on the Coulomb-Attenuating Method</i>
CBS	Conjunto Completo de funciones de Base / <i>Complete Basis Set</i>
CC	– / <i>Coupled Cluster method</i>
CCSD	– / <i>Coupled Cluster method with Single and Double excitations</i>
CHARMM	– / <i>Chemistry at HARvard Macromolecular Mechanics</i>
CI	Método de Interacción de Configuraciones / <i>Configuration Interaction method</i>
CIS	Método de Interacción de Configuraciones con excitaciones Simples / <i>Configuration Interaction method with Single excitations</i>
DFT	Teoría del Funcional de la Densidad / <i>Density Functional Theory</i>
FBP	Flurbiprofeno / <i>Flurbiprofen</i>
FRET	Transferencia de Energía de Resonancia de Förster / <i>Förster Resonance Energy Transfer</i>
GAFF	– / <i>General AMBER Force Field</i>
GBSA	Modelo Generalizado de Born aumentado con el término hidrofóbico de Área Superficial accesible al disolvente/ <i>Generalized Born model augmented with the hydrophobic solvent accessible Surface Area term</i>
HF	Método Hartree-Fock / <i>Hartree-Fock method</i>
HSA	Albúmina Sérica Humana / <i>Human Serum Albumin</i>
IEFPCM	– / <i>Integral Equation Formalism of the Polarizable Continuum Model</i>
LR	Respuesta Lineal / <i>Linear Response</i>
MD	Dinámica Molecular / <i>Molecular Dynamics</i>
MM	Mecánica Molecular / <i>Molecular Mechanics</i>
MMPol	Mecánica Molecular Polarizable / <i>Polarizable Molecular Mechanics</i>
MST	Modelo de solvatación Miertus Scrocco Tomasi / <i>Miertus Scrocco Tomasi solvation model</i>
PCM	Modelo Continuo Polarizable / <i>Polarizable Continuum Model</i>
PDB	– / <i>Protein Data Bank</i>

QM	Mecánica Cuántica / <i>Quantum Mechanics</i>
QM/MM	Mecánica Cuántica Mecánica Molecular / <i>Quantum Mechanics Molecular Mechanics</i>
QM/MMPol	Mecánica Cuántica Mecánica Molecular Polarizable / <i>Quantum Mechanics Polarizable Molecular Mechanics</i>
RMSD	Desviación Cuadrática Media / <i>Root-Mean Square Deviation</i>
SMD	Modelo de Solvatación universal basado en la Densidad electrónica del soluto / <i>Universal Solvation Model based on solute electron Density</i>
SCS-MP2	Teoría perturbativa de Møller-Plesset a segundo orden con Componentes de Espín Escaladas / <i>Spin-Component-Scaled Møller-Plesset second order perturbation theory</i>
SAPT	– / <i>Symmetry-Adapted Perturbation Theory</i>
TD-DFT	Teoría del Funcional de la Densidad Dependiente del Tiempo / <i>Time-Dependent Density Functional Theory</i>
TIP3P	– / <i>Three-site Transferrable Intermolecular Potential</i>
ZINDO	– / <i>Zerner's Intermediate Neglect of Differential Overlap</i>



Introducción

Capítulo
1

1. INTRODUCCIÓN

Las herramientas de la química teórica y computacional comprenden, básicamente, los métodos basados en la mecánica cuántica (*quantum mechanics*, QM), que incluyen métodos *ab initio*, semi-empíricos y la teoría del funcional de la densidad (DFT), y aquellos basados en la mecánica clásica (*molecular mechanics*, MM). Estas herramientas juegan un papel cada vez más relevante en la comprensión del funcionamiento de sistemas químicos a nivel microscópico, tal y como ilustra el crecimiento en sus aplicaciones en las últimas décadas. Esta progresión se debe al aumento en la calidad de los resultados obtenidos para la descripción de propiedades moleculares o de reacciones químicas, así como a la posibilidad de abordar sistemas de gran tamaño cada vez más realistas. Este último punto es especialmente relevante en el estudio de sistemas biológicos, ya que actualmente es posible realizar cálculos en sistemas de centenares o incluso miles de átomos usando metodologías de alto nivel (*ab initio* y DFT). En tales sistemas, sin embargo, la necesidad de obtener predicciones promediadas estadísticamente sobre las fluctuaciones estructurales del sistema hace que las técnicas MM aún sean la opción más utilizada para su estudio en aplicaciones, por ejemplo, de Dinámica Molecular (MD). Cuando se estudian reacciones químicas o propiedades espectroscópicas que requieren una descripción explícita de la estructura electrónica, la técnica habitual, sin embargo, suele combinar la potencialidad de ambas descripciones: la química cuántica (QM), pero también el uso de campos de fuerzas clásicos (MM) a fin de describir el impacto del entorno mediante métodos mixtos QM/MM (Mecánica Cuántica/Mecánica Molecular) que permitan un costo computacional razonable. Tales modelos multiescala permiten así explorar, por ejemplo, reacciones de catálisis enzimática o procesos fotoquímicos en ambientes biológicos. Para aplicar con criterio este tipo de metodologías, es primordial conocer los fundamentos teóricos en los que se basan tales descripciones, así como sus principales limitaciones.

Un problema ampliamente debatido en la aplicación de métodos MM, por ejemplo en simulaciones de MD o en cálculos QM/MM, es el uso mayoritario de campos de fuerzas aditivos que no incluyen explícitamente los efectos de polarización electrónica.^{1,2} Paradójicamente, tales efectos ya se introdujeron en la presentación pionera de un método QM/MM multiescala por Warshel y Levitt en 1976,³ autores que han sido recientemente galardonados con el premio Nobel de Química del año 2013 por el desarrollo de métodos

CAPÍTULO 1. INTRODUCCIÓN

multiescala. En las últimas décadas se han desarrollado diversas formulaciones de campos de fuerzas polarizables.¹ No obstante, la complejidad relacionada con su correcta parametrización, su considerable coste computacional y la constatación de que su impacto en la mayoría de aplicaciones no es determinante, explican en parte que los campos de fuerzas polarizables sean aún poco usados.

En este contexto, esta tesis pretende explorar el impacto de los efectos de la polarización electrónica en dos casos paradigmáticos de interacciones donde su efecto es primordial, tanto a nivel de estado fundamental como de estado excitado, ambos evaluados en contextos biológicos. Además de analizar el impacto de la polarización, la presente tesis pretende contribuir en la mejora de los modelos MM polarizables existentes para su modelización.

En primer lugar, esta memoria aborda el estudio de un tipo de interacción de estado fundamental donde el impacto de la polarización es relevante, interacciones catión- π -catión a tres cuerpos, sobre las cuales se sabe aún muy poco. En interacciones catión- π a dos cuerpos, donde el catión es capaz de polarizar de forma significativa al sistema aromático, la importancia de la polarización está bien establecida.⁴⁻⁷ En motivos catión- π -catión, este impacto es más complejo debido a la posible cancelación parcial del campo eléctrico producido por los dos cationes a cada lado del sistema aromático. Además, su naturaleza a tres cuerpos da lugar a contribuciones no aditivas que no pueden ser descritas mediante campos de fuerzas aditivos basados en interacciones de pares. Recientemente, se ha demostrado que las interacciones catión- π -catión pueden estabilizar la estructura terciaria de proteínas,⁸ y podrían ser importantes en la unión de ligandos con sistemas π a dianas terapéuticas. La presente tesis pretende en este sentido profundizar en la comprensión de las características energéticas de tales interacciones, de su ocurrencia en estructuras de proteínas y el papel que éstas pueden tener en su estabilización, así como desarrollar un conjunto de interacciones catión- π -catión y de cálculos de su energía de interacción de referencia, que pueda ser usada para explorar y validar la capacidad de campos de fuerzas clásicos para describirlas.

La segunda parte de la tesis, se centra en la aplicación de tales campos de fuerzas polarizables en el estudio de procesos fotoinducidos en sistemas biológicos. En particular, se estudia el impacto de la polarización heterogénea del entorno en interacciones entre

estados excitados a partir del modelo QM/MM polarizable desarrollado por el grupo de Barcelona en colaboración con el grupo de la Prof. Mennucci en Pisa.⁹⁻¹² A diferencia de propiedades tales como desplazamientos solvatocrómicos, donde su papel es menor comparado con los efectos puramente electrostáticos, la polarización tiene un impacto dramático en este contexto, pues apantalla los acoplamientos electrónicos entre estados excitados que median reacciones de transferencia de energía electrónica resonante (*Förster resonance energy transfer*, FRET).¹²⁻¹⁶ Tal fenómeno tiene una gran aplicación en el campo de la biomedicina, ya que permite medir en tiempo real y a nivel de moléculas individuales cambios estructurales en biomoléculas.¹⁷⁻¹⁹ En la derivación pionera de su teoría para describir procesos FRET en 1948,¹³ T. Förster ya reconoció la importancia de los efectos de polarización introduciendo un factor de apantallamiento del entorno igual a $1/n^2$, donde n es el índice de refracción del medio, factor que modula hasta ~4 veces la velocidad de FRET en ambientes biológicos. A pesar de su impacto, este factor ampliamente usado en la interpretación de fenómenos FRET representa una aproximación muy drástica y asume que la polarización electrónica del entorno juega un papel uniforme en tales fenómenos. En los últimos años, se ha demostrado que tal aproximación introduce importantes desviaciones en la velocidad de FRET en comparación con predicciones que tienen en cuenta el ambiente dieléctrico heterogéneo que rodea a los pigmentos en complejos fotosintéticos.^{10,12}

Esta tesis aborda el impacto que la naturaleza polarizable heterogénea juega en la interpretación de fenómenos FRET en sistemas proteína-ligando, así como las implicaciones que tiene la aproximación dipolo-dipolo subyacente a la teoría de Förster en este contexto. Tal investigación, además, pretende explorar el potencial de la metodología MD-QM/MM polarizable desarrollada en el grupo a fin de caracterizar uniones ligando-receptor y ayudar al descubrimiento de nuevos sitios de unión. Tal aplicación, por ejemplo, puede ser relevante en la caracterización de sitios de unión crípticos que aparecen transitoriamente por un cambio conformacional en la estructura de la proteína, y que por tanto son difíciles de descubrir a partir de la estructura cristal dada por rayos X.

La tesis, por tanto, está estructurada en diversos capítulos. En este Capítulo 1 se describen los métodos habitualmente usados para estudiar sistemas biológicos en modelización molecular, desde campos de fuerzas clásicos MM hasta modelos multiescala QM/MM, haciendo especial hincapié en sus variantes que permiten un tratamiento

CAPÍTULO 1. INTRODUCCIÓN

explícito de la polarización electrónica. También se describen las características de los modelos continuos de solvatación, donde la polarización del entorno se describe mediante sus propiedades dieléctricas macroscópicas. Tales modelos, complementarios a modelos QM/MM, tienen formulaciones acopladas tanto a descripciones MM como QM del soluto, y también han sido usadas en la presente tesis. A continuación, se describen las características de las interacciones catión- π -catión, y el impacto que tiene la polarización electrónica en las mismas, así como la metodología usada en la presente tesis para estimar la energética de tales interacciones a partir de cálculos QM de alto nivel. Finalmente, se termina este Capítulo 1 describiendo las bases teóricas de procesos de transferencia de energía, como esta modifica la fluorescencia de sistemas biológicos, así como las características de los acoplamientos electrónicos responsables de tales procesos y como estos pueden ser estimados mediante modelos teóricos. En el Capítulo 2 se enumeran los objetivos generales y específicos de la presente tesis doctoral, en el Capítulo 3 se presentan los resultados de la investigación realizada, y en el Capítulo 4 su discusión. Finalmente, en el Capítulo 5 se presentan las conclusiones y posibles perspectivas futuras derivadas de la tesis.

1.1 Descripción de la polarización en la modelización de biosistemas

1.1.1 Campos de fuerzas clásicos

Muchos de los sistemas en modelación molecular poseen dimensiones apreciables como para poder ser abordados por métodos QM *ab initio*, semi-empíricos o DFT. Estos métodos describen explícitamente los electrones del sistema y, aún cuando algunos de estos son ignorados (como sucede en los métodos semi-empíricos), su aplicación acaba consumiendo un tiempo de cálculo que los torna inviables.²⁰ Los campos de fuerzas clásicos, también conocidos como métodos MM, ignoran los movimientos de los electrones y calculan la energía de un sistema a partir de una función que depende únicamente de las posiciones nucleares. Esta simplificación permite que los métodos MM sean aplicables a sistemas con un número grande de átomos. En algunos casos, y mediante parametrizaciones minuciosas, los campos de fuerzas son capaces de describir propiedades con una precisión comparable a los más elevados niveles de cálculo QM pero en un tiempo computacional mucho menor.²⁰

Debido a su simplicidad computacional y eficiencia, los campos de fuerzas han sido ampliamente utilizados para modelar macromoléculas en diversas áreas de la química y bioquímica, tales como proteínas o ácidos nucleicos. Su fundamento teórico se basa en la aproximación de Born-Oppenheimer, que permite tratar separadamente los movimientos de núcleos y electrones. En concreto, los modelos MM consideran las moléculas como una colección de átomos unidos entre sí por fuerzas elásticas o harmónicas, siendo estas fuerzas descritas por funciones de energía potencial parametrizadas empíricamente o por comparación con datos QM de alto nivel. El conjunto de estas funciones es conocido como campo de fuerzas (*force field*).²¹⁻²³

De esta forma, los modelos MM describen las moléculas como una colección de esferas conectadas por resortes, donde las esferas representan a los núcleos y los resortes a los enlaces (Figura 1).

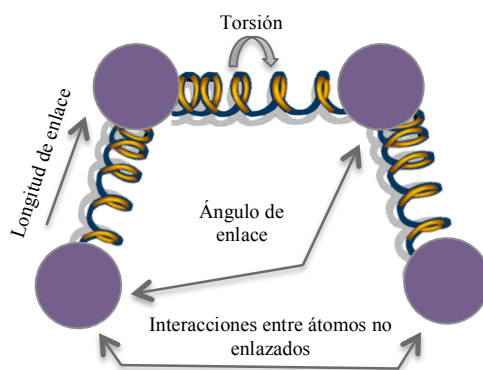


Figura 1. Representación de las contribuciones enlazantes y no enlazantes en un campo de fuerzas clásico. Las esferas representan a los núcleos y los resortes a los enlaces.

Un campo de fuerzas es por tanto una función matemática con parámetros ajustables que describe la energía potencial de un sistema de partículas, con el objetivo de representar distintas propiedades moleculares (por ejemplo conformaciones, energías de interacción, etc). Aunque la ecuación matemática adoptada varía entre distintos campos de fuerzas, de manera general se incluyen los siguientes términos:²⁰

- **Interacciones no enlazantes:** intervienen átomos no unidos de manera directa por enlaces. Comprenden los términos electrostáticos y de interacciones de van der Waals. Comúnmente solo se consideran las interacciones entre átomos separados por 3 o más enlaces, ya que las interacciones 1-2 y 1-3 se consideran implícitamente incorporadas en los términos enlazantes.
- **Interacciones enlazantes:** intervienen átomos unidos por enlaces químicos. Se trata de los términos de tensión de enlace, flexión de ángulo, torsión de enlace y torsiones impropias. Además, se pueden incluir términos de interacciones cruzadas que reflejan el acoplamiento entre estos términos: acoplamientos ángulo-enlace, ángulo-ángulo, enlace-enlace, etc.

$$E_{total} = E_{enlazantes} + E_{no\ enlazantes} \quad (1)$$

Por tanto, se considera que los núcleos están sometidos a un campo de fuerzas que se compone de la suma de una serie de términos que modelizan las tensiones de los enlaces, de sus ángulos, sus torsiones y las interacciones de van der Waals y electrostáticas.

$$E_{total} = E_{ent} + E_{ang} + E_{tor} + E_{vdw} + E_{ele} \quad (2)$$

Existe una gran variedad de campos de fuerzas, los cuales han sido diseñados y parametrizados para describir distintos tipos de sistemas químicos, en función de los grupos de moléculas empleados como referencia en su parametrización. Estos campos de fuerzas van desde modelos generales aplicables en principio a cualquier sistema químico (UFF) hasta modelos más habituales diseñados para el estudio de moléculas orgánicas e inorgánicas (DREIDING), compuestos con metales de transición (SHAPES, VALBOND), moléculas orgánicas pequeñas y medianas (MM2, MM3, MM4, GAFF), o macromoléculas biológicas (AMBER, CHARMM, GROMOS, OPLS).²⁰

1.1.1.1 Campos de fuerzas aditivos

El objetivo de un buen campo de fuerzas es describir, tanto cuanto sea posible, diferentes clases de moléculas con una calidad razonable. La exactitud de un cálculo MM, en este sentido, depende tanto de las funciones de energía potencial adoptadas como de la calidad de los parámetros incorporados a estas funciones. Así, es fácil comprender que el uso de parámetros poco adecuados impacten de forma importante los resultados. Para evitar esta situación, un campo de fuerzas adecuado a una investigación en particular debe ser cuidadosamente seleccionado.²⁴ Además, el desarrollo de nuevos campos de fuerzas parametrizados sobre la base de conjuntos de datos tanto experimentales como teóricos más extensos conllevan un aumento de la variedad de sistemas y propiedades que pueden ser analizadas.²⁵

De la misma forma que los parámetros intrínsecos a un campo de fuerzas limitan su aplicabilidad, la función de energía de potencial adoptada también impone ciertas limitaciones. Los campos de fuerzas habitualmente usados en el estudio de sistemas biológicos se basan en funciones similares a la ilustrada en la Ec. 3, es decir, no consideran términos cruzados, y las interacciones no enlazantes se limitan a términos de dispersión-repulsión y electrostáticos aditivos basados en interacciones de pares de átomos. En este contexto, quizá los campos de fuerzas más frecuentemente utilizados son los desarrollados en el contexto del código AMBER²⁶ y CHARMM,²⁷ los cuales están parametrizados para simulaciones de ácidos nucleicos, proteínas o lípidos. En esta tesis, en particular, se ha usado el campo de fuerzas aditivo de Amber llamado ff12SB.^{28,29} Este campo de fuerzas adopta la siguiente función de energía, basada en un potencial Lennard-Jones (LJ) para la descripción de las interacciones de dispersión-repulsión, y de la asignación de cargas puntuales para la descripción del término electrostático .

CAPÍTULO 1. INTRODUCCIÓN

$$E_{par} = \sum_{enl} k_r (r - r_{eq})^2 + \sum_{ang} k_\theta (\theta - \theta_{eq})^2 + \sum_{tor} \frac{v_n}{2} \times [1 + \cos(n\phi - \gamma)] \\ + \sum_{i < j} \left[\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^6} + \frac{q_i q_j}{\epsilon R_{ij}} \right] \quad (3)$$

donde, r_{eq} y θ_{eq} son distancias y ángulos de equilibrio; k_r , k_θ , v_n son constantes de fuerza; y n es la multiplicidad y γ el ángulo de fase de las torsiones. Los parámetros A y B caracterizan los potenciales de LJ y las cargas q el término electrostático.

Los campos de fuerzas aditivos son utilizados en la gran mayoría de simulaciones moleculares de biosistemas, pues han demostrado ser capaces de describir las características estructurales y energéticas esenciales de sistemas bioquímicos complejos de forma muy eficiente. Sin embargo, hay evidencias de que tales campos de fuerzas, los cuales no incluyen de forma explícita efectos de polarización, presentan limitaciones cuando tales efectos de inducción son significativos. Por ejemplo, para describir la partición de moléculas entre medios polares y apolares³⁰ o la descripción de interacciones catión- π .^{5,31} Con el objetivo de superar estas limitaciones, se han desarrollado diversos modelos para introducir los efectos de polarización en los campos de fuerzas clásicos, modelos que se describen a continuación.

1.1.1.2 Campos de fuerzas polarizables

La inclusión explícita de efectos de polarización en campos de fuerzas es necesaria para describir la polarización variable que tienen las moléculas en función de su entorno más inmediato, y superar de esta forma la aproximación de campo medio, implícita en el uso de cargas puntuales fijas que describen un "estado de polarización" constante en fase condensada. Los campos de fuerzas polarizables, en este contexto, permiten describir las variaciones en la distribución de carga en respuesta a los campos eléctricos locales creados por el entorno. Dichos efectos de inducción constituyen el principal reto en el desarrollo de nuevos campos de fuerzas debido a la dificultad que presenta la descripción de la no aditividad y la cooperatividad de la energía de inducción a nivel clásico. En consecuencia, se han desarrollado varias estrategias para incorporar explícitamente la polarización en campos de fuerza clásicos.^{1,2,32-34} Los más extendidos se pueden clasificar en tres

categorías principales: modelo de dipolos puntuales, modelo de carga fluctuante, y modelo del oscilador clásico de Drude.

Modelo de dipolos puntuales

Este modelo, formulado por Applequist en 1972,³⁵ asigna dipolos polarizables puntuales a un conjunto de sitios distribuidos a través del sistema molecular. Es decir, asigna un tensor de polarizabilidad a cada átomo. Asumiendo un modelo de polarizabilidad atómica isotrópica, el campo eléctrico experimentado por cada átomo da lugar a un dipolo inducido según la ecuación:

$$\mu_{ind,i} = \alpha_i E_i \quad (4)$$

donde α_i es la polarizabilidad atómica, y el campo eléctrico E_i sobre el átomo i es la suma del campo eléctrico estático y el creado por los dipolos inducidos de los otros átomos del sistema. De esta forma, la determinación de los momentos dipolares inducidos sobre cada átomo debe resolverse iterativamente hasta alcanzar convergencia. Una vez alcanzada, la energía de interacción entre un dipolo y el campo eléctrico (la energía de inducción) se calcula según:

$$E_{pol} = -\frac{1}{2} \mu_{ind,i} \cdot E_{i,0} \quad (5)$$

donde $E_{i,0}$ es el campo eléctrico estático generado únicamente por las cargas puntuales.

Esta contribución se calcula para todos los dipolos inducidos, y sumada al término electrostático, describe las interacciones existentes entre las distribuciones de carga de las diferentes partes del sistema y nos permite obtener la energía electrostática total del sistema.

El modelo de dipolos inducidos ha sido adoptado en el desarrollo de diversos campos de fuerzas, tales como OPLS/PPF,^{36,37} AMOEBA^{23,33,38} y AMBER ff02³⁰ y ff12pol.^{32,39-41} En estos campos de fuerzas, de esta forma, se agrega un término de energía adicional, E_{pol} , a la energía total del sistema:

$$E_{tot} = E_{bonded} + E_{nonbonded} + E_{pol} \quad (6)$$

CAPÍTULO 1. INTRODUCCIÓN

Dicho término, E_{pol} , describe la interacción entre las cargas parciales permanentes y los dipolos inducidos, y se calcula a partir de la siguiente fórmula:

$$E_{pol} = -\frac{1}{2} \sum_i \mu_i E_{i,0} = -\frac{1}{2} \sum_i \alpha_i E_i E_{i,0} \quad (7)$$

donde la suma se extiende sobre todos los sitios que interaccionan, i , y el término α_i denota la polarizabilidad isotrópica del átomo i . $E_{i,0}$ es el campo eléctrico en el átomo i debido a las cargas parciales de los otros átomos. E_i es el campo eléctrico en el átomo i debido a cargas y dipolos inducidos.

Finalmente, cabe destacar que el modelo de dipolos inducidos puede sufrir la popularmente llamada "catástrofe de polarización", la cual consiste en una sobrepolarización artificial de los dipolos debido a la descripción aproximada de la distribución de carga en forma de dipolo puntual. A fin de evitar tal problema, por ejemplo, el campo de fuerzas Amber ff02, siguiendo la misma filosofía adoptada para el término electrostático, ignora interacciones entre sitios polarizables 1-2 y 1-3.³⁰ Otra alternativa consiste en adoptar un tensor dipolo-dipolo modificado con funciones de apantallamiento lineales o exponenciales, desarrolladas por Thole, las cuales han sido adoptadas en distintos campos de fuerzas tales como Amber ff12pol o AMOEBA.^{32,33}

Modelo de fluctuación de carga

El modelo de fluctuación de carga se basa en el principio de igualdad de electronegatividad, de forma que una carga fluye entre átomos hasta que las electronegatividades de los átomos se igualan. Mediante la adición de este efecto, la distribución de carga de una molécula se acopla a su entorno, proporcionando una manera de incorporar los efectos de polarización. La idea de este modelo es tratar estas cargas parciales como variables que responden a sus ambiente de una manera similar a la respuesta de polarización de las moléculas reales.

Para un conjunto de moléculas en una configuración molecular dada, el conjunto óptimo de cargas se obtiene por minimización de la energía electrostática total del sistema según:

$$\begin{aligned}
U_{ele} = \sum_i \sum_\alpha \left[X_{i\alpha}^o q_{i\alpha} + \frac{1}{2} J_{i\alpha i\alpha}^o q_{i\alpha}^2 \right] + \sum_i \sum_{i\alpha < i\beta} J_{i\alpha i\beta} (r_{i\alpha i\beta}) q_{i\alpha} q_{i\beta} \\
+ \sum_{i < j} \sum_{i\alpha < j\beta} J_{i\alpha j\beta} (r_{i\alpha j\beta}) q_{i\alpha} q_{j\beta}
\end{aligned} \tag{8}$$

donde i y j representan moléculas, α y β representan átomos, $X_{i\alpha}^o$ y $J_{i\alpha i\alpha}^o$ son la electronegatividad de Mulliken y la dureza absoluta del átomo α , y $J_{i\alpha i\beta}$ describe la interacción de Coulomb entre átomos α y β .^{1,2,42-44} Este modelo, por ejemplo, ha sido usado en el campo de fuerzas UFF⁴⁵ o en el campo de fuerzas de CHARMM.⁴⁶

Modelo de osciladores clásicos de Drude

El modelo del oscilador de Drude representa la inducción electrónica mediante la introducción de una partícula auxiliar cargada unida al átomo polarizable por un resorte armónico. La polarización de un sitio atómico, de esta forma, se explica mediante la unión al átomo polarizable de una partícula móvil, que puede o no tener masa y tiene asignada una carga, a través de un resorte armónico caracterizado por una constante de fuerza. Es decir, para modelizar la polarizabilidad de un átomo se asocia a este un “oscilador de Drude”.⁴⁷ Los métodos de oscilador de Drude incorporan la polarizabilidad electrónica al representar un átomo o un ión como un sistema de dos partículas: un núcleo cargado con carga $q_{i,0}$ y una cargada con carga $q_{i,D}$.

Así, la polarizabilidad isotrópica α_i al átomo está relacionada con la constante de fuerza k_D del resorte armónico que conecta el núcleo y el *shell* y está determinada por:^{1,48}

$$\alpha_i = \frac{q_{i,D}^2}{k_D} \tag{9}$$

Para una polarizabilidad α dada, se escoge la constante de fuerza k_D de forma que la partícula Drude se mantenga a una distancia d mucho menor que cualquier distancia interatómica. De esta manera, el dipolo inducido resultante es casi equivalente a un dipolo puntual.^{1,47-49} El campo de fuerzas polarizable Drude, desarrollado en el contexto del campo de fuerzas CHARMM, es seguramente el más completo para estudiar biosistemas basado en este modelo.³⁴

1.1.2 Métodos híbridos QM/MM

A menudo, en simulaciones de proteínas, ácidos nucleicos o solutos en disolución, la propiedad de interés o la reacción objeto de estudio se localiza en una parte relativamente pequeña del sistema, mientras que el entorno restante juega un papel indirecto perturbando las propiedades de esta región. En este contexto, los métodos QM son adecuados para describir la energía o las propiedades de la región del sistema donde tenemos especial interés. Sin embargo, se torna inviable el uso de este método para tratar sistemas grandes como proteínas, debido al elevado costo computacional asociado. Los métodos clásicos MM, por otra parte, pueden ser utilizados como alternativa a los métodos QM para sistemas de gran tamaño. Sin embargo, tales modelos no son adecuados para simular un proceso químico que involucre la ruptura o formación de enlaces o bien procesos fotoquímicos o propiedades espectroscópicas que necesitan de una descripción explícita de distintos estados electrónicos. Las limitaciones y ventajas de los métodos MM y QM han dado lugar, pues, a estrategias alternativa basadas en la combinación de estos dos métodos, creando así lo que se conoce como el método híbrido de mecánica cuántica y mecánica molecular (QM/MM).^{3,50,51}

Así, el método híbrido QM/MM es una técnica multiescala que se caracteriza por la partición del sistema en una región QM, donde se localizan los átomos responsables de la propiedad espectroscópica de interés o donde tiene lugar la reacción, y una región MM que incluye el resto del sistema, y que se describe mediante un método de mecánica molecular⁵¹ (ver Figura 2).

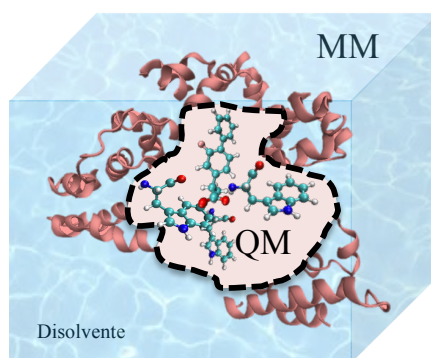


Figura 2. Representación esquemática de las particiones clásica (MM) y cuántica (QM). La región en salmón corresponde al centro activo.

De esta manera, el método QM/MM es una metodología muy potente pues combina las ventajas de ambas descripciones QM y MM. En particular, permite introducir la perturbación del entorno MM en el sistema QM dando lugar a un Hamiltoniano efectivo que puede ser usado en la resolución de la ecuación de Schrödinger. Mediante la introducción de la aproximación de Born-Oppenheimer, la función de onda electrónica satisfará la siguiente ecuación:

$$\hat{H}^{\text{eff}}|\Psi\rangle = (\hat{H}^{QM} + \hat{H}^{QM/MM} + \hat{H}^{MM})|\Psi\rangle = E|\Psi\rangle \quad (10)$$

Donde H^{QM} es el Hamiltoniano del sistema QM aislado, el operador $H^{QM/MM}$ introduce el acoplamiento entre ambas descripciones QM y MM, y el término H^{MM} describe la energía de la región MM basada en el campo de fuerzas correspondiente. En las formulaciones más comunes, el término del acoplamiento entre la región QM y MM se limita a interacciones electrostáticas y de van der Waals (dispersión-repulsión). Estas últimas se suelen basar en un término empírico, por ejemplo basado en un potencial de Lennard-Jones, el cual no depende de la densidad electrónica de la región QM. Así pues, mediante la resolución de la Ec. 10, la función de onda del sistema QM solo se ve afectada por el término electrostático, que polariza la densidad electrónica.

1.1.2.1 Modelos QM/MM basados en cargas puntuales

En general, los modelos QM/MM limitan el acoplamiento entre las regiones QM y MM a un término electrostático, en el cual el entorno se describe mediante cargas puntuales distribuidas en los átomos de la región MM:

$$\hat{H}^{QM/MM} \rightarrow \hat{H}^{\text{ele}} = \sum_m q_m(r_m) \hat{V}(r_m) \quad (11)$$

Donde $\hat{V}(r_m)$ es el operador de potencial electrostático debido a los electrones y núcleos del subsistema QM que actúan sobre el "ambiente" representado en términos de cargas puntuales q_m .¹²

En estos modelos, el entorno MM polariza la función de onda del subsistema QM. Sin embargo, no tiene en cuenta la polarización que a su vez ejerce el sistema QM en su entorno, ya que la distribución de carga del entorno es fija. Tal descripción, de esta forma, es incapaz de describir, por ejemplo, los efectos de reorganización electrónica que tienen

lugar cuando se da una transición electrónica en la región QM debido a la absorción de un fotón de luz, los llamados efectos de solvatación de no-equilibrio.

1.1.2.2 Modelos QM/MM polarizables

En modelos QM/MM polarizables, la descripción del entorno se basa en alguno de los campos de fuerzas clásicos polarizables descritos anteriormente en la Sección 1.1.1.2 de esta tesis: modelo de dipolos inducidos, modelo de fluctuación de carga, o modelo basado en osciladores clásicos de Drude. En la última década, se han desarrollado múltiples modelos de este tipo, que difieren principalmente en el campo de fuerzas adoptado así como en el método usado para describir la región QM. En esta tesis nos centramos en modelos desarrollados para el estudio de estados excitados.^{9,52-62} En particular, se ha aplicado el modelo QM/MM polarizable desarrollado en el grupo,^{9,10} el cual se basa en el modelo de dipolos inducidos, por ello de aquí en adelante nos centraremos en este modelo en particular. Dicho modelo asigna a los átomos de la región MM una carga parcial y una polarizabilidad dipolar atómica isotrópica.

La estrategia QM/MMPol combina de esta forma una descripción QM con una descripción clásica polarizable (MMPol) del entorno circundante, e incluye los efectos de polarización mutua entre las regiones QM y MM tanto durante el proceso SCF como en el cálculo de la respuesta del entorno a una excitación electrónica (Figura 3). Cabe destacar, sin embargo, que la inclusión explícita de la polarización en regiones MM muy distantes de la región QM implican cambios despreciables en las propiedades estudiadas.

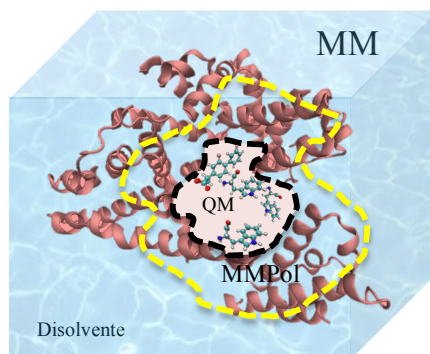


Figura 3. Representación esquemática de las particiones clásica (MM), clásica polarizable (MMPol) y cuántica (QM).

De esta forma, este modelo en la práctica usa una descripción a tres capas, donde los átomos MM situados más allá de una distancia escogida son descritos con un campo de fuerzas no polarizable.

El método QM/MMPol, por tanto, incluye tanto términos electrostáticos como de polarización en el Hamiltoniano efectivo descrito en la Ec. 10. En particular, la interacción entre la región QM y MM, $\hat{H}_{QM/MM}$, y el término \hat{H}_{MM} que describe la energía de la región MM adoptan la siguiente forma:

$$\hat{H}_{QM/MM} = \hat{H}_{QM/MM}^{el} + \hat{H}_{QM/MM}^{pol} = \sum_m q_m \hat{V}(r_m) - \frac{1}{2} \sum_a \mu_a^{ind} \hat{E}_a^{solute}(r_a) \quad (12)$$

$$\hat{H}_{MM} = \hat{H}_{MM}^{el} + \hat{H}_{MM}^{pol} = \sum_m \sum_{n>m} \frac{q_m q_n}{r_{mn}} - \frac{1}{2} \sum_a \mu_a^{ind} \sum_m \frac{q_m (r_a - r_m)}{|r_a - r_m|^3} \quad (13)$$

donde $\hat{V}(r_m)$ y $\hat{E}_a^{solute}(r_a)$ representan los operadores de potencial electrostático y campo eléctrico debido a electrones y núcleos del sistema QM en los sitios MM, y los índices m , n y a designan las cargas q_m y los dipolos inducidos μ_a^{ind} situados en r_m y r_a , respectivamente. En la Ec. 13 los términos $\hat{H}_{QM/MM}^{el}$ y $\hat{H}_{QM/MM}^{pol}$ describen la interacción entre el sistema QM y las cargas MM y los dipolos inducidos, respectivamente. Por otro lado, en la Ec. 14, \hat{H}_{MM}^{el} describe la energía propia electrostática de las cargas MM mientras que \hat{H}_{MM}^{pol} representa la interacción de polarización entre dichas cargas y los dipolos inducidos. Aquí cabe destacar que el término \hat{H}_{MM}^{el} entra en el Hamiltoniano efectivo sólo como una cantidad energética constante, mientras que la contribución \hat{H}_{MM}^{pol} es considerada explícitamente en el operador de Fock correspondiente debido a la dependencia explícita de los dipolos inducidos respecto a la función de onda QM.

Este modelo se ha desarrollado e implementado en una versión local modificada del código Gaussian, con el objetivo de investigar propiedades espectroscópicas y procesos de transferencia de energía de excitación entre cromóforos situados en entornos complejos.^{9,10,12} Además, tal modelo ha sido extendido para acoplarse a una descripción dieléctrica continua del entorno basada en el Modelo Continuo Polarizable (PCM), dando lugar así al método multiescala QM/MMPol/PCM que incluye la mutua polarización entre las tres regiones.⁶³

CAPÍTULO 1. INTRODUCCIÓN

1.1.2.3 Extensión a TD-DFT

En la actualidad, los métodos DFT se han consolidado como la opción más adecuada para el estudio de sistemas de gran tamaño, ya que permiten obtener resultados de calidad notable a un coste computacional mucho menor que el de métodos *ab initio* de alto nivel. De forma similar, la teoría del funcional de la densidad dependiente del tiempo (TD-DFT), y en particular su extensión al cálculo de estados electrónicos excitados en su formulación basada en respuesta lineal (LR),^{64,65} se ha consolidado como una opción eficiente y de calidad para el estudio de la fotofísica de moléculas orgánicas.⁶⁶ TD-DFT suele describir muy bien estados excitados de valencia. Sin embargo, también son bien conocidos sus problemas para describir correctamente estados excitados de tipo Rydberg o de transferencia de carga.⁶⁷ En la presente tesis, se ha aplicado TD-DFT al estudio de procesos de transferencia de energía de excitación entre cromóforos en ambientes biológicos, donde su aplicación, por ejemplo, a complejos fotosintéticos, ha tenido una gran popularidad en los últimos años.^{12,68} En este sentido, estudios previos han mostrado que TD-DFT representa la metodología óptima en términos de coste computacional y calidad de los resultados ya sea para la estimación de las energías relativas de cromóforos en una proteína,⁶⁹ como para la estimación de los acoplamientos electrónicos responsables de procesos de transferencia de energía.⁷⁰

En la presente tesis, en concreto, se ha usado la extensión de la metodología QM/MMPol descrita en la Sección 1.1.2.2 a nivel LR TD-DFT, extensión que permite la incorporación de los efectos del entorno biológico en la estimación tanto de las energías de los estados excitados como de los acoplamientos electrónicos entre ellos.^{9,12} Respecto a un método QM/MM estándar, cabe destacar que la estrategia QM/MMPol permite obtener una descripción más próxima a la obtenida mediante una descripción QM del ambiente, como se ha demostrado recientemente para el cromóforo de la proteína verde fluorescente⁷¹ o para la rodopsina.⁷² Aún más relevante, sin embargo, es su impacto en el cálculo de acoplamientos electrónicos, ya que los efectos de polarización son enteramente responsables de los efectos de apantallamientos inducidos por el entorno en éstos.^{9,10}

1.1.3 Modelos continuos de disolvente

Los modelos continuos de solvatación representan una alternativa simple y eficiente a las descripciones MM atomísticas del ambiente, ya que a partir de un sólo

cálculo permiten obtener el efecto del entorno, evitando así tener que promediar estadísticamente los resultados sobre los grados de libertad del disolvente. En un modelo continuo, se sitúa la molécula objeto de interés dentro de una cavidad generada a partir de esferas centradas en los átomos y rodeada por un medio dieléctrico continuo, caracterizado por las propiedades dieléctricas macroscópicas del disolvente (ver Figura 4).

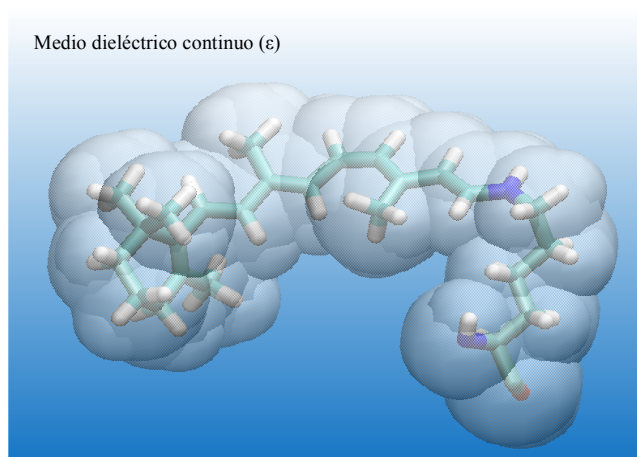


Figura 4. Ilustración de la cavidad que aloja al soluto dentro del medio dieléctrico en un modelo continuo.

Estas propiedades dieléctricas, entonces, permiten calcular la polarización del disolvente como respuesta a la distribución de carga del soluto. Se han desarrollado múltiples variantes de modelos continuos, que difieren básicamente en la descripción usada para el soluto (QM o MM), y en el formalismo matemático usado para resolver la polarización del disolvente.⁷³⁻⁷⁵ A continuación se describen los dos modelos usados en la presente tesis, el Modelo Continuo Polarizable (PCM) acoplado a una descripción QM del soluto, así como el modelo Generalizado de Born (GB) acoplado a una descripción clásica MM del sistema objeto de estudio y complementado con términos no-electrostáticos dependientes del área superficial (SA).

1.1.3.1 QM/PCM

El modelo PCM, formulado originalmente por Tomasi y colaboradores en 1981,⁷⁶ representa uno de los métodos más difundidos para incluir los efectos del disolvente en cálculos QM,⁷⁷ y es el ejemplo más antiguo de modelo continuo basado en cargas aparentes superficiales.⁷⁵ En este tipo de modelos, la polarización del disolvente se describe mediante cargas puntuales localizadas a lo largo de la superficie de la cavidad. En

CAPÍTULO 1. INTRODUCCIÓN

su formulación QM/PCM, basado en una descripción QM del soluto, la inclusión de la polarización del entorno en el Hamiltoniano del soluto implica la necesidad de iterar hasta lograr la mutua polarización entre las regiones QM y MM, ya que las cargas superficiales dependen de la densidad electrónica del soluto:

$$\hat{H}^{\text{eff}}|\Psi\rangle = (\hat{H}^{\text{QM}} + \hat{H}^{\text{PCM}})|\Psi\rangle = E|\Psi\rangle \quad (14)$$

$$\hat{H}^{\text{PCM}} = \sum_m q_m(r_m) \hat{V}(r_m) \quad (15)$$

Si bien las Ecs. 14 y 15 son muy similares a las Ecs. 10 y 11 correspondientes a un esquema QM/MM estándar, la dependencia de las cargas aparentes superficiales en la Ec. 15 respecto a la función de onda del soluto tiene distintas implicaciones en el modo de resolver las Ecs. 10 y 14. De forma similar a la descrita en la Sección 1.1.2.3 para la metodología QM/MMPol, el modelo PCM ha sido extendido al cálculo de estados excitados en su formulación LR TD-DFT, así como al cálculo de acoplamientos electrónicos entre estados excitados,⁷⁸ mostrándose como una alternativa eficiente y precisa para el estudio de procesos de transferencia de energía cuando los cromóforos involucrados se encuentran en disolución.⁹ En ambientes heterogéneos como una proteína, sin embargo, la inclusión del entorno local introduce importantes diferencias respecto a una descripción homogénea continua.¹⁰

1.1.3.2 MM/GBSA

En el modelo GB, el soluto se describe mediante cargas puntuales, y la energía libre de solvatación se obtiene a partir de la solvatación individual de cada átomo (carga) obtenida según el modelo de Born, corregida por el efecto perturbador de los otros átomos.^{74,79-81} En su versión MM/GBSA usada en la presente tesis para estimar el efecto del entorno proteico en interacciones catión- π -catión,⁸¹ el modelo GB, que describe el término electrostático a la energía libre de solvatación, se complementa con términos no-electrostáticos que dependen de la superficie accesible al disolvente (SA), dando lugar al llamado modelo MM/GBSA. Este método es ampliamente utilizado para estimar la energía libre de unión de pequeños ligandos a macromoléculas en estudios de diseño de fármacos. En este caso, el cálculo de las energías libres de unión contiene los distintos términos fisicoquímicos que intervienen en el proceso de unión de un ligando a una proteína, combinando energías de interacción calculadas a nivel MM con la contribución de

solvatación descrita a nivel GBSA, además de eventuales términos entrópicos, que sin embargo son frecuentemente omitidos por la complejidad asociada a su cálculo.⁸² Así, el método MM/GBSA se utiliza para determinar la energía libre (G) de distintas especies (ligando, proteína y complejo). Esta energía se subdivide en diferentes términos: entálpico en fase gas (E_{gas}), de solvatación (G_{solv}) y entrópico (TS), según se detalla en la ecuación siguiente:

$$G = E_{gas} + G_{solv} - TS = E_{int} + E_{ele} + E_{vdW} + G_{solv-p} + G_{solv-np} - TS \quad (16)$$

donde E_{int} , E_{ele} y E_{vdW} son los términos energéticos internos, de Coulomb y de van der Waals que consideran las contribuciones energéticas de enlace, polares y no polares en fase gaseosa, respectivamente. La energía de solvatación se divide en componentes polares (G_{solv-p}) y no polares ($G_{solv-np}$). El término polar describe el cambio en la energía libre electrostática asociado a la transferencia desde la fase gaseosa al disolvente acuoso a partir de la expresión GB:

$$G_{solv-p} = -\frac{1}{2} \left(1 - \frac{1}{\epsilon_{out}}\right) \sum_{i,j} \frac{q_i q_j}{f_{GB}} \quad (17)$$

donde q_i (q_j) se refiere a las cargas del soluto, ϵ_{out} es la constante dieléctrica del disolvente, y f_{GB} es una función que depende de la distancia interatómica r_{ij} :

$$f_{GB}(r_{ij}) = \left[r_{ij}^2 + \alpha_i \alpha_j \exp \left(-r_{ij}^2 / (4\alpha_i \alpha_j) \right) \right]^{1/2} \quad (18)$$

donde el término α_i es el radio efectivo de Born del átomo i .

El segundo término ($G_{solv-np}$), por su parte, considera las contribuciones no polares y se calcula mediante una relación lineal con la superficie accesible al disolvente:

$$G_{solv-np} = \gamma SAS + \beta \quad (19)$$

donde γ es la tensión superficial y β es un término de corrección.^{82,83}

1.2 Impacto de la polarización en interacciones a nivel de estado electrónico fundamental

1.2.1 Interacciones intermoleculares y descomposición en términos físicos

Las interacciones intermoleculares están estrechamente relacionadas con las propiedades termodinámicas de líquidos, sólidos y gases. Por lo tanto, su comprensión es de gran relevancia a fin de entender el comportamiento de sistemas químicos. Tales interacciones, además, juegan un papel fundamental en sistemas biológicos, dado que determinan las estructuras tridimensionales de las biomoléculas (ácidos nucleicos y proteínas) y otros sistemas químicos. Dado que las energías de interacción entre moléculas son mucho menores que las energías de enlace, una manera natural de comprender tales interacciones es empezar por monómeros aislados y describir las interacciones como pequeñas perturbaciones al sistema. Esta es la base de la teoría SAPT (symmetry-adapted perturbation theory), que permite descomponer energías de interacción en términos con un significado físico bien definido, básicamente en términos electrostáticos, de inducción (polarización), y de dispersión y repulsión (intercambio).⁸⁴ Tal descomposición, por tanto, permite una comparación con los términos no enlazantes descritos en campos de fuerzas clásicos en la Sección 1.1.1, aunque la descomposición SAPT incluye también términos cruzados, por ejemplo, de repulsión-inducción.

1.2.2 Interacciones catión- π -catión

La interacción catión- π , en la cual un catión interactúa con la nube electrónica de un sistema π , es esencialmente de origen electrostático y representa una de las interacciones no covalente más fuertes. Además, son un ejemplo paradigmático de interacciones donde la polarización electrónica juega un papel importante (ver Figura 5), dado que el catión puede distorsionar de forma considerable los electrones del sistema π . Junto a los enlaces de hidrógeno y las interacciones de apilamiento π - π , las interacciones catión- π constituyen las fuerzas más relevantes que controlan la estructura y función de macromoléculas.^{4,5,85,86}

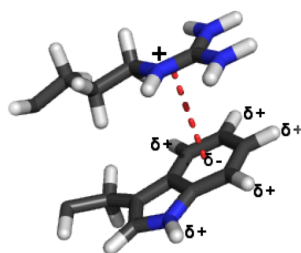


Figura 5. Interacción entre un grupo de átomos cargados positivamente y un sistema π .

Estudios pioneros realizados por Luque y Orozco en Barcelona demostraron que la contribución de polarización en este tipo de interacciones es muy importante, y que ésta puede incluso superar a la contribución electrostática, proveniente de un monopolo (catión) interactuando con un cuádruplo (anillo aromático).⁴ Por tanto, la presencia de una carga positiva cerca de un sistema π altamente polarizable conduce a efectos de polarización significativos, no descritos adecuadamente por campos de fuerzas MM estándar no polarizables.^{4,5,87-89} Esto a dado lugar a numerosos estudios incluyendo contribuciones significativas de los grupos del Prof. Luque en Barcelona y del Prof. Chipot en Nancy, con el objetivo de mejorar la descripción de tales interacciones mediante campos de fuerzas polarizables.⁸⁷⁻⁸⁹

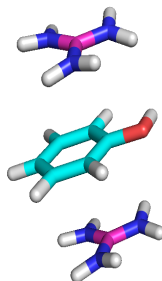


Figura 6. Interacción entre el sistema π aromático de una tyrosina y el grupo guanidinio catiónico de las argininas.

Aunque el papel y las características energéticas de las interacciones catión- π están bien establecidos, hay un caso especial y muy poco estudiado, que involucra a tres especies moleculares, denominada interacción catión- π -catión, por ejemplo, una interacción Arg-Tyr-Arg (ver Figura 6). En proteínas, por ejemplo, tal caso se da cuando interactúan simultáneamente dos cationes (las cadenas laterales de residuos como Arg o Lys) con un sistema π como el que se puede encontrar en la Phe, Tyr o Trp. En estas interacciones, el efecto de polarización es más complejo que en interacciones catión- π debido a su naturaleza a tres cuerpos. Esto da lugar a contribuciones de polarización no aditivas, más

CAPÍTULO 1. INTRODUCCIÓN

complejas de describir en simulaciones clásicas que las interacciones catión- π a dos cuerpos, y que no pueden ser descritas por campos de fuerzas aditivos. La posible cancelación parcial del campo eléctrico debido a la presencia de los dos cationes a cada lado del sistema π , sin embargo, también supone un reto complicado para campos de fuerzas polarizables, por ejemplo, basados en el modelo de dipolos inducidos puntuales o del oscilador de Drude. En la presente tesis, se ha investigado la ocurrencia y características geométricas de tales interacciones en las estructuras proteicas contenidas en el Protein Data Bank (PDB), así como su energética y la posibilidad de que tales interacciones contribuyan a la estabilización de estructuras terciarias de proteínas. Tal análisis energético se ha realizado mediante la metodología que se explicará en la siguiente sección.

1.2.3 Cálculo de energías de interacción

El cálculo de energías de interacción precisas, al igual que muchas otras propiedades, presenta dos principales problemas mediante la aplicación de métodos QM *ab initio* en los cuales los orbitales moleculares se representan como una combinación lineal de orbitales atómicos.⁹⁰ El primero reside en el tratamiento de la correlación electrónica mediante el método más preciso posible, mientras que el segundo está relacionado con el conjunto de funciones de base usado para describir los orbitales atómicos, tal y como ilustra la Figura 7.

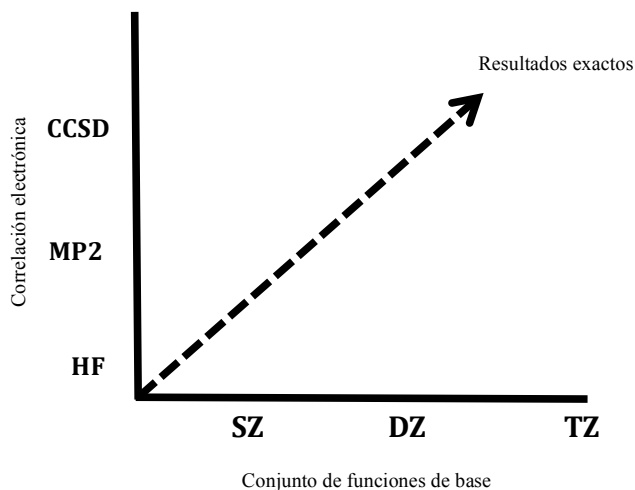


Figura 7. Esquema de los dos principales problemas asociados al cálculo de energías de interacción precisas: el tratamiento de la correlación electrónica y el conjunto de funciones de base usado.

El método MP2, que incluye los efectos de correlación electrónica basado en la teoría perturbativa de Møller-Plesset de segundo orden,⁹¹ es uno de los métodos más sencillos y populares capaz de incluir efectos de correlación electrónica mediante la aplicación de la teoría de perturbaciones a la solución del método Hartree-Fock (HF). A pesar del elevado coste computacional de la metodología MP2 comparado con la teoría HF, otros métodos post-HF para el tratamiento de la correlación electrónico, como los de interacción de configuraciones (CI) o el método *Coupled Cluster* (CC) exigen aún más recursos computacionales y su aplicación a sistemas de tamaño mediano-grande es inviable. La energía MP2 se puede descomponer en los términos descritos en las Ecs. 20 y 21, donde la energía de correlación electrónica contiene términos asociados a electrones con el mismo espín, $E_{corr(\alpha-\alpha)}$ y $E_{corr(\beta-\beta)}$, y de espín opuesto, $E_{corr(\alpha-\beta)}$:

$$E_{MP2} = E_{HF} + E_{corr} \quad (20)$$

$$E_{MP2} = E_{HF} + E_{corr(\alpha-\alpha)} + E_{corr(\beta-\beta)} + E_{corr(\alpha-\beta)} \quad (21)$$

En la presente tesis, se ha usado la variante SCS-MP2 (*spin-component-scaled* MP2) propuesta por Grimme en 2003,⁹² para estudiar la energética de interacciones catión- π -catión, ya que ésta ha demostrado mejorar significativamente los resultados de cálculos MP2. Este método consiste simplemente en escalar las componentes de la energía de correlación asociadas a electrones con espín igual y opuesto:⁹²

$$E_{SCS-MP2} = E_{HF} + \frac{1}{3}(E_{corr(\alpha-\alpha)} + E_{corr(\beta-\beta)}) + \frac{6}{5}E_{corr(\alpha-\beta)} \quad (22)$$

A fin de minimizar el impacto del conjunto limitado de funciones de base adoptado, además, las energías de interacción SCS-MP2 se han extrapolado hasta el límite del conjunto de funciones de base completo (*complete basis set*, CBS) utilizando el esquema propuesto por Truhlar⁹³ combinando cálculos SCS-MP2 realizados con las bases cc-pVDZ y cc-pVTZ de acuerdo con la fórmula siguiente:

$$E_{SCS-MP2/CBS} = \frac{3\alpha}{3\alpha-2\alpha}E_3^{HF} - \frac{2\alpha}{3\alpha-2\alpha}E_2^{HF} + \frac{3\beta}{3\beta-2\beta}E_3^{corr} - \frac{2\beta}{3\beta-2\beta}E_2^{corr} \quad (23)$$

donde E_2 y E_3 representan energías HF y de correlación obtenidas usando bases double- ζ y triple- ζ , siendo $\alpha = 3.4$ y $\beta = 2.2$.

CAPÍTULO 1. INTRODUCCIÓN

La adopción de conjuntos de funciones de base limitados, además, presenta un problema específico al cálculo de energías de interacción. Normalmente, las funciones de base están centradas en los núcleos, lo que permite que sean más compactas, pero se introduce un error porque la calidad del conjunto de las funciones no es la misma en los distintos cálculos de monómeros/dímeros debido a que la densidad electrónica alrededor de un núcleo puede ser descrita por funciones centradas en otros núcleos. Esto es, dentro de un complejo, el conjunto de funciones de una molécula puede ayudar a compensar el conjunto incompleto de la otra molécula, y viceversa. De esta manera su energía se ve rebajada de una manera artificial y la energía de interacción es sobreestimada. Este efecto es conocido como el error de superposición de base (*basis set superposition error*, BSSE). En el límite teórico que corresponde al uso de un conjunto de funciones de base completo, este error es cero. En esta tesis, el error BSSE se ha corregido en las energías de interacción de dímeros y trímeros calculadas a nivel SCS-MP2/CBS utilizando la corrección *counterpoise* (CP),⁹⁴ donde los cálculos de monómeros utilizan el conjunto de bases completo de los sistemas diméricos o triméricos con el cual interaccionan. Por ejemplo, las energías de interacción de los trímeros calculadas son susceptibles al error BSSE según la Ec. 24, y la corrección CP se realiza utilizando el conjunto de base del trímero, indicado mediante los subíndices presentes en la Ec. 25:

$$\Delta E_{int(ABC)} = E(ABC)_{(ABC)} - E(A)_{(A)} - E(B)_{(B)} - E(C)_{(C)} \quad (24)$$

$$\Delta E_{int(ABC)} = E(ABC)_{(ABC)} - E(A)_{(ABC)} - E(B)_{(ABC)} - E(C)_{(ABC)} \quad (25)$$

1.3 Impacto de la polarización en interacciones entre estados excitados

1.3.1 Transferencia de energía electrónica

La transferencia de energía electrónica es el proceso mediante el cual un dador inicialmente excitado se relaja al estado fundamental transfiriendo su energía a un aceptor cercano de forma no radiativa, es decir, sin la emisión de un fotón.¹⁴ En 1948, Theodor Förster desarrollo una teoría elegante y práctica que aún hoy representa el marco conceptual básico para comprender estos procesos, hecho por el cual este proceso fotofísico también es comúnmente llamado transferencia de energía de resonancia de Förster (*Förster resonance energy transfer*, FRET).¹³ Este acrónimo, sin embargo, también es usado para referirse específicamente a la técnica FRET, que combina el potencial de este fenómeno con la fluorescencia, técnica espectroscópica de gran importancia para el estudio de la estructura de sistemas biológicos debido su alta sensibilidad y a la posibilidad de estudiar moléculas individuales.

La relevancia de la fluorescencia en el estudio de biosistemas, por ejemplo, queda reflejada en los premios Nobel de Química otorgados en el 2008 y el 2014 por el descubrimiento de la proteína verde fluorescente o el desarrollo de la microscopía de fluorescencia de alta resolución. La técnica FRET, en particular, se basa en detectar mediante cambios en la fluorescencia la eventual transferencia de energía de excitación entre moléculas, ya sean estas fluoróforos intrínsecos (por ejemplo el triptófano) o fluoróforos extrínsecos incorporados a una biomolécula. Debido a que la eficiencia de FRET es muy sensible a la separación entre dador/aceptor, esta técnica permite medir distancias y detectar interacciones en sistemas biológicos incluso a nivel de moléculas individuales y a tiempo real.⁹⁵⁻⁹⁷

Esta técnica, propuesta por Stryer y Haugland a finales de los 60,⁹⁸ se usa, por ejemplo, para observar la cinética de plegamiento de proteínas o para observar el funcionamiento de máquinas moleculares tales como transportadores, además de muchas otras aplicaciones clave en la comprensión a nivel molecular de mecanismo biológicos.

En un experimento FRET, típicamente se unen covalentemente dos moléculas fluorescentes a la biomolécula de interés. Una de ellas, el dador (D), se excita selectivamente mediante absorción de luz, y en ciertas condiciones, tal energía de excitación se puede

CAPÍTULO 1. INTRODUCCIÓN

transferir de forma no radiactiva al aceptor (A) a distancias de hasta 100 Å. La medición de la fluorescencia del D y A, de esta forma, permite cuantificar la eficiencia de FRET, eficiencia que puede ser convertida en distancia D/A mediante la teoría de Förster, permitiendo así obtener información acerca de la estructura de la biomolécula o de sus interacciones con otras biomoléculas. La exactitud de tal información estructural, por tanto, depende básicamente de las aproximaciones usadas en la teoría de Förster para convertir los datos espectroscópicos en información estructural. A continuación, se describe esta teoría, así como las aproximaciones fundamentales en las cuales se basa.

1.3.2 Teoría de Förster

La teoría de Förster, introducida hace más de 60 años,¹³ fue clave en la comprensión de la dinámica de procesos FRET y en la explosión de aplicaciones basadas en este fenómeno, ya que permite estimar la velocidad de FRET a partir de datos puramente espectroscópicos. Según ésta, la velocidad de FRET entre un dador (D) y un aceptor (A) débilmente acoplados depende de su separación, R , expresada en cm, su orientación relativa (a través del factor de orientación κ), y el solapamiento espectral I entre los espectros de emisión de D y de absorción de A (en unidades de $M^{-1}cm^3$):

$$k_{FRET} = \frac{1}{\tau_D} \frac{9000(\log 10)\kappa^2\phi_D I}{128\pi^5 N_A n^4} \frac{1}{R^6} \quad (26)$$

donde N_A es el número de Avogadro (en unidades de mol^{-1}), n es el índice de refracción del medio, ϕ_D es el rendimiento cuántico de fluorescencia y τ_D el tiempo de vida del dador (en las mismas unidades que $1/k_{FRET}$).

Otra forma de expresar la velocidad es introduciendo R_0 , el radio crítico de Förster, en el cual la eficiencia es de 0.5:

$$k_{FRET} = \frac{1}{\tau_D} \left(\frac{R_0}{R}\right)^6 \quad (27)$$

Esta teoría, pues, asume que la dinámica de FRET sigue la conocida dependencia R^{-6} con la distancia entre D y A, relación que tiene su origen en la aproximación dipolo-dipolo usada para describir el acoplamiento electrónico entre D y A responsable de la transferencia. Para comprender mejor el origen de esta dependencia y de la Ec. 26, así como para estimar tal

velocidad a partir de cálculos teóricos, es preferible sin embargo expresar la teoría de Förster mediante la regla de oro de la mecánica cuántica:

$$k_{FRET} = \frac{2\pi}{\hbar} V^2 J \quad (28)$$

donde V representa el acoplamiento D/A, y J la integral de solapamiento entre los espectros normalizados de emisión del dador $f_D(\epsilon)$ y de absorción del aceptor $a_A(\epsilon)$, ilustrada en la Figura 8:

$$J = \int_0^\infty f_D(\epsilon) a_A(\epsilon) d\epsilon \quad (29)$$

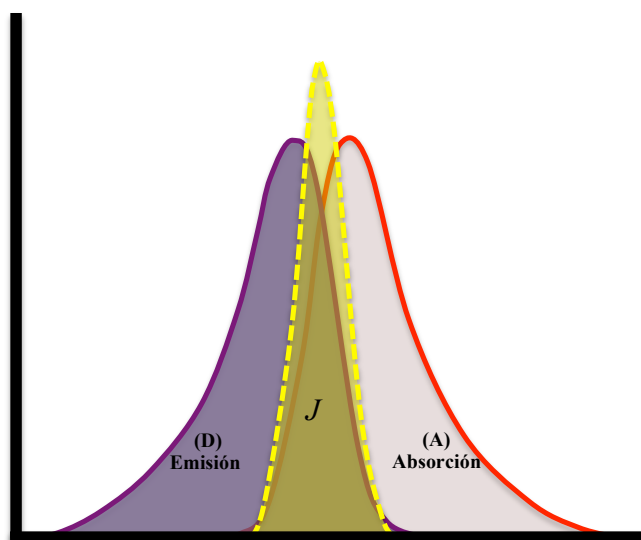


Figura 8. Solapamiento espectral entre el espectro de emisión del dador y el de absorción del aceptor.

La derivación de la Ec. 28 asume que el acoplamiento entre D/A es pequeño comparado con la energía de reorganización de los cromóforos, mientras que el factor de solapamiento espectral en la Ec. 29 asume que las excitaciones en D y A están acopladas a vibraciones independientes del ambiente.⁹⁹ Actualmente hay un gran interés en la derivación de expresiones válidas también para el régimen de acoplamiento intermedio y fuerte, especialmente en el estudio de complejos fotosintéticos, aunque en la mayoría de sistemas donde se aplica FRET para derivar información estructural la aproximación de acoplamiento débil es válida.¹⁰⁰ La dependencia R^{-6} de la velocidad, sin embargo, está encapsulada principalmente en el término V^2 de la Ec. 28, que la teoría de Förster aproxima mediante una aproximación dipolo-dipolo y una descripción continua de las propiedades dieléctricas del medio mediante su índice de refracción:

CAPÍTULO 1. INTRODUCCIÓN

$$V \approx sV_{dip-dip} \quad (30)$$

$$s = \frac{1}{n^2} \quad (31)$$

$$V_{dip-dip} = \frac{\kappa \mu_D^T \mu_A^T}{R^3} \quad (32)$$

donde el factor de orientación corresponde a

$$\kappa = \hat{\mu}_D^T \cdot \hat{\mu}_A^T - 3(\hat{\mu}_D^T \cdot \hat{R})(\hat{\mu}_A^T \cdot \hat{R}) \quad (33)$$

donde μ_D^T/μ_A^T indican los dipolos de transición de las moléculas D y A y R su separación ($\hat{\mu}_D^T, \hat{\mu}_A^T$ y \hat{R} son los correspondientes vectores unitarios).

Como se puede observar en las ecuaciones descritas anteriormente, la velocidad de FRET depende de varios factores: la integral de solapamiento J , la distancia D/A, y la orientación y magnitud de los dipolos de transición de D y A. En la última década, sin embargo, avances en el estudio de sistemas naturales fotosintéticos y de nanomateriales con aplicaciones en energía solar han mostrado la importancia de distintas contribuciones a la dinámica de FRET más allá del modelo de Förster, tales como desviaciones en el factor de apantallamiento inducidos por el entorno o desviaciones significativas en la aproximación dipolo-dipolo, que modifican la dependencia con la distancia de la velocidad de FRET, avances pues que tienen importantes implicaciones en la derivación de datos estructurales a partir de la técnica FRET.⁹⁹ La incorporación de tales efectos, en particular, se puede realizar mediante cálculos que tengan en cuenta las densidades de transición completas en D/A, evitando la aproximación de dipolos puntuales, así como la incorporación del impacto del ambiente en la estimación de los acoplamientos electrónicos, metodologías usadas en esta tesis que se explican a continuación.

1.3.3 Acoplamiento electrónico

El acoplamiento electrónico es el parámetro clave que impulsa un proceso FRET y explica su dependencia con la separación y la orientación mutua entre D y A. En la última década, ha habido importantes avances en el desarrollo de métodos teóricos QM capaces de estimar con precisión esta cantidad.¹² En esta tesis se ha usado la metodología QM/MMPol desarrollada en el grupo, que permite incluir el efecto del entorno polarizable heterogéneo que

rodea a los cromóforos en el cálculo de los acoplamientos mediante la combinación del método TD-DFT con campos de fuerzas polarizables.^{9,10}

El acoplamiento electrónico se puede descomponer en una contribución coulombica de largo alcance, V^{largo} , que corresponde al mecanismo de Förster y al término dipolo-dipolo descrito en la Ec. 30, y un término de corto alcance, V^{corto} , que depende de la superposición entre las funciones de onda de D y A, y depende exponencialmente de la distancia D/A, asociado al mecanismo Dexter de transferencia de energía entre estados triplete, o sin fuerza de oscilador donde el término Coulomb es cero:^{101,102}

$$V = V^{largo} + V^{corto} \quad (34)$$

En procesos de transferencia de energía singulete-singulete entre estados permitidos, el término Coulomb de largo alcance es dominante y se puede ignorar el término V^{corto} , que solo es importante a distancias de contacto entre D/A.

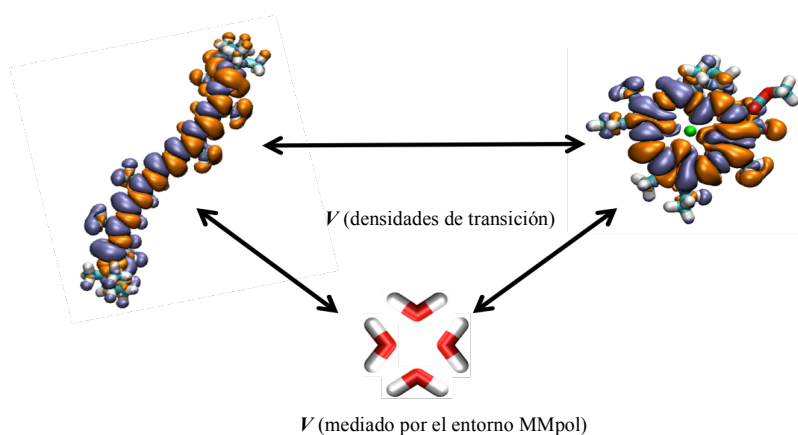


Figura 9. Ilustración de las contribuciones directa y mediada por el entorno al acoplamiento coulombico entre dos cromóforos.

A separaciones comparables a las dimensiones de los cromóforos, la aproximación dipolo-dipolo de Förster presenta desviaciones importantes. En estos casos, se puede calcular directamente la interacción coulombica entre densidades de transición derivadas, por ejemplo, de un cálculo TD-DFT, sin tener que recurrir a una expansión multipolar. En la estrategia desarrollada por el grupo de Barcelona en colaboración con la Prof. Mennucci en Pisa, además, es posible incorporar el impacto de la polarización del ambiente en el cálculo de los acoplamientos entre las densidades de transición ya sea mediante el modelo continuo PCM o un modelo MM polarizable.¹² En este modelo, de esta forma, los acoplamientos electrónicos

CAPÍTULO 1. INTRODUCCIÓN

se calculan a partir de las densidades de transición derivadas de los cálculos de estado excitado QM/PCM o QM/MMPol, y los efectos de apantallamiento del entorno, descritos por el término $1/n^2$ en la Ec. 31 por Förster, se incorporan naturalmente mediante el cálculo explícito de un término mediado por el ambiente al acoplamiento coulombico entre D y A:

$$V = V_{coul} + V_{env} \quad (35)$$

donde V_{coul} se refiere al termino coulombico de interacción directa entre D/A, y V_{env} al término mediado por el ambiente. Las siguientes ecuaciones describen las expresiones usadas para calcular estos términos ya sea con el modelo PCM o MM polarizable:⁹⁻¹¹

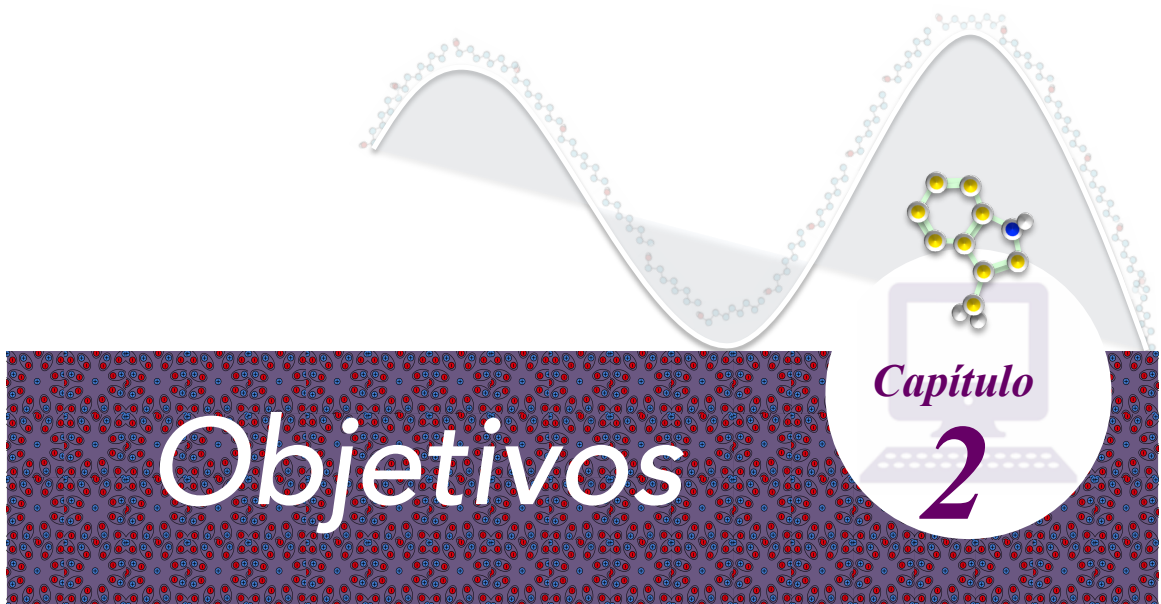
$$V_{coul} = \int d\mathbf{r} \int d\mathbf{r}' \rho_A^{T*}(\mathbf{r}') \frac{1}{|\mathbf{r}' - \mathbf{r}|} \rho_D^T(\mathbf{r}) \quad (36)$$

$$V_{PCM} = \sum_l \int d\mathbf{r} \rho_A^{T*}(\mathbf{r}) \frac{1}{|s_l - \mathbf{r}|} q_l^{PCM}(\rho_D^T) \quad (37)$$

$$V_{MMPol} = - \sum_k \int d\mathbf{r} \rho_A^{T*}(\mathbf{r}) \frac{(\mathbf{r}_k - \mathbf{r})}{|\mathbf{r}_k - \mathbf{r}|^3} \mu_k^{MMPol}(\rho_D^T) \quad (38)$$

donde $\rho_{D/A}^T$ indica las densidades de transición de los cromóforos, y μ_k^{MMPol} y q_l^{PCM} los dipolos inducidos MMPol y las cargas superficiales aparentes PCM que describen la polarización del entorno en respuesta a una densidad de transición.

El modelo QM/PCM, por ejemplo, permite incluir el impacto de la forma y orientación de los cromóforos, efecto que introduce una nueva contribución exponencial a la dependencia con la distancia de la velocidad de FRET.^{103,104} Por otra parte, el modelo QM/MMPol permite tener en cuenta a nivel molecular los efectos de apantallamiento ejercidos por el entorno polarizable heterogéneo en la velocidad de transferencia de energía en ambientes complejos, por ejemplo un cromóforo en el interior de una proteína. Distintos estudios de la dinámica de FRET en complejos fotosintéticos han mostrado que tal naturaleza dieléctrica heterogénea puede introducir variaciones de hasta 4 veces en la velocidad de FRET comparados con estimaciones basadas en una descripción dieléctrica continua.^{10,105} En esta tesis, se pretende investigar el impacto de tales efectos en la comprensión de procesos FRET en complejos ligando-proteína, donde distintos modos de unión caracterizados con distancias D/A similares pueden sin embargo dar lugar a velocidades FRET distintas debido a cambios en el entorno polarizable del ligando, un efecto ignorado en la teoría de Förster.



Objetivos

Capítulo

2

2. OBJETIVOS

El objetivo principal de la presente tesis doctoral es profundizar en la comprensión del impacto de las fuerzas de polarización electrónica en interacciones intermoleculares de sistemas biológicos. En particular, se han estudiado dos casos paradigmáticos de interacciones donde su efecto es primordial, tanto a nivel de estado fundamental como de estado excitado, ambos evaluados en contextos biológicos.

2.1 Impacto de las fuerzas de polarización electrónica en interacciones catión- π -catión

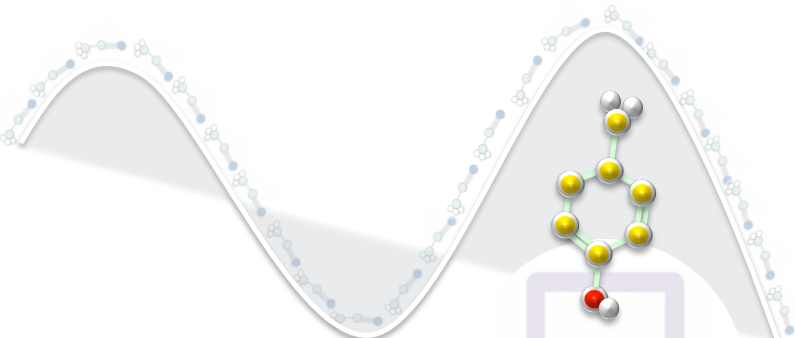
Esta línea de investigación tiene como objetivo profundizar en la comprensión de un tipo de interacciones biológicas a tres cuerpos donde se espera un impacto significativo de la polarización, las interacciones catión- π -catión. Además, también pretende profundizar en distintos aspectos relativos a este tipo de interacciones, de las cuales se sabe muy poco:

- Determinar la frecuencia, composición y geometría de las interacciones catión- π -catión presentes en un conjunto de proteínas no redundante de la base de datos PDB.
- Analizar el grado de conservación de los aminoácidos que forman tales interacciones.
- Estimar las energías de interacción a dos y tres cuerpos de los motivos catión- π -catión más relevantes identificados en el PDB mediante cálculos QM de alto nivel.
- Evaluar la contribución no aditiva en las interacciones catión- π -catión mediante comparación de las energías a dos y tres cuerpos.
- Determinar el impacto del entorno proteico en la fuerza de las interacciones catión- π -catión y analizar su potencial papel estabilizador de la estructura terciaria de proteínas.

2.2 Impacto de la polarización en procesos de transferencia de energía

Esta línea de investigación de la tesis tiene como objetivo profundizar en la comprensión del impacto que la naturaleza polarizable heterogénea juega en la interpretación de fenómenos de transferencia de energía de excitación en sistemas proteína-ligando. En particular, se investiga el *quenching* dinámico enantioselectivo observado en la fluorescencia del flurbiprofeno (FBP) en su complejo con la albúmina sérica humana (HSA), así como en un sistema modelo FBP-TrpMe en disolución. El *quenching* ha sido atribuido a un proceso de transferencia de energía entre el ligando y el Trp214 de la HSA (o el grupo TrpMe en el sistema modelo). Esta investigación, además, pretende determinar el potencial de la metodología MD-QM/MM polarizable desarrollada en el grupo, pionera en la incorporación de los efectos de polarización del entorno en este contexto, a fin de caracterizar uniones ligando-receptor y ayudar al descubrimiento de nuevos sitios de unión:

- Evaluar el apantallamiento ejercido por el entorno polarizable heterogéneo en los acoplamientos excitónicos entre el estado $\pi \rightarrow \pi^*$ del FBP y el estado *La* del Trp en los complejos FBP/HSA y el sistema FBP-TrpMe.
- Cuantificar la velocidad de transferencia de energía y la proporción de velocidades entre los enantiómeros (R) y (S) del FBP y validar la hipótesis experimental sobre el origen del *quenching* enantioselectivo.
- Comparar de forma crítica la capacidad de distintas metodologías de química cuántica (métodos ZINDO, CIS y TD-DFT) para describir las velocidades de transferencia así como su enantioselectividad.
- Determinar el impacto que supone la aproximación dipolo-dipolo y la adopción de un factor de apantallamiento constante en el cálculo de los acoplamientos electrónicos en la teoría de Förster ampliamente usada en este contexto.
- Validar los modos de unión propuestos para los enantiómeros del FBP en el complejo con la HSA a partir de las simulaciones de MD mediante comparación de las velocidades de transferencia estimadas y derivadas de los *quenchings* experimentales.



Capítulo
3

Resultados

3. RESULTADOS

La investigación desarrollada en esta tesis doctoral ha dado lugar a dos publicaciones científicas. A continuación, se indican las referencias correspondientes, y se reproducen las publicaciones acompañadas de su resumen, las cuales detallan los resultados obtenidos:

Silvana Pinheiro and Carles Curutchet, *Can Förster Theory Describe Stereoselective Energy Transfer Dynamics in a Protein-Ligand Complex?* *Journal of Physical Chemistry B*, **2017**, 121 (10), 2265–2278.

Silvana Pinheiro, Ignacio Soteras, Josep Lluís Gelpí, François Dehez, Christophe Chipot, F. Javier Luque, and Carles Curutchet, *Structural and energetic study of cation- π -cation interactions in proteins*, *Physical Chemistry Chemical Physics*, **2017**, 19, 9849-9861.

CAPÍTULO 3. RESULTADOS

Structural and energetic study of cation– π –cation interactions in proteins

Physical Chemistry Chemical Physics, **2017**, 19, 9849-9861

Silvana Pinheiro,¹ Ignacio Soteras,² Josep Lluís Gelpí,³ François Dehez,⁴

Christophe Chipot,^{4,5} F. Javier Luque,² and Carles Curutchet¹

¹*Departament de Farmàcia i Tecnologia Farmacèutica i Fisicoquímica and Institut de Biomedicina (IBUB), Facultat de Farmàcia, Universitat de Barcelona, Barcelona Spain*

²*Departament de Nutrició, Ciències de l'Alimentació i Gastronomia and Institut de Biomedicina (IBUB), Facultat de Farmàcia, Universitat de Barcelona, Santa Coloma de Gramenet, Spain.*

³*Departament de Bioquímica i Biomedicina Molecular, Facultat de Biologia, Universitat de Barcelona, Spain*

⁴*Laboratoire International Associé Centre National de la Recherche Scientifique et University of Illinois at Urbana –Champaign, Unité Mixte de Recherche No. 7565, Université de Lorraine, Vandoeuvre-lès-Nancy cedex, France*

⁵*Beckman Institute for Advanced Science and Engineering, University of Illinois at Urbana — Champaign, Urbana, Illinois (USA).*

CAPÍTULO 3. RESULTADOS

3.1 Artículo 1: Structural and energetic study of cation- π -cation interactions in proteins.**Resumen**

Las estructuras biológicas se estabilizan mediante una variedad de interacciones no covalentes, tales como enlaces de hidrógeno, apilamiento π , puentes salinos o interacciones hidrofóbicas. Además de enlaces de hidrógeno y apilamiento π , las interacciones catión- π entre anillos aromáticos y grupos cargado positivamente han surgido como una de las más importantes interacciones en biología estructural. Aunque el papel y las características energéticas de estas interacciones está bien establecido, un caso especial de ésta que involucra a tres especies moleculares, denominada interacción catión- π -catión, es aún poco comprendida. Consiste en la interacción simultánea entre un sistema π como la fenilalanina (Phe, F), tirosina (Tyr, Y) o triptófano (Trp, W), y dos cationes como las cadenas laterales de lisina (Lys, K) o arginina (Arg, R). En comparación con interacciones catión- π , en estos motivos el efecto de polarización es más complejo debido a su naturaleza a tres cuerpos, hecho que puede dar lugar a efectos no aditivos que no pueden ser descritos por campos de fuerzas aditivos. En este estudio, hemos avanzado en la comprensión de las interacciones catión- π -catión y su papel en la estructura y estabilidad de biosistemas a través de dos enfoques complementarios. El primero consiste en un estudio estadístico de la ocurrencia, composición y geometría de las interacciones catión- π -catión identificadas en un conjunto no redundante de estructuras proteicas del PDB, que demuestra que las interacciones catión- π -catión son de hecho comunes en proteínas. La segunda parte del estudio consiste en un análisis energético de las interacciones más relevantes estimadas mediante cálculos SCS-MP2/CBS acoplados a modelos continuos de solvatación. Además de la conocida deficiencia de los campos de fuerzas aditivos para describir la contribución de polarización, muy relevante en interacciones catión- π , nuestros resultados indican que las contribuciones no aditivas a tres cuerpos son significativas en este caso. De esta forma, se espera que las interacciones catión- π -catión no sean descritas correctamente por campos de fuerzas aditivos. Nuestros resultados demuestran que este motivo estructural se observa en una frecuencia pequeña, aunque no despreciable, en proteínas, y sugiere una preferencia para establecer motivos catión- π -catión con Trp, seguido de Tyr y Phe. Además, demostramos que los aminoácidos que forman las interacciones catión- π -catión identificadas tienden a estar altamente conservados, lo que apoya un posible papel estructural o funcional. El análisis energético de un subconjunto representativo de los complejos catión- π -catión seleccionados

CAPÍTULO 3. RESULTADOS

señala además que el entorno proteico puede apantallar fuertemente la repulsión catión-catión, dando lugar a una interacción atractiva en el 64% de los complejos analizados. Junto con el alto grado de conservación observado, estos resultados sugieren un posible papel estabilizador en el plegamiento de proteínas. Desde un punto de vista computacional, la contribución significativa de los términos no aditivos a tres cuerpos cuestiona la idoneidad de los campos de fuerzas aditivos estándar para describir los motivos catión- π -catión en simulaciones moleculares.



Cite this: *Phys. Chem. Chem. Phys.*,
2017, **19**, 9849

Structural and energetic study of cation– π –cation interactions in proteins†

Silvana Pinheiro, ^a Ignacio Soteras, ^b Josep Lluís Gelpi, ^c François Dehez, ^d
Christophe Chipot,^{de} F. Javier Luque ^b and Carles Curutchet ^{*a}

Cation– π interactions of aromatic rings and positively charged groups are among the most important interactions in structural biology. The role and energetic characteristics of these interactions are well established. However, the occurrence of cation– π –cation interactions is an unexpected motif, which raises intriguing questions about its functional role in proteins. We present a statistical analysis of the occurrence, composition and geometrical preferences of cation– π –cation interactions identified in a set of non-redundant protein structures taken from the Protein Data Bank. Our results demonstrate that this structural motif is observed at a small, albeit non-negligible frequency in proteins, and suggest a preference to establish cation– π –cation motifs with Trp, followed by Tyr and Phe. Furthermore, we have found that cation– π –cation interactions tend to be highly conserved, which supports their structural or functional role. Finally, we have performed an energetic analysis of a representative subset of cation– π –cation complexes combining quantum-chemical and continuum solvation calculations. Our results point out that the protein environment can strongly screen the cation–cation repulsion, leading to an attractive interaction in 64% of the complexes analyzed. Together with the high degree of conservation observed, these results suggest a potential stabilizing role in the protein fold, as demonstrated recently for a miniature protein (Craven *et al.*, *J. Am. Chem. Soc.* 2016, **138**, 1543). From a computational point of view, the significant contribution of non-additive three-body terms challenges the suitability of standard additive force fields for describing cation– π –cation motifs in molecular simulations.

Received 10th December 2016,
Accepted 13th March 2017

DOI: 10.1039/c6cp08448f

rsc.li/pccp

1. Introduction

The comprehension of intermolecular forces is key to understanding the behavior of chemical systems at the molecular level. This is well exemplified in biological systems, given that the molecules of life (DNA, RNA and proteins) are held in their three-dimensional structures by a variety of intermolecular noncovalent interactions, and that mutational changes in proteins can affect the structure and dynamics, and even ultimately the evolution toward novel functions in enzymes.¹ Because the

three-dimensional molecular structure of these molecules is responsible for their specific biological activity, understanding noncovalent interactions is fundamental to rationalize the relationships between structure, dynamics, and function. Furthermore, such knowledge is essential for engineering proteins toward the design of enzymes with novel catalytic roles.^{2,3}

In the last few decades, there has been a continuous effort aimed at identifying and characterizing the many types of noncovalent interactions observed in chemical systems, including, for example, non-conventional hydrogen bonds, π -stacking, cation– π and anion– π contacts, as well as halogen bonding.^{4–19} Cation– π interactions have emerged as one of the most important interactions in determining specific recognition motifs in biomolecules.^{4–7,11,12} The term cation– π describes the interaction between a positively charged atom or group and a π -electron system. Theoretical studies have elucidated the nature of cation– π interactions, establishing that the leading term involves the electrostatic attraction between a monopole (the cation charge) and the quadrupole created by the π -electron cloud of the aromatic ring.^{6,7,13,20–22} The presence of a positive charge near a highly polarizable π -system, however, leads also to important induction contributions to the interaction energy.^{6,7,13,20} In addition, dispersion contributions, although smaller, can also

^a Departament de Farmàcia i Tecnologia Farmacèutica i Físicoquímica and Institut de Biomedicina (IBUB), Facultat de Farmàcia i Ciències de l'Alimentació, Universitat de Barcelona, Barcelona, Spain. E-mail: carles.curutchet@ub.edu

^b Departament de Nutrició, Ciències de l'Alimentació i Gastronomia and Institut de Biomedicina (IBUB), Facultat de Farmàcia i Ciències de l'Alimentació, Universitat de Barcelona, Santa Coloma de Gramenet, Spain

^c Departament de Bioquímica i Biomedicina Molecular, Facultat de Biologia, Universitat de Barcelona, Spain

^d Laboratoire International Associé Centre National de la Recherche Scientifique et University of Illinois at Urbana – Champaign, Unité Mixte de Recherche No. 7565, Université de Lorraine, Vandoeuvre-lès-Nancy cedex, France

^e Beckman Institute for Advanced Science and Engineering, University of Illinois at Urbana – Champaign, Urbana, Illinois, USA

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c6cp08448f

be important depending on the nature of the cation involved in the interaction.^{13,20} On the other hand, solvation and more generally environment effects may have a crucial influence on the stability of cation- π complexes given the charged nature of the cation.²³ Cation- π interactions are not properly described in standard pairwise additive force fields used in biomolecular simulations, where polarization effects are only taken into account in an implicit way. This shortcoming can be naturally overcome by switching to an explicit polarizable force field,^{24,25} though some research groups have also proposed empirical corrections to additive force fields to improve the description of cation- π interactions.^{7,26–30} Such corrections, nevertheless, cannot account for nonadditive effects, which are important to describe such interactions.³¹

There is a large body of evidence illustrating the common occurrence of cation- π interactions in biological systems, including peptides, nucleic acids, enzymes, antibody-binding interactions or ligand-protein binding.¹² Statistical analysis of the structures deposited in the Protein Data Bank (PDB)³² has unveiled the occurrence and geometric characteristics of cation- π interactions in biomolecules.⁷ In addition, some research groups have shown the existence of cooperative effects among cation- π interactions and hydrogen bonds or π -stacking.^{9,33–42} Despite all these advances in characterizing cation- π motifs and their interplay with other interactions, a special case of cation- π interaction involving three molecular species, termed cation- π -cation motif, has hitherto received little attention and is still poorly understood. The ability of cation- π -cation interactions to stabilize tertiary structures of proteins has been recently demonstrated for a miniature protein, which emulates a network of cation- π interactions found in fibronectin type III domain containing proteins.⁴³

In this contribution, we aim at advancing our understanding of cation- π -cation interactions and their role in the structure and stability of biomolecular systems *via* two complementary approaches. First, a thorough statistical analysis of the frequency, composition and geometrical features of cation- π -cation complexes identified in a set of non-redundant protein structures from the PDB is presented. Specifically, interactions between the π -electron system of the aromatic residues with not only cationic amino acids, but also different monovalent and divalent cations have been considered. We have also analyzed the degree of conservation of the interactions based on the probability of residue replacement as well as the relations among the protein and its homologs. Second, the most relevant contributions to the interaction energy in cation- π -cation interactions from high-level quantum mechanical (QM) calculations have been analyzed. This study demonstrates that cation- π -cation interactions are unexpectedly quite common in proteins, as about 7% of the structures investigated have one or more cation- π -cation interactions. The results also indicate that the stability of a cation- π -cation complex, which is intrinsically repulsive, is largely dependent on the local environment. Finally, non-additive three-body contributions are significant in cation- π -cation interactions, implying that the proper description of these motifs in standard additive pairwise force fields is even more challenging from a theoretical perspective compared to cation- π pairs.

The present contribution is organized as follows. First, we discuss the methodological approach used in this work and the computational details of the QM calculations. Second, we report the results of the statistical analysis performed on the set of non-redundant proteins of the PDB, as well as the analysis of the conservation of the most relevant cation- π -cation motifs. Then, we discuss the energetic analysis of the most representative subset of these interactions. Finally, we close with future prospects and possible extensions of the work.

2. Methods

2.1 Statistical analysis

The statistical analysis of cation- π -cation interactions was performed over a non-redundant protein structure dataset consisting of *ca.* 21 000 protein structures taken from the PDB's Cluster90.^{32,44} We performed an exhaustive search for cation- π -cation motifs established between the side chains of aromatic residues (Phe, Tyr, Trp) with cationic amino acids (Lys and Arg), as well as with different monovalent and divalent cations (Na^+ , Ca^{2+} , Mg^{2+} , K^+ , Li^+ , Cu^{2+}). The criteria to detect a cation- π -cation complex were based on the distance from the cation to the aromatic ring, as well as on the angle formed by the two cations and the aromatic ring. The latter is defined as the angle between the normal of the aromatic ring and the vector connecting the centroid of the ring and the positively charged site (Fig. 1). In the case of Lys and Arg, the distance was taken to the N_ζ atom of the protonated amine and to the guanidine C_ζ atom, respectively. In the case of Trp, both six- and five-membered aromatic rings were considered. The set of interactions initially identified was then filtered in order to remove possible duplicates in each interaction arising from symmetry considerations within the PDB.

For the sake of completeness, we have also extended the analysis to complexes that include His as a potential partner. Whereas most Lys and Arg residues are expected to be cationic, it is well known that His residues generally populate the neutral state, even though they may be protonated depending on their local environment in proteins. Due to the lack of information regarding the protonation states in the PDB structures, and the

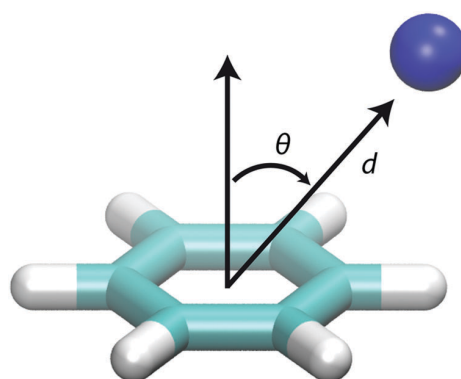


Fig. 1 Definition of the geometrical parameters that characterize the cation- π interaction between a π ring and a positively charged group.

range of uncertainty in predicting the pK_a of His residues,^{45,46} we have decided to extend our statistical analysis to cation- π -His and His- π -His complexes, since even a neutral His might *a priori* be an unexpected motif to assist in the formation and stability of the complex through hydrogen bonding to the aromatic ring. The protonation state of the His residues participating in the complexes identified was then investigated using the PROPKA server.^{45,46} In the case of complexes involving His, the geometrical parameters (Fig. 1) were defined to determine the relative position of the nitrogen atoms of the His residue relative to the aromatic ring.

Taking into account the preceding definitions for the center of the aromatic rings and the location of the charged sites, we considered a threshold value for the maximum distance (d) between the cation and the aromatic ring of 5 Å, and an angle cutoff (θ) for the cation-aromatic ring-cation arrangement of 45 degrees (Fig. 1). The initial set of interactions was then refined to identify the most relevant ones by imposing a more stringent criterion, consisting of a combined distance and angle cutoff for both cations, such that $d_1 + d_2 < 8$ Å and $\theta_1 + \theta_2 < 40^\circ$, respectively. From this new subset, we removed (i) complexes containing His residues, due to the uncertainty related to its protonation state, (ii) Lys complexes where the ammonium group was largely displaced from the normal axis to the aromatic ring, and (iii) proteins with a high sequence similarity with other proteins in the set. Some proteins of this set featured multiple interactions (PDB IDs 1GXS, 1T72, 2H39, 2J47, 2JKV, 2PAM and 3G9V), which were kept, although they did not meet our geometric criteria, as we deemed it interesting to investigate the energetic characteristics of these special cases. Moreover, we also included in the final set of interactions a particularly interesting system, namely the human growth hormone receptor (PDB entry 1A22), where three consecutive cation- π -cation interactions are established involving amino acids Lys379, Trp386, Arg411, Phe425, Arg413, Tyr422, and Lys415, as shown in Fig. 2. The final refined dataset included 53 interactions (see Results section). For this final set of interactions, we analyzed the degree of conservation of the amino acids involved in the complex, and performed the energetic analysis described in the following section.

The degree of conservation of the cation- π -cation interactions in each protein was determined with the aid of the ConSurf web server,^{47,48} which accounts for the probability of residue replacement, as well as the evolutionary relations among the protein and its homologs. Finally, we also computed the pK_a and the fraction of buried surface area of the titratable amino acids in the cation- π -cation interactions using the PROPKA server.^{45,46} In all cases, the Arg and Lys amino acids were found to be protonated at neutral pH (see Table S7 in the ESI†).

2.2 Energetic analysis

For each complex featured in the subset of cation- π -cation motifs, a simplified model was built up to perform the energetic analysis using high-level QM calculations. To this end, Lys and Arg amino acids were represented as ammonium and guanidinium ions, whereas Phe, Tyr, and Trp were represented as benzene,

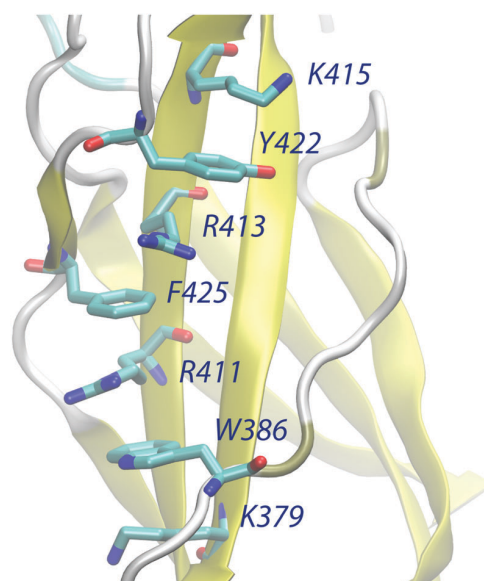


Fig. 2 Multiple cation- π -cation interactions identified in the human growth hormone receptor (PDB entry 1A22).

phenol, and indole units, respectively. The geometries of these prototypical molecules were fully optimized at the MP2/aug-cc-pVDZ level using the Gaussian 09 suite of programs.⁴⁹ The optimized geometries were then aligned onto the corresponding fragments of the amino-acid side chains, thereby forming the cation- π -cation motif as found in the original PDB structures. For the models involving Lys, the orientation of the hydrogen atoms in the ammonium group was then optimized through MP2/aug-cc-pVDZ calculations, while keeping the rest of the atoms frozen in the model system.

The trimer interaction energy, $\Delta E_{XArY} = E_{XArY} - E_X - E_{Ar} - E_Y$, as well as the corresponding pair interaction energies, $\Delta E_{XAr} = E_{XAr} - E_X - E_{Ar}$, $\Delta E_{ArY} = E_{ArY} - E_{Ar} - E_Y$ and $\Delta E_{XY} = E_{XY} - E_X - E_Y$, was determined to assess the contribution of three-body non-additive effects to the interaction according to the following expression:

$$\Delta E_{3\text{-body}} = \Delta E_{XArY} - \Delta E_{XAr} - \Delta E_{ArY} - \Delta E_{XY} \quad (1)$$

where Ar is the aromatic amino acid and X, Y denote the cations involved in the X-Ar-Y cation- π -cation interaction.

Single-point energies for monomers, dimers and trimers were calculated using the spin-component-scaled Møller-Plesset second order (SCS-MP2) method proposed by Grimme, which has been shown to improve significantly the accuracy of MP2 calculations.⁵⁰ Essentially, the SCS modification involves scaling parallel and anti-parallel-spin components of the MP2 correlation energy according to the following expression:

$$E_{\text{SCS-MP2}} = E_{\text{HF}} + \frac{1}{3}(E_{\text{corr}(\alpha-\alpha)} + E_{\text{corr}(\beta-\beta)}) + \frac{6}{5}E_{\text{corr}(\alpha-\beta)} \quad (2)$$

In addition, all SCS-MP2 energies were extrapolated to the complete basis set limit (CBS) combining the results obtained

using Dunning's cc-pVDZ and cc-pVTZ basis sets according to the formula proposed by Truhlar:⁵¹

$$E_{\text{SCS-MP2/CBS}} = \frac{3^\alpha}{3^\alpha - 2^\alpha} E_3^{\text{HF}} - \frac{2^\alpha}{3^\alpha - 2^\alpha} E_2^{\text{HF}} + \frac{3^\beta}{3^\beta - 2^\beta} E_3^{\text{corr}} - \frac{2^\beta}{3^\beta - 2^\beta} E_2^{\text{corr}} \quad (3)$$

where E_2 and E_3 denote Hartree-Fock and correlation energies obtained using double- ζ and triple- ζ basis sets, and $\alpha = 3.4$ and $\beta = 2.2$.

For the 3 cation- π -cation interactions identified in the human growth hormone receptor (PDB entry 1A22), interaction energies were also computed at the coupled cluster with single and double excitations (CCSD) level of theory using the aug-cc-pVDZ basis set. In all cases, dimer and trimer interaction energies were corrected for basis-set superposition errors (BSSE) using the counterpoise correction (CP),⁵² in which monomer calculations were performed using the complete basis set of the interacting dimer or trimer systems.

Finally, the effect of the environment on the cation- π -cation complex was examined using two strategies. First, preliminary insight into the effect of the environment was gained from QM continuum solvation calculations in water and in *n*-octanol using the MST-IEFPCM B3LYP/6-31G(d) model, following the scheme described in previous studies.⁵³⁻⁵⁵ Next, we estimated the contribution of the protein environment by computing the free-energy change associated to transferring the trimer from water to the protein scaffold using the classical MM-GBSA method, as implemented in the Amber12 software.⁵⁶ In particular, we used the modified GB model developed by Onufriev, Bashford and Case using a dielectric constant of 78.3.⁵⁷ In order to avoid steric clashes between the protein and the model trimer in the latter calculations, Arg and Lys residues participating in the interaction were mutated into Gly, and the aromatic residue (Trp, Phe or Tyr) was removed, capping the neighboring amino acids with a hydrogen atom. In addition, all cofactors were removed from the structures, and the protonation states of the residues in the polypeptide chain were assigned according to PROPKA calculations.^{45,46} In the particular case of the phosphate transport phoU protein (PDB entry 1T72), the low resolution of the crystal structure led to artificially large van der Waals contributions in the MM-GBSA calculations of its two cation- π -cation complexes. Therefore, before performing the MM-GBSA calculations, we minimized the protein in vacuum keeping the amino acids involved in the interactions restrained to their initial position. All MM calculations were based on the parm99SB Amber force field,⁵⁸ and the parameters for the ammonium, guanidinium, benzene, phenol and indole units were assigned using the Antechamber module based on the Amber force field with RESP charges derived at the HF/6-31G(d) level of theory.⁵⁹

3. Results and discussion

3.1 Statistical analysis of cation- π -cation motifs in proteins

We performed an initial analysis of the non-redundant protein dataset in search for cation- π -cation, cation- π -His and His- π -His motifs. Interactions were identified when the distance and angle

formed by the cationic moieties relative to the aromatic ring were below 5 Å and 45°, respectively (Fig. 1). These criteria led to the identification of a total of 2898 interactions located in 2328 protein structures, which correspond to 1675 cation- π -cation, 1029 cation- π -His and 194 His- π -His motifs. Investigation of the protonation state of the His residues involved in the cation- π -cation motifs suggested a cationic His participating in 24 cation- π -His⁺ and 10 His⁺- π -His motifs, whereas no His⁺- π -His⁺ complexes were found. Thus, our results indicate an overall number of 1699 cation- π -cation complexes found in 1450 proteins (7.0% of the structures analyzed). This result stresses the unexpectedly significant population of these interactions in proteins.

Fig. 3 shows the statistical distribution of cation- π and His- π distances and angles for the 2898 complexes. The aromatic ring corresponds to Tyr in 47% of the cases, whereas a similar population is found for Trp (25%) and Phe (28%). Additional details about the particular composition of each interaction identified are supplied in Tables S1-S3 of the ESI.† Moreover, Table 1 shows the minimum and mean values of the distance (d) and angle (θ) between the cation and the aromatic ring. The distributions of average distances and angles are very similar for the three aromatic amino acids, with the mean distance ranging from 4.1 to 4.4 Å and the average angle ranging from 27 to 30 degrees.

The analysis of the complexes formed by Tyr as the aromatic residue, which is the most frequent cation- π -cation motif within the dataset, primarily involves the mixed interaction with the ammonium group of Lys and the guanidinium moiety of Arg, the Lys-Tyr-Arg motif being detected in 26% of the X-Tyr-Y complexes. Complexes involving the same charged amino acid residue (X-Tyr-X) were found in 38% of the cases, the most favored one being Lys-Tyr-Lys (19%). In very few instances the cation- π -cation interaction involved a metal cation (Ca²⁺, Na⁺ or Mg²⁺), generally forming Lys-Tyr-Mg²⁺ complexes. On the other hand, no interaction motifs were found involving the simultaneous presence of two metal cations.

With regard to Trp, the preferred motif with charged amino acids was also the mixed complex formed with Lys and Arg (28% of the cases), thus mimicking the preferences observed for Tyr. Similar trends were observed for Phe, as the preferred motif (23%) is the Lys-Phe-Arg complex. Moreover, concerning cases in which the same amino acid is found on each side of the aromatic ring, the interaction motifs Arg-Trp-Arg (20%) and Arg-Phe-Arg (16%) were found to be the most common, in contrast with the Tyr case. Again, only a few complexes were detected involving the metal cations Na⁺ and Mg²⁺ in the case of Trp, and Ca²⁺ in the case of Phe. As has been noted for Tyr, the simultaneous interaction of both Trp and Phe with two metal cations was not observed in the present analysis.

Overall, we can see that all the considered aromatic amino acids (Trp, Tyr, Phe) show a clear preference to interact with charged amino acids in cation- π -cation interactions rather than monovalent metal cations. This trend can presumably be attributed to the larger hydration free energies of the metal cations compared to the ammonium and guanidinium groups,⁶⁰⁻⁶³

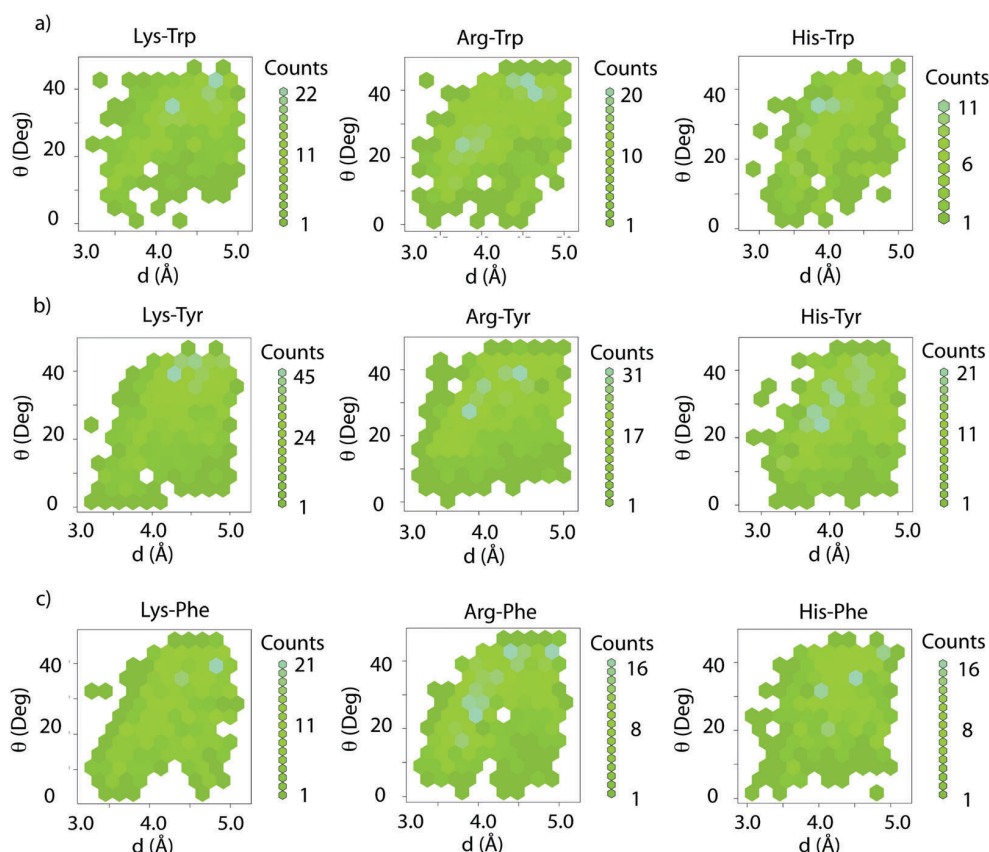


Fig. 3 Statistical distribution of pairwise cation(Lys, Arg)- π and His- π distances and angles of approach observed for the 2898 interactions identified in the Cluster90 database involving (a) Trp, (b) Tyr and (c) Phe.

Table 1 Minimum and mean pairwise cation(Lys, Arg)- π and His- π distances and angles of approach observed for the 2898 interactions identified in the PDB

		Lys		Arg		His	
		d (Å)	θ (°)	d (Å)	θ (°)	d (Å)	θ (°)
Trp	Min	3.23	0.9	3.24	1.0	2.91	2.5
	Mean	4.29	29.2	4.21	27.8	4.11	27.4
Tyr	Min	3.22	1.6	3.29	0.4	2.78	1.2
	Mean	4.35	29.7	4.22	29.1	4.16	27.3
Phe	Min	3.28	3.4	3.28	1.4	3.07	1.7
	Mean	4.33	27.5	4.22	28.0	4.25	28.4

and the larger volume accessibility of the metal cations compared to the pre-organized arrangement of the charged residues in the protein fold, but also to the limited number of protein structures involving metal cations in the dataset. If we discard those complexes involving a neutral His, the preferred amino acid in such interactions is found to be Tyr (48%), followed by Trp (26%) and Phe (25%). Despite its two rings, the fact that Trp appears in a small number of cation- π -cation complexes can be explained by the smaller natural occurrence of Trp (1.3%) compared to Tyr (3.3%) or Phe (3.9%).⁶⁴ Indeed, if the number of interactions found is weighted by their natural abundance, our results suggest a preference to establish cation- π -cation motifs with

Trp, followed by Tyr and Phe (approximate ratios 3:2:1). If we consider charged amino acids on each side of the aromatic ring, the dominant motif contains Lys and Arg, and these residues appear with a similar frequency among all the possible motifs. Again, this finding can be explained by the similar abundance of Lys (5.8%) and Arg (5.2%). The smaller number of complexes involving His probably stems not only from its lower natural abundance (2.2%), but also from the intrinsic preference for its neutral form (see above).

Given the large amount of cation- π -cation complexes found in the dataset, we performed a selection procedure in order to identify the most relevant interactions to be used in the energetic analysis. To this end, a more stringent geometrical analysis was carried out looking at the distribution of distances and angles found in the trimeric (X- π -Y) complexes, shown in Fig. 4-6.

As can be noted in Fig. 4-6, most of the complexes are clustered in the upper right corners, corresponding to total distances ($d_1 + d_2$) and angles ($\theta_1 + \theta_2$) in the range of 9.0-9.5 Å and 60-80 degrees, respectively. However, there are cases with total distances in the range of 7.0-8.0 Å, and angles lower than 40 degrees. For our purposes here, we filtered out complexes using such criteria (see Methods). These additional criteria led to a selection of 175 complexes. As indicated in the Methods

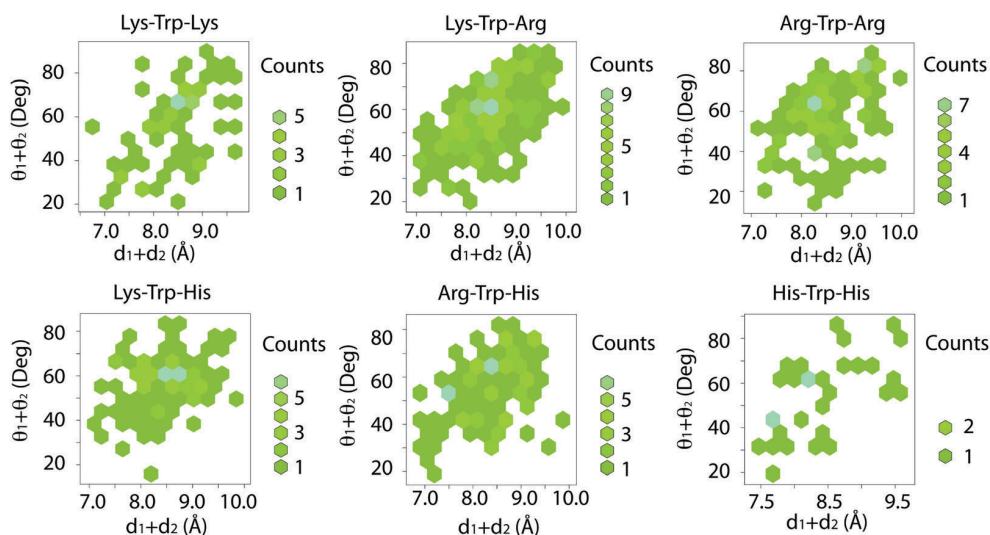


Fig. 4 Statistical distribution of distances and angles of approach observed for the trimeric (X- π -Y) complexes identified involving Trp as the aromatic residue.

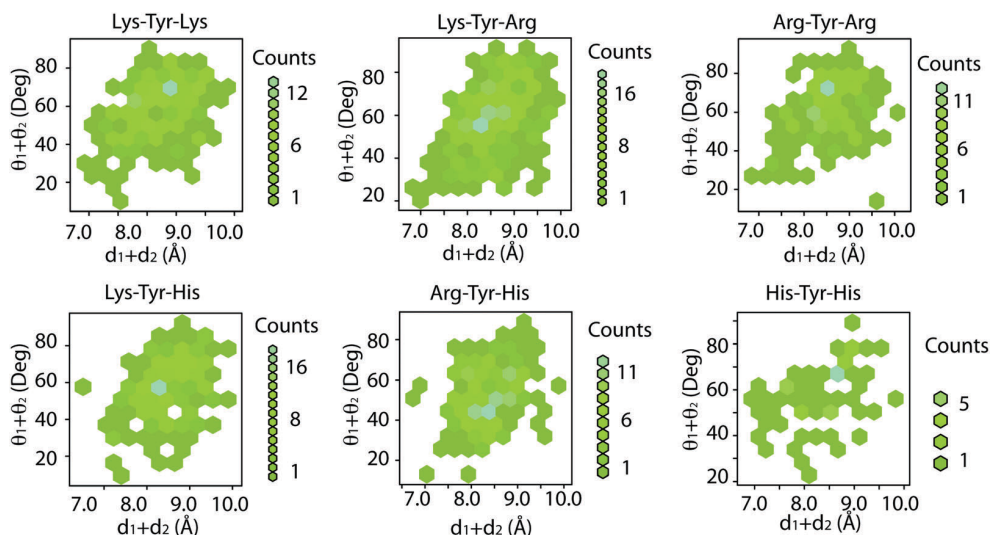


Fig. 5 Statistical distribution of distances and angles of approach observed for the trimeric (X- π -Y) complexes identified involving Tyr as the aromatic residue.

section, this subset was further refined by excluding (i) complexes with His residues, (ii) Lys complexes where the ammonium group was largely displaced from the normal axis to the aromatic ring, and (iii) highly similar proteins. In addition, we kept in our subset specific complexes taken from proteins that featured multiple cation- π -cation interactions, although they did not strictly meet our geometric criteria.

These refinements led to a final selection of 53 complexes (Tables S4–S6 of the ESI[†]). Then, we performed a multiple structural alignment based on the central aromatic residue in order to investigate the three-dimensional distribution of charged residues around the aromatic ring, which is shown in Fig. 7. This analysis reveals three trends: (i) the nitrogen

atom of Lys does not tend to be faced towards the center of the ring, an effect that may reflect a balance between the cation- π -cation and other interactions formed by the protonated amine with neighboring residues, (ii) the distribution of Arg (as measured from the position of the guanidinium C atom) is closer to the normal to the centroid of the aromatic ring, which reflects both the larger delocalization of the positive charge in the guanidinium moiety and the enhanced contribution of dispersion forces between the guanidinium group and the aromatic ring,^{13,20} and finally (iii) in the case of Trp, the preferential ring found in cation- π -cation interactions seems to be the five-membered one, likely reflecting the preference for the negative charge borne by the nitrogen atom.

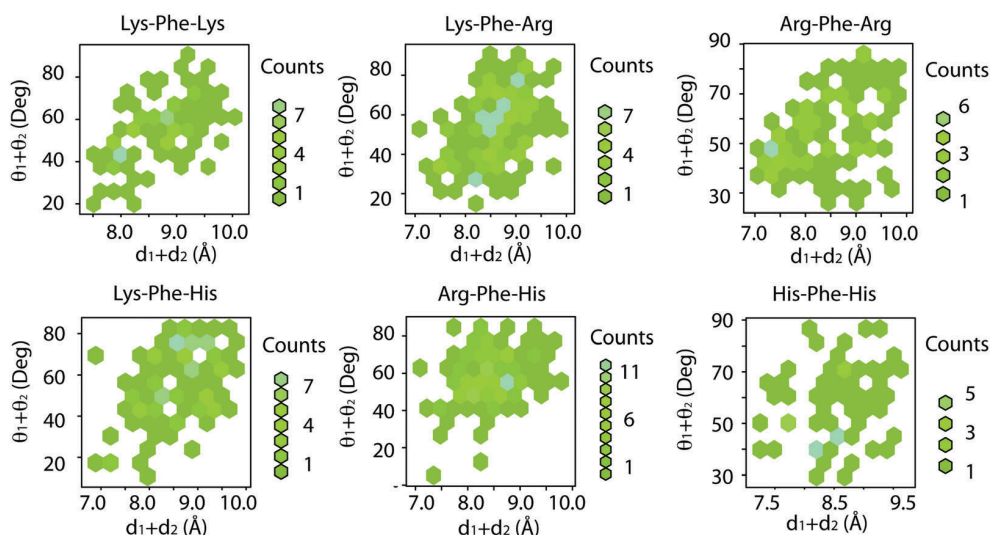


Fig. 6 Statistical distribution of distances and angles of approach observed for the trimeric ($X-\pi-Y$) complexes identified involving Phe as the aromatic residue.

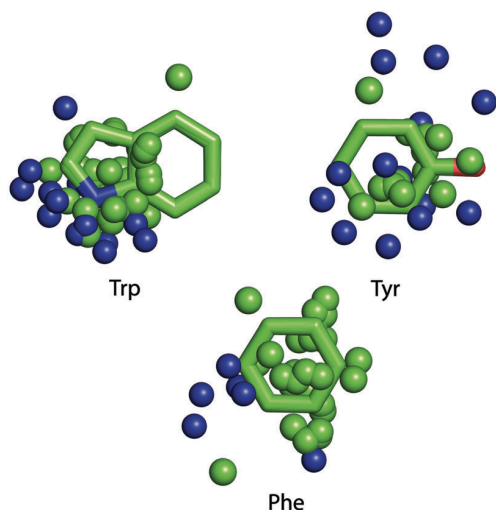


Fig. 7 Spatial distribution of Lys and Arg cations around Trp (upper left), Tyr (upper right) and Phe (bottom center). The C_z and the N_z atoms of Arg and Lys are represented by green and blue spheres, respectively.

We also examined the relative orientation of the Arg guanidinium moiety and the aromatic ring. Fig. 8 shows the distribution of the angles formed by the molecular planes of these two moieties. As expected, our results show that both planes are roughly parallel, with peaks centered in the range of 0–20 degrees for the three aromatic residues. However, in two specific cases involving Trp and Phe complexes, respectively, the guanidinium moiety and the aromatic ring adopted a T-shaped orientation (PDB entries 1CN4 and 1GQK).

Finally, we also investigated the degree of conservation of the amino acids found in the set of 53 cation– π –cation interactions with the aid of the ConSurf web server,^{47,48} which accounts for the probability of residue replacement as well as the evolutionary relations among the protein and its homologs. Fig. 9 shows the distribution of conservation scores for the set of 53 complexes,

obtained as an average over the individual values of each amino acid. Such individual scores, as well as the corresponding residue variety (an index that defines the degree of mutational diversity in ConSurf), can be found in Table S7 of the ESI.[†] The scoring scale ranges from 1 to 9, with 9 indicating the highest degree of conservation. Interestingly, the interactions involving Trp and Phe seem to be strongly conserved, with average scores between 6 and 9 in most cases. On the other hand, interactions involving Tyr tend to be significantly less conserved.

3.2 Energetic analysis

The common occurrence of cation– π –cation motifs identified in our statistical analysis warrants investigating the underlying energetics of such interactions. A particularly challenging question is whether or not they can be expected to stabilize the tertiary structure of a protein. On the other hand, it is of interest to investigate the cooperativity among simultaneous two-body cation– π interactions, which we estimate here by computing three-body non-additive effects in the interaction energies, effects that are envisioned to be nonnegligible, given the well-known impact of polarization effects on cation– π interactions.^{6,7,13,20}

Fig. 10 shows the distribution of pair interaction energies obtained for the complexes containing Trp, Tyr and Phe at the SCS-MP2/CBS level of theory in the gas phase. In addition, we show the overall distribution of pair interaction energies, as well as the trimer interaction energies and the corresponding three-body contributions (data available in Table S8 of the ESI[†]).

As expected, our results indicate that the interaction energies involving two cations are strongly repulsive, in the range of ~ 30 – 50 kcal mol⁻¹, with similar energetic distributions observed for Lys–Lys, Arg–Arg and Lys–Arg pairs in Trp, Tyr or Phe complexes. On the other hand, the distribution of cation– π interaction energies shows larger differences depending on the specific nature of the aromatic ring involved in the complex. A notable difference is that the distribution of Lys–Phe and Arg–Phe interaction

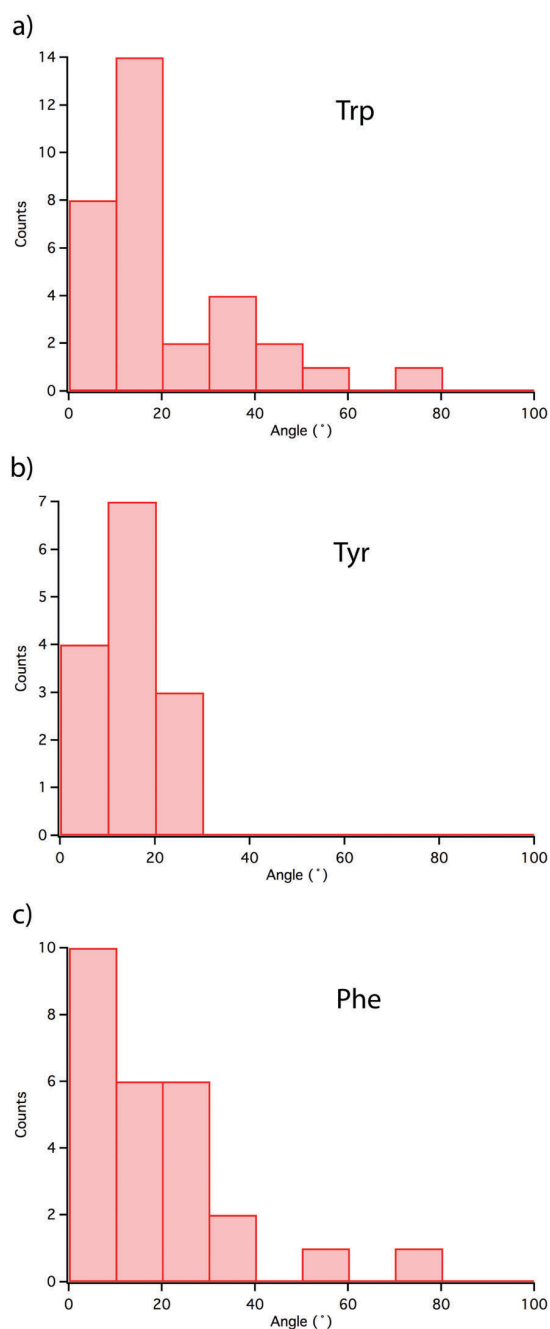


Fig. 8 Histograms of the relative angle between the arginine and aromatic planes for cation- π -cation interactions involving (a) Trp, (b) Tyr and (c) Phe.

energies is considerably narrower than their counterparts involving Trp or Tyr, probably due to the greater symmetry of Phe. Indeed, the distributions involving Trp, where the cations can interact with both the five- and six-membered rings, exhibit the broadest distributions. Interestingly, regardless of the aromatic amino acid involved, the strongest interaction energies in the distribution always arise from cation- π pairs involving Lys rather than Arg.

Because cation-cation repulsion dominates the interaction energy of the cation- π -cation complexes, the overall trimer

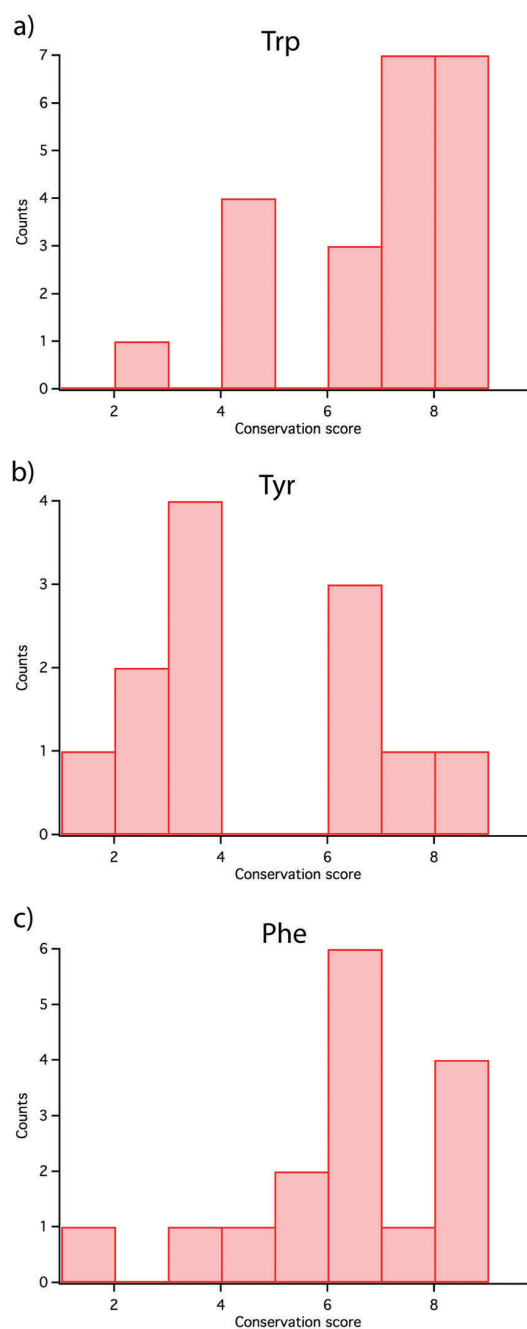


Fig. 9 Histograms of the conservation score averaged over the three amino acids for cation- π -cation interactions involving (a) Trp, (b) Tyr and (c) Phe. The scoring scale ranges from 1 to 9, with 9 indicating the highest degree of conservation.

interaction energies shown in Fig. 10 are strongly positive in vacuum, ranging from 21 to 35 kcal mol⁻¹. Such energies, however, do not match the sum of pair interaction energies. In other words, nonadditivity plays an important role in these interactions, as the three-body contribution amounts typically to 3–5 kcal mol⁻¹, and, in some cases, up to 7 kcal mol⁻¹. We examined the impact of higher-order electron correlation

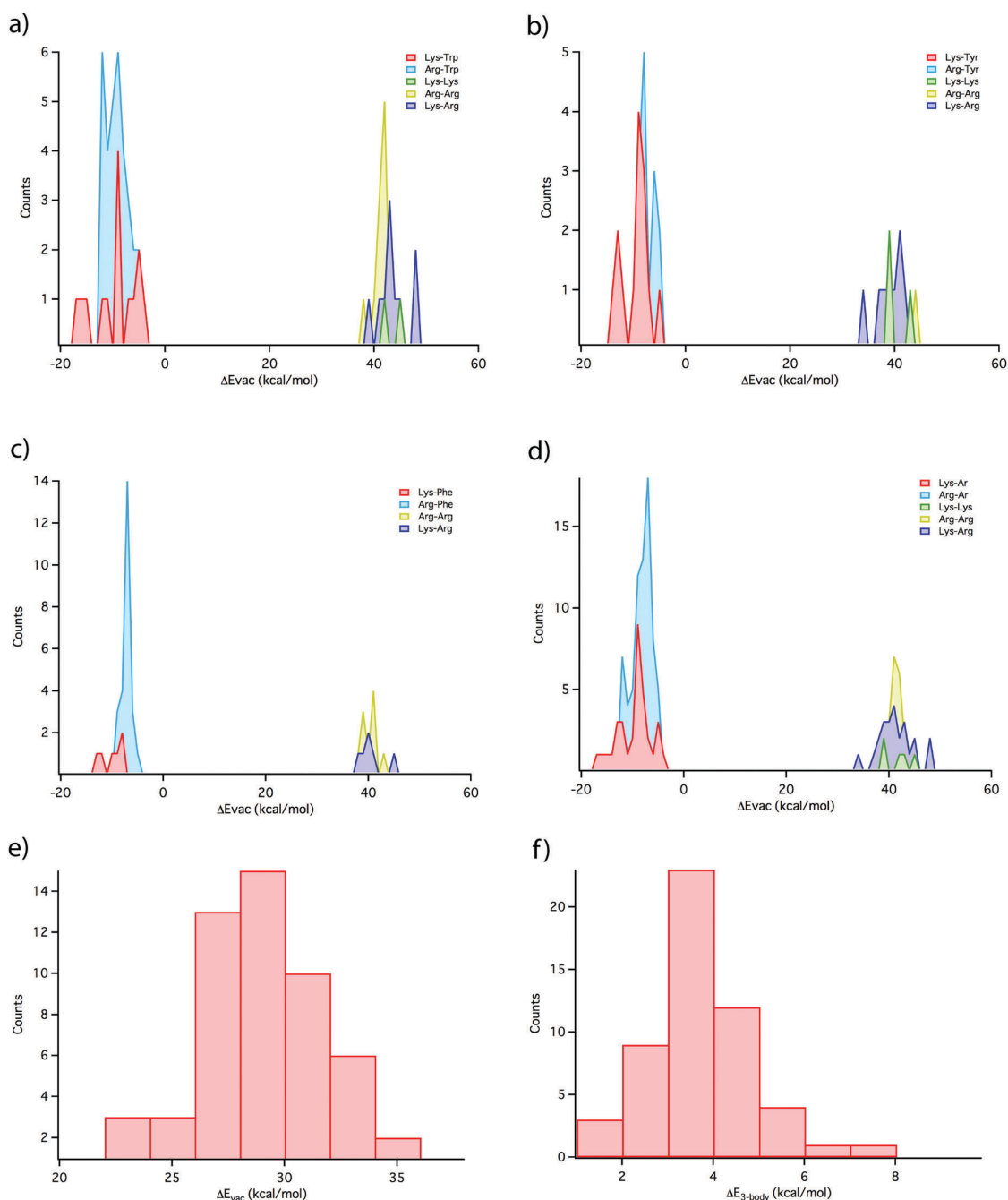


Fig. 10 Interaction energies computed for cation- π -cation complexes in vacuum. (a) Pair interactions in Trp complexes, (b) pair interactions in Tyr complexes, (c) pair interactions in Phe complexes, (d) all pair interactions, (e) trimer interactions and (f) three-body non-additive contributions.

effects on the three-body contribution by performing CCSD/aug-cc-pVDZ calculations on the 3 complexes identified in the human growth hormone receptor (see Fig. 2). Our results pointed, however, to negligible corrections less than $0.1 \text{ kcal mol}^{-1}$ to the SCS-MP2 estimates, as reported in Table S10 of the ESI.†

The finding of significant nonadditivity has important consequences for the description of cation- π -cation motifs in widely used pairwise additive, notably multipurpose macromolecular

force fields for molecular simulations of proteins, as nonadditivity is neglected in such formalism. For instance, in order to improve the description of cation- π interactions in pairwise, additive force fields, Minoux and Chipot introduced a short-range 4–12 potential in the context of the Amber force field able to reproduce induction phenomena in an average sense.⁷ Recently, Khan and co-workers showed that the Drude polarizable force field improves the description of tyrosine-choline cation- π energetics compared

to the CHARMM additive force field, and proposed a modification of the Lennard-Jones terms for the latter in order to recover the missing induction effects.⁶⁵ Although such *ad hoc* modifications provide efficient and simple corrections, the implicit treatment of polarization effects in such formalisms becomes problematic in a cation- π -cation motif, given that the approach of the two cations on each side of the aromatic ring will partially zero out the induction contribution, whereas the 4-12 potential or the modified Lennard-Jones one would predict similar interaction energies as those found in the individual cation- π complexes. This fact reflects the neglect of non-additive contributions, like those estimated here for cation- π -cation motifs, in a pairwise additive force field.

A natural alternative to improve the description of cation- π -cation interactions would be to use an explicit polarizable force field.^{24,25} In this case, common formalisms in this context, like those based on Drude oscillators or induced dipole moments, would likely provide good descriptions of cation- π interactions, as shown in the recent study by Khan and co-workers.⁶⁵ When two cations approach the aromatic ring like in a cation- π -cation motif, however, most likely the Drude oscillators will reorient in the plane of the π -electron cloud (or the induced dipoles will partially vanish), so their ability to describe such interactions is not warranted. A rigorous model to account for such multibody effects would probably need the inclusion of charge flows and dipolar polarizabilities, but the development of such models for solvated macromolecules is rather complex. Nevertheless, it would be interesting to explore the ability of non-additive polarizable force fields, for example Drude⁶⁶ or AMOEBA,⁶⁷ to describe the energetics of cation- π -cation interactions, and whether they are able to account for such nonadditivity.

In a protein, the local environment can significantly impact these energies by significantly screening the electrostatic forces at play.⁶⁸ In order to investigate the impact of the environment, we adopted two strategies. First, we evaluated the impact that a polar or an apolar environment can have on cation- π -cation interactions by computing the solvation free energies in water and *n*-octanol for the individual monomers, and the trimers, based on continuum solvent MST-IEFPCM calculations⁵³⁻⁵⁵ performed at the B3LYP/6-31G(d) level of theory. This allows us to estimate the effective interaction energy of a given cation- π -cation complex in solution, which is estimated as the addition of the gas phase interaction energy and the change in the solvation free energies of the complex and the separated monomers (eqn (4)). Next, we estimated the contribution of the protein environment by computing the free-energy change associated to transferring the trimer from water to the protein scaffold using the classical MM-GBSA method^{56,57} based on the Generalized Born continuum solvation model (eqn (5)).

$$\Delta E_X = \Delta E_{\text{vac}} + \Delta \Delta G_{\text{solv}}^X \quad (\text{X: water, } n\text{-octanol}) \quad (4)$$

$$\Delta E_{\text{prot}} = \Delta E_{\text{wat}} + \Delta G_{\text{wat} \rightarrow \text{prot}} \quad (5)$$

The corresponding interaction energies in *n*-octanol, water and the protein, obtained combining the vacuum SCS-MP2/CBS interaction energies with the solvation energies computed using the QM MST solvation model and the transfer free

energies computed using MM-GBSA, are shown in Fig. 11 (see also Table S9 of the ESI†). As expected, the environment has a strong impact on the interaction energies, stabilizing them by 14–28 kcal mol⁻¹ in *n*-octanol ($\Delta \Delta G_{\text{solv}}^{\text{Oct}}$) and by 15–31 kcal mol⁻¹ in water ($\Delta \Delta G_{\text{solv}}^{\text{Wat}}$). In *n*-octanol, the resulting interaction energies ΔE_{Oct} are still predicted to be positive in all cases, whereas in water solvation leads to attractive interactions (ΔE_{wat}) in several complexes, even though only by up to -1.3 kcal mol⁻¹. Such calculations, nevertheless, ignore the local environment of the cation- π -cation interaction, where the presence of anionic amino acids could substantially further stabilize it. Indeed, the water to protein transfer free energies computed for the complexes in their native environment, $\Delta G_{\text{wat} \rightarrow \text{prot}}$, show an important impact of the protein environment on the interaction energies, with values ranging from -45 to 30 kcal mol⁻¹. In 40 out of the 53 interactions considered, however, the protein environment stabilizes the interaction, whereas in 13 complexes its contribution is positive.

By adding the transfer free energy to the interaction energies in water, we, thus, provide an estimate of the interaction energies for the cation- π -cation motifs in their context, ΔE_{prot} . The resulting values range from -43 to 33 kcal mol⁻¹, 34 out of the 53 complexes resulting in negative interaction energies. This trend is similar for the complexes involving Trp, Tyr and Phe, wherein the fraction of interactions that are predicted to be attractive is 61%, 64% and 69%. As noted previously, the presence of nearby anionic amino acids can considerably stabilize the cation- π -cation motifs by establishing favorable electrostatic interactions with Lys and Arg residues, and thus attenuate the repulsion between the cations. The impact of nearby anions on cation- π interactions has been the focus of several studies.⁶⁹⁻⁷⁵ Moreover, the term salt-bridge- π interaction has been proposed to designate the contact between an aromatic ring and a planar salt-bridge, as opposed to cation- π and anion- π contacts where the cation or the anion is located above/below the aromatic ring.⁷⁵ In the set of 53 cation- π -cation motifs analyzed here, there are often one or more anionic amino acids near the complex. However, in the few cases where the anion-ring separation is comparable to the cation-ring separation, there is a large displacement of the anion from the normal axis to the aromatic ring. This is probably caused by the geometric criteria used to select the cation- π -cation motifs, which favors the location of the cations over the ring.

The impact of negatively charged amino acids is exemplified by the network of interactions established between Glu and Asp residues and the cation- π -cation motifs in the structures of hydroxynitrile lyase and deblocking aminopeptidase, as shown in Fig. 12, which translate into ΔE_{prot} values of -25.0 and -18.3 kcal mol⁻¹ for these particular interactions. These results suggest that the local environment may confer to cation- π -cation interactions a stabilizing role in the tertiary structure of proteins. This possibility is supported by the high degree of conservation found for these interactions, in particular for those involving Trp and Phe. We have not found direct correlations among the average conservation score of each interaction and their interaction energies, but other functional roles could explain this finding.

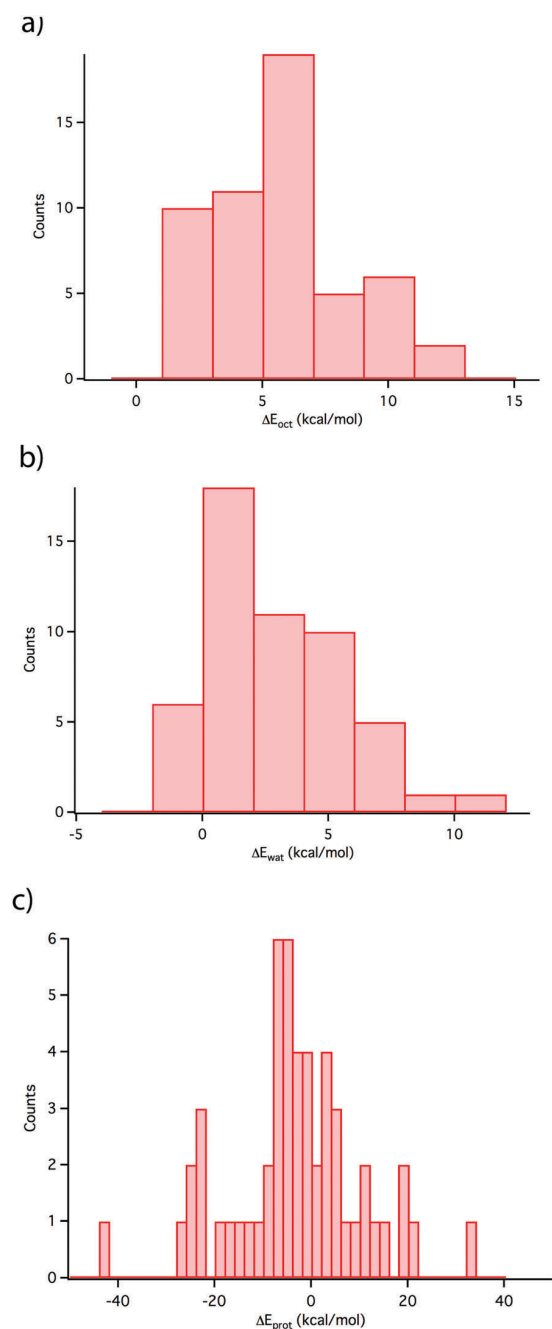


Fig. 11 Effective interaction energies computed for cation- π -cation complexes in different environments. (a) Apolar (*n*-octanol), (b) polar (water) and (c) protein environments.

Beyond individual interactions, networks of solvent-exposed cation- π interactions formed by two consecutive Arg-Trp-Arg motifs have recently been demonstrated to stabilize a miniature protein tertiary structure.⁴³ Here, we have explored an extended cation- π network, also exposed to the solvent, identified in the human growth hormone receptor (PDB entry 1A22), where three consecutive cation- π -cation interactions are established: Lys-Trp-Arg, Arg-Phe-Arg and Arg-Tyr-Lys (see Fig. 2). We estimate the

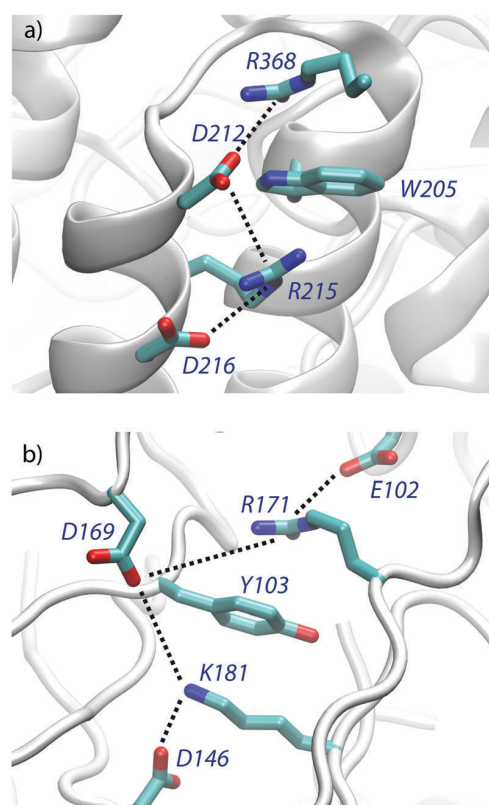


Fig. 12 Illustration of the stabilizing interactions established between cation- π -cation motifs and negatively charged (Asp and Glu) amino acids in (a) hydroxynitrile lyase (code 1GXS-2) and (b) deblocking aminopeptidase (code 2GRE-1). More details on each interaction code can be found in Tables S4 and S5 of the ESI.†

interaction energies to be equal to -16.7 , -1.0 and -7.8 kcal mol⁻¹ for these complexes, supporting the notion that diverse cation- π networks can be engineered in order to stabilize a tertiary structure. Future studies are required to discern the potential implications of these motifs for the intrinsic conformational flexibility and functional role of the protein.

4. Conclusions

We have presented a statistical analysis of the occurrence, the composition and the geometrical preferences of cation- π -cation interactions identified in a non-redundant set of protein structures taken from the PDB. Our analysis indicates that this structural motif is common in proteins, 7% of them containing at least one cation- π -cation interaction. The composition of the interactions identified, when weighted by the natural abundance of amino acids, suggests a preference to establish cation- π -cation motifs with Trp, followed by Tyr and Phe, with approximate ratios 3 : 2 : 1, whereas no particular preference among Lys and Arg is observed regarding the cations. We have also found that cation- π -cation interactions tend to be highly conserved, hence suggesting a relevant structural or functional role. In this case, interactions

involving Trp, and to a lesser extent Phe, seem to be more strongly conserved compared to complexes involving Tyr.

We have also performed an energetic analysis of a representative subset of cation- π -cation complexes combining quantum chemical SCS-MP2/CBS calculations and MST and MM-GBSA continuum solvation models. Our results point out that, whereas the interaction energy in vacuum is strongly positive, an apolar (*n*-octanol) or polar (water) environment can strongly screen the cation-cation repulsion, although in most cases the interaction is still predicted to be repulsive. The impact of the particular protein environment is, however, predicted to be stronger, and leads to an attractive interaction in 64% of the complexes analyzed. This result, together with the high degree of conservation of the amino acids involved in the interactions, suggests a potential stabilizing role in protein tertiary structures, as demonstrated recently for a miniature protein.⁴³ Finally, we found a significant degree of cooperativity among the two cation- π interactions at play. From a computational point of view, the significant contribution of such non-additive, three-body terms, which can amount up to 7 kcal mol⁻¹, challenges the suitability of standard additive biomolecular force fields for describing cation- π -cation motifs in molecular simulations of proteins.

Acknowledgements

We are grateful to the Consorci de Serveis Universitaris de Catalunya for providing access to computational resources. Financial support from the Spanish Ministerio de Economía y Competitividad (MINECO; grants CTQ2012-36195, SAF2014-57094-R and RYC2011-08918) and the Agència de Gestió d'Ajuts Universitaris i de Recerca from Generalitat de Catalunya (GENCAT; SGR2014-1189) is acknowledged. C. C. is a Serra Hünter Fellow. S. P. is a fellow of the Ciência Sem Fronteiras program of the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) de Brasil (246791/2012-8). F. J. L. acknowledges the financial support from ICREA Academia.

References

- 1 E. Campbell, M. Kaltenbach, G. J. Correy, P. D. Carr, B. T. Porebski, E. K. Livingstone, L. Afriat-Jurnou, A. M. Buckle, M. Weik, F. Hollfelder, N. Tokuriki and C. J. Jackson, *Nat. Chem. Biol.*, 2016, **12**, 944–950.
- 2 G. Kiss, N. Çelebi-Ölçüm, R. Moretti, D. Baker and K. N. Houk, *Angew. Chem., Int. Ed.*, 2013, **52**, 5700–5725.
- 3 H. Kries, R. Blomberg and D. Hilvert, *Curr. Opin. Chem. Biol.*, 2013, **17**, 221–228.
- 4 D. A. Dougherty, *Science*, 1996, **271**, 163–168.
- 5 J. C. Ma and D. A. Dougherty, *Chem. Rev.*, 1997, **97**, 1303–1324.
- 6 E. Cubero, F. J. Luque and M. Orozco, *Proc. Natl. Acad. Sci. U. S. A.*, 1998, **95**, 5976–5980.
- 7 H. Minoux and C. Chipot, *J. Am. Chem. Soc.*, 1999, **121**, 10366–10372.
- 8 K. E. Riley, M. Pitoňák, P. Jurečka and P. Hobza, *Chem. Rev.*, 2010, **110**, 5023–5063.
- 9 I. Alkorta, F. Blanco, P. M. Deyà, J. Elguero, C. Estarellas, A. Frontera and D. Quiñonero, *Theor. Chem. Acc.*, 2010, **126**, 1–14.
- 10 A. Frontera, D. Quiñonero and P. M. Deyà, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2011, **1**, 440–459.
- 11 D. A. Dougherty, *Acc. Chem. Res.*, 2013, **46**, 885–893.
- 12 A. S. Mahadevi and G. N. Sastry, *Chem. Rev.*, 2013, **113**, 2100–2138.
- 13 C. D. Sherrill, *Acc. Chem. Res.*, 2013, **46**, 1020–1028.
- 14 J. J. Fiol, M. Barceló-Oliver, A. Tasada, A. Frontera, À. Terrón and Á. García-Raso, *Coord. Chem. Rev.*, 2013, **257**, 2705–2715.
- 15 T. Clark, P. Politzer and J. S. Murray, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2015, **5**, 169–177.
- 16 A. Bauzá, T. J. Mooibroek and A. Frontera, *ChemPhysChem*, 2015, **16**, 2496–2517.
- 17 M. H. Kolář and P. Hobza, *Chem. Rev.*, 2016, **116**, 5155–5187.
- 18 A. S. Mahadevi and G. N. Sastry, *Chem. Rev.*, 2016, **116**, 2775–2825.
- 19 J. M. Andrić, M. Z. Misini-Ignjatović, J. S. Murray, P. Politzer and S. D. Zarić, *ChemPhysChem*, 2016, **17**, 2035–2042.
- 20 A. A. Rodríguez-Sanz, E. M. Cabaleiro-Lago and J. Rodríguez-Otero, *Phys. Chem. Chem. Phys.*, 2014, **16**, 22499–22512.
- 21 C. Rapp, E. Goldberger, N. Tishbi and R. Kirshenbaum, *Proteins: Struct., Funct., Bioinf.*, 2014, **82**, 1494–1502.
- 22 D. Vijay and G. N. Sastry, *Phys. Chem. Chem. Phys.*, 2008, **10**, 582–590.
- 23 B. U. Emenike, S. N. Bey, R. A. Spinelle, J. T. Jones, B. Yoo and M. Zeller, *Phys. Chem. Chem. Phys.*, 2016, **18**, 30940–30945.
- 24 P. Cieplak, F.-Y. Y. Dupradeau, Y. Duan and J. M. Wang, *J. Phys.: Condens. Matter*, 2009, **21**, 333102.
- 25 F. J. Luque, F. Dehez, C. Chipot and M. Orozco, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2011, **1**, 844–854.
- 26 I. Soteras, C. Curutchet, A. Bidon-Chanal, F. Dehez, J. G. Ángyán, M. Orozco, C. Chipot and F. J. Luque, *J. Chem. Theory Comput.*, 2007, **3**, 1901–1913.
- 27 F. Dehez, J. G. Ángyán, I. S. Gutiérrez, F. J. Luque, K. Schulten and C. Chipot, *J. Chem. Theory Comput.*, 2007, **3**, 1914–1926.
- 28 I. Soteras, M. Orozco and F. J. Luque, *Phys. Chem. Chem. Phys.*, 2008, **10**, 2616–2624.
- 29 F. Archambault, C. Chipot, I. Soteras, F. J. Luque, K. Schulten and F. Dehez, *J. Chem. Theory Comput.*, 2009, **5**, 3022–3031.
- 30 K. Ansorg, M. Tafipolsky and B. Engels, *J. Phys. Chem. B*, 2013, **117**, 10093–10102.
- 31 J. W. Caldwell and P. a. Kollman, *J. Am. Chem. Soc.*, 1995, **117**, 4177–4178.
- 32 H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov and P. E. Bourne, *Nucleic Acids Res.*, 2000, **28**, 235–242.
- 33 C. Biot, R. Wintjens and M. Rooman, *J. Am. Chem. Soc.*, 2004, **126**, 6220–6221.
- 34 D. Quiñonero, A. Frontera, C. Garau, P. Ballester, A. Costa and P. M. Deyà, *ChemPhysChem*, 2006, **7**, 2487–2491.
- 35 A. S. Reddy, D. Vijay, G. M. Sastry and G. N. Sastry, *J. Phys. Chem. B*, 2006, **110**, 2479–2481.
- 36 A. Frontera, D. Quiñonero, A. Costa, P. Ballester and P. M. Deyà, *New J. Chem.*, 2007, **31**, 556–560.
- 37 Q. Li, W. Li, J. Cheng, B. Gong and J. Sun, *J. Mol. Struct.: THEOCHEM*, 2008, **867**, 107–110.

- 38 D. Escudero, A. Frontera, D. Quiñonero and P. M. Deyà, *Chem. Phys. Lett.*, 2008, **456**, 257–261.
- 39 C. Estarellas, A. Frontera, D. Quiñonero and P. M. Deyà, *Chem. Phys. Lett.*, 2009, **479**, 316–320.
- 40 D. Vijay and G. N. Sastry, *Chem. Phys. Lett.*, 2010, **485**, 235–242.
- 41 R. Li, Q. Li, J. Cheng, Z. Liu and W. Li, *ChemPhysChem*, 2011, **12**, 2289–2295.
- 42 A. Campo-Cacharrón, E. M. Cabaleiro-Lago, J. a. Carrazana-García and J. Rodríguez-Otero, *J. Comput. Chem.*, 2014, **35**, 1290–1301.
- 43 T. W. Craven, M.-K. Cho, N. J. Traaseth, R. Bonneau and K. Kirshenbaum, *J. Am. Chem. Soc.*, 2016, **138**, 1543–1550.
- 44 T. Meyer, M. D'Abramo, A. Hospital, M. Rueda, C. Ferrer-Costa, A. Pérez, O. Carrillo, J. Camps, C. Fenollosa, D. Repchevsky, J. L. Gelpí and M. Orozco, *Structure*, 2010, **18**, 1399–1409.
- 45 M. H. M. Olsson, C. R. Søndergaard, M. Rostkowski and J. H. Jensen, *J. Chem. Theory Comput.*, 2011, **7**, 525–537.
- 46 C. R. Søndergaard, M. H. M. Olsson, M. Rostkowski and J. H. Jensen, *J. Chem. Theory Comput.*, 2011, **7**, 2284–2295.
- 47 M. Landau, I. Mayrose, Y. Rosenberg, F. Glaser, E. Martz, T. Pupko and N. Ben-Tal, *Nucleic Acids Res.*, 2005, **33**, W299–W302.
- 48 H. Ashkenazy, S. Abadi, E. Martz, O. Chay, I. Mayrose, T. Pupko and N. Ben-Tal, *Nucleic Acids Res.*, 2016, **44**, W344–W350.
- 49 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, N. J. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski and D. J. Fox, *Gaussian 09, Revision A.2.*, Gaussian Inc., Wallingford CT, 2009.
- 50 S. Grimme, *J. Chem. Phys.*, 2003, **118**, 9095–9102.
- 51 D. G. Truhlar, *Chem. Phys. Lett.*, 1998, **294**, 45–48.
- 52 S. Boys and F. Bernardi, *Mol. Phys.*, 1970, **19**, 553–566.
- 53 C. Curutchet, M. Orozco and F. J. Luque, *J. Comput. Chem.*, 2001, **22**, 1180–1193.
- 54 C. Curutchet, A. Bidon-Chanal, I. Soteras, M. Orozco and F. J. Luque, *J. Phys. Chem. B*, 2005, **109**, 3565–3574.
- 55 I. Soteras, C. Curutchet, A. Bidon-Chanal, M. Orozco and F. J. Luque, *THEOCHEM*, 2005, **727**, 29–40.
- 56 D. A. Case, T. A. Darden, I. T. E. Cheatham, C. L. Simmerling, J. Wang, R. E. Duke, R. Luo, R. C. Walker, W. Zhang, K. M. Merz, B. Roberts, S. Hayik, A. Roitberg, G. Seabra, J. Swails, A. W. Götz, I. Kolossváry, K. F. Wong, F. Paesani, J. Vanicek, R. M. Wolf, J. Liu, X. Wu, S. R. Brozell, T. Steinbrecher, H. Gohlke, Q. Cai, X. Ye, J. Wang, M.-J. Hsieh, G. Cui, D. R. Roe, D. H. Mathews, M. G. Seetin, R. Salomon-Ferrer, C. Sagui, V. Babin, T. Luchko, S. Gusarov, A. Kovalenko and P. A. Kollman, *AMBER 12*, University of California, San Francisco, 2012.
- 57 A. Onufriev, D. Bashford and D. A. Case, *Proteins: Struct., Funct., Bioinf.*, 2004, **55**, 383–394.
- 58 V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg and C. Simmerling, *Proteins: Struct., Funct., Bioinf.*, 2006, **65**, 712–725.
- 59 C. I. Bayly, P. Cieplak, W. Cornell and P. A. Kollman, *J. Phys. Chem.*, 1993, **97**, 10269–10280.
- 60 P. E. Mason, G. W. Neilson, C. E. Dempsey, A. C. Barnes and J. M. Cruickshank, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**, 4557–4561.
- 61 A. Grossfield, P. Ren and J. W. Ponder, *J. Am. Chem. Soc.*, 2003, **125**, 15671–15682.
- 62 D. Jiao, C. King, A. Grossfield, T. A. Darden and P. Ren, *J. Phys. Chem. B*, 2006, **110**, 18553–18559.
- 63 J. Liu, C. P. Kelly, A. C. Goren, A. V. Marenich, C. J. Cramer, D. G. Truhlar and C.-G. Zhan, *J. Chem. Theory Comput.*, 2010, **6**, 1109–1117.
- 64 G. Trinquier and Y. H. Sanejouand, *Protein Eng., Des. Sel.*, 1998, **11**, 153–169.
- 65 H. M. Khan, C. Grauffel, R. Broer, A. D. MacKerell, R. W. A. Havenith and N. Reuter, *J. Chem. Theory Comput.*, 2016, **12**, 5585–5595.
- 66 J. A. Lemkul, J. Huang, B. Roux and A. D. MacKerell, *Chem. Rev.*, 2016, **116**, 4983–5013.
- 67 Y. Shi, Z. Xia, J. Zhang, R. Best, C. Wu, J. W. Ponder and P. Ren, *J. Chem. Theory Comput.*, 2013, **9**, 4046–4063.
- 68 A. Warshel, P. K. Sharma, M. Kato and W. W. Parson, *Biochim. Biophys. Acta, Proteins Proteomics*, 2006, **1764**, 1647–1676.
- 69 Y. Inoue, S. Sugio, J. Andzelm and N. Nakamura, *J. Phys. Chem. A*, 1998, **102**, 646–648.
- 70 P.-O. Norrby and T. Liljefors, *J. Am. Chem. Soc.*, 1999, **121**, 2303–2306.
- 71 S. D. Zarić, D. M. Popović and E.-W. Knapp, *Chemistry*, 2000, **6**, 3935–3942.
- 72 S. E. Thompson and D. B. Smithrud, *J. Am. Chem. Soc.*, 2002, **124**, 442–449.
- 73 V. Dvornikovs and D. B. Smithrud, *J. Org. Chem.*, 2002, **67**, 2160–2167.
- 74 M. A. Anderson, B. Ogbay, R. Arimoto, W. Sha, O. G. Kisselev, D. P. Cistola and G. R. Marshall, *J. Am. Chem. Soc.*, 2006, **128**, 7531–7541.
- 75 M. Mitra, P. Manna, S. K. Seth, A. Das, J. Meredith, M. Helliwell, A. Bauzá, S. R. Choudhury, A. Frontera and S. Mukhopadhyay, *CrystEngComm*, 2013, **15**, 686–696.

**Can Förster theory describe stereoselective energy transfer
dynamics in a protein-ligand complex?**

Journal of Physical Chemistry B, **2017**, 121 (10), 2265–2278

Silvana Pinheiro¹ and Carles Curutchet¹

¹Departament de Farmàcia i Tecnologia Farmacèutica i Fisicoquímica and Institut de Biomedicina (IBUB), Facultat de Farmàcia i Ciències de l'Alimentació, Universitat de Barcelona, Barcelona Spain

CAPÍTULO 3. RESULTADOS

3.2 Artículo 2: Can Förster theory describe stereoselective energy transfer dynamics in a protein-ligand complex?

Resumen

La fluorescencia de las proteínas es un proceso complejo que frecuentemente involucra reacciones de transferencia de energía de resonancia de Förster (FRET) entre aminoácidos aromáticos como el triptófano. Además, en complejos proteína-ligando, la participación del ligando en tales procesos de transferencia puede afectar sustancialmente a las propiedades de fluorescencia del complejo, un impacto íntimamente relacionado con el modo de unión correspondiente. La caracterización estructural de tales eventos de unión en términos de distancias intermoleculares se puede hacer a través de la conocida expresión de Förster, que describe una dependencia R^{-6} de la velocidad de FRET con la distancia entre cromóforos. Sin embargo, tal interpretación sufre de incertidumbres subyacentes a la teoría de Förster en la descripción del acoplamiento electrónico que promueve procesos FRET, principalmente relacionadas con el factor de orientación dipolo-dipolo, desviaciones de la aproximación ideal de dipolos puntuales y efectos de apantallamiento dieléctrico del entorno. En este estudio, investigamos el impacto de las aproximaciones asumidas en la teoría de Förster en la predicción de la dinámica de FRET en un sistema modelo flurbiprofeno-triptófano (FBP-TrpMe) y en el complejo de este fármaco con la albúmina sérica humana (FBP/HSA), en los cuales se ha observado recientemente un *quenching* dinámico de la fluorescencia enantioselectivo que ha sido atribuido a un proceso FRET del FBP al Trp. A fin de simular la velocidad de FRET, combinamos técnicas de dinámica molecular clásica con cálculos de las propiedades de transición y de los acoplamientos electrónicos responsables del FRET mediante la metodología QM/MM polarizable desarrollada en el grupo, la cual combina una descripción químico-cuántica de los pigmentos (TD-DFT, CIS o ZINDO) con una descripción MM polarizable del entorno. Este método, pues, permite superar las aproximaciones dipolo-dipolo y en el factor de apantallamiento constante asumidas en la teoría de Förster. Los resultados obtenidos describen satisfactoriamente la enantioselectividad observada cuando el acoplamiento se calcula a partir de cálculos TD-DFT basados en densidades de transición, mientras que la aproximación dipolo-dipolo de Förster o el uso de métodos más aproximados como CIS o ZINDO se muestran significativamente menos robustos. El excelente acuerdo en términos de enantioselectividad, así como respecto a las velocidades de FRET en términos absolutos, apoyan firmemente los modos de unión

CAPÍTULO 3. RESULTADOS

propuestos para los enantiómeros (S) y (R) del flurbiprofeno en la HSA. En particular, demostramos que el *quenching* dinámico surge de una pequeña fracción de fármaco unido al sitio secundario de HSA en la interfase entre los subdominios IIA y IIB, mientras que la enantioselectividad surge de la mayor flexibilidad del enantiómero (S) del flurbiprofeno en el sitio de unión. En este caso, la polarizabilidad heterogénea del entorno no contribuye a la enantioselectividad, ya que encontramos factores de apantallamientos similares para ambos enantiómeros. Nuestros resultados, sin embargo, muestran como tal efecto fluctúa de forma muy significativa, hecho que sugiere que el factor constante asumido en la teoría de Förster es una aproximación importante y puede conducir a desviaciones significativas en otros sistemas.

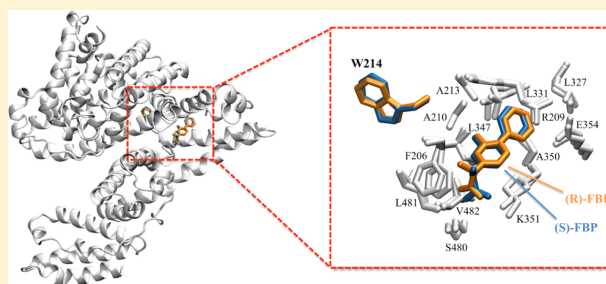
Can Förster Theory Describe Stereoselective Energy Transfer Dynamics in a Protein–Ligand Complex?

Silvana Pinheiro and Carles Curutchet*[✉]

Departament de Farmàcia i Tecnologia Farmacèutica i Físicquímica and Institut de Biomedicina (IBUB), Facultat de Farmàcia i Ciències de l'Alimentació, Universitat de Barcelona, Av. Joan XXIII s/n, Barcelona 08028, Spain

Supporting Information

ABSTRACT: Förster resonance energy transfer (FRET) reactions involving ligands and aromatic amino acids can substantially impact the fluorescence properties of a protein–ligand complex, an impact intimately related to the corresponding binding mode. Structural characterization of such binding events in terms of intermolecular distances can be done through the well-known R^{-6} distance-dependent Förster rate expression. However, such an interpretation suffers from uncertainties underlying Förster theory in the description of the electronic coupling that promotes FRET, mostly related to the dipole–dipole orientation factor, dielectric screening effects, and deviations from the ideal dipole approximation. Here, we investigate how Förster approximations impact the prediction of energy transfer dynamics in the complex between flurbiprofen (FBP) and human serum albumin (HSA), as well as a model FBP–Trp dyad, in which recent observation of enantioselective fluorescence quenching has been ascribed to energy transfer from FBP to Trp. To this end, we combine classical molecular dynamics simulations with polarizable quantum mechanics/molecular mechanics calculations that allow overcoming Förster approximations. On the basis of our results, we discuss the potential of structure-based simulations in the characterization of drug-binding events through fluorescence techniques. Overall, we find an excellent agreement between theory and experiment both in terms of enantioselectivity and FRET times, thus strongly supporting the reliability of the binding modes proposed for the (S) and (R) enantiomers of FBP. In particular, we show that the dynamic quenching arises from a small fraction of drug bound to the secondary site of HSA at the interface between subdomains IIA and IIB, whereas the enantioselectivity arises from the larger flexibility of the (S)-FBP enantiomer in the binding pocket.



1. INTRODUCTION

The fluorescence of proteins is exquisitely dependent on structural changes, a property that has led to a plethora of techniques aimed at assessing the protein structure and dynamics.^{1–10} Such fluorescence generally arises from Trp residues in the structure, whose emission wavelength, quantum yield, and lifetimes are exquisitely sensitive to the surrounding environment.^{2,8,11,12} Indeed, the protein scaffold not only tunes the emission wavelength of Trp by electrostatic interactions but also mediates proton, electron, and energy transfer reactions that can efficiently quench Trp emission, thus modifying its quantum yield and lifetime.¹³ In a protein–ligand complex, the bound ligand can further modify the fluorescence properties of the protein by participating in Förster resonance energy transfer (FRET) reactions with Trp, an impact that is intimately related to the particular binding mode adopted by the ligand. Conversely, the ligand fluorescence can also be drastically modulated upon protein binding.¹⁴

Structural characterization of ligand-binding events from FRET data based on the popular Förster expression, which predicts a R^{-6} distance decay of the transfer rate with donor–acceptor separation, suffers however from uncertainties related

to Förster description of the donor–acceptor electronic coupling. Such limitations are related to the orientation factor, dielectric screening effects exerted by the surrounding environment, and potential deviations from the ideal dipole approximation.¹⁵ For example, because a ligand has a restricted rotational freedom in the protein-binding site, an isotropic approximation for the dipole–dipole orientation factor, $\langle \kappa^2 \rangle = 2/3$, will likely break down.¹⁶ In addition, if the ligand is close to the Trp residue, the coupling can significantly deviate from the ideal dipole description, especially if the rotational freedom of the ligand is restricted.¹⁷ Finally, dielectric screening effects strongly depend on the nature of the amino acids that surround the chromophores involved in a FRET event.¹⁸ The potential of FRET experiments to characterize ligand-binding events, its exquisite sensitivity to subtle structural changes in the underlying binding mode, is thus underexploited under a simple Förster-type interpretation in terms of intermolecular distances.

Received: January 8, 2017

Revised: February 23, 2017

Published: February 24, 2017

Different research groups have used molecular dynamics (MD) simulations to characterize FRET data in terms of structure.^{16,19–40} Krueger and co-workers, for example, have shown that κ and R can be strongly correlated for a pair of fluorescent probes attached to lysozyme, thus modifying the predicted rates by 60% compared to the assumptions of independent κ and R common in FRET studies.¹⁶ On the other hand, by combining MD and polarizable quantum mechanics/molecular mechanics (QM/MM) calculations, we have recently examined the impact of dielectric screening in several photosynthetic pigment–protein complexes. Our results showed that the heterogeneous nature of the environment can modulate FRET rates by a factor of up to ~ 4 with respect to those predicted for a homogeneous medium, as assumed in Förster theory.^{18,41}

In this study, we apply the MD-QM/MMpol strategy we have developed to describe energy migration in complex environments⁴² to examine the validity of Förster assumptions in FRET studies of ligand binding. In particular, we investigate the enantioselective fluorescence dynamic quenching recently observed by Vayá and co-workers in the complex between human serum albumin (HSA) and the (S) and (R) enantiomers of flurbiprofen (FBP), which has been ascribed to excitation transfer from 1FBP* to the Trp214 residue in HSA.¹⁴ Characterizing the binding of drugs to HSA is important to understand their pharmacodynamic and pharmacokinetic profile, given that HSA is the major transport plasma protein.^{43,44} In this case, we study the binding of FBP, a chiral nonsteroidal anti-inflammatory drug, in which the (S) enantiomer shows most of the anti-inflammatory activity, although both enantiomers possess analgesic activity. In addition, FBP exhibits stereoselectivity in its pharmacokinetics,⁴⁵ and stereoselective binding to HSA has been shown to occur.^{46,47} It is known that FBP binds preferentially to site II (benzodiazepine-binding site) of HSA but also binds with remarkable affinity to another site.^{46,48,49} Often this secondary binding site has been identified with well-known site I (warfarin-binding site) of HSA. However, the crystal structure of ibuprofen bound to HSA,⁵⁰ a drug with a rather similar structure and binding characteristics,⁴⁸ indicates that the secondary site is not site I. Rather, it is a novel site located between subdomains IIA and IIB. This finding explains why there is only a partial quenching of HSA phosphorescence in the presence of FBP,⁴⁷ as FBP in that secondary site is not in direct contact with Trp214 as would be in site I, but at a distance ~ 10 Å.⁵⁰

In addition to the FBP/HSA complex, we also examine covalently linked dyads formed by (S) and (R) enantiomers of FBP and (S)-tryptophan methyl ester (TrpMe), in which the enantioselectivity is reversed.¹⁴ For the FBP/HSA complexes, we explore the flexibility and binding modes of the enantiomers using classical MD simulations, then we estimate the FRET properties of the system from polarizable QM/MM excited-state calculations performed along the MD trajectories.⁵¹ For the model dyads, we find that the classical force field is unable to describe their conformational properties in acetonitrile, so we determine the main conformers and evaluate the EET properties of the dyad based on static QM calculations coupled to a continuum solvation model. Interestingly, we find that the enantioselectivity in the model dyad arises from an almost orthogonal arrangement of the transition dipole moments of FBP and Trp in the (S,S) dyad, which explains the 2–3 times slower dynamic quenching observed. On the other hand, we

find that the dynamic quenching in the FBP/HSA complex arises from the fraction of FBP bound to the secondary site of HSA at the interface between subdomains IIA and IIB, rather than the main population bound at site II. In this case, the enantioselective quenching arises from the increased flexibility of the (S)-FBP diastereomer compared to that of the (R)-FBP one, which induces large fluctuations in electronic couplings, in line with the faster reorientational time derived from fluorescence anisotropy for that enantiomer.¹⁴ In both model dyad and FBP/complexes, we find that a precise determination of the transition dipole orientation of Trp L_a state through time-dependent density functional theory (TD-DFT), as well as an account of higher order multipolar effects beyond the dipole approximation based on three-dimensional (3D) transition densities, is necessary to capture the subtle changes in FRET rates observed.

The present contribution is organized as follows. First, we describe the methods used in this work. Then, we discuss the results for the model dyads followed by those obtained for the FBP/HSA complexes. Finally, we close with our conclusions and future prospects.

2. METHODS

2.1. Förster Energy Transfer Theory. The interpretation of FRET studies is typically performed based on Förster theory, introduced more than 60 years ago,^{52,53} which describes the rate of nonradiative electronic excitation energy transfer from a donor (D) molecule to a weakly coupled acceptor (A)

$$k_{\text{FRET}} = \frac{2\pi}{\hbar} V^2 J \quad (1)$$

where V describes the donor–acceptor electronic coupling, and J is the overlap factor between donor emission $f_{\text{D}}(\epsilon)$ and acceptor absorption lineshapes $a_{\text{A}}(\epsilon)$, both normalized to unit area on an energy scale

$$J = \int_0^\infty f_{\text{D}}(\epsilon) a_{\text{A}}(\epsilon) d\epsilon \quad (2)$$

The distance and orientation dependence of the rate is mainly encapsulated in the V^2 term, which is approximated as

$$V \approx s V_{\text{dip-dip}} \quad (3)$$

$$s = \frac{1}{n^2} \quad (4)$$

$$V_{\text{dip-dip}} = \frac{\kappa \mu_{\text{D}}^{\text{T}} \mu_{\text{A}}^{\text{T}}}{R^3} \quad (5)$$

where the orientation factor is given by

$$\kappa = \hat{\mu}_{\text{D}}^{\text{T}} \cdot \hat{\mu}_{\text{A}}^{\text{T}} - 3(\hat{\mu}_{\text{D}}^{\text{T}} \cdot \hat{R})(\hat{\mu}_{\text{A}}^{\text{T}} \cdot \hat{R}) \quad (6)$$

and n indicates the refractive index of the medium, $\mu_{\text{D}}^{\text{T}}/\mu_{\text{A}}^{\text{T}}$ the D/A transition dipole moments, and R the intermolecular center-to-center separation ($\hat{\mu}_{\text{D}}^{\text{T}}$, $\hat{\mu}_{\text{A}}^{\text{T}}$, and \hat{R} are the corresponding unit vectors).

The derivation of Förster theory relies on the weak-coupling approximation, according to which the electronic coupling is small compared to the reorganization energy of the chromophores. Although there is a vivid debate on theories valid for other coupling regimes,⁵⁴ in particular in studies of light harvesting in photosynthesis, typically FRET studies are performed under conditions of weak coupling. In this case, the

most problematic assumptions related to Förster theory¹⁵ are the ideal dipole approximation used for the coupling in eq 5, as well as the rather crude dielectric screening factor based on the refractive index of the medium in eq 4, which does not depend on the local environment or the relative D/A arrangement. Moreover, the impossibility of determining the relative orientation of D/A molecules often leads to the adoption of an isotropic average for the orientation factor, $\langle \kappa^2 \rangle = 2/3$.

2.2. Polarizable QM/MM and QM/Continuum Models for Energy Transfer. To overcome the limitations of Förster dielectric screening and dipole–dipole approximations discussed above, we use two multiscale approaches developed to study energy transfer in condensed phases. These methods combine a quantum mechanical (QM) description of the chromophores (in this case, FBP and Trp) with either a classical polarizable molecular mechanics (MMPol) description of the surrounding environment based on point charges and induced dipoles⁵¹ or a continuum solvation description based on the Polarizable Continuum Model (PCM).^{55,56} Both approaches take into account mutual polarization effects among the QM and MM regions during the self-consistent-field process and the excited-state environment response. The electronic couplings are then estimated from the transition densities derived from the QM/MMPol or QM/PCM excited-state calculations, and dielectric screening effects are naturally incorporated by explicitly computing an environment-mediated term to the Coulomb coupling among D and A

$$V = V_{\text{coul}} + V_{\text{env}} \quad (7)$$

The Coulomb and the environment-mediated terms in the QM/MMPol and QM/PCM formulations are given by

$$V_{\text{coul}} = \int \mathrm{d}\mathbf{r} \int \mathrm{d}\mathbf{r}' \rho_{\text{A}}^{\text{T}*}(\mathbf{r}') \frac{1}{|\mathbf{r}' - \mathbf{r}|} \rho_{\text{D}}^{\text{T}}(\mathbf{r}) \quad (8)$$

$$V_{\text{MMPol}} = - \sum_k \int \mathrm{d}\mathbf{r} \rho_{\text{A}}^{\text{T}*}(\mathbf{r}) \frac{(\mathbf{r}_k - \mathbf{r})}{|\mathbf{r}_k - \mathbf{r}|^3} \boldsymbol{\mu}_k^{\text{MMPol}}(\rho_{\text{D}}^{\text{T}}) \quad (9)$$

$$V_{\text{PCM}} = \sum_i \int \mathrm{d}\mathbf{r} \rho_{\text{A}}^{\text{T}*}(\mathbf{r}) \frac{1}{|\mathbf{s}_i - \mathbf{r}|} \mathbf{q}_i^{\text{PCM}}(\rho_{\text{D}}^{\text{T}}) \quad (10)$$

where $\rho_{\text{D}/\text{A}}^{\text{T}}$ indicates the transition densities of the interacting pigments, and $\boldsymbol{\mu}_k^{\text{MMPol}}$ and $\mathbf{q}_i^{\text{PCM}}$ are the MMPol-induced dipoles and PCM apparent surface charges describing the polarization response of the environment to a given transition density. The integrals in eqs 8–10 are solved using the efficient DFT grid-based integration methods implemented in the Gaussian code.^{42,55} The V_{coul} term describes the Coulomb interaction between the transition densities, thus representing an extension of Förster dipole–dipole coupling when higher multipolar contributions are included, that is, when the shape of the interacting molecules is fully taken into account. Both PCM and MMPol terms depend on the distance and mutual orientation of the chromophores and typically counteract the Coulomb term, thus leading to an overall screening of the interactions. One can thus define an effective screening factor, s , that can be directly compared to the Förster $\frac{1}{n^2}$ term in eq 4

$$s = \frac{(V_{\text{coul}} + V_{\text{env}})}{V_{\text{coul}}} \quad (11)$$

2.3. MD Simulations. We performed MD simulations of the (S)- and (R)-FBP/HSA complexes, as well as the (S,S)- and (R,S)-FBP–TrpMe dyads. The (S,S)- and (R,S)-FBP–TrpMe

dyads were solvated in a CH₃CN truncated octahedron box (buffer zone of 25 Å) using the Leap tool of the Amber 12 software.⁵⁷ The FBP/HSA systems, on the other hand, were initially prepared from the crystal structure of HSA complexed with ibuprofen (PDB 2BXG), solved at a 2.7 Å resolution,⁵⁰ where the structures of the (S) and (R) enantiomers of anionic FBP were aligned to the structure of ibuprofen. In particular, ibuprofen is bound to the well-known site II of HSA but also to a secondary site of HSA located between subdomains IIA and IIB.⁵⁰ We align FBP enantiomers to place them at this secondary site, as an energy transfer in the subnanosecond time scale is only compatible with that binding mode, populated by 30–40% of FBP enantiomers under the experimental conditions we aim at simulating,^{14,46} as will be discussed in the Results section. We assumed a standard protonation state for all residues except for Glu244, which was in the neutral state, as indicated by pK_a estimates at neutral pH computed using the PROPKA3 server.⁵⁸ The system was then solvated in a TIP3P⁵⁹ water truncated octahedron box (buffer zone of 20 Å). We used the force field for CH₃CN derived by Grabuleda et al.⁶⁰ and the Amber ff12SB force field for the protein,^{61,62} whereas the parameters for FBP and the FBP–TrpMe dyad were derived on the basis of the Amber ff12SB force field^{61,62} using the Antechamber module of Amber 12. The torsional parameters describing the rotation of the bond connecting the FBP phenyl units were explicitly parametrized by comparing fully relaxed torsion profiles computed at 5° intervals, in which the dihedral angle was frozen, obtained by minimizing the FBP structure at the B3LYP-D3/6-31++G(d,p) level of theory (adopting D3 Grimme's dispersion correction⁶³) and using the classical force field. The atomic charges for the FBP and the FBP–TrpMe enantiomers were computed following the restrained electrostatic potential (RESP) approach⁶⁴ based on the HF/6-31G(d) electrostatic potential computed on B3LYP/6-31++G(d,p) (dyads) and MP2/6-311++G(d,p) (FBP) optimized geometries. The FBP and FBP–TrpMe force-field parameters are reported in the Supporting Information.

The (S,S)- and (R,S)-FBP–TrpMe dyads and the (S)- and (R)-FBP/HSA systems were first energy-minimized for 1000 steps (only solvent and ions) and then for 2500 steps allowing the complete systems to relax. Then, the systems were gradually thermalized by running five 100 ps simulations at constant volume to increase the temperature from 0 to 298 K and subsequently simulated in the isothermal–isobaric ensemble (1 atm and 298 K) for a total of 100 ns. For the (S)- and (R)-FBP/HSA systems, the equilibration plus production runs starting from the energy-minimized structures were repeated for a total of five replicas. All simulations were done using the Amber 12 code,⁵⁷ with periodic boundary conditions, the particle-mesh Ewald approach to account for long-range electrostatics, a nonbonded cutoff of 8 Å, a 2 fs integration time step, the SHAKE algorithm to restrain bonds involving hydrogen, and the Langevin dynamics method to control the temperature (collision frequency 1 ps^{−1}). In constant pressure simulations, we used the isotropic position scaling algorithm (pressure relaxation time 2 ps). From the last 50 ns of each replica, a total of 100 snapshots were extracted every 0.5 ns to be used in the QM/classical calculations.

2.4. QM/MMPol and QM/PCM Calculations. QM geometry optimizations of the model dyad structures extracted from the MD simulations were done at the B3LYP/6-31G(d) level of theory adopting D3 Grimme's dispersion correction⁶³ and the SMD solvation model⁶⁵ in acetonitrile. We then

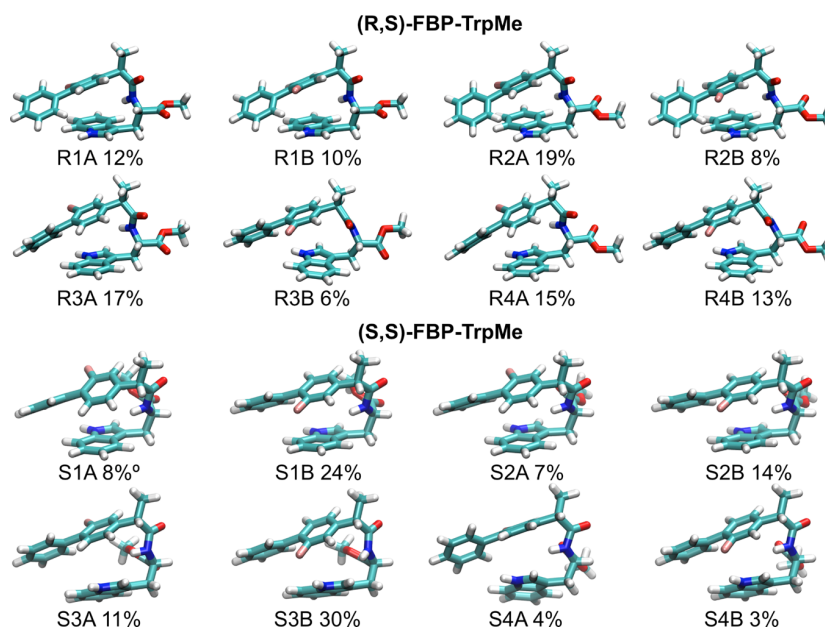


Figure 1. Main conformers for the (*R,S*)- and (*S,S*)-FBP–TrpMe dyads and their corresponding Boltzmann populations predicted by SMD-B3LYP-D3/6-31G(d) calculations in acetonitrile solution.

performed a frequency analysis to verify the nature of the minima and to compute the total free energies of each conformer and their corresponding Boltzmann populations. All frequencies were scaled by 0.977, as recommended for that level of theory.⁶⁶

QM/MMPol and QM/PCM excited-state calculations, on the other hand, were performed at different QM levels of theory: semiempirical Zerner's intermediate neglect of differential overlap (ZINDO),⁶⁷ configuration interaction of single excitations (CIS),⁶⁸ and TD-DFT⁶⁹ using the CAMB3LYP functional.⁷⁰ In CIS and TD-CAMB3LYP calculations, we used the 6-31G(d) basis set. All calculations were performed with a locally modified version of the Gaussian 09 code.⁷¹ In the PCM calculations, we used the default cavity parameters in Gaussian. On the other hand, QM/MMPol calculations were performed using the Amber pol12 AL polarizable force field to describe the protein and water solvent,^{72,73} where water atomic charges were computed at the MP2/aug-cc-pVTZ level of theory on the TIP3P geometry using the RESP approach. Calculations for the model FBP–TrpMe dyads were performed either on the single Trp and FBP chromophoric units (the indole ring of Trp and the two phenyl rings of FBP, D–A system) or on the complete FBP–TrpMe dyad, including the bridge linking those units (D–B–A system). In the D–B–A models, the D/A units were separated at the bond between the C_α and the N atoms of TrpMe to compute the transition densities for each fragment. In both D–A and D–B–A model fragments, we capped the appropriate carbon atoms by introducing hydrogens at standard bond lengths. In all D–B–A and D–A monomers and coupling calculations, we used the PCM cavity corresponding to the complete FBP–TrpMe molecule.

3. RESULTS AND DISCUSSION

In this study, we aim at investigating the dynamic fluorescence quenching observed in FBP–TrpMe model dyads and FBP–HSA complexes by Vayá and co-workers.¹⁴ In particular, a fast stereoselective decay of the fluorescence at 310 nm, where

emission is dominated by FBP, was suggested to indicate fast energy transfer from 1FBP* to the Trp214 residue in HSA. By simulating the energy transfer dynamics in these systems, we want to validate this hypothesis and critically assess the ability of Förster theory and its underlying approximations to discriminate the FRET properties of the FBP enantiomers bound to HSA. In turn, this allows us to critically examine the theoretical binding modes simulated by MD and the potential of the MD-QM/MMPol strategy proposed to describe FRET processes in protein–ligand complexes.

3.1. FBP–TrpMe Model Dyads. The dynamic quenchings observed in (*S,S*)- and (*R,S*)-FBP–TrpMe are expected to strongly depend on the conformational preferences adopted by the dyad in acetonitrile solution. To investigate such preferences, in particular the mutual arrangement between the FBP and Trp units, we performed classical MD simulations for each dyad. The structures sampled along a 100 ns MD trajectory indicated that both dyads adopt an extended conformation most of the time. To investigate whether the conformations sampled by the MD are realistic, we then determined the most stable conformations of the dyads in acetonitrile solution by performing DFT (B3LYP-D3) geometry optimizations coupled to the SMD continuum solvation model.⁶⁵ In particular, we optimized the geometry of the dyads starting from a set of 100 structures extracted from the MD optimizations, which lead in many cases to equivalent conformers. Analysis of the free energies of the final unique optimized structures indicated that several stacked conformations in which the FBP and the Trp aromatic rings are in close contact mainly populate the ensemble of both dyads, whereas extended conformations are destabilized by several kcal/mol. Thus, the classical force field used in the MD was clearly misrepresenting the conformational preferences of the dyads by overstabilizing extended structures. This probably arises from an underestimation of dispersion interactions among the aromatic rings or an unbalanced description of solvation effects among extended and stacked conformations. We thus focused

on the most stable conformations identified from DFT calculations, contributing >1% to the ensemble, which displayed a similar stacked orientation differing in (i) the orientation of the Trp indole ring with respect to the FBP rings, (ii) the orientation of the methyl ester group, and (iii) the position of the F atom on one side or the other of the FBP central ring. In Figure 1, we report these conformations, together with the corresponding Boltzmann populations, following a notation in which R/S indicates the (R,S) or the (S,S) enantiomer, and A/B is used for equivalent conformations, where the position of the F atom is located on opposite sides of the FBP central ring.

Accordingly, we computed the excitonic couplings for the different conformers of the dyads shown in Figure 1, describing the surrounding acetonitrile solvent through the PCM continuum solvation model, as given by eq 10. In particular, we performed calculations for both D–A models, where only the aromatic rings of the chromophores are accounted for, and for D–B–A models, where the bridge connecting them is also included. In Table 1, we report the orientation factors and the

Table 1. Orientation Factors and Electronic Couplings Calculated for the (S,S)- and (R,S)-FBP–TrpMe Model Dyads at Different QM/PCM Levels of Theory

		$\langle \kappa^2 \rangle^{a,b}$	$\langle V^2(R) \rangle / \langle V^2(S) \rangle^a$	
			dip–dip ^b	QM/PCM
ZINDO				
D–A ^c	R	0.10	0.68	0.25
	S	0.13		
D–B–A ^d	R	0.10	1.05	0.30
	S	0.08		
CIS				
D–A ^c	R	0.07	0.39	0.36
	S	0.13		
D–B–A ^d	R	0.06	0.39	0.61
	S	0.09		
TD-CAMB3LYP				
D–A ^c	R	0.09	0.44	0.53
	S	0.18		
D–B–A ^d	R	0.07	0.52	1.12
	S	0.12		
exp $k_q(R)/k_q(S)^e$			2.21	

^aValues averaged over the Boltzmann populations of the conformers. ^bDipole–dipole couplings calculated from the transition dipoles of the corresponding QM/PCM calculations based on the center of the indole ring (Trp) and the center of the two phenyl units (FBP). ^cCalculations performed on the single Trp and FBP chromophoric units (aromatic rings). ^dCalculations performed on the complete FBP–TrpMe dyad. ^eDynamic quenching rates measured from fluorescence upconversion.¹⁴

ratios of squared coupling values obtained for each enantiomer, which are averaged over the Boltzmann populations of the conformers. The actual values of squared couplings, screening factors, and the average transition energies and dipoles of FBP and Trp are reported in Tables S1 and S2, whereas in Table 2, we report the coupling data for the individual D–B–A conformers of the dyads obtained at the TD-CAMB3LYP level of theory.

In line with the absorption spectra of the dyads,¹⁴ our calculations predict almost no changes in transition energy or transition dipole strength among enantiomers, indicating that

Table 2. Boltzmann Populations, Interchromophoric Distances, Orientation Factors, and Electronic Couplings Calculated for the Complete D–B–A Conformers of the (S,S)- and (R,S)-FBP–TrpMe Model Dyads at the PCM TD-CAMB3LYP/6-31G(d) Level of Theory

	% ^a	$\langle R \rangle$ (Å)	$\langle \kappa^2 \rangle$	$\langle V_{\text{dip-dip}}^2 \rangle^b$ (cm ⁻¹)	$\langle V_{\text{QM/PCM}}^2 \rangle$ (cm ⁻¹)
(R,S)-FBP–TrpMe					
R1A	12	4.09	0.06	1.50×10^4	5.04×10^4
R1B	10	3.99	0.09	2.75×10^4	4.05×10^4
R2A	19	4.08	0.05	1.29×10^4	4.87×10^4
R2B	8	4.17	0.02	3.50×10^3	3.15×10^4
R3A	17	3.97	0.06	1.92×10^4	2.42×10^4
R3B	6	4.68	0.11	1.25×10^4	6.90×10^3
R4A	15	3.99	0.05	1.58×10^4	2.49×10^4
R4B	13	4.69	0.14	1.57×10^4	6.64×10^3
(S,S)-FBP–TrpMe					
S1A	8	3.87	0.25	7.64×10^4	4.52×10^4
S1B	24	4.03	0.22	5.24×10^4	5.24×10^4
S2A	7	3.94	0.21	6.05×10^4	4.29×10^4
S2B	14	4.09	0.23	4.86×10^4	5.31×10^4
S3A	11	4.36	0.02	3.94×10^3	4.06×10^3
S3B	30	4.25	0.00	3.80×10^2	1.54×10^2
S4A	4	4.91	0.01	1.38×10^3	1.37×10^4
S4B	3	4.35	0.00	3.67×10^2	2.72×10^3

^aBoltzmann weights derived from free energies calculated at the SMD-B3LYP-D3/6-31G(d) level of theory. ^bDipole–dipole couplings include Förster screening factor $s^2 = 1/n^4 = 0.31$ ($n^2 = 1.806$ in acetonitrile).

changes in energy transfer properties arise from the V^2 electronic factor in eq 1, instead of the spectral overlap factor J . Interestingly, the mutual arrangement of the chromophores leads to an almost orthogonal configuration of the transition dipole moments of the L_a state of Trp and the bright $\pi \rightarrow \pi^*$ state of FBP. If we ignore the bridge that links the Trp and the FBP chromophoric units in the calculations (D–A models), we find average squared orientation factors ($\langle \kappa^2 \rangle \sim 0.07$ – 0.10 and 0.13 – 0.18 for the (R,S) and (S,S) enantiomers, respectively, in contrast to the ~ 2 faster dynamic quenching observed for the (R,S) diastereomer. This result is very similar regardless of whether the transition dipole moments are derived from semiempirical ZINDO, CIS, or TD-DFT calculations. Because of the almost orthogonal orientation of the respective dipoles, the coupling in the complex however becomes extremely sensitive to the precise orientation of such dipoles. If we add the bridge linking the Trp and FBP units (D–B–A models), the Trp L_a dipole slightly reorients, leading to a $\sim 30\%$ decrease in the $\langle \kappa^2 \rangle$ value for the (S,S) dyad, whereas it remains very similar for the (R,S) one. The resulting ratio among $\langle \kappa^2 \rangle$ values, however, still suggests a faster energy transfer for the (S,S) enantiomer. Indeed, the ratios among squared coupling values $\langle V^2(R) \rangle / \langle V^2(S) \rangle$ shown in Table 1, obtained either using the ideal dipole approximation or from the QM/PCM model based on transition densities, are in most cases smaller than one.

As expected, our results show that Förster dipole–dipole approximation introduces significant deviations in the coupling estimates, given the close separation between FBP and Trp in the dyad, especially in the D–B–A models. If we focus on the most accurate estimates based on transition densities, interestingly, only the TD-DFT data based on the complete D–B–A models are able to describe the correct enantioselectivity, whereas the simple D–A models or the CIS and

ZINDO calculations predict a $\sim 2\text{--}3$ faster transfer for the (S,S) dyad, in contrast to the observed ratio $k_q(\text{R})/k_q(\text{S}) = 2.21$ among dynamic quenching rates experimentally observed in the dyads.¹⁴ The TD-DFT calculations, however, still predict a ratio $\langle V^2(\text{R}) \rangle / \langle V^2(\text{S}) \rangle = 1.12$, smaller than the experimental value of 2.21.

Here, it is worth noting that in the Förster dipole–dipole picture, the ratio among enantiomers arises entirely from the changes in the orientation factors, as interchromophoric distances and transition dipole strengths are very similar, as reported in Tables S1 and S2. In the more accurate picture obtained from QM-PCM calculations, however, not only are the full transition densities considered but the shape of the chromophores is also accounted for while evaluating screening effects. The screening factors reported in Table S2 indicate that Förster overestimates dielectric screening in the dyad, as the s^2 values range from ~ 0.33 to 0.53 in the D–B–A models, depending on the QM method of choice, whereas Förster approximation leads to a value $1/n^4 = 0.31$. However, we find similar screening effects in both dyads, so solvent screening does not contribute to the observed enantioselective energy transfer.

To gain further insights into the origin of the coupling values, in Table 2, we report the results obtained for each individual conformer of the dyads. In the case of (R,S)-FBP–TrpMe, most conformers display similar coupling values and orientation factors $\langle \kappa^2 \rangle \sim 0.05\text{--}0.10$. In contrast, the results for the (S,S) enantiomer clearly show two different behaviors. Whereas conformers S1A/B and S2A/B are characterized by an orientation factor $\langle \kappa^2 \rangle \sim 0.25$, the rotated orientation of the indole group of Trp in conformers S3A/B and S4A/B makes both transitions virtually orthogonal with orientation factors $\langle \kappa^2 \rangle \sim 0.01$. Thus, the population of S3A/B and S4A/B conformers is the origin of the enantioselective dynamic quenching, making energy transfer in the (R,S) enantiomer about 2 times faster. Our TD-DFT calculations, however, only predict an energy transfer 12% faster, probably owing to the extreme sensitivity of our results to the relative free energies calculated for the S3A/B and S4A/B conformers compared to those for the S1A/B and S2A/B ones, the two groups having a population of about 50%. In this case, a further stabilization of the S3A/B and S4A/B conformers by just 0.7 kcal/mol would increase their overall population to 75% and lead to an excellent agreement with experiments. Of course, another reason for this discrepancy could be related to the intrinsic limitations of the TD-DFT method.⁷⁴

Overall, the results found for the dyad strongly support an energy transfer process from the 1FBP* state to the Trp L_a state as the origin of the observed dynamic quenching. In addition, they show that capturing the proper orientation for the Trp L_a dipole, as well as the proper balance among (S,S) conformers, is key to explain the observed enantioselectivity. Indeed, both dipole–dipole and transition density-based estimates of the coupling values lead in most cases to a wrong answer that predict similar transfer rates for the enantiomers or even a faster transfer for the (S,S) one. On the basis of the spectral overlaps computed from the experimental spectra of FBP and TrpMe in acetonitrile, in Table 3, we also report a comparison between the actual values of the experimental and predicted energy transfer rates. Despite the fact that the TD-DFT simulations are able to predict the enantioselectivity, although with moderate accuracy, the actual rates are very fast by approximately 1 order of magnitude. This

Table 3. Electronic Energy Transfer Times Calculated at the PCM TD-CAMB3LYP/6-31G(d) Level of Theory for the (R,S)- and (S,S)-FBP–TrpMe Dyads

		τ_{FRET} (ps) ^a		
		dip–dip ^b	QM/PCM ^b	exp ^c
D–B–A	R	3.60	1.81	28
	S	1.86	2.03	62

^aEnergy transfer times defined as the inverse of the rate $\tau_{\text{FRET}} = 1/k_{\text{FRET}}$. ^bEnergy transfer rates calculated using a spectral overlap $J = 0.1208 \text{ eV}^{-1}$ obtained from the emission spectra of FBP and the absorption spectra of TrpMe in acetonitrile/air¹⁴ according to eq 2. ^cDynamic quenching rates measured from fluorescence upconversion in ref.¹⁴

discrepancy could arise from the neglect of thermal effects in the coupling values, which are very sensitive to the actual geometry of the dyad due to the almost orthogonal orientation of the transition dipoles. In addition, the significant coupling between FBP and TrpMe could lead to potential deviations from the Förster ideal weak-coupling regime, although in this case the fact that the absorption spectrum of FBP–TrpMe is approximately the sum of the FBP and TrpMe ones supports the weak coupling assumption.¹⁴

3.2. FBP/HSA Complexes. The first question to address to explain the subnanosecond stereoselective dynamic quenching of FBP concerns the location of the drug in the HSA complex. The experiments were performed at an FBP/HSA ratio 0.7:1,¹⁴ and at this ratio, 40 and 30% of (S)- and (R)-FBP are bound to the secondary binding site of HSA, whereas the remaining drug is bound to the preferred site II.⁴⁶ As discussed in the Introduction, we argue that this secondary binding corresponds to a site of HSA at the interface between subdomains IIA and IIB.⁵⁰ In contrast to the model dyads, where a >90% quenching was observed, in the FBP/HSA complexes, the quenching was around $\sim 50\%$. We thus performed exploratory calculations of the approximate time scales of 1FBP* to Trp L_a excitation transfer by aligning both enantiomers of FBP at the location of ibuprofen at both sites of HSA, as found in the crystal structure (PDB 2BXG).⁵⁰ In these calculations, we estimated the Coulomb coupling using the full transition densities (eq 8) obtained at the ZINDO level and accounted for dielectric screening through a simple Förster screening factor $1/n^2$ (eq 4), where n^2 was taken to be 2 for a protein environment. Then, we computed a spectral overlap factor $J = 0.1321 \text{ eV}^{-1}$ from the emission spectra of FBP and the absorption spectra of HSA¹⁴ according to eq 2. The estimated energy transfer rates for the (R)- and (S)-FBP to Trp214 energy transfer were found to be $\sim 500\text{--}800$ ps for the enantiomers located at the secondary site and much slower, $\sim 30\text{--}150$ ns, when located at site II, in line with the much shorter distance ~ 10 Å between FBP and Trp214 in the secondary site compared to ~ 20 Å in site II. The time scales predicted from the secondary site are in excellent agreement with the subnanosecond quenching times of 180 and 250 ps measured for (S)- and (R)-FBP, whereas the rates from site II are clearly very slow. This explains why only a $\sim 50\%$ quenching of FBP fluorescence is observed, as only the fraction of FBPs located in the secondary site contribute to the subnanosecond dynamic quenching through energy transfer to Trp214. We have thus focused only on the binding of FBP at the secondary site of HSA, which is postulated as the origin of the dynamic quenching.

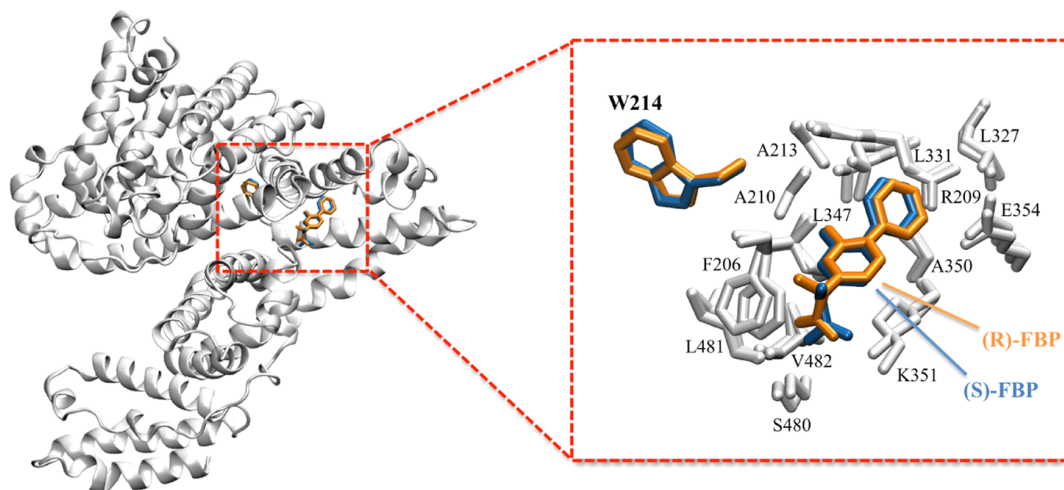


Figure 2. Illustration of the binding modes predicted from MD simulations for the (R) and (S) enantiomers of FBP at the secondary site of HSA located at the interface between subdomains IIA and IIB. The inset shows the FBP ligand, the Trp214 residue participating in the energy transfer process, and the amino acids surrounding the ligand in the binding pocket.

In the FBP/HSA complexes, as in the model dyads discussed above, the energy migration properties between the FBP enantiomers and Trp214 exquisitely depend on the separation and mutual orientation between them. We have thus run MD simulations for a total time of 0.5 μ s, divided in 5 replicas, for each enantiomer of FBP. In Figure S1, we show the root-mean-squared deviations (RMSD) of atomic positions for the complex and for the ligand along each MD replica. As can be observed from the RMSD plots, the ligand is rather flexible in the binding pocket. In particular, the (S)-FBP enantiomer often changes its configuration in the binding site, enlarging the RMSD values to ~ 4 – 5 Å, although it returns rapidly to the stable binding mode with an RMSD of ~ 2 – 3 Å. Such transitions are also observed for the (R) enantiomer, although to a smaller extent. Overall, the ligand and the protein RMSD values indicate that the trajectories are stable.

In Figure 2, we show the arrangement of (R)- and (S)-FBP in the secondary binding site of HSA at the end of the trajectories. As can be observed, the binding mode predicted for both enantiomers is very similar. Moreover, the mutual orientation between FBP and Trp214 is also comparable, thus at first sight unable to explain the 30% faster quenching observed in the fluorescence of (S)-FBP compared to that of the other enantiomer. However, as indicated by the RMSD plots, the ligands are rather flexible, and such flexibility could have a strong impact on the actual electronic coupling between Trp214 and FBP. In Figure 3, we show the fluctuations of the electronic couplings, calculated at the QM/MMPol TD-CAMB3LYP/6-31G(d) level of theory, for each enantiomer and MD replica. As can be observed, there are large fluctuations around the average coupling values. Indeed, the standard deviations found for the coupling fluctuations, $\sigma \sim 5$ cm^{-1} , are comparable to the mean values $\langle V \rangle \sim 7$ cm^{-1} for both enantiomers. Such dramatic oscillations, for example, are in sharp contrast to results found for photosynthetic pigment–protein complexes, where the pigments are tightly held to the protein scaffold and standard deviations typically amount to only $\sim 20\%$ of the mean coupling values.¹⁸ Thus, whereas in photosynthetic complexes it is common to ignore coupling fluctuations and evaluate the pigment–pigment interactions based on the crystal structure,⁴² such a strategy is clearly not

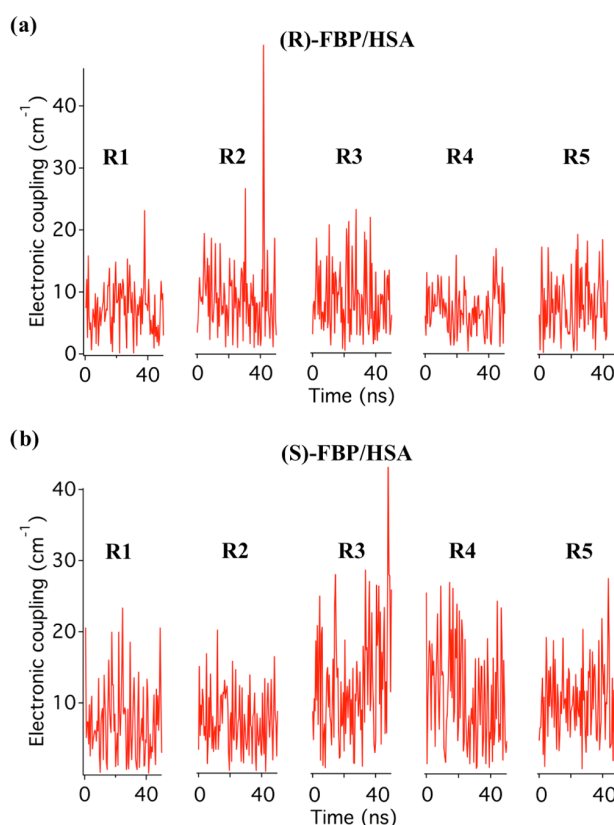


Figure 3. Fluctuations of electronic couplings along the five MD replicas calculated at the QM/MMPol TD-CAMB3LYP/6-31G(d) level of theory between Trp214 and the (R) and (S) enantiomers of FBP in the HSA complex.

justified in the case of a drug–protein complex, where the bound ligand has more flexibility, even if that flexibility depends on each particular ligand and binding site. We thus average the coupling values over the MD trajectories. Note, however, that we do not explicitly account for the coupling fluctuations in the estimate of FRET rates. Explicit account of such fluctuations

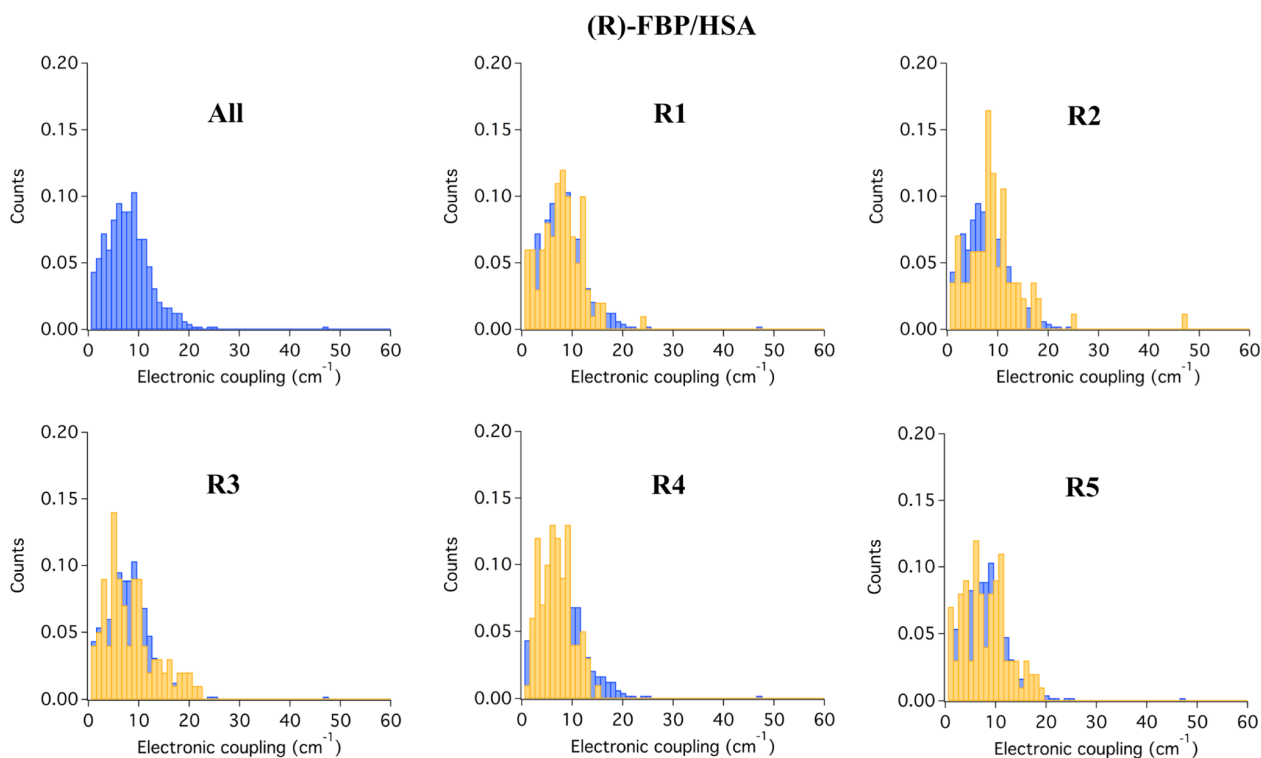


Figure 4. Distribution of electronic couplings calculated at the MD-QM/MMPol TD-CAMB3LYP/6-31G(d) level of theory for the (R)-FBP/HSA complex: total distribution (blue) and distributions computed for each MD replica (yellow).

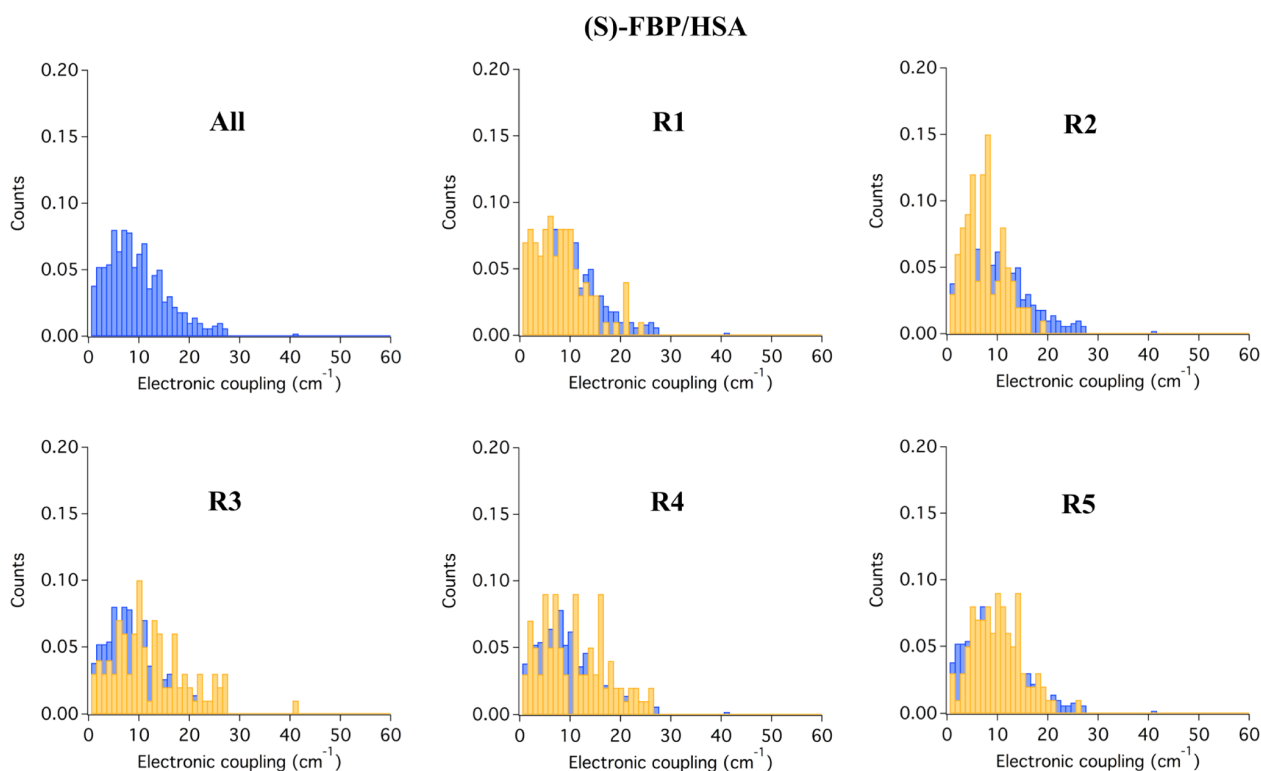


Figure 5. Distribution of electronic couplings calculated at the MD-QM/MMPol TD-CAMB3LYP/6-31G(d) level of theory for the (S)-FBP/HSA complex: total distribution (blue) and distributions computed for each MD replica (yellow).

has been recently shown to introduce negligible changes in the exciton dynamics of a photosynthetic complex.⁷⁵ However,

when the chromophores involved in the transfer display a larger degree of structural dynamics, as happens in common FRET

experiments, it can be important to explicitly account for such coupling fluctuations, for example, by processing the coupling trajectory through Markov chain approaches.^{29,34} Here, the adoption of an average coupling value is deemed a reasonable approximation, given that the FBP ligand stays most of the time in a stable binding mode.

Thus, the large degree of coupling fluctuations makes it necessary to average the energy transfer properties over a large number of structures. In Figures 4 and 5, we show the total distributions of coupling values, as well as those derived from each MD replica. As can be observed, the total distributions, which include a total of 500 structures extracted along the 0.5 μ s of MD, are reasonably well converged, whereas the individual distributions derived for the individual replicas are clearly not representative of the ensemble. Interestingly, the distribution of coupling values has a similar maximum around ~ 7 – 8 cm^{-1} for both enantiomers, but the (S) one displays a larger tail in the distribution toward larger values, thus leading to a mean coupling $\langle V \rangle = 9.2$ cm^{-1} compared to the smaller value of 7.4 cm^{-1} found for (R)-FBP. The tail toward large coupling values indicates the larger degree of flexibility of the (S)-FBP diastereomer compared to that of the (R) one, in line with the larger RMSD fluctuations from our simulations discussed previously and the faster reorientational time derived from fluorescence anisotropy.¹⁴ We have analyzed the RMSD values of the structures that give rise to the tail toward high-coupling values, that is, those leading to coupling > 22 cm^{-1} , and found no particular increase in the RMSD values for the high-coupling structures. Thus, the temporary changes in binding mode, illustrated by ~ 4 – 5 Å RMSD values, are not the reason for the increased coupling strength. Rather, it is the *dynamic* characteristics of the stable binding mode characterized by RMSD of ~ 2 – 3 Å that lead to the wider coupling distribution for the (S) enantiomer.

To gain further insights into the origin of the larger coupling distribution obtained for (S)-FBP, in Table 4, we provide the corresponding FBP–Trp214 distances and orientation factors,

Table 4. Interchromophoric Distances, Orientation Factors, and Electronic Couplings Calculated for the (S)- and (R)-FBP/HSA Complexes at Different QM/MMPol Levels of Theory

	$\langle R \rangle^a$ (Å)	$\langle \kappa^2 \rangle^{a,b}$	$\langle V^2(S) \rangle / \langle V^2(R) \rangle^a$	
			dip–dip ^b	QM/MMPol
ZINDO				
R	11.01 ± 0.56	0.12 ± 0.14	4.89	1.89
S	10.52 ± 0.88	0.27 ± 0.31		
CIS				
R	11.01 ± 0.56	0.23 ± 0.20	2.14	1.77
S	10.52 ± 0.88	0.26 ± 0.28		
TD-CAMB3LYP				
R	11.01 ± 0.56	0.28 ± 0.24	1.38	1.54
S	10.52 ± 0.88	0.26 ± 0.27		
exp $k_q(S)/k_q(R)^c$			1.39	

^aValues averaged over the MD trajectories. For distances and orientation factors, standard deviations are also reported. ^bDipole–dipole couplings calculated from the transition dipoles of the corresponding QM/MMPol calculations based on the center of the indole ring (Trp) and the center of the two phenyl units (FBP). ^cDynamic quenching rates measured from fluorescence upconversion.¹⁴

as well as the ratio $\langle V^2(S) \rangle / \langle V^2(R) \rangle$, calculated at the ZINDO, CIS, and TD-CAMB3LYP levels of theory. The corresponding values of transition energies and dipoles, as well as the averaged values of squared electronic couplings and screening factors, are provided in Tables S3 and S4. As in the model dyads discussed in the previous section, we find no significant differences in transition energies or dipole strengths for the states of FBP and Trp in the (R)- and (S)-FBP/HSA complexes, in agreement with similar absorption spectra measured for both complexes.¹⁴ Thus, variations in the 1FBP* to Trp energy transfer arise from the V^2 electronic factor in eq 1, as the spectral overlap factor J does not change. The $\langle V^2(S) \rangle / \langle V^2(R) \rangle$ ratios predicted from our QM/MMPol simulations indeed nicely explain the ratio of quenching rates $k_q(S)/k_q(R)$ measured from fluorescence upconversion, indicating 89, 77, and 54% faster energy transfer for the (S) enantiomer at the ZINDO, CIS, and TD-DFT levels of theory, respectively. Thus, as expected, and in line with the findings discussed in the previous section for the model FBP–TrpMe dyads, again TD-DFT performs slightly better than ZINDO or CIS by predicting a 54% faster quenching for (S)-FBP compared to that of the (R) form, very close to the 30% experimental value. The better performance of TD-DFT is even amplified when the couplings are derived based on the ideal dipole approximation, as in this case, only the dipole data derived from TD-DFT are able to explain the experimental ratio giving a rather accurate estimate of 38% increase, whereas the adoption of transition dipoles derived from CIS and ZINDO calculations predict faster rates in the order of 214 and 489%. The important conclusion from these data is that the dipole–dipole couplings assumed in Förster theory are too sensitive to small changes in the actual orientation of the transition dipoles, whereas the estimates based on full 3D transition densities are remarkably more robust to such changes and make the choice of QM method less critical.

If we turn again to the origin of the larger coupling distribution obtained for the (S) enantiomer of FBP, the data in Table 4 allow us to dissect the variation of the coupling values in terms of distances and orientation factors. In addition, in Figure 6, we report the distribution of screening factors derived from MD-QM/MMPol calculations compared to the $s = 1/n^2$ Förster approximation computed by assuming a value $n^2 = 2$ for the protein environment.¹⁸ Regarding the Trp214 to FBP distances, we find a slightly shorter average separation of 10.5 Å for (S)-FBP compared to the value of 11.0 Å found for (R)-FBP. Such a change suggests a larger FRET rate for the (S) enantiomer by 31% based on the Förster R^{-6} distance-dependent rate expression, very close to the 30% increase observed. Other factors beyond chromophore separation however come into play, namely, the orientation between chromophores and changes in dielectric screening effects. On the basis of the most accurate TD-DFT data, these two effects, in contrast to the FBP–Trp distance, tend to increase the FRET rate of the (R) enantiomer by 8% each, as $\langle \kappa^2 \rangle$ passes from 0.26 to 0.28 and $\langle s^2 \rangle$ from 0.35 to 0.38 (see Table S4). Thus, the *average* conformation of the (S) enantiomer in the binding pocket explains only about 15% of the total 54% increase in transfer rate compared to that of the (R) enantiomer, as predicted by the TD-DFT data based on transition densities, an increase caused by a shorter interchromophoric distance but counteracted by a less favorable orientation and modulation of the coupling by dielectric screening effects. The remaining increase, thus, can be associated with the *dynamic* characteristics of the stable binding

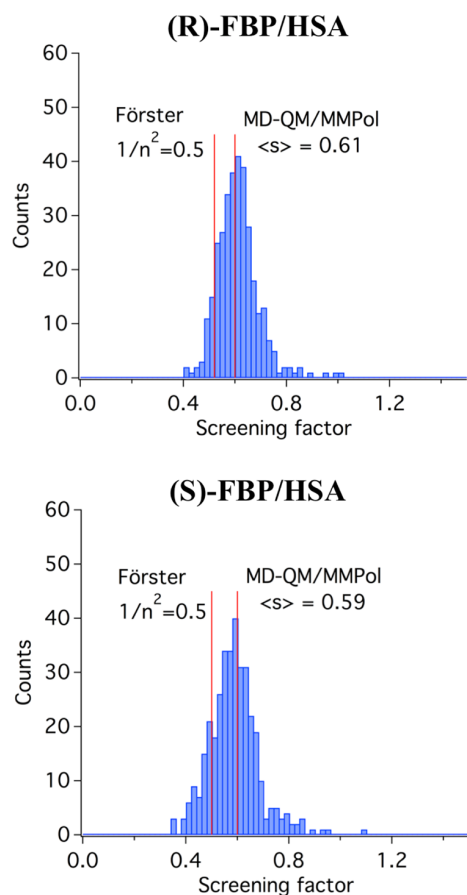


Figure 6. Distribution of dielectric screening factors calculated at the MD-QM/MMPol TD-CAMB3LYP/6-31G(d) level of theory for the (R)- and (S)-FBP/HSA complexes.

mode adopted by (S)-FBP, which seems to be more flexible, thus leading to a wider distribution of coupling values, as shown in Figure 5 and discussed before. Indeed, the distance fluctuations for the (S) enantiomer, with a $\sigma = 0.88$ Å, are significantly larger than the value of 0.56 Å found for the (R) form. It is also interesting to note the large distribution of screening factors obtained along the MD simulations shown in Figure 6, which underscores the strong variation that can arise in transfer rates due to differences in the local environment and pigment orientation, an effect completely ignored when the Förster $1/n^2$ screening factor is assumed.

Finally, in Table 5, we provide the predicted energy migration rates for the (R)- and (S)-FBP enantiomers,

estimated using the MD-QM/MMPol couplings at the ZINDO, CIS, and TD-CAMB3LYP levels of theory, together with the experimental spectral overlap factor. As can be seen, although TD-DFT provides the best enantioselective quenching ratios shown in Table 4, CIS and ZINDO methods in this case provide a slightly better estimate of the energy transfer times. In particular, TD-DFT predicts the energy transfer to occur in 666 and 434 ps for (R)- and (S)-FBP, respectively, compared to the measured quenching times of 250 and 180 ps,¹⁴ whereas ZINDO and CIS predict values of ~ 300 – 400 and ~ 150 – 200 ps. Here, we remark that the comparison of the predicted absolute FRET times with the measured quenching times is considerably more delicate than the comparison of the relative rate among enantiomers. First, deviations in the transfer rates are expected to arise from inaccuracies in the predicted transition dipole strengths of FBP and Trp that impact the actual coupling values. Indeed, whereas TD-DFT, and in particular the CAMB3LYP functional, is expected to provide reasonable estimates of oscillator strengths,⁷⁶ CIS has the well-known trend to overestimate transition dipoles, and thus, the electronic couplings and transfer rates,⁷⁷ in line with the observed faster rates obtained with CIS compared with TD-DFT for the FBP/HSA complexes. Here, we do not attempt to compare our predicted transition dipole strengths of FBP and Trp, reported in the Supporting Information, with experiments for two reasons. The main one is that the Trp L_a band overlaps the weaker L_b absorption band, making experimental derivation of the L_a transition dipole strength from the intensity of the absorption band not direct. In addition, the comparison between theory and experiment is not straightforward even if the band is well characterized, as local field effects need to be accounted for, and different models lead to significant differences.^{78,79}

Another reason that could explain the somewhat slow transfer times predicted by TD-DFT, which otherwise predicts the most accurate enantioselective ratio, is the potential participation of the Trp L_b state in the transfer, which could slightly increase the FRET rates. In this study, we also computed the couplings of the FBP $\pi \rightarrow \pi^*$ state with the L_b state of Trp (data not shown), which indicated that such transfer should occur 2–6 times slower than the transfer to the Trp L_a state based on the coupling values. However, derivation of the spectral overlap for this weak band, which overlaps with the brighter L_a band, is not clear. In addition, the experimental quenching times 250/180 ps correspond to the fastest contribution of the multiexponential fluorescence decay of the FBP/HSA complexes at 310 nm, where the fluorescence mainly originates from FBP, with additional slower components at ~ 0.7 and ~ 3 ns, so transfer to the L_b state most probably

Table 5. Electronic Energy Transfer Times Calculated for the (R)- and (S)-FBP/HSA Complexes at Different QM/MMPol Levels of Theory

	τ_{FRET} (ps) ^a					
	(R)-FBP/HSA			(S)-FBP/HSA		
	dip–dip ^b	QM/MMPol ^b	exp ^c	dip–dip ^b	QM/MMPol ^b	exp ^c
ZINDO	316	301	250	64	159	180
CIS	354	398		166	225	
TD-CAMB3LYP	403	666		293	434	

^aEnergy transfer times defined as the inverse of the rate $\tau_{\text{FRET}} = 1/k_{\text{FRET}}$. ^bEnergy transfer rates calculated using a spectral overlap $J = 0.1321$ eV⁻¹ obtained from the emission spectra of FBP and the absorption spectra of HSA¹⁴ according to eq 2. ^cDynamic quenching rates measured from fluorescence upconversion in ref.¹⁴

affects such slower components. Interestingly, the average lifetimes accounting for all components in the multiexponential decay, 700 and 540 ps, are in excellent agreement with the TD-DFT predicted rates, and still describe a quenching 30% faster for the (S) enantiomer, very similar to the ratio 38% derived only from the fastest components.

Overall, this analysis illustrates the fact that it is rather difficult to predict absolute accurate transfer times for a ligand–protein complex. The prediction of relative transfer times among the different enantiomers in a binding site, or for a given ligand in different binding sites, appears however to be much more robust, as the uncertainties related to the prediction of transition dipole moments, and thus in the coupling estimates, are expected to cancel in that case.

Nevertheless, our simulations strongly support a picture where energy transfer from FBP bound at the secondary site of HSA causes the subnanosecond enantioselective dynamic quenching observed. We find that the binding modes for both enantiomers at this site of HSA are fairly similar, the most apparent change in the binding pocket being a displacement of Phe206 due to the different orientation of the carboxylic group of FBP. Such similar binding modes, however, translate into a significantly larger flexibility of the (S) enantiomer of FBP, and this appears to be the main reason for the enantioselective dynamic quenching observed.¹⁴

4. CONCLUSIONS

In this contribution, we have explored the reliability of Förster theory to describe electronic energy transfer in ligand–protein complexes. We investigated the enantioselective dynamic quenching observed for the (R) and (S) enantiomers of FBP complexes with HSA, as well as parallel observations in a solvated model FBP–TrpMe dyad, by combining MD simulations and QM/PCM and QM/MMPol multiscale calculations of the energy transfer properties of the system. Our results strongly support the hypothesis that the subnanosecond dynamic quenching in these systems originates from energy transfer from the 1FBP* state to the L_a state of Trp. In the model dyads, we show that the enantioselectivity arises from a virtually orthogonal arrangement of the transitions of FBP and Trp in several conformers of (S)-FBP–TrpMe in acetonitrile solution. In FBP/HSA complexes, our simulations show that such transfer occurs from the fraction of FBP molecules bound to the secondary site of HSA, located at the interface between subdomains IIA and IIB, and that fairly similar binding modes for both enantiomers are found, such enantioselective energy transfer mostly arising from the increased flexibility of (S)-FBP in the binding pocket compared to that of (R)-FBP.

We show that the MD-QM/MMPol and QM/PCM strategies are able to describe, with almost quantitative accuracy, the enantioselectivity in the FBP/HSA complexes and the model dyads, when a TD-DFT description is used to model the FBP and Trp excitations. This seems to be particularly important to describe the proper orientation of the L_a transition dipole of Trp. In turn, they show that accounting for the shape of the interacting chromophores and the local environment in the excitonic couplings through the QM/MMPol model based on transition densities provides significantly more robust estimates of the transfer rates than the Förster ideal dipole approximation, where the molecular shape of the chromophores is ignored both in the calculation of the Coulomb term of the coupling and in the estimation of

screening effects. For example, using transition densities, we find a relatively good enantioselective quenching ratio among enantiomers at the ZINDO, CIS, and TD-DFT levels of theory, whereas only dipole–dipole estimates based on TD-DFT calculations provide results of similar quality.

Compared to photosynthetic pigment–protein complexes, where rather detailed theoretical studies on energy transfer have been presented in the last decade, we show that fluctuations in electronic couplings are much larger in a ligand–protein complex due to the considerable flexibility of the ligand, at least in the FBP/HSA complex. This finding makes it important to average coupling predictions over the thermal motions of the protein–ligand complex, whereas calculations based on a static structure, for instance a docking pose predicted for a ligand, are expected to introduce too much noise in the predictions. Regarding dielectric screening effects, we find large thermal fluctuations that translate into strong changes in the corresponding couplings. Although the average values found are quite consistent with the Förster simple $1/n^2$ factor, such fluctuations illustrate the fact that screening can vary significantly depending on the local environment and the donor–acceptor orientation, a factor neglected in Förster theory that can translate into wrong structural interpretations of FRET data that arise from changes in dielectric screening.

Overall, this study supports the potential of theoretical simulations to link variations in the fluorescence of ligand–protein complexes to subtle changes in the underlying structure originated by energy transfer, simulations that can be used to characterize ligand-binding events both from a static and a dynamic perspective at the molecular level.

■ ASSOCIATED CONTENT

📄 Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jpcc.7b00217.

Root-mean-squared deviation of atomic positions along the MD simulations of FBP/HSA complexes; tables of transition energies, transition dipoles, interchromophoric distances, orientation factors, screening factors, and electronic couplings for the model dyads and FBP/HSA complexes; force field parameters derived for (R)- and (S)-FBP and for (R,S)- and (S,S)-FBP–TrpMe (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: carles.curutchet@ub.edu.

ORCID

Carles Curutchet: 0000-0002-0070-1208

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We are grateful to the Consorci de Serveis Universitaris de Catalunya for providing access to computational resources. Financial support from the Agència de Gestió d'Ajuts Universitaris i de Recerca from Generalitat de Catalunya (GENCAT; SGR2014-1189) and from the Spanish Ministerio de Economía y Competitividad (MINECO; grants RYC2011-08918 and CTQ2012-36195) is acknowledged. C.C. is a Serra Hünter Fellow. S.P. is a fellow of the Ciências Sem Fronteiras

program of the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) de Brasil (246791/2012-8).

REFERENCES

- (1) *Protein Fluorescence*; Lakowicz, J. R., Ed.; Topics in Fluorescence Spectroscopy; Kluwer Academic Publishers: New York, 2002; Vol. 6.
- (2) Royer, C. A. Probing Protein Folding and Conformational Transitions with Fluorescence. *Chem. Rev.* **2006**, *106*, 1769–1784.
- (3) Michalet, X.; Weiss, S.; Jäger, M. Single-Molecule Fluorescence Studies of Protein Folding and Conformational Dynamics. *Chem. Rev.* **2006**, *106*, 1785–1813.
- (4) Schuler, B.; Eaton, W. A. Protein Folding Studied by Single-Molecule FRET. *Curr. Opin. Struct. Biol.* **2008**, *18*, 16–26.
- (5) Ciruela, F. Fluorescence-Based Methods in the Study of Protein-Protein Interactions in Living Cells. *Curr. Opin. Biotechnol.* **2008**, *19*, 338–343.
- (6) Yengo, C. M.; Berger, C. L. Fluorescence Anisotropy and Resonance Energy Transfer: Powerful Tools for Measuring Real Time Protein Dynamics in a Physiological Environment. *Curr. Opin. Pharmacol.* **2010**, *10*, 731–737.
- (7) Basak, S.; Chattopadhyay, K. Studies of Protein Folding and Dynamics Using Single Molecule Fluorescence Spectroscopy. *Phys. Chem. Chem. Phys.* **2014**, *16*, 11139.
- (8) Ghisaidoobe, A. B.; Chung, S. Intrinsic Tryptophan Fluorescence in the Detection and Analysis of Proteins: A Focus on Förster Resonance Energy Transfer Techniques. *Int. J. Mol. Sci.* **2014**, *15*, 22518–22538.
- (9) Banerjee, P. R.; Deniz, A. A. Shedding Light on Protein Folding Landscapes by Single-Molecule Fluorescence. *Chem. Soc. Rev.* **2014**, *43*, 1172–1188.
- (10) Czar, M. F.; Jockusch, R. A. Sensitive Probes of Protein Structure and Dynamics in Well-Controlled Environments: Combining Mass Spectrometry with Fluorescence Spectroscopy. *Curr. Opin. Struct. Biol.* **2015**, *34*, 123–134.
- (11) Callis, P. R. Binding Phenomena and Fluorescence Quenching. I: Descriptive Quantum Principles of Fluorescence Quenching Using a Supermolecular Approach. *J. Mol. Struct.* **2014**, *1077*, 14–21.
- (12) Callis, P. R. Binding Phenomena and Fluorescence Quenching. II: Photophysics of Aromatic Residues and Dependence of Fluorescence Spectra on Protein Conformation. *J. Mol. Struct.* **2014**, *1077*, 22–29.
- (13) Callis, P. R. Simulating Electrostatic Effects on Electronic Transitions in Proteins. *Mol. Simul.* **2015**, *41*, 190–204.
- (14) Vayá, I.; Bonancía, P.; Jiménez, M. C.; Markovitsi, D.; Gustavsson, T.; Miranda, M. A. Excited State Interactions between Flurbiprofen and Tryptophan in Drug–protein Complexes and in Model Dyads. Fluorescence Studies from the Femtosecond to the Nanosecond Time Domains. *Phys. Chem. Chem. Phys.* **2013**, *15*, 4727.
- (15) Beljonne, D.; Curutchet, C.; Scholes, G. D.; Silbey, R. J. Beyond Förster Resonance Energy Transfer in Biological and Nanoscale Systems. *J. Phys. Chem. B* **2009**, *113*, 6583–6599.
- (16) VanBeek, D. B.; Zwier, M. C.; Shorb, J. M.; Krueger, B. P. Fretting about FRET: Correlation between Kappa and R. *Biophys. J.* **2007**, *92*, 4168–4178.
- (17) Muñoz-Losa, A.; Curutchet, C.; Hartsell, L. R.; Krueger, B. P.; Mennucci, B. Fretting About FRET: Failure of the Ideal Dipole Approximation. *Biophys. J.* **2009**, *96*, 4779–4788.
- (18) Curutchet, C.; Kongsted, J.; Muñoz-Losa, A.; Hossein-Nejad, H.; Scholes, G. D.; Mennucci, B. Photosynthetic Light-Harvesting Is Tuned by the Heterogeneous Polarizable Environment of the Protein. *J. Am. Chem. Soc.* **2011**, *133*, 3078–3084.
- (19) Schuler, B.; Lipman, E. A.; Steinbach, P. J.; Kumke, M.; Eaton, W. A. Polyproline and the “Spectroscopic Ruler” Revisited with Single-Molecule Fluorescence. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 2754–2759.
- (20) Jean, J. M.; Krueger, B. P. Structural Fluctuations and Excitation Transfer between Adenine and 2-Aminopurine in Single-Stranded Deoxytrinucleotides. *J. Phys. Chem. B* **2006**, *110*, 2899–2909.
- (21) Beierlein, F. R.; Othersen, O. G.; Lanig, H.; Schneider, S.; Clark, T. Simulating FRET from Tryptophan: Is the Rotamer Model Correct? *J. Am. Chem. Soc.* **2006**, *128*, 5142–5152.
- (22) Merchant, K. A.; Best, R. B.; Louis, J. M.; Gopich, I. V.; Eaton, W. A. Characterizing the Unfolded States of Proteins Using Single-Molecule FRET Spectroscopy and Molecular Simulations. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 1528–1533.
- (23) Dolgih, E.; Roitberg, A. E.; Krause, J. L. Fluorescence Resonance Energy Transfer in Dye-Labeled DNA. *J. Photochem. Photobiol., A* **2007**, *190*, 321–327.
- (24) Wozniak, A. K.; Schroeder, G. F.; Grubmueller, H.; Seidel, C. A. M.; Oesterhelt, F. Single-Molecule FRET Measures Bends and Kinks in DNA. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 18337–18342.
- (25) Allen, L. R.; Paci, E. Orientational Averaging of Dye Molecules Attached to Proteins in Förster Resonance Energy Transfer Measurements: Insights from a Simulation Study. *J. Chem. Phys.* **2009**, *131*, No. 065101.
- (26) Dolgih, E.; Ortiz, W.; Kim, S.; Krueger, B. P.; Krause, J. L.; Roitberg, A. E. Theoretical Studies of Short Polyproline Systems: Recalibration of a Molecular Ruler. *J. Phys. Chem. A* **2009**, *113*, 4639–4646.
- (27) Stevens, J. A.; Link, J. J.; Kao, Y. T.; Zang, C.; Wang, L.; Zhong, D. Ultrafast Dynamics of Resonance Energy Transfer in Myoglobin: Probing Local Conformation Fluctuations. *J. Phys. Chem. B* **2010**, *114*, 1498–1505.
- (28) Callis, P. R. *Predicting Fluorescence Lifetimes and Spectra of Biopolymers*, 1st ed.; Elsevier Inc., 2011; Vol. 487.
- (29) Speelman, A. L.; Muñoz-Losa, A.; Hinkle, K. L.; VanBeek, D. B.; Mennucci, B.; Krueger, B. P. Using Molecular Dynamics and Quantum Mechanics Calculations to Model Fluorescence Observables. *J. Phys. Chem. A* **2011**, *115*, 3997–4008.
- (30) Beauchamp, K. A.; Ensign, D. L.; Das, R.; Pande, V. S. Quantitative Comparison of Villin Headpiece Subdomain Simulations and Triplet–Triplet Energy Transfer Experiments. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 12734–12739.
- (31) Kalinin, S.; Peulen, T.; Sindbert, S.; Rothwell, P. J.; Berger, S.; Restle, T.; Goody, R. S.; Gohlke, H.; Seidel, C. A. M. A Toolkit and Benchmark Study for FRET-Restrained High-Precision Structural Modeling. *Nat. Methods* **2012**, *9*, 1218–1227.
- (32) Stevens, J. A.; Link, J. J.; Zang, C.; Wang, L.; Zhong, D. Ultrafast Dynamics of Nonequilibrium Resonance Energy Transfer and Probing Globular Protein Flexibility of Myoglobin. *J. Phys. Chem. A* **2012**, *116*, 2610–2619.
- (33) Milas, P.; Gamari, B. D.; Parrot, L.; Krueger, B. P.; Rahmanseresht, S.; Moore, J.; Goldner, L. S. Indocyanine Dyes Approach Free Rotation at the 3' Terminus of A-RNA: A Comparison with the 5' Terminus and Consequences for Fluorescence Resonance Energy Transfer. *J. Phys. Chem. B* **2013**, *117*, 8649–8658.
- (34) Hoefling, M.; Grubmüller, H. In Silico FRET from Simulated Dye Dynamics. *Comput. Phys. Commun.* **2013**, *184*, 841–852.
- (35) Kellner, R.; Hofmann, H.; Barducci, A.; Wunderlich, B.; Nettels, D.; Schuler, B. Single-Molecule Spectroscopy Reveals Chaperone-Mediated Expansion of Substrate Protein. *Proc. Natl. Acad. Sci. U.S.A.* **2014**, *111*, 13355–13360.
- (36) Walczewska-Szewc, K.; Corry, B. Accounting for Dye Diffusion and Orientation When Relating FRET Measurements to Distances: Three Simple Computational Methods. *Phys. Chem. Chem. Phys.* **2014**, *16*, 12317–12326.
- (37) Reif, M. M.; Oostenbrink, C. Molecular Dynamics Simulation of Configurational Ensembles Compatible with Experimental FRET Efficiency Data through a Restraint on Instantaneous FRET Efficiencies. *J. Comput. Chem.* **2014**, *35*, 2319–2332.
- (38) Shoura, M. J.; Ranatunga, R. J. K. U.; Harris, S. A.; Nielsen, S. O.; Levene, S. D. Contribution of Fluorophore Dynamics and Solvation to Resonant Energy Transfer in Protein-DNA Complexes: A Molecular-Dynamics Study. *Biophys. J.* **2014**, *107*, 700–710.
- (39) Søndergaard, S.; Aznauryan, M.; Hastrup, E. K.; Schiøtt, B.; Birkedal, V.; Corry, B. Dynamics of Fluorescent Dyes Attached to G-

Quadruplex DNA and Their Effect on FRET Experiments. *ChemPhysChem* **2015**, *16*, 2562–2570.

(40) Kulesza, A.; Daly, S.; MacAleese, L.; Antoine, R.; Dugourd, P. Structural Exploration and Förster Theory Modeling for the Interpretation of Gas-Phase FRET Measurements: Chromophore-Grafted Amyloid- β Peptides. *J. Chem. Phys.* **2015**, *143*, No. 025101.

(41) Jurinovich, S.; Curutchet, C.; Mennucci, B. The Fenna-Matthews-Olson Protein Revisited: A Fully Polarizable (TD)DFT/MM Description. *ChemPhysChem* **2014**, *15*, 3194–3204.

(42) Curutchet, C.; Mennucci, B. Quantum Chemical Studies of Light Harvesting. *Chem. Rev.* **2017**, *117*, 294–343.

(43) Curry, S. Lessons from the Crystallographic Analysis of Small Molecule Binding to Human Serum Albumin. *Drug Metab. Pharmacokinet.* **2009**, *24*, 342–357.

(44) Zhivkova, Z. D. Studies on Drug – Human Serum Albumin Binding: The Current State of the Matter. *Curr. Pharm. Des.* **2015**, *21*, 1817–1830.

(45) Davies, N. M. Clinical Pharmacokinetics of Flurbiprofen and Its Enantiomers. *Clin. Pharmacokinet.* **1995**, *28*, 100–114.

(46) Vayá, I.; Bueno, C. J.; Jiménez, M. C.; Miranda, M. A. Use of Triplet Excited States for the Study of Drug Binding to Human and Bovine Serum Albumins. *ChemMedChem* **2006**, *1*, 1015–1020.

(47) Lammers, I.; Lhiaubet-Vallet, V.; Consuelo Jiménez, M.; Ariese, F.; Miranda, M. A.; Gooijer, C. Stereoselective Binding of Flurbiprofen Enantiomers and Their Methyl Esters to Human Serum Albumin Studied by Time-Resolved Phosphorescence. *Chirality* **2012**, *24*, 840–846.

(48) Deschamps-Labat, L.; Péhourcq, F.; Jagou, M.; Bannwarth, B. Relationship between Lipophilicity and Binding to Human Serum Albumin of Arylpropionic Acid Non-Steroidal Anti-Inflammatory Drugs. *J. Pharm. Biomed. Anal.* **1997**, *16*, 223–229.

(49) Østergaard, J.; Schou, C.; Larsen, C.; Heegaard, N. H. H. Effect of Dextran as a Run Buffer Additive in Drug–Protein Binding Studies Using Capillary Electrophoresis Frontal Analysis. *Anal. Chem.* **2003**, *75*, 207–214.

(50) Ghuman, J.; Zunszain, P. A.; Petitpas, I.; Bhattacharya, A. A.; Ottagiri, M.; Curry, S. Structural Basis of the Drug-Binding Specificity of Human Serum Albumin. *J. Mol. Biol.* **2005**, *353*, 38–52.

(51) Curutchet, C.; Muñoz-Losa, A.; Monti, S.; Kongsted, J.; Scholes, G. D.; Mennucci, B. Electronic Energy Transfer in Condensed Phase Studied by a Polarizable QM/MM Model. *J. Chem. Theory Comput.* **2009**, *5*, 1838–1848.

(52) Forster, T. Zwischenmolekulare Energiewanderung Und Fluoreszenz. *Ann. Phys.* **1948**, *437*, 55–75.

(53) Braslavsky, S. E.; Fron, E.; Rodríguez, H. B.; Román, E. S.; Scholes, G. D.; Schweitzer, G.; Valeur, B.; Wirz, J. Pitfalls and Limitations in the Practical Use of Förster's Theory of Resonance Energy Transfer. *Photochem. Photobiol. Sci.* **2008**, *7*, 1444.

(54) Chenu, A.; Scholes, G. D. Coherence in Energy Transfer and Photosynthesis. *Annu. Rev. Phys. Chem.* **2015**, *66*, 69–96.

(55) Iozzi, M. F.; Mennucci, B.; Tomasi, J.; Cammi, R. Excitation Energy Transfer (EET) between Molecules in Condensed Matter: A Novel Application of the Polarizable Continuum Model (PCM). *J. Chem. Phys.* **2004**, *120*, 7029–7040.

(56) Curutchet, C.; Mennucci, B. Toward a Molecular Scale Interpretation of Excitation Energy Transfer in Solvated Bichromophoric Systems. *J. Am. Chem. Soc.* **2005**, *127*, 16733–16744.

(57) Case, D. A.; Darden, T. A.; Cheatham, T. E., III; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Walker, R. C.; Zhang, W.; Merz, K. M. et al. *AMBER 12*; University of California: San Francisco, 2012.

(58) Olsson, M. H. M.; Søndergaard, C. R.; Rostkowski, M.; Jensen, J. H. PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical pKa Predictions. *J. Chem. Theory Comput.* **2011**, *7*, 525–537.

(59) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926–935.

(60) Grabuleda, X.; Jaime, C.; Kollman, P. A. Molecular Dynamics Simulation Studies of Liquid Acetonitrile: New Six-Site Model. *J. Comput. Chem.* **2000**, *21*, 901–908.

(61) Wang, J.; Cieplak, P.; Kollman, P. A. How Well Does a Restrained Electrostatic Potential (RESP) Model Perform in Calculating Conformational Energies of Organic and Biological Molecules? *J. Comput. Chem.* **2000**, *21*, 1049–1074.

(62) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of Multiple Amber Force Fields and Development of Improved Protein Backbone Parameters. *Proteins: Struct., Funct., Bioinf.* **2006**, *65*, 712–725.

(63) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A Consistent and Accurate Ab Initio Parametrization of Density Functional Dispersion Correction (DFT-D) for the 94 Elements H–Pu. *J. Chem. Phys.* **2010**, *132*, No. 154104.

(64) Bayly, C. I.; Cieplak, P.; Cornell, W.; Kollman, P. A. A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges: The RESP Model. *J. Phys. Chem.* **1993**, *97*, 10269–10280.

(65) Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. Universal Solvation Model Based on Solute Electron Density and on a Continuum Model of the Solvent Defined by the Bulk Dielectric Constant and Atomic Surface Tensions. *J. Phys. Chem. B* **2009**, *113*, 6378–6396.

(66) Alecu, I. M.; Zheng, J.; Zhao, Y.; Truhlar, D. G. Computational Thermochemistry: Scale Factor Databases and Scale Factors for Vibrational Frequencies Obtained from Electronic Model Chemistries. *J. Chem. Theory Comput.* **2010**, *6*, 2872–2887.

(67) Zerner, M. C. Semi Empirical Molecular Orbital Methods. In *Reviews of Computational Chemistry*; Lipkowitz, K. B., Boyd, D. B., Eds.; VCH: New York, 1991; Vol. 2, pp 313–366.

(68) Foresman, J. B.; Headgordon, M.; Pople, J. A.; Frisch, M. J. Toward a Systematic Molecular-Orbital Theory for Excited-States. *J. Phys. Chem.* **1992**, *96*, 135–149.

(69) Casida, M. E.; Huix-Rotllant, M. Progress in Time-Dependent Density-Functional Theory. *Annu. Rev. Phys. Chem.* **2012**, *63*, 287–323.

(70) Yanai, T.; Tew, D. P.; Handy, N. C. A New Hybrid Exchange–correlation Functional Using the Coulomb-Attenuating Method (CAM-B3LYP). *Chem. Phys. Lett.* **2004**, *393*, 51–57.

(71) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; et al. *Gaussian 09*, revision A.2; Gaussian, Inc.: Wallingford, CT, 2009.

(72) Wang, J.; Cieplak, P.; Li, J.; Hou, T.; Luo, R.; Duan, Y. Development of Polarizable Models for Molecular Mechanical Calculations I: Parameterization of Atomic Polarizability. *J. Phys. Chem. B* **2011**, *115*, 3091–3099.

(73) Wang, J.; Cieplak, P.; Li, J.; Wang, J.; Cai, Q.; Hsieh, M.; Lei, H.; Luo, R.; Duan, Y. Development of Polarizable Models for Molecular Mechanical Calculations II: Induced Dipole Models Significantly Improve Accuracy of Intermolecular Interaction Energies. *J. Phys. Chem. B* **2011**, *115*, 3100–3111.

(74) Wanko, M.; Hoffmann, M.; Strodel, P.; Koslowski, A.; Thiel, W.; Neese, F.; Frauenheim, T.; Elstner, M. Calculating Absorption Shifts for Retinal Proteins: Computational Challenges. *J. Phys. Chem. B* **2005**, *109*, 3606–3615.

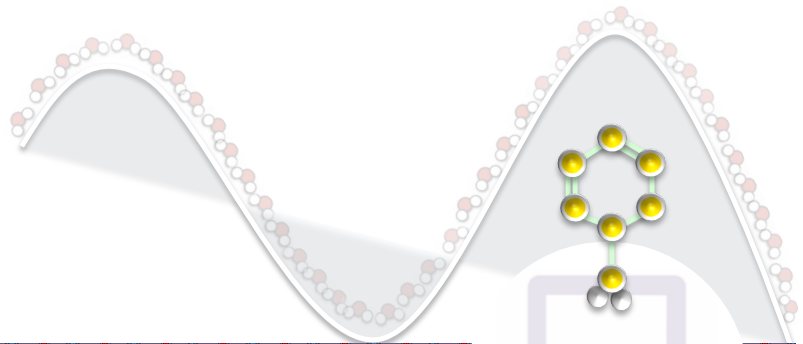
(75) Aghtar, M.; Kleinekathöfer, U.; Curutchet, C.; Mennucci, B. Impact of Electronic Fluctuations and Their Description on the Exciton Dynamics in the Light-Harvesting Complex PE545. *J. Phys. Chem. B* **2017**, *121*, 1330–1339.

(76) Caricato, M.; Trucks, G. W.; Frisch, M. J.; Wiberg, K. B. Oscillator Strength: How Does TDDFT Compare to EOM-CCSD? *J. Chem. Theory Comput.* **2011**, *7*, 456–466.

(77) Muñoz-Losa, A.; Curutchet, C.; Galván, I. F.; Mennucci, B. Quantum Mechanical Methods Applied to Excitation Energy Transfer: A Comparative Analysis on Excitation Energies and Electronic Couplings. *J. Chem. Phys.* **2008**, *129*, No. 034104.

(78) Knox, R. S. Dipole and Oscillator Strengths of Chromophores in Solution. *Photochem. Photobiol.* **2003**, *77*, 492–496.

(79) Knox, R. S.; Spring, B. Q. Dipole Strengths in the Chlorophylls. *Photochem. Photobiol.* **2003**, *77*, 497–501.



Capítulo
4

Discusión

4. DISCUSIÓN

En esta tesis doctoral hemos investigado el impacto de las fuerzas de polarización electrónica en interacciones intermoleculares de sistemas biológicos. En este sentido, discutiremos dos líneas de investigación complementarias. En la primera, se ha investigado el impacto de la polarización en interacciones a nivel de estado electrónico fundamental, donde se explora su impacto en interacciones catión- π -catión a tres cuerpos. En la segunda, se investigó el impacto de la polarización en interacciones entre estados excitados, es decir, en los acoplamientos electrónicos entre ellos, responsables de procesos tales como la transferencia de energía.

De manera general, en las dos líneas se ha explorado el impacto de la polarizabilidad, que es la tendencia relativa que tienen distribuciones de cargas tales como la nube electrónica de una molécula a distorsionarse de su estado normal a causa de un campo eléctrico externo, lo que puede ocurrir por la existencia de un dipolo o un ión cercano. En este sentido, la polarización electrónica describe la redistribución de la densidad electrónica de una molécula cuando esta experimenta el efecto de un campo eléctrico externo. Por ejemplo, en el modelo de dipolos inducidos que se ha aplicado en esta tesis, se establece que el medio es un dieléctrico en el cual los átomos dan lugar a dipolos puntuales caracterizados por tener un momento dipolar con un comportamiento isotrópico y lineal con el campo total, es decir, $\mu_{ind,i} = \alpha_i E_i$, donde el dipolo inducido $\mu_{ind,i}$ sobre el átomo i depende tanto del campo eléctrico aplicado E_i como de la polarizabilidad isotrópica del átomo, α_i .

En las últimas dos décadas hemos sido testigos de importantes desarrollos en la sofisticación de los campos de fuerzas,¹⁰⁶ permitiendo la identificación de distintas deficiencias en los modelos de cargas fijas usados habitualmente. La forma en que las moléculas responden a las condiciones del entorno, tales como la presencia de un disolvente polar o apolar, el pH o la concentración de iones, son difíciles de capturar sin tener en cuenta el efecto de la polarización.¹ En ausencia de términos explícitos de polarización, los modelos de cargas puntuales fijas sólo pueden describir tales efectos de una manera promediada, a través de la re-parametrización de los términos de energía por pares que tienen a su disposición. Esta descripción rudimentaria puede ser adecuada en condiciones suficientemente cercanas a aquellas para las que se han parametrizado, pero, a menudo, es poco transferible a diferentes fases o entornos diferentes, como las interfases.¹⁰⁷ Además, la naturaleza no aditiva de las interacciones de polarización impone una limitación fundamental

CAPÍTULO 4. DISCUSIÓN

a tales modelos aditivos. En consecuencia, tal y como se describe en la Sección 1.1.1.2, han surgido una gran variedad de estrategias a fin de incorporar la polarización de forma explícita en campos de fuerzas.^{1,48} Esencialmente estos modelos permiten i) el flujo de la densidad de carga entre átomos (modelo de fluctuación de carga) o ii) la aparición de dipolos inducidos localizados en los átomos (modelos de dipolos inducidos y del oscilador de Drude). Mientras que un modelo completo debería permitir ambos efectos, los campos de fuerzas polarizables desarrollados hasta ahora suelen incorporar uno u otro efecto, pero no ambos.

En esta tesis hemos abordado, en primer lugar, el estudio de interacciones catión- π -catión, ejemplo paradigmático de impacto de la polarización por la capacidad de los cationes de polarizar sistemas π así como por la posibilidad de tener efectos no aditivos importantes. Debido a la posible cancelación parcial del campo eléctrico generado por los dos cationes a cada lado del sistema π , además, tales interacciones suponen un gran reto para modelos clásicos de polarización, tales como el modelo de dipolos inducidos o de fluctuación de carga, este último siendo incapaz de describir la polarización fuera del plano del sistema π . Esta línea de estudio, por tanto, pretende, además de profundizar en este tipo de interacciones poco estudiadas, permitir la evaluación y eventual mejora de los modelos habitualmente usados para incorporar la polarización en campos de fuerzas.

En segundo lugar, esta tesis ha abordado el impacto de la polarización electrónica, mediante modelos híbridos QM/MMpol basados en el modelo de dipolos inducidos, en procesos de transferencia de energía. Esta segunda línea pretende investigar de forma crítica las aproximaciones subyacentes a la teoría de Förster en la descripción de tales procesos, y evaluar el impacto de la polarización en la derivación de información estructural a partir de datos espectroscópicos de FRET.

A continuación, se presenta la discusión de los principales resultados obtenidos en ambas líneas de investigación.

4.1 Impacto de la polarización en interacciones a nivel de estado electrónico fundamental

En la primera línea de investigación, hemos realizado un extenso análisis estadístico de las características estructurales, energéticas, de conservación y de ocurrencia de motivos catión- π -catión sobre un conjunto no redundante de estructuras de proteínas del PDB. Este estudio tiene como objetivo cuantificar y caracterizar la presencia de tales interacciones, así como su energética, usando una estrategia donde hemos combinado tanto datos experimentales como teóricos por medio de cálculos de alto nivel SCS-MP2/CBS combinados con cálculos de solvatación continuos QM/PCM y MM-GBSA, estos últimos usados para estimar el impacto del disolvente o el entorno proteico en las energías de interacción. Los cálculos de energías de interacción han permitido profundizar en la comprensión de los efectos de polarización en este caso especial de interacciones, donde los efectos de polarización son considerablemente complejos por su naturaleza de interacción a tres cuerpos y que por tanto pueden dar lugar a efectos no aditivos. Tales interacciones representan un reto para los campos de fuerzas clásicos actuales, incluidos los que incorporan efectos de polarización de forma explícita. Este estudio demuestra que en general la contribución no aditiva de polarización es muy importante en este tipo de sistemas y que estas interacciones, poco estudiadas, son bastante comunes en las estructuras de proteínas presentes en el PDB.

4.1.1 Análisis estructural de los motivos catión- π -catión en proteínas

Nuestros resultados demuestran que las interacciones catión- π -catión son de hecho bastante comunes en proteínas, ya que cerca del 7% de las estructuras investigadas poseen una o más interacciones catión- π -catión. Entre las interacciones catión- π -catión identificadas en el PDB, la frecuencia de aminoácidos aromáticos encontrados (Tyr > Phe > Trp) es similar a la frecuencia observada para complejos catión- π a dos cuerpos. Debido a que el aminoácido Trp es menos frecuente en proteínas que Tyr o Phe, el número comparativamente grande de interacciones catión- π -catión identificadas sugieren una preferencia por Trp sobre Tyr o Phe. Por otra parte, hemos observado que los aminoácidos aromáticos (Trp, Tyr, Phe) muestran una clara preferencia para interactuar con los aminoácidos cargados en las interacciones catión- π -catión en lugar de cationes metálicos monovalentes. Esta tendencia presumiblemente se puede atribuir a las energías libres de hidratación de los cationes metálicos, mayores que las de los grupos amonio y guanidinio, y a la accesibilidad a un mayor volumen de los

CAPÍTULO 4. DISCUSIÓN

cationes metálicas en comparación con la disposición pre-organizada de los residuos cargados en la estructura terciaria de la proteína. Este resultado, sin embargo, también se debe al número limitado de estructuras de proteína con cationes metálicas en el PDB. En general, si ponderamos el número de interacciones encontradas por su abundancia natural, nuestros resultados sugieren una preferencia para establecer las interacciones catión- π -catión con Trp, seguido por Tyr y Phe. Además, si tenemos en cuenta los aminoácidos cargados situados en cada lado del anillo aromático, el motivo dominante contiene Lys y Arg, encontrándose estos dos residuos con una frecuencia similar en las posibles interacciones consideradas. Una vez más, este hallazgo puede explicarse por la abundancia natural similar de Lys y Arg.

Otro punto importante a discutir en este análisis estructural es el grado de conservación de los aminoácidos involucrados en las interacciones catión- π -catión. Los aminoácidos presentes en las interacciones que implican Trp y Phe parecen estar fuertemente conservados en la mayoría de los casos investigados. Por otro lado, las que están presentes en interacciones que implican Tyr tienden a estar significativamente menos conservados. En general, los aminoácidos presentes en las interacciones catión- π -catión detectadas en el PDB tienden a estar altamente conservados, sugiriendo un papel estructural o funcional relevante. En este caso, los residuos en interacciones que poseen Trp, y, en menor medida, Phe, parecen estar más fuertemente conservados en comparación con complejos que poseen Tyr.

4.1.2 Análisis energético de los motivos catión- π -catión en proteínas

Mediante cálculos QM de alto nivel realizados sobre el subconjunto de interacciones catión- π -catión más relevantes encontrados, hemos observado que la contribución no aditiva a tres cuerpos es significativa, siendo esta de hasta 7 kcal/mol, lo que implica que las interacciones catión- π -catión constituyen presumiblemente un desafío aún mayor para campos de fuerzas clásicos en comparación con interacciones catión- π a dos cuerpos. Desde un punto de vista computacional, la contribución significativa de los términos no aditivos a tres cuerpos cuestionan la idoneidad de los campos de fuerzas aditivos estándar para describir los motivos catión- π -catión en simulaciones moleculares.

Las energías de interacción obtenidas para los complejos que contienen Trp, Tyr y Phe a nivel SCS-MP2/CBS en fase gas, además, muestran que tales energías son fuertemente repulsivas, hecho que indica que sus energías están dominadas por la repulsión entre los dos cationes en fase gas, término que alcanza valores en el rango ~ 30 -50 kcal/mol. Por otra parte,

la distribución de energías de interacción catión- π calculadas para los dímeros que conforman los complejos muestran grandes diferencias dependiendo de la naturaleza específica del anillo aromático implicado en el complejo. Por ejemplo, la distribución de energías de interacción de dímeros Lys-Phe y Arg-Phe es considerablemente más estrecha que la de dímeros que implican Trp o Tyr, probablemente debido a la mayor simetría del anillo bencénico de la Phe. En este sentido, es interesante recalcar que las distribuciones de energías Trp-catión, donde los cationes pueden interactuar tanto con el anillo de cinco como con el de seis del grupo indol, presentan las distribuciones más amplias. Un detalle destacable de este análisis, por otra parte, es el hecho de que, independientemente del aminoácido aromático involucrado, las energías de interacción más fuertes en la distribución siempre surgen de pares catión- π que implican a la Lys en lugar de la Arg. Tal y como se ha comentado, debido a que la repulsión catión-catión domina las interacciones, las energías de interacción de los trímeros son fuertemente positivas en vacío, con valores de 21 a 35 kcal/mol. Tales energías, sin embargo, no coinciden con la suma de las energías de interacción de los pares, es decir, la no aditividad juega un papel importante en estas interacciones, y la contribución a tres cuerpos asciende normalmente a 3-5 kcal/mol llegando hasta 7 kcal/mol en algunos complejos.

Por otra parte, la fuerza de las interacciones catión- π -catión se espera que estén moduladas fuertemente en un ambiente proteico, donde la repulsión catión-catión estará apantallada. Es decir, el entorno local puede afectar significativamente a estas energías mediante un apantallamiento de las fuerzas electrostáticas en juego. Nuestros cálculos QM/MST en disolución indican efectivamente que el ambiente tiene un fuerte impacto en las energías de interacción, estabilizándolas en 14-28 kcal/mol en *n*-octanol y 15-31 kcal/mol en agua. En *n*-octanol, las energías de interacción resultantes son positivas en todos los casos, mientras que la solvatación en agua conduce a interacciones atractivas en algunos de los motivos catión- π -catión estudiados, aunque sólo en hasta -1.3 kcal/mol. Por otra parte, las energías libres de transferencia de agua a proteína calculadas a nivel MM/GBSA para los complejos en su entorno nativo muestran un importante impacto del entorno proteico en las energías de interacción, con valores que van desde -45 a 30 kcal/mol. En la mayoría de las interacciones investigadas, el entorno proteico estabiliza la interacción, mientras que en una minoría su contribución es positiva. En este sentido, la presencia de aminoácidos aniónicos cercanos puede estabilizar considerablemente los motivos catión- π -catión estableciendo interacciones electrostáticas favorables con los residuos de Lys y Arg y atenuando así la repulsión entre los cationes.

CAPÍTULO 4. DISCUSIÓN

Por lo tanto, un ambiente apolar (*n*-octanol) o polar (agua) apantalla fuertemente la repulsión catión-catión, aunque en la mayoría de los casos la interacción aún sea repulsiva. El impacto del ambiente proteico es, sin embargo, más fuerte, y conduce a una interacción atractiva en 64% de los complejos investigados. Este resultado, juntamente con el elevado grado de conservación de los aminoácidos implicados en estas interacciones, sugiere que éstas pueden tener un papel estabilizador de la estructura terciaria de proteínas. Esta posibilidad está respaldada por el alto grado de conservación observado para los residuos que forman estas interacciones, en particular para aquellas que implican Trp y Phe. Tal papel estabilizador, por ejemplo, ha sido demostrado recientemente para una pequeña proteína que emula una red de interacciones catión- π encontradas en las proteínas que contienen el dominio fibronectina tipo III.⁸ En esta tesis, por ejemplo, hemos estudiado el receptor de la hormona del crecimiento humano (PDB:1A22), la cual presenta una red catión- π extendida donde se establecen tres interacciones catión- π -catión consecutivas (Lys-Trp-Arg, Arg-Phe-Arg y Arg-Tyr-Lys), para las cuales hemos estimado energías de interacción de -16.7, -1.0 y -7.8 kcal/mol, apoyando así la idea de que redes catión- π extendidas pueden contribuir a estabilizar una estructura terciaria.

El hecho de que las energías de interacción tengan una componente no aditiva considerable supone un problema para la descripción de motivos catión- π -catión en simulaciones de biomoléculas basadas en campos de fuerzas aditivos. Distintos grupos, por ejemplo, han propuesto modificaciones *ad hoc* tales como la introducción de un potencial de corto alcance,⁵ o la modificación de los parámetros de dispersión-repulsión,³¹ a fin de mejorar la descripción de interacciones catión- π en campos de fuerzas aditivos. Aunque estas modificaciones permiten mejorar la descripción de contactos catión- π a dos cuerpos de forma simple y eficiente, el tratamiento implícito de los efectos de polarización en tales formalismos se vuelve problemático en interacciones catión- π -catión, dado que la presencia de los dos cationes a cada lado del anillo aromático anularán parcialmente la contribución de inducción. Este problema de hecho refleja la ausencia de contribuciones no aditivas, como las estimadas en este estudio para los motivos catión-catión, en un campo de fuerzas aditivo basado en la interacción de pares.

La alternativa natural a fin de mejorar la descripción de motivos catión- π -catión es utilizar un campo de fuerzas explícitamente polarizable,^{1,2} como los basados en los osciladores de Drude o dipolos inducidos. Tales modelos, por ejemplo, se ha demostrado que permiten mejorar la descripción de interacciones catión- π .³¹ Cuando dos cationes se

aproximan al anillo aromático desde lados opuestos, tal y como sucede en los motivos catión- π -catión, lo más probable, sin embargo, es que los osciladores de Drude se reorienten en el plano de la nube de electrones π , o que los dipolos inducidos desaparezcan parcialmente, así que su capacidad para describir tales interacciones tampoco está garantizada. Un modelo riguroso para explicar los efectos de polarización no aditivos necesita probablemente de la inclusión tanto de flujos de cargas como polarizabilidades dipolares, pero el desarrollo de tales modelos para macromoléculas solvatadas es aún más complejo que el desarrollo de campos de fuerzas que solo incluyen uno de estos efectos.

En este sentido, actualmente estamos explorando la capacidad de los campos de fuerzas polarizables no aditivos, por ejemplo Drude³⁴ o AMOEBA,^{33,108} así como otros campos de fuerzas clásicos aditivos comúnmente usados en simulaciones de biomoléculas, para describir la energía de las interacciones catión- π -catión, y si éstos son realmente capaces de describir la no aditividad. Además, se ha realizado una descomposición de las energías de interacción a nivel SAPT.⁸⁴ Nuestros resultados preliminares muestran, por ejemplo, que la contribución no aditiva se origina en un 90% por efectos de inducción, siendo los efectos de dispersión y intercambio mucho menores. Además, muestran que el campo de fuerzas polarizable Drude mejora la predicción de interacciones catión- π en comparación con el campo de fuerzas aditivo CHARMM, en línea con resultados recientes de Khan y colaboradores.³¹ Sin embargo, paradójicamente la inclusión de polarización mediante el modelo de Drude empeora la descripción de las interacciones catión- π -catión, debido a una incorrecta descripción del efecto a tres cuerpos. Estos hallazgos, que serán discutidos en futuros trabajos, sugieren que es necesario el desarrollo de campos de fuerzas polarizables más rigurosos en el futuro para el estudio de biomoléculas.

4.2 Impacto de la polarización en interacciones entre estados excitados

En esta segunda línea de investigación se ha investigado el impacto de la polarización electrónica del entorno en interacciones entre estados excitados, en particular, en los acoplamientos electrónico responsables de procesos de transferencia de energía de excitación. El impacto de la polarización del entorno se ha explorado mediante la metodología MD-QM/MMPol desarrollada en el grupo, que combina simulaciones de dinámica molecular (MD) con cálculos de química cuántica/mecánica molecular (QM/MM) basados en campos de fuerzas polarizables para el estudio de procesos de transferencia de energía. La técnica FRET, basada en este fenómeno, es ampliamente usada en la actualidad, y representa una de las

CAPÍTULO 4. DISCUSIÓN

herramientas más potentes para la determinación de distancias en sistemas biológicos, por ejemplo, para observar la dinámica estructural de proteínas o su cinética de plegamiento. A partir de medidas de fluorescencia que permiten cuantificar la eficiencia de FRET, tales distancias se derivan aplicando la teoría de Förster, que asume que la dinámica de FRET sigue una dependencia R^{-6} con la distancia entre el dador y el aceptor, relación basada en asumir una interacción dipolo-dipolo y un factor de apantallamiento dieléctrico constante igual a $1/n^2$ dependiente del índice de refracción del medio, tal y como se ha explicado en la Sección 1.3.2. Esta teoría, pues, asocia cambios en la eficiencia de FRET *únicamente* a cambios de distancias dador-aceptor, mientras que ignora variaciones relacionadas con los distintos entornos locales que conforman las propiedades dieléctricas heterogéneas de una proteína. En comparación con su impacto en propiedades tales como desplazamientos solvatocrómicos, donde la polarización tiene un papel menor comparado con efectos puramente electrostáticos, ésta juega un papel fundamental en la modulación de los acoplamientos electrónicos entre estados excitados que median reacciones FRET, pues es capaz de apantallarlos induciendo variaciones de hasta 4 veces en la velocidad de transferencia.^{12,13,15}

El factor de apantallamiento $1/n^2$ de Förster, ampliamente usado en la interpretación de fenómenos FRET, representa pues una aproximación importante ya que asume que la polarización electrónica del entorno juega un papel uniforme en tales fenómenos, independientemente de las moléculas involucradas, su disposición relativa, así como las propiedades locales del entorno que las rodea. En esta línea de investigación, en estas tesis hemos explorado el impacto que la naturaleza polarizable heterogénea juega en la interpretación de fenómenos FRET en sistemas proteína-ligando, donde cambios sutiles del modo de unión pueden llevar a cambios significativos en la velocidad de FRET que potencialmente no sean capturados por la simple relación R^{-6} subyacente a la teoría de Förster. Esta investigación, además, pretende explorar el potencial de la metodología MD-QM/MM polarizable desarrollada en el grupo para caracterizar uniones ligando-receptor mediante el análisis de datos FRET, con el objetivo de ayudar al descubrimiento de nuevos sitios de unión. Tal aplicación, por ejemplo, es relevante en la caracterización de sitios de unión crípticos que aparecen transitoriamente por un cambio conformacional en la estructura de la proteína, y que por tanto son difíciles de descubrir a partir de estructuras cristal resueltas por rayos X.

En la tesis, en particular, hemos estudiado el complejo formado por la albúmina sérica humana (HSA) y el flurbiprofeno (FBP), donde se ha observado que distintos enantiómeros del flurbiprofeno sufren un *quenching* distinto en su fluorescencia cuando se unen a la HSA, *quenching* que ha sido atribuida a una eventual transferencia de energía enantioselectiva entre el FBP y el Trp214 de la HSA. También hemos investigado un sistema modelo FBP-TrpMe, formado por la unión del FBP al éster metílico del Trp, donde se ha observado un *quenching* parecido que, sin embargo, muestra una enantioselectividad opuesta.¹⁰⁹

4.2.1 Sistema modelo FBP-TrpMe

En primer lugar, se ha estudiado la hipótesis de que un proceso de transferencia de energía explique el *quenching* dinámico enantioselectivo observado en el sistema modelo FBP-TrpMe. Las simulaciones de MD realizadas para investigar las preferencias conformacionales adoptadas por el sistema en acetonitrilo indicaron que tanto el enantiómero (R,S) como el (S,S) adoptan una conformación extendida la mayor parte del tiempo. A fin de evaluar si estas conformaciones obtenidas a partir de la simulación eran realistas, determinamos las conformaciones más estables del sistema optimizando un conjunto de geometrías exploradas en la simulación de MD a nivel DFT, donde los efectos de solvatación se incluyeron mediante el modelo de solvatación continua SMD. En contraste con los resultados de MD, las energías libres conformacionales estimadas a nivel DFT-SMD indicaron que varias conformaciones plegadas, donde los anillos aromáticos del FBP y Trp están en estrecho contacto, eran las más estables y determinan las conformaciones más pobladas del sistema FBP-TrpMe en disolución. Este resultado indica que el campo de fuerzas clásico utilizado en la simulación de MD no es adecuado para el estudio de este sistema, ya que sobreestabiliza las estructuras extendidas, problema debido probablemente a una descripción errónea de las interacciones de dispersión entre anillos aromáticos apilados o a una descripción no balanceada de los efectos de solvatación entre conformaciones apiladas y extendidas. Por esta razón, se prefirió investigar las propiedades de transferencia de energía calculadas a partir de los conformeros obtenidos a nivel DFT-SMD, promediando las propiedades obtenidas mediante las poblaciones de Boltzmann de cada conformero.

Respecto a la estimación de las energías de transición de los cromóforos, y de acuerdo con sus espectros de absorción experimentales, nuestros resultados indican que ambos enantiómeros tienen energías de transición y dipolos de transición muy similares. Este primer resultado indica pues que los cambios en las propiedades de transferencia de energía se originan en variaciones del acoplamiento electrónico V^2 en lugar del factor de superposición

CAPÍTULO 4. DISCUSIÓN

espectral J . En este sentido, la orientación de los cromóforos conduce a una disposición casi ortogonal de los dipolos de transición del estado L_a del Trp y del estado $\pi \rightarrow \pi^*$ del FBP, hecho que origina un acoplamiento relativamente pequeño y extremadamente sensible a la orientación precisa entre las unidades cromofóricas en el sistema. Además, tal disposición implica que, presumiblemente, una descripción a nivel de dipolos puntuales de las transiciones involucradas conlleve desviaciones importantes en los valores de acoplamiento estimados. Nuestros resultados muestran que, como era de esperar, la aproximación dipolo-dipolo de Förster introduce desviaciones significativas en las estimaciones del acoplamiento en comparación con los valores más rigurosos derivados a partir de las densidades de transición.

Si nos centramos en las estimaciones más precisas de acoplamientos obtenidas a partir del modelo QM/PCM basado en densidades de transición, curiosamente sólo los datos TD-DFT basados en los modelos D-B-A completos (que incluyen el puente que une D y A en los cálculos) son capaces de describir correctamente la enantioselectividad en los que *quenchings* dinámicos observados. La adopción de una descripción más simple que incluye solamente los anillos aromáticos del Trp o el FBP, o la estimación de las propiedades mediante métodos menos rigurosos tales como CIS o ZINDO, en efecto predicen una transferencia ~ 2 -3 veces más rápida para el enantiómero (S,S), en contraste con la enantioselectividad opuesta observada. El análisis de los resultados TD-DFT/PCM más precisos obtenidos para los sistemas D-B-A, que reflejan correctamente la enantioselectividad en el proceso de transferencia, indican que ésta se origina por el término coulombico de interacción directa entre los cromóforos, mientras que los factores de apantallamientos estimados para los dos enantiómeros son similares en ambos sistemas. Es decir, la polarización del disolvente (acetonitrilo) no contribuye a la transferencia de energía enantioselectiva observada.

Los resultados encontrados para el sistema modelo, pues, apoyan fuertemente la hipótesis de un proceso de transferencia de energía desde el estado 1FBP^* al estado L_a del Trp como origen del *quenching* dinámico observado. Además, muestran que es fundamental una descripción adecuada de la orientación precisa del dipolo de transición correspondiente al estado L_a de Trp, que se ve ligeramente modificado si se incluye el puente que une al Trp con el FBP en los cálculos, así como una descripción rigurosa del acoplamiento no limitada a la aproximación dipolo-dipolo. Además, muestran que el equilibrio entre los distintos conformeros del enantiómero (S,S) son la clave que explica la enantioselectividad.

A partir de los factores de solapamiento espectral estimados con los espectros de emisión del FBP y de absorción del Trp en acetonitrilo, también hemos calculado las velocidades de FRET absolutas para cada enantiómero. En este sentido, a pesar de que las simulaciones TD-DFT son capaces de predecir correctamente la enantioselectividad, las velocidades estimadas son demasiado rápidas comparadas con las velocidades de *quenching* experimentales. Esta discrepancia podría tener su origen en la no inclusión de las fluctuaciones térmicas del sistema en los valores de los acoplamientos, muy sensibles a la geometría del sistema debido a la disposición casi ortogonal de los dipolos de transición del FBP y del Trp antes comentada. Por otro lado, el hecho de que los cromóforos estén a una distancia de contacto y experimenten un fuerte acoplamiento entre ellos podría conducir a posibles desviaciones del régimen de acoplamiento débil asumido en la teoría de Förster. En este sentido, sin embargo, la constatación de que el espectro de absorción del sistema FBP-TrpMe sea aproximadamente la suma de los espectros del FBP y del TrpMe apoya el supuesto de acoplamiento débil.

4.2.2 Complejo FBP-HSA

El transporte y la distribución de drogas en el cuerpo es un mecanismo complejo regido por muchos procesos diferentes. Un papel importante es el desempeñado por proteínas de transporte específicas, responsables de la fijación y liberación de los fármacos. Por lo tanto, es de gran interés obtener una mejor comprensión de los factores que rigen la estructura y la dinámica de los complejos entre fármacos y tales proteínas de transporte. Entre las diversas técnicas utilizadas para estudiar la interacción fármaco-proteína, la espectroscopia de fluorescencia es frecuentemente utilizada cuando la proteína contiene algún Trp, debido a la alta sensibilidad de su fluorescencia respecto al entorno local. Si además se usan técnicas de fluorescencia resuelta en el tiempo, es posible obtener información sobre las interacciones a nivel de estado excitado que se puedan producir entre el fármaco y la proteína, interacciones que reflejan también la dinámica estructural del complejo.

En esta tesis hemos simulado las interacciones a nivel de estados excitados entre el FBP, un fármaco antiinflamatorio no esteroideo quiral, y el Trp214 presente en la HSA, principal proteína de transporte en el cuerpo humano. El mecanismo de unión de del FBP con la albumina sérica humana (HSA) es relevante para una mejor comprensión del transporte y biodistribución de este fármaco en el organismo. A fin de interpretar los *quenches* dinámicos enantioselectivos observados para el complejo FBP-HSA, hemos combinado

CAPÍTULO 4. DISCUSIÓN

técnicas de MD clásicas con la metodología QM/MM polarizable desarrollada en el grupo y descrita en las Secciones 1.1.2 y 1.3.3 de esta tesis.

En general, nuestros resultados MD-QM/MMpol calculados a nivel TD-DFT describen una enantioselectividad en el proceso FRET postulado para explicar los *quenchtings* en excelente acuerdo con los experimentos, hecho que nos ha permitido validar los modos de unión propuestos. Por otra parte, desde un punto de vista de aplicación futura de la estrategia MD-QM/MMPol al descubrimiento de sitios de unión, hemos observado que la aproximación dipolo-dipolo subyacente a la teoría de Förster solamente describe correctamente la enantioselectividad de FRET cuando los dipolos de transición se obtienen a nivel TD-DFT, mientras que el método QM/MMPol basado en densidades de transición también da resultados precisos a nivel ZINDO y CIS, mostrando por tanto una mayor robustez.

En este sentido, para explicar el *quenching* dinámico del FBP, una de las cuestiones discutidas es la localización del fármaco en el complejo HSA. Nuestros resultados indican que el *quenching* dinámico observado proviene de una pequeña fracción del fármaco unido a un sitio secundario de la HSA localizado en la interfase entre los subdominios IIA y IIB, mientras que la enantioselectividad surge de la mayor flexibilidad del enantiómero (S) del FBP en el sitio de unión respecto a la forma (R). Es decir, la fracción mayoritaria de FBP que se une al sitio de unión II (sitio de unión de benzodiazepinas) no contribuye al *quenching* observado. Respecto al sitio de unión secundario, además, cabe señalar que este ha sido identificado frecuentemente con el sitio I de la HSA (sitio de unión de warfarina). Sin embargo, la estructura cristal de ibuprofeno unido a HSA,¹¹⁰ un fármaco con estructura y propiedades de unión a la HSA muy parecidas respecto al FBP, sugiere que dicho sitio secundario no es el sitio I, sino otro sitio localizado entre los subdominios IIA y IIB de la HSA. Nuestras simulaciones han demostrado que efectivamente el FBP se une a este sitio secundario, y no al sitio I, además de la unión mayoritaria bien conocida al sitio II. Esto explica, por ejemplo, el hecho de que sólo se haya observado un *quenching* parcial del ~50% y no completo en la fosforescencia de la HSA, ya que el FBP en ese sitio secundario no está en contacto directo con el Trp214 como estaría en el sitio I, pero a una distancia de ~10 Å. En este estudio, pues, nos hemos centrado en la unión del FBP al sitio secundario, que se postula como origen del *quenching* dinámico. Una estimación preliminar de las velocidades de transferencia de energía del FBP al Trp214, por ejemplo, indican tiempos de FRET del orden de ~500-800ps para este sitio secundario, y mucho más lentos, de ~30-150ns, cuando el FBP está situado en el sitio II de unión mayoritario, en línea con la distancia mucho más corta

de ~ 10 Å entre FBP y Trp214 en el sitio secundario en comparación con ~ 20 Å en el sitio II.

Las simulaciones de MD del complejo, por otra parte, indican que el ligando FBP es bastante flexible en este sitio de unión secundario. En particular, el enantiómero (S)-FBP a menudo cambia su configuración, mostrando valores de desviación cuadrática media (RMSD) respecto a la posición inicial de los átomos de ligando de ~ 4 -5 Å, aunque rápidamente regresa al modo de unión estable caracterizado por valores de RMSD de ~ 2 -3 Å. Tales transiciones también se observan para el enantiómero (R), aunque en menor medida. En general, las trayectorias de las simulaciones son estables y el modo de unión predicho para ambos enantiómeros es muy similar. Por otra parte, la orientación mutua entre las unidades cromofóricas del FBP y el Trp214 son también comparables, siendo por lo tanto difícil explicar el *quenching* caracterizado por una velocidad un $\sim 30\%$ más rápida observado para el (S)-FBP en comparación con el enantiómero (R) a primera vista. Un análisis más detallado de las simulaciones, sin embargo, sugiere que la flexibilidad de los ligandos juega un papel importante en la enantioselectividad observada.

En este sentido, el análisis de las fluctuaciones de los acoplamientos electrónicos calculados a lo largo de las simulaciones MD a nivel QM/MMpol TD-CAMB3LYP/6-31G(d) indica que la flexibilidad del sistema conlleva fluctuaciones muy grandes alrededor de los valores medios de acoplamiento. La magnitud de estas fluctuaciones, por ejemplo, es drásticamente mayor a la encontrada anteriormente en complejos pigmento-proteína fotosintéticos,^{10,105} donde es común evaluar las interacciones pigmento-pigmento a partir de la estructura cristal. En este caso, nuestros resultados muestran que tal estrategia no está justificada para un complejo fármaco-proteína donde el ligando tiene cierta flexibilidad, aunque este aspecto obviamente dependerá de cada complejo ligando-proteína, y en particular, de sus propiedades de flexibilidad. En el complejo FBP-HSA, es importante tener en cuenta las fluctuaciones térmicas en los acoplamientos a fin de obtener un valor promedio correcto.

Cabe remarcar aquí, sin embargo, que nuestras estimaciones de las velocidad de FRET a través de la expresión de Förster (Ec. 28) se basan en un único valor de acoplamiento, aunque este está promediado a lo largo de la simulación de MD. Es decir, no consideramos explícitamente las fluctuaciones de los acoplamientos en la predicción de la velocidad de FRET. Recientemente, el grupo ha mostrado que la inclusión explícita de tales fluctuaciones en el estudio de procesos FRET en complejos fotosintéticos conlleva cambios muy pequeños

CAPÍTULO 4. DISCUSIÓN

en tales predicciones.¹¹¹

Si retornamos al análisis de las fluctuaciones inducidas por la flexibilidad del FBP, cabe destacar que nuestros resultados indican que el enantiómero (S), en efecto, experimenta un mayor acoplamiento promedio, responsable del *quenching* enantioselectivo observado, efecto originado principalmente por el mayor grado de flexibilidad del enantiómero (S) del FBP comparado con el (R). En concreto, las características dinámicas del modo de unión explican el aumento del 30% en la velocidad de FRET observada en el complejo (S)-FBP/HSA, mientras que las características estáticas (orientación y separación promedio respecto al Trp214), contribuyen de forma menor, más o menos la mitad.

Finalmente, si analizamos las energías de transición estimadas para cada enantiómero, de nuevo estas sugieren, como para el sistema modelo FBP-TrpMe, que la enantioselectividad se debe al factor electrónico V^2 , ya que el factor de solapamiento espectral J no cambia, hecho corroborado por los espectros de absorción experimentales correspondientes a cada enantiómero. Si comparamos las velocidades estimadas a nivel ZINDO, CIS y TD-DFT, nuestros resultados indican una transferencia de energía un 89, 77 y 54%, respectivamente, más rápida para el enantiómero (S). En este sentido, y de forma similar a los resultados obtenidos para el modelo FBP-TrpMe, de nuevo TD-DFT se comporta como el método más preciso, dando un resultado de enantioselectividad del 54% muy cercano al valor experimental del 30%. La superioridad de las predicciones TD-DFT se amplifican aún más cuando los acoplamientos se basan en la aproximación dipolo-dipolo de Förster, ya que en este caso sólo los datos TD-DFT son capaces de dar una estimación razonable, del 38% mientras que los dipolos calculados a nivel CIS y ZINDO predicen una enantioselectividad mucho mayor del 214 y 489%. La conclusión importante de estos datos es que los acoplamientos dipolo-dipolo asumidos en la teoría de Förster son muy sensibles a pequeños cambios en la orientación real de los dipolos de transición, mientras que las estimaciones basadas en densidades de transición son notablemente más robustas a tales cambios.

Globalmente, nuestros resultados apoyan firmemente la hipótesis de que un proceso FRET que involucra a la fracción minoritaria de FBP unido al sitio secundario de HSA origina el *quenching* dinámico enantioselectivo observado. Además, apoyan firmemente los modos de unión simulados para ambos enantiómeros del FBP en este sitio de la HSA. Tales modos de unión son bastante similares, siendo el cambio más evidente un desplazamiento del aminoácido Phe206 debido a la diferente orientación del grupo carboxílico del FBP. Tales

modos de unión similares, sin embargo, se traducen en una flexibilidad significativamente mayor del enantiómero (S) de FBP, y ésta parece ser la principal razón del *quenching* dinámico enantioselectivo observado.

Desde el punto de vista de aplicaciones futuras, la estrategia MD-QM/MMPol ha demostrado que es capaz de describir la enantioselectividad observada en los complejos FBP/HSA y en los sistemas FBP-TrpMe originada por cambios sutiles en la estructura y disposición relativa de los cromóforos. Tales cambios incluyendo las características dinámicas de tal disposición relacionadas con la flexibilidad de los cromóforos, aspecto clave a fin de comprender la enantioselectividad observada en el complejo FBP/HSA. Sin embargo, también muestra que, en general, es necesario usar una descripción precisa basada en densidades de transición obtenidas a nivel TD-DFT para modelar las excitaciones del FBP y el Trp, aspecto particularmente importante para describir la orientación correcta del dipolo de transición del estado L_a del Trp.

Respecto al impacto de la polarización heterogénea del entorno, nuestros resultados muestran que no contribuye a la enantioselectividad observada en este caso, ya que los factores de apantallamiento promedio estimados para los dos enantiómeros son muy parecidos. Sin embargo, los cálculos realizados a lo largo de las trayectorias de MD muestran fluctuaciones muy importantes en este factor de apantallamiento, hecho que demuestra las limitaciones importantes que presenta el factor de apantallamiento constante de Förster en este contexto.

En conclusión, nuestros resultados apoyan el potencial de simulaciones teóricas con el fin de relacionar variaciones en la fluorescencia de complejos ligando-proteína originados por procesos FRET entre ligandos y triptófanos con cambios sutiles en la estructura del modo de unión. Esta estrategia se puede utilizar para caracterizar eventos de unión de ligandos tanto desde una perspectiva estática como dinámica, y puede ser de gran interés, por ejemplo, para validar sitios de unión crípticos no detectables en estructuras cristal. Nuestros resultados, sin embargo, también indican que, debido a la extrema sensibilidad de la velocidad de FRET a las fluctuaciones estructurales, es importante tener en cuenta la flexibilidad del sistema y promediar los acoplamientos electrónicos calculados con las fluctuaciones térmicas del sistema. Es decir, la validación, por ejemplo, de modos de unión obtenidos a partir de técnicas de *docking* mediante comparación de la eficiencia de FRET observada y calculada se espera que conlleve errores importantes, si tales cálculos ignoran la flexibilidad del ligando.

CAPÍTULO 4. DISCUSIÓN



Conclusiones

Capítulo

5

5. CONCLUSIONES

En esta tesis hemos explorado el impacto de los efectos de la polarización electrónica en dos casos paradigmáticos de interacciones biomoleculares donde su efecto es primordial, ya sea en el estado fundamental como involucrando estados excitados. A continuación enumeramos las principales conclusiones derivadas de la presente tesis.

5.1 Impacto de las fuerzas de polarización electrónica en interacciones catión- π -catión

1. Los motivos estructurales catión- π -catión son bastante comunes en las proteínas, 7% de ellas contienen al menos una interacción catión- π -catión. La composición de las interacciones identificadas, cuando se ponderan por la abundancia natural de aminoácidos, sugiere una preferencia para establecer motivos catión- π -catión con el Trp, seguido por Tyr y Phe, mientras que ninguna preferencia en particular entre Lys y Arg se observa con respecto a los cationes.
2. Los aminoácidos de las interacciones catión- π -catión tienden a estar altamente conservados, hecho que sugiere un papel estructural o funcional relevante. En este caso, las interacciones que implican Trp, y en menor medida Phe, parecen estar más conservadas que los motivos que implican Tyr.
3. La energía de interacción en el vacío es fuertemente positiva. Un ambiente apolar (n-octanol) o polar (agua) apantalla fuertemente la repulsión catión-catión, aunque en la mayoría de los casos se prevé que la interacción resultante aún sea repulsiva. Sin embargo, el impacto del ambiente proteico es más fuerte, y conduce a una interacción atractiva en el 64% de los motivos analizados.
4. El alto grado de conservación de los aminoácidos involucrados en las interacciones y el impacto del ambiente proteico sugiere un potencial papel estabilizador en las estructuras terciarias de proteínas, como se demostró recientemente para una proteína en miniatura.
5. La contribución no aditiva de polarización es muy significativa en las interacciones catión- π -catión estudiadas, llegando hasta valores de 7 kcal/mol y desafiando por tanto la capacidad de los campos de fuerzas aditivos para describir motivos catión- π -catión en simulaciones de proteínas.

5.2 Impacto de la polarización del entorno en interacciones entre estados excitados

1. El *quenching* dinámico en los sistemas FBP/HSA y FBP-TrpMe se origina por una transferencia de energía desde el estado $^1\text{FBP}^*$ al estado L_a del Trp.
2. En el sistema modelo FBP-TrpMe, la enantioselectividad surge de una disposición prácticamente ortogonal de las transiciones de FBP y Trp en varios conformeros del (S)-FBP-TrpMe en disolución de acetonitrilo.
3. En los complejos FBP/HSA, la transferencia de energía se produce a partir de la fracción de moléculas de FBP unidas al sitio secundario de HSA localizado en la interfase entre los subdominios IIA y IIB. Tal transferencia de energía enantioselectiva surge principalmente de la mayor flexibilidad de (S)-FBP en el sitio de unión comparada con (R)-FBP.
4. Las estrategias MD-QM/MMPol y QM/PCM son capaces de describir satisfactoriamente la enantioselectividad en el complejo FBP/HSA y el sistema FBP-TrpMe cuando se utiliza una descripción precisa a nivel TD-DFT para modelar las excitaciones FBP y Trp, siendo especialmente importante una descripción correcta de la orientación del dipolo de transición del estado L_a de Trp. Los métodos CIS y ZINDO, sin embargo, así como la aproximación dipolo-dipolo de Förster, se muestran menos robustos.
5. En el complejo FBP/HSA las fluctuaciones de los acoplamientos electrónicos son mucho mayores que las observadas, por ejemplo, en complejos fotosintéticos, introduciendo la necesidad de promediar sobre un número considerable de estructuras. Así, la validación de modos de unión mediante comparación de fluorescencia y simulaciones de FRET basadas en una estructura estática, por ejemplo a partir de estructuras obtenidas por *docking* para un ligando, se espera que introduzcan un ruido considerable en las predicciones.
6. Las fluctuaciones observadas en los acoplamiento surgen principalmente de la orientación entre los cromóforos, pero, en contraste con estudios previos en complejos fotosintéticos, se observan fluctuaciones muy importantes también en el factor de apantallamiento inducido por la respuesta polarizable heterogénea del entorno, hecho que ilustra las limitaciones del factor de apantallamiento constante $1/n^2$ asumido en la teoría de Förster.

7. Los resultados apoyan el potencial de las simulaciones teóricas con el fin de vincular variaciones en la fluorescencia de complejos ligando-proteína a cambios sutiles en la estructura subyacente originada por la transferencia de energía, simulaciones que se pueden utilizar para caracterizar eventos de unión del ligando tanto desde una perspectiva estática y dinámica a nivel molecular.

CAPÍTULO 5. CONCLUSIONES

Bibliografía

Bibliografía

1. Cieplak, P., Dupradeau, F.-Y., Duan, Y. & Wang, J. Polarization effects in molecular mechanical force fields. *J. Phys. Condens. Matter* 21, 333102 (2009).
2. Luque, F. J., Dehez, F., Chipot, C. & Orozco, M. Polarization effects in molecular interactions. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* 1, 844–854 (2011).
3. Warshel, A. & Levitt, M. Theoretical studies of enzymic reactions: Dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *J. Mol. Biol.* 103, 227–249 (1976).
4. Cubero, E., Luque, F. J. & Orozco, M. Is polarization important in cation- π interactions? *Proc. Natl. Acad. Sci.* 95, 5976–5980 (1998).
5. Minoux, H. & Chipot, C. Cation- π interactions in proteins: Can simple models provide an accurate description? *J. Am. Chem. Soc.* 121, 10366–10372 (1999).
6. Sherrill, C. D. Energy component analysis of π interactions. *Acc. Chem. Res.* 46, 1020–1028 (2013).
7. Rodríguez-Sanz, A., Cabaleiro-Lago, E. M. & Rodríguez-Otero, J. Interaction between the guanidinium cation and aromatic amino acids. *Phys. Chem. Chem. Phys.* 16, 22499–22512 (2014).
8. Craven, T. W., Cho, M. K., Traaseth, N. J., Bonneau, R. & Kirshenbaum, K. A miniature protein stabilized by a cation- π interaction network. *J. Am. Chem. Soc.* 138, 1543–1550 (2016).
9. Curutchet, C. *et al.* Electronic energy transfer in condensed phase studied by a polarizable QM/MM model. *J. Chem. Theory Comput.* 5, 1838–1848 (2009).
10. Curutchet, C. *et al.* Photosynthetic light-harvesting is tuned by the heterogeneous polarizable environment of the protein. *J. Am. Chem. Soc.* 133, 3078–3084 (2011).
11. Curutchet, C. *et al.* Energy flow in the cryptophyte PE545 antenna is directed by bilin pigment conformation. *J. Phys. Chem. B* 117, 4263–4273 (2013).
12. Curutchet, C. & Mennucci, B. Quantum chemical studies of light harvesting. *Chem. Rev.* 117, 294–343 (2017).
13. Förster, T. Zwischenmolekulare energiewanderung und fluoreszenz. *Ann. Phys.* 437, 55–75 (1948).
14. Scholes, G. D. Long-range resonance energy transfer in molecular systems. *Annu. Rev. Phys. Chem.* 54, 57–87 (2003).
15. Braslavsky, S. E. *et al.* Pitfalls and limitations in the practical use of Förster's theory of resonance energy transfer. *Photochem. Photobiol. Sci.* 7, 1444–1448 (2008).
16. Jares-Erijman, E. A. & Jovin, T. FRET imaging. *Nat. Biotechnol.* 21, 1387–1395 (2003).
17. Förster, T. Deslocalized excitation and excitation transfer. In: *Modern Quantum Chemistry, Section III B, Action of Light and Organic Crystals*, ed. O. Sinanoglu. Academic Press, New York. 93-137 (1965).
18. Clegg, R. M. Fluorescence resonance energy transfer. In: *Fluorescence Imaging Spectroscopy and Microscopy*, ed. Wang, X. F. and Herman, B. John Wiley, New York. 179–252 (1996).

BIBLIOGRAFÍA

19. Lakowicz, J. R. Introduction to fluorescence. In: *Principles of Fluorescence Spectroscopy*, 2nd ed. Plenum Publishing Corp., New York. 1-23 (1999).
20. Cramer, C. J. Molecular Mechanics. In: *Essentials of Computational Chemistry Theories and Models*, 2nd ed. John Wiley & Sons Ltd., West Sussex. 17-67 (2004).
21. MacKerell, A. D. Atomistic models and force fields. In: *Computational Biochemistry and Biophysics*, eds. Becker, O. M., MacKerell, A. D., Benoit Roux, B., Watanabe, M. CRC Press, New York. 7-38 (2001).
22. Hünenberger, P. H. & van Gunsteren, W. F. Empirical classical interaction functions for molecular simulation. In: *Computer Simulation of Biomolecular Systems*, eds. van Gunsteren, W. F., Weiner, P. K. & Wilkinson, A. J. Springer, Netherlands, 3–82 (1997).
23. Ponder, J. W. & Case, D. A. Force fields for protein simulations. *Adv. Protein Chem.* 66, 27–85 (2003).
24. Höltje, H.-D., Sippl, W., Rognan, D. & Folkers, G. *Molecular Modeling: Basic Principles and Applications*, 3rd ed. Wiley-VCH, Weinheim (2008).
25. Mackerell, A. D. Empirical force fields for biological macromolecules: Overview and issues. *J. Comput. Chem.* 25, 1584–1604 (2004).
26. Weiner, S. J. *et al.* A new force field for molecular mechanical simulation of nucleic acids and proteins. *J. Am. Chem. Soc.* 106, 765–784 (1984).
27. Brooks, B. R. *et al.* CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.* 4, 187–217 (1983).
28. Wang, J., Cieplak, P. & Kollman, P. A. How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *J. Comput. Chem.* 21, 1049–1074 (2000).
29. Hornak, V. *et al.* Comparison of multiple AMBER force fields and development of improved protein backbone parameters. *Proteins* 15, 712–725 (2006).
30. Cieplak, P., Caldwell, J. & Kollman, P. Molecular mechanical models for organic and biological systems going beyond the atom centered two body additive approximation: Aqueous solution free energies of methanol and *N*-methyl acetamide, nucleic acid base, and amide hydrogen bonding and chloroform. *J. Comput. Chem.* 22, 1048–1057 (2001).
31. Khan, H. M. *et al.* Improving the force field description of tyrosine-choline cation- π interactions: QM investigation of phenol-N(Me)₄⁺ interactions. *J. Chem. Theory Comput.* 12, 5585–5595 (2016).
32. Wang, J. *et al.* Development of polarizable models for molecular mechanical calculations I: Parameterization of atomic polarizability. *J. Phys. Chem. B* 115, 3091–3099 (2011).
33. Shi, Y. *et al.* Polarizable atomic multipole-based AMOEBA force field for proteins. *J. Chem. Theory Comput.* 9, 4046–4063 (2013).
34. Lemkul, J. A., Huang, J., Roux, B. & Mackerell, A. D. An empirical polarizable force field based on the classical Drude oscillator model: Development history and recent applications. *Chem. Rev.* 116, 4983–5013 (2016).
35. Applequist, J., Carl, J. R. & Fung, K.-K. An atom dipole interaction model for molecular polarizability. Application to polyatomic molecules and determination of atom polarizabilities. *J. Am. Chem. Soc.* 94, 2952–2960 (1972).

36. Kaminski, G. A. *et al.* Development of a polarizable force field for proteins via ab initio quantum chemistry: First generation model and gas phase tests. *J. Comput. Chem.* 23, 1515–1531 (2002).
37. Friesner, R. A. Modeling polarization in proteins and protein-ligand complexes: Methods and preliminary results. *Adv. Protein Chem.* 72, 79–104 (2005).
38. Ren, P. & Ponder, J. W. Consistent treatment of inter- and intramolecular polarization in molecular mechanics calculations. *J. Comput. Chem.* 23, 1497–1506 (2002).
39. Wang, J. *et al.* Development of polarizable models for molecular mechanical calculations II: Induced dipole models significantly improve accuracy of intermolecular interaction energies. *J. Phys. Chem. B* 115, 3100–3111 (2011).
40. Wang, J. *et al.* Development of polarizable models for molecular mechanical calculations. 3. Polarizable water models conforming to thole polarization screening schemes. *J. Phys. Chem. B* 116, 7999–8008 (2012).
41. Wang, J. *et al.* Development of polarizable models for molecular mechanical calculations. 4. van der Waals parametrization. *J. Phys. Chem. B* 116, 7088–7101 (2012).
42. Rick, S. W., Stuart, S. J. & Berne, B. J. Dynamical fluctuating charge force fields: Application to liquid water. *J. Chem. Phys.* 101, 6141–6156 (1994).
43. Patel, S., Mackerell, A. D. & Brooks, C. L. CHARMM fluctuating charge force field for proteins: II protein/solvent properties from molecular dynamics simulations using a nonadditive electrostatic model. *J. Comput. Chem.* 25, 1504–1514 (2004).
44. Savelyev, A., Roux, B. & MacKerell, A. Explicit inclusion of induced polarization in atomistic force fields based on the classical Drude oscillator model. In: *Many-Body Effects and Electrostatics in Biomolecules*, eds. Cui, Q., Meuwly, M. & Ren, P. CRC Press, Boca Raton, 191–232 (2016).
45. Rappe, A. K., Casewit, C. J., Colwell, K. S., Goddard, W. A. & Skiff, W. M. UFF, a full periodic table force field for molecular mechanics and molecular dynamics simulations. *J. Am. Chem. Soc.* 114, 10024–10035 (1992).
46. Patel, S. & Brooks, C. L. CHARMM fluctuating charge force field for proteins: I parameterization and application to bulk organic liquid simulations. *J. Comput. Chem.* 25, 1–15 (2004).
47. Drude, P. *The theory of optics*. Longmans, Green and Co, New York (1902).
48. Rick, S. W. & Stuart, S. J. Potentials and algorithms for incorporating polarizability in computer simulations. *Rev. Comput. Chem.* 18, 89 (2002).
49. Boulanger, E. & Thiel, W. Solvent boundary potentials for hybrid QM/MM computations using classical drude oscillators: A fully polarizable model. *J. Chem. Theory Comput.* 8, 4527–4538 (2012).
50. van Der Kamp, M. W. & Mulholland, A. J. Combined quantum mechanics/molecular mechanics (QM/MM) methods in computational enzymology. *Biochemistry* 52, 2708–2728 (2013).
51. Senn, H. M. & Thiel, W. QM/MM methods for biomolecular systems. *Angew. Chem. Int. Ed.* 48, 1198–1229 (2009).
52. Thompson, M. A. & Schenter, G. K. Excited states of the bacteriochlorophyll b dimer of rhodospseudomonas viridis: A QM/MM study of the photosynthetic reaction center that includes MM polarization. *J. Phys. Chem.* 99, 6374–6386 (1995).

BIBLIOGRAFÍA

53. Osted, A., Kongsted, J. & Mikkelsen, K. V. Statistical mechanically averaged molecular properties of liquid water calculated using the combined coupled cluster/molecular dynamics method. *J. Chem. Phys.* 124, 1–16 (2006).
54. Nielsen, C. B., Christiansen, O., Mikkelsen, K. V & Kongsted, J. Density functional self-consistent quantum mechanics/molecular mechanics theory for linear and nonlinear molecular properties: Applications to solvated water and formaldehyde. *J. Chem. Phys.* 126, 154112 (2007).
55. Öhrn, A. & Karlström, G. A theoretical study of the solvent shift to the transition in formaldehyde with an effective discrete quantum chemical solvent model including non-electrostatic perturbation. *Mol. Phys.* 104, 3087–3099 (2006).
56. Lin, Y. L. & Gao, J. Solvatochromic shifts of the $n \rightarrow \pi^*$ transition of acetone from steam vapor to ambient aqueous solution: A combined configuration interaction QM/MM simulation study incorporating solvent polarization. *J. Chem. Theory Comput.* 3, 1484–1493 (2007).
57. Losa, A., Galván, I., Aguilar, M. & Martín, E. A CASPT2//CASSCF study of vertical and adiabatic electron transitions of acrolein in water solution. *J. Phys. Chem. B* 111, 9864–9870 (2007).
58. Söderhjelm, P., Husberg, C., Strambi, A., Olivucci, M. & Ryde, U. Protein influence on electronic spectra modeled by multipoles and polarizabilities. *J. Chem. Theory Comput.* 5, 649–658 (2009).
59. Sneskov, K., Schwabe, T., Christiansen, O. & Kongsted, J. Scrutinizing the effects of polarization in QM/MM excited state calculations. *Phys. Chem. Chem. Phys.* 13, 18551–18560 (2011).
60. Slipchenko, L. V. Solvation of the excited states of chromophores in polarizable environment: Orbital relaxation versus polarization. *J. Phys. Chem. A* 114, 8824–8830 (2010).
61. Lipparini, F., Cappelli, C. & Barone, V. Linear response theory and electronic transition energies for a fully polarizable QM/Classical Hamiltonian. *J. Chem. Theory Comput.* 8, 4153–4165 (2012).
62. Thellamurege, N. M. *et al.* QuanPol: A full spectrum and seamless QM/MM program. *J. Comput. Chem.* 34, 2816–2833 (2013).
63. Caprasecca, S., Curutchet, C. & Mennucci, B. Dissecting the nature of exciton interactions in ethyne-linked tetraarylporphyrin arrays. *J. Phys. Chem. C* 117, 12423–12431 (2013).
64. Casida, M. E. Time-dependent density-functional response theory for molecules. In: *Recent Advances in Density Functional Methods, Part I*, ed. Chong, D.P. World Scientific, Singapore, 155–192 (1995).
65. Casida, M. E. & Huix-Rotllant, M. Progress in time-dependent density-functional theory. *Annu. Rev. Phys. Chem.* 63, 287–323 (2012).
66. Adamo, C. & Jacquemin, D. The calculations of excited-state properties with time-dependent density functional theory. *Chem. Soc. Rev.* 42, 845–856 (2013).
67. Dreuw, A. & Head-Gordon, M. Single-reference ab initio methods for the calculation of excited states of large molecules. *Chem. Rev.* 105, 4009–4037 (2005).

68. König, C. & Neugebauer, J. Quantum chemical description of absorption properties and excited-state processes in photosynthetic systems. *ChemPhysChem*. 13, 386–425 (2012).
69. List, N. H., Curutchet, C., Knecht, S., Mennucci, B. & Kongsted, J. Toward reliable prediction of the energy ladder in multichromophoric systems: A benchmark study on the FMO light-harvesting complex. *J. Chem. Theory Comput.* 9, 4928–4938 (2013).
70. Losa, A., Curutchet, C., Galvan, I. & Mennucci, B. Quantum mechanical methods applied to excitation energy transfer: A comparative analysis on excitation energies and electronic couplings. *J. Chem. Phys.* 129, 034104 (2008).
71. Daday, C., Curutchet, C., Sinicropi, A., Mennucci, B. & Filippi, C. Chromophore-protein coupling beyond nonpolarizable models: Understanding absorption in green fluorescent protein. *J. Chem. Theory Comput.* 11, 4825–4839 (2015).
72. Guareschi, R., Valsson, O., Curutchet, C., Mennucci, B. & Filippi, C. Electrostatic versus resonance interactions in photoreceptor proteins: The case of rhodopsin. *J. Phys. Chem. Lett.* 7, 4547–4553 (2016).
73. Cramer, C. J. & Truhlar, D. G. Implicit solvation models: Equilibria, structure, spectra, and dynamics. *Chem. Rev.* 99, 2161–2200 (1999).
74. Orozco, M. & Luque, F. J. Theoretical methods for the description of the solvent effect in biomolecular systems. *Chem. Rev.* 100, 4187–4225 (2000).
75. Tomasi, J., Benedetta, M. & Cammi, R. Quantum mechanical continuum solvation models. *Chem. Rev.* 105, 2999–3093 (2005).
76. Miertuš, S., Scrocco, E. & Tomasi, J. Electrostatic interaction of a solute with a continuum. A direct utilization of ab initio molecular potentials for the prevision of solvent effects. *Chem. Phys.* 55, 117–129 (1981).
77. Mennucci, B. Polarizable continuum model. *WIREs Comput. Mol. Sci.* 2, 386–404 (2012).
78. Iozzi, M. F., Mennucci, B., Tomasi, J. & Cammi, R. Excitation energy transfer (EET) between molecules in condensed matter: A novel application of the Polarizable Continuum Model (PCM). *J. Chem. Phys.* 120, 7029 (2004).
79. Still, W. C., Tempczyk, A., Hawley, R. C. & Hendrickson, T. Semianalytical treatment of solvation for molecular mechanics and dynamics. *J. Am. Chem. Soc.* 112, 6127–6129 (1990).
80. Srinivasan, J., Trevathan, M. W., Beroza, P. & Case, D. A. Application of a pairwise generalized Born model to proteins and nucleic acids: Inclusion of salt effects. *Theor. Chem. Acc.* 101, 426–434 (1999).
81. Onufriev, A., Bashford, D. & Case, D. A. Exploring protein native states and large-scale conformational changes with a modified generalized Born model. *Proteins* 55, 383–394 (2004).
82. Genheden, S. & Ryde, U. The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert Opin. Drug Discov.* 10, 449–461 (2015).
83. Zamora, W., Campanera, J. M. & Luque, F. J. Implicit solvation methods in the study of ligand–protein interactions. In: *Silico Drug Discovery and Design: Theory, Methods, Challenges, and Applications*, ed. Cavasotto, C. N. CRC Press, New York, 247–272 (2015).

BIBLIOGRAFÍA

84. Szalewicz, K. Symmetry-adapted perturbation theory of intermolecular forces. *WIREs Comput. Mol. Sci.* 2, 254–272 (2012).
85. Ma, J. C. & D. The cation– π interaction. *Chem. Rev.* 97, 1303–1324 (1997).
86. Mahadevi, A. S. & Sastry, G. N. Cation– π interaction: Its role and relevance in chemistry, biology, and material science. *Chem. Rev.* 113, 2100–2138 (2013).
87. Soteras, I. *et al.* Derivation of distributed models of atomic polarizability for molecular simulations. *J. Chem. Theory Comput.* 3, 1901–1913 (2007).
88. Dehez, F. *et al.* Modeling induction phenomena in intermolecular interactions with an ab initio force field. *J. Chem. Theory Comput.* 3, 1914–1926 (2007).
89. Soteras, I. *et al.* Induction effects in metal cation–benzene complexes. *Phys. Chem. Chem. Phys.* 10, 2616–2624 (2008).
90. Riley, K. E., Pitonák, M., Jurecka, P. & Hobza, P. Stabilization and structure calculations for noncovalent interactions in extended molecular systems based on wave function and density functional theories. *Chem. Rev.* 110, 5023–5063 (2010).
91. Møller, C. & Plesset, M. S. Note on an approximation treatment for many-electron systems. *Phys. Rev.* 46, 618–622 (1934).
92. Grimme, S. Improved second-order Møller–Plesset perturbation theory by separate scaling of parallel- and antiparallel-spin pair correlation energies. *J. Chem. Phys.* 118, 9095 (2003).
93. Truhlar, D. G. Basis-set extrapolation. *Chem. Phys. Lett.* 294, 45–48 (1998).
94. Boys, S. F. & Bernardi, F. The calculation of small molecular interactions by the differences of separate total energies. Some procedures with reduced errors. *Mol. Phys.* 19, 553–566 (1970).
95. Schuler, B. & Eaton, W. A. Protein folding studied by single molecule FRET. *Curr. Opin. Struct. Biol.* 18, 16–26 (2008).
96. Roy, R., Hohng, S. & Ha, T. A practical guide to single-molecule FRET. *Nat. Methods* 5, 507–516 (2008).
97. Ha, T. & Tinnefeld, P. Photophysics of fluorescent probes for single-molecule biophysics and super-resolution imaging. *Annu. Rev. Phys. Chem.* 595–617 (2012).
98. Stryer, L. Fluorescence energy transfer as a spectroscopic ruler. *Annu. Rev. Biochem.* 47, 819–846 (1978).
99. Beljonne, D., Curutchet, C., Scholes, G. D. & Silbey, R. J. Beyond Förster resonance energy transfer in biological and nanoscale systems. *J. Phys. Chem. B* 113, 6583–6599 (2009).
100. Chenu, A. & Scholes, G. D. Coherence in energy transfer and photosynthesis. *Annu. Rev. Phys. Chem.* 66, 69–96 (2015).
101. Dexter, D. L. A Theory of sensitized luminescence in solids. *J. Chem. Phys.* 21, 836–850 (1953).
102. Hsu, C. P. The electronic couplings in electron transfer and excitation energy transfer. *Acc. Chem. Res.* 42, 509–518 (2009).
103. Scholes, G. D., Curutchet, C., Mennucci, B., Cammi, R. & Tomasi, J. How solvent controls electronic energy transfer and light harvesting. *J. Phys. Chem. B* 111, 6978–6982 (2007).

104. Curutchet, C., Scholes, G. D., Mennucci, B. & Cammi, R. How solvent controls electronic energy transfer and light harvesting: Toward a quantum-mechanical description of reaction field and screening effects. *J. Phys. Chem. B* 111, 13253–65 (2007).
105. Jurinovich, S., Curutchet, C. & Mennucci, B. The Fenna-Matthews-Olson protein revisited: A fully polarizable (TD) DFT/MM description. *ChemPhysChem* 15, 3194–3204 (2014).
106. Jorgensen, W. L. Special issue on polarization. *J. Chem. Theory Comput.* 3, 1877 (2007).
107. Demerdash, O., Yap, E.-H. & Head-Gordon, T. Advanced potential energy surfaces for condensed phase simulation. *Annu. Rev. Phys. Chem.* 65, 149–174 (2014).
108. Ren, P. Y. & Ponder, J. W. Temperature and pressure dependence of the AMOEBA water model. *J. Phys. Chem. B* 108, 13427–13437 (2004).
109. Vayá, I. *et al.* Excited state interactions between flurbiprofen and tryptophan in drug–protein complexes and in model dyads. Fluorescence studies from the femtosecond to the nanosecond time domains. *Phys. Chem. Chem. Phys.* 15, 4727–4734 (2013).
110. Ghuman, J. *et al.* Structural basis of the drug-binding specificity of human serum albumin. *J. Mol. Biol.* 353, 38–52 (2005).
111. Aghtar, M., Kleinekathöfer, U., Curutchet, C. & Mennucci, B. Impact of electronic fluctuations and their description on the exciton dynamics in the light-harvesting complex PE545. *J. Phys. Chem. B* 121, 1330–1339 (2017).

BIBLIOGRAFÍA

Apéndices

Apéndice I. Información suplementaria: Structural and energetic study of cation– π –cation interactions in proteins.

Structural and energetic study of cation– π –cation interactions in proteins

Physical Chemistry Chemical Physics, **2017**, 19, 9849-9861

Silvana Pinheiro,¹ Ignacio Soteras,² Josep Lluís Gelpí,³ François Dehez,⁴

Christophe Chipot,^{4,5} F. Javier Luque,² and Carles Curutchet¹

¹*Departament de Farmàcia i Tecnologia Farmacèutica i Físicoquímica and Institut de Biomedicina (IBUB), Facultat de Farmàcia, Universitat de Barcelona, Barcelona Spain*

²*Departament de Nutrició, Ciències de l'Alimentació i Gastronomia and Institut de Biomedicina (IBUB), Facultat de Farmàcia, Universitat de Barcelona, Santa Coloma de Gramenet, Spain.*

³*Departament de Bioquímica i Biomedicina Molecular, Facultat de Biologia, Universitat de Barcelona, Spain*

⁴*Laboratoire International Associé Centre National de la Recherche Scientifique et University of Illinois at Urbana –Champaign, Unité Mixte de Recherche No. 7565, Université de Lorraine, Vandoeuvre-lès-Nancy cedex, France*

⁵*Beckman Institute for Advanced Science and Engineering, University of Illinois at Urbana — Champaign, Urbana, Illinois (USA).*

APÉNDICE I

Supporting Information for

Structural and energetic study of cation– π –cation interactions in proteins

Silvana Pinheiro,¹ Ignacio Soteras,² Josep Lluís Gelpí,³ François Dehez,⁴ Christophe Chipot,^{4,5} F. Javier Luque,² and Carles Curutchet^{1,*}

¹*Departament de Farmàcia i Tecnologia Farmacèutica i Fisicoquímica and Institut de Biomedicina (IBUB), Facultat de Farmàcia i Ciències de l'Alimentació, Universitat de Barcelona, Barcelona Spain*

²*Departament de Nutrició, Ciències de l'Alimentació i Gastronomia and Institut de Biomedicina (IBUB), Facultat de Farmàcia i Ciències de l'Alimentació, Universitat de Barcelona, Santa Coloma de Gramenet, Spain.*

³*Departament de Bioquímica i Biomedicina Molecular, Facultat de Biologia, Universitat de Barcelona, Spain*

⁴*Laboratoire International Associé Centre National de la Recherche Scientifique et University of Illinois at Urbana –Champaign, Unité Mixte de Recherche No. 7565, Université de Lorraine, Vandoeuvre-lès-Nancy cedex, France*

⁵*Beckman Institute for Advanced Science and Engineering, University of Illinois at Urbana — Champaign, Urbana, Illinois (USA).*

Table S4: Protein structures corresponding to the cation– π –cation interactions selected for the energetic analysis.

ID	Classification	Organism
1A22	Growth hormone bound to single receptor	Homo sapiens
1B35	Virus	Cricket paralysis virus
1CN4	Haematopoietic Cytokine	Homo sapiens
1EGJ	Immune System	Homo sapiens
1FG9	Immune System	Homo sapiens
1GQK	Hydrolase	Cellvibrio japonicus
1GXS	Lyase	Sorghum bicolor
1IWB	Lyase	Klebsiella oxytoca
1KXV	Hydrolase	Camelus dromedarius
1MDX	Transferase	Salmonella typhimurium
1SQV	Oxidoreductase	Bos taurus
1T72	Transport Protein	Aquifex aeolicus
1TJ0	Oxidoreductase	Escherichia coli
1TZ9	Lyase	Enterococcus faecalis v583
1UA4	Transferase	Pyrococcus furiosus
1VFF	Hydrolase	Pyrococcus horikoshii
2AK4	Immune System	Homo sapiens
2CQQ	Membrane Protein	Homo sapiens
2CVF	DNA Binding Protein	Thermococcus kodakarensis
2CYC	Ligase	Pyrococcus horikoshii ot3
2GRE	Hydrolase	Bacillus cereus
2H39	Transferase	Arabidopsis thaliana
2HCS	Transferase	Kunjin virus
2J47	Inhibitor	Bacteroides thetaiotaomicron
2J7A	Oxidoreductase	Desulfovibrio vulgaris
2JKV	Oxidoreductase	Human
2OT4	Oxidoreductase	Thioalkalivibrio nitratireducens
2PAM	Isomerase	Aneurinibacillus thermoaerophilus
2PI0	Transcription Activator/dna	Homo sapiens

2QOW	Ribosome	Escherichia coli
2QWN	Chaperone	Bos taurus
2VLR	Immune System	Homo sapiens
2YU9	Transcription	Saccharomyces cerevisiae
2ZTG	Ligase	Archaeoglobus fulgidus
3C5Z	Sugar Binding Protein/immune System	Mus musculus
3CES	RNA Binding Protein	Escherichia coli
3DGC	Cytokine/signaling Protein	Homo sapiens
3E0M	Oxidoreductase	Streptococcus pneumoniae
3FQD	Hydrolase/protein Binding	Schizosaccharomyces pombe
3G9V	Cytokine/cytokine Receptor	Homo sapiens
3GVP	Hydrolase	Homo sapiens
3HUI	Immune System	Homo sapiens

Table S5: Amino acids corresponding to the cation– π –cation interactions selected for the energetic analysis.

ID	X		Ar		Y	
X-Trp-Y						
1A22-1	LYS	B379	TRP	B386	ARG	B411
1CN4-1	ARG	A197	TRP	A209	ARG	A199
1EGJ-1	ARG	A377	TRP	A383	ARG	A411
1EGJ-2	ARG	A413	TRP	A425	ARG	A415
1FG9-1	LYS	C47	TRP	C56	ARG	C84
1GXS-1	ARG	A215	TRP	A205	ARG	D368
1GXS-2	ARG	B368	TRP	C205	ARG	C215
1IWB-2	LYS	A395	TRP	A538	LYS	A542
1KXV-1	ARG	C45	TRP	C111	LYS	C96
1MDX-1	ARG	A110	TRP	A353	ARG	A357
1SQV-2	LYS	C311	TRP	C379	ARG	F33
1UA4-1	LYS	A37	TRP	A103	ARG	A47
1VFF-1	ARG	A75	TRP	A49	LYS	A90
2CQQ-1	ARG	A43	TRP	A11	LYS	A51
2HCS-1	ARG	A742	TRP	A800	ARG	A797
2PI0-1	LYS	A77	TRP	A38	LYS	A105
2QOW-1	ARG	I112	TRP	N100	ARG	J48
2ZTG-1	ARG	A89	TRP	A619	ARG	A679
3DGC-1	LYS	R58	TRP	R65	ARG	R95
3FQD-1	ARG	A853	TRP	B204	ARG	B271
3G9V-1	LYS	A65	TRP	A72	ARG	A102
3G9V-2	ARG	A102	TRP	A114	ARG	A104
3GVP-1	LYS	A228	TRP	A302	ARG	A254
X-Tyr-Y						
1A22-2	ARG	B413	TYR	B422	LYS	B415
1TJ0-1	ARG	A100	TYR	A99	ARG	A409
1TZ9-2	ARG	A80	TYR	A188	ARG	A192
2GRE-1	ARG	G171	TYR	A103	LYS	G181

2J47-1	ARG	A235	TYR	A230	LYS	A269
2J47-2	ARG	A219	TYR	A260	LYS	A264
2J47-4	ARG	A519	TYR	A523	LYS	A527
2JKV-1	LYS	B76	TYR	A481	LYS	B261
2JKV-2	ARG	C136	TYR	C137	LYS	F309
2PAM-1	LYS	A32	TYR	A113	LYS	B25
2PAM-2	LYS	A25	TYR	B113	LYS	B32
2QWN-1	ARG	A100	TYR	A115	LYS	A102
2YU9-1	ARG	C84	TYR	C82	LYS	C161
3E0M-1	ARG	A259	TYR	A246	ARG	B115
<hr/>						
X-Phe-Y						
1A22-3	ARG	B411	PHE	B425	ARG	B413
1B35-1	ARG	B78	PHE	B75	LYS	B115
1GQK-2	ARG	A550	PHE	A558	ARG	A560
1T72-1	ARG	A107	PHE	A189	ARG	B192
1T72-2	ARG	A192	PHE	B189	ARG	B107
2AK4-1	ARG	E190	PHE	E124	ARG	E230
2CVF-1	LYS	A165	PHE	A157	ARG	A181
2CYC-1	LYS	A247	PHE	A243	ARG	A285
2H39-1	ARG	A26	PHE	A240	ARG	B116
2H39-2	ARG	A116	PHE	B240	ARG	B26
2J7A-1	ARG	A133	PHE	A111	LYS	A295
2OT4-1	ARG	A131	PHE	A109	LYS	A358
2VLR-1	ARG	E187	PHE	E121	ARG	E227
3C5Z-1	ARG	B183	PHE	B117	ARG	B223
3CES-1	LYS	A145	PHE	A2	ARG	A147
3HUI-1	ARG	F190	PHE	F124	ARG	F230
<hr/>						

Table S6: Distances, angles, and dihedral angles between the aromatic amino acid and the arginine guanidinium planes characterizing the cation– π –cation interactions selected for the energetic analysis.

ID	d_{XAr} (Å)	θ_{XAr} (°)	d_{ArY} (Å)	θ_{ArY} (°)	$\Psi_{\text{XAr planes}}$ (°)	$\Psi_{\text{ArY planes}}$ (°)
X-Trp-Y						
1A22-1	3.7	25.5	4.1	33.7		14.7
1CN4-1	3.6	15.2	3.8	20.6	25.8	74.0
1EGJ-1	3.8	19.6	3.7	15.0	42.3	9.4
1EGJ-2	3.8	14.7	3.7	20.2	18.9	44.4
1FG9-1	4.0	12.7	3.7	11.5		12.8
1GXS-1	3.9	16.0	3.7	26.0	36.9	13.3
1GXS-2	3.6	24.5	3.8	12.7	12.3	32.2
1IWB-2	3.6	11.7	3.5	16.3		
1KXV-1	3.5	20.2	4.0	6.4	15.3	
1MDX-1	3.7	13.3	3.6	23.5	32.4	5.4
1SQV-2	3.6	6.5	3.5	29.4		9.6
1UA4-1	3.3	8.4	3.6	28.3		22.0
1VFF-1	3.8	17.9	3.8	11.6	7.2	
2CQQ-1	4.2	17.4	3.6	21.5	37.5	
2HCS-1	3.6	10.9	4.4	19.8	10.5	6.2
2PI0-1	3.7	29.3	3.5	6.7		
2QOW-1	3.9	18.4	3.3	17.7	17.4	12.2
2ZTG-1	3.5	10.7	3.7	21.1	2.2	13.9
3DGC-1	4.6	17.5	3.2	7.4		13.2
3FQD-1	3.9	19.4	3.5	1.0	6.5	1.4
3G9V-1	4.2	19.2	3.4	8.2		14.9
3G9V-2	3.5	12.4	3.9	32.4	10.8	12.6
3GVP-1	4.3	30.4	3.5	8.7		57.4
X-Tyr-Y						
1A22-2	4.4	36.1	4.0	25.3	10.2	
1TJ0-1	3.9	7.3	3.6	19.9	26.2	16.2
1TZ9-2	3.4	2.9	3.5	21.1	5.2	28.7
2GRE-1	4.1	11.2	3.8	14.1	19.8	
2J47-1	3.6	21.7	4.0	4.2	10.3	
2J47-2	4.1	39.9	4.4	41.6	8.2	
2J47-4	4.2	24.1	4.4	24.8	9.9	
2JKV-1	3.7	16.7	3.6	12.9		
2JKV-2	5.0	28.1	3.6	6.9	10.4	
2PAM-1	3.9	37.7	4.7	14.7		
2PAM-2	4.1	8.5	3.8	27.2		
2QWN-1	3.9	27.3	3.7	3.1	14.8	
2YU9-1	3.7	14.5	4.0	23.2	23.6	
3E0M-1	3.5	3.5	3.6	24.3	5.0	12.7
X-Phe-Y						

1A22-3	4.1	29.0	4.7	43.4	21.4	24.9
1B35-1	3.8	1.4	3.7	18.2	23.0	
1GQK-2	3.7	19.4	3.6	19.8	76.2	3.4
1T72-1	4.2	31.8	3.7	25.6	3.4	20.0
1T72-2	3.9	23.1	3.9	14.6	23.3	18.9
2AK4-1	3.6	10.5	3.9	25.3	6.4	9.9
2CVF-1	3.8	20.9	4.0	18.8		53.7
2CYC-1	3.9	12.2	3.9	22.1		28.5
2H39-1	4.2	30.5	3.7	8.3	4.4	3.8
2H39-2	3.6	7.1	4.5	32.4	10.0	12.4
2J7A-1	4.1	22.5	3.9	8.6	5.2	
2OT4-1	4.0	17.2	3.5	14.6	33.7	
2VLR-1	3.8	15.8	4.0	23.7	4.1	10.1
3C5Z-1	3.6	16.6	3.6	19.5	39.4	14.2
3CES-1	3.8	23.4	3.6	11.6		13.0
3HUI-1	3.5	5.8	3.9	27.2	9.2	8.7

Table S7: Conservation, residue variety, pKa and percent of buried surface computed using the PropKa and the ConSurf servers for the cation- π -cation interactions selected for energetic analysis.

ID	Residue	pKa	Buried (%)	Score	Residue variety
X-Trp-Y					
1A22-1	TRP B386			9	W,P
	LYS B379	11.2	0	7	S,Q,T,N,R,K
	ARG B411	11.5	0	8	H,Q,D,R,E,V
1CN4-1	TRP A209			7	W,P,R
	ARG A197	11.6	32	9	T,R,L
	ARG A199	11.0	38	8	H,Q,N,K,R,L
1EGJ-1	TRP A383			6	R,W,E,L
	ARG A377	11.6	8	6	W,K,T,R,M
	ARG A411	10.7	30	9	R,K
1EGJ-2	TRP A425			8	W,Y
	ARG A413	11.6	23	9	R
	ARG A415	11.8	10	5	E,M,R,K
1FG9-1	TRP C56			9	W
	LYS C47	10.5	30	8	K,I,L,M
	ARG C84	12.7	29	8	K,Q,R
1GXS-1	TRP A205			9	W
	ARG A215	12.5	62	5	N,R,A,Y,F,L,H,V,I
	ARG D368	10.2	100	8	M,K,I,R,Q
1GXS-2	TRP C205			9	W
	ARG B368	10.4	100	8	M,I,K,R,Q
	ARG C215	12.7	62	5	L,I,F,V,N,Y,R,H,A
1IWB-2	TRP A538			7	Q,W,R
	LYS A395	8.7	72	1	S,A,Q,M,R,K,G,E
	LYS A542	11.0	43	6	S,Q,D,N,I,K,E,V
1KXV-1	TRP C111			9	W
	ARG C45	11.0	31	9	R,P,L
	LYS C96	8.2	73	7	T,K,E,S,V,G,R,M,A
1MDX-1	TRP A353			1	L,F,V,D,S,T,R,H,E,W,K,N,A,Q
	ARG A110	13.1	1	2	Q,M,A,N,E,Y,S,R,D,L,K,G,I,P,H,V,C
	ARG A357	12.9	19	4	S,T,R,H,G,E,K,N,Q,A
1SQV-2	TRP C379			8	F,W
	LYS C311	9.4	60	8	N,K
	ARG F33	12.8	73	4	F,A,S,W,T,N,K,Y,H,Q,M,R,I,G,L
1UA4-1	TRP A103			2	A,S,F,T,W,N,K,E,Y,V,H,I,R,L
	LYS A37	9.0	45	5	A,T,N,K,Y,V,H,M,C,R,I
	ARG A47	13.9	28	7	F,M,I,R,Y,L,V
1VFF-1	TRP A49			8	F,W,R,Y
	ARG A75	13.8	96	9	S,T,R,G,L
	LYS A90	9.7	41	6	H,F,R,K,Y

2CQQ-1	TRP A11			9	W
	ARG A43	12.8	2	9	K,R,H
	LYS A51	10.1	0	7	K,F,V,R,H,A,Q,M
2HCS-1	TRP A800			8	W,E
	ARG A742	12.5	15	9	T,K,R
	ARG A797	11.7	20	9	K,R
2PI0-1	TRP A38			9	W
	LYS A77	9.8	75	9	K,E
	LYS A105	9.3	72	8	P,R,K,E
2QOW-1	TRP N100			9	W
	ARG I112	13.0	0	9	R
	ARG J48	12.0	0	1	H,I,K,R,L,V
2ZTG-1	TRP A619			4	A,S,T,W,N,K,E,V,H,Q,M,I,R,G,L
	ARG A89	9.1	100	9	K,R
	ARG A679	9.8	85	8	S,A,T,K,E,V,Q,M,D,C,R,I,G
3DGC-1	TRP R65			9	W
	LYS R58	10.2	26	6	F,A,S,T,N,K,Y,Q,M,I,L
	ARG R95	12.3	23	7	S,Q,D,N,K,R
3FQD-1	TRP B204			4	A,F,W,K,E,Y,V,Q,C,I,R,L
	ARG A853	10.4	98	4	A,S,N,K,P,E,V,H,Q,C,D,I,R
	ARG B271	10.9	66	6	A,T,N,K,Y,E,H,Q,D,R,I,L
3G9V-1	TRP A72			9	W
	LYS A65	8.9	52	9	S,K,F,L
	ARG A102	10.1	50	8	S,K,R
3G9V-2	TRP A114			8	C,T,W
	ARG A102	10.1	50	8	S,K,R
	ARG A104	11.9	14	3	S,K,T,M,R,H,Q
3GVP-1	TRP A302			1	A,W,K,P,E,D,G,L
	LYS A228	10.1	0	6	Q,T,N,R,K,V
	ARG A254	12.2	55	6	S,A,T,K,Y,V,Q,R,L

X-Tyr-Y

1A22-2	TYR B422			6	A,F,W,R,K,Y
	ARG B413	12.8	0	9	S,H,Q,R
	LYS B415	10.3	1	6	A,T,N,K,V,Q,M,R,I
1TJ0-1	TYR A99			4	A,S,F,T,W,N,P,K,Y,H,M,C,R,I,G
	ARG A100	12.1	34	7	A,S,T,N,Y,V,H,Q,M,C,D,I,R,G,L
	ARG A409	10.7	89	9	S,R
1TZ9-2	TYR A188			2	A,F,W,E,Y,V,H,D,L
	ARG A80	12.0	22	6	A,F,W,C,R,Y
	ARG A192	12.7	11	2	A,Q,N,K,R,G,E
2GRE-1	TYR A103			8	S,T,N,P,K,E,Y,V,H,Q,D,C,R,G
	ARG G171	12.5	100	6	S,A,T,N,P,K,E,H,Q,M,C,D,R,G,L
	LYS G181	9.0	89	8	A,S,T,N,K,V,M,R,I,G,L
2J47-1	TYR A230			9	Y
	ARG A235	13.0	31	9	R
	LYS A269	11.4	0	9	H,K

2J47-2	TYR A260			6	R,Y,F,D,K
	ARG A219	13.1	0	8	M,Y,S,R
	LYS A264	10.5	0	4	V,T,R,A,E,K,H,N,Q
2J47-4	TYR A523			1	V,Y,S,R,T,K,G,E,L,F,A,N,H,D,Q
	ARG A519	12.7	0	1	Q,D,P,N,H,K,G,E,S,R
	LYS A527	10.7	15	1	Q,S,R,T,Y,V,N,I,H,G,K,L,A
2JKV-1	TYR A481			9	H,Y
	LYS B76	10.1	0	7	S,Q,T,M,R,K,E
	LYS B261	7.2	70	9	K
2JKV-2	TYR C137			1	F,W,T,N,K,E,Y,H,Q,M,I,R,L
	ARG C136	12.9	26	9	R,L
	LYS F309	11.5	0	1	F,S,T,N,K,E,V,Q,M,C,L,A,P,H,D,I,R,G
2PAM-1	TYR A113			1	S,F,T,N,K,E,Y,V,Q,M,L,A,P,H,D,I,R,G
	LYS A32	11.2	3	6	S,A,N,K,E,V,H,Q,C,R,L
	LYS B25	11.5	0	1	F,S,T,N,K,Y,E,V,Q,M,C,L,A,W,P,H,D,R,I,G
2PAM-2	TYR B113			1	S,F,T,N,K,E,Y,V,Q,M,L,A,P,H,D,I,R,G
	LYS A25	10.4	0	1	F,S,T,N,K,Y,E,V,Q,M,C,L,A,W,P,H,D,R,I,G
	LYS B32	11.5	13	6	S,A,N,K,E,V,H,Q,C,R,L
2QWN-1	TYR A115			5	A,S,F,T,N,K,Y,H,Q,R
	ARG A100	10.4	21	5	A,S,T,N,K,Y,E,V,Q,M,D,R,I,G,L
	LYS A102	10.4	11	1	S,A,W,T,N,K,E,Y,V,H,Q,M,D,R,I,G,L
2YU9-1	TYR C82			4	S,F,T,N,K,E,Y,M,L,A,W,P,H,D,R,I,G
	ARG C84	11.8	50	6	S,T,N,K,Y,E,V,Q,M,A,W,P,H,D,R,I,G
	LYS C161	9.2	57	9	M,P,R,K,L
3E0M-1	TYR A246			1	S,F,T,N,K,E,Y,V,Q,L,A,W,P,H,D,R,I
	ARG A259	13.0	60	8	S,T,I,K,R,L,V
	ARG B115	11.2	59	1	A,S,T,N,P,K,E,H,Q,D,R,G,L
X-Phe-Y					
1A22-3	PHE B425			9	F,W
	ARG B411	11.5	0	8	H,Q,D,R,E,V
	ARG B413	12.8	0	9	S,H,Q,R
1B35-1	PHE B75			8	F
	ARG B78	12.1	0	9	R
	LYS B115	10.2	0	9	K
1GQK-2	PHE A558			9	F,T,K,Y,V,M,C,R,I
	ARG A550	11.8	29	6	F,S,N,K,Y,H,Q,R,G
	ARG A560	12.4	100	9	R
1T72-1	PHE A189			5	S,F,T,N,Y,V,Q,M,C,L,A,W,H,D,R,I
	ARG A107	11.6	82	6	S,F,T,N,K,Y,E,V,Q,M,L,A,H,D,I,R,G
	ARG B192	9.6	100	8	S,N,K,E,H,Q,R,G,L
1T72-2	PHE B189			5	S,F,T,N,Y,V,Q,M,C,L,A,W,H,D,R,I
	ARG A192	12.5	100	8	S,N,K,E,H,Q,R,G,L
	ARG B107	10.5	76	6	S,F,T,N,K,Y,E,V,Q,M,L,A,H,D,I,R,G
2AK4-1	PHE E124			3	S,H,F,Q,T,N,I,R,K,E
	ARG E190	11.6	42	6	S,T,N,K,Y,V,H,Q,M,R,L
	ARG E230	11.3	17	7	S,W,N,I,R

2CVF-1	PHE A157			3	F,Y
	LYS A165	9.9	12	5	K,R
	ARG A181	12.1	44	9	R
2CYC-1	PHE A243			3	S,F,N,I,E,Y
	LYS A247	9.8	47	6	A,S,T,N,K,Y,V,Q,D,G,L
	ARG A285	12.8	0	3	A,S,F,N,P,K,E,Q,D,R,I,G
2H39-1	PHE A240			5	W,F,Y
	ARG A26	13.7	50	9	H,R
	ARG B116	11.6	36	5	N,C,M,D,Q,W,K,S,A,V,R,F,G,T,E,L
2H39-2	PHE B240			5	Y,F,W
	ARG A116	11.5	38	5	K,L,N,E,S,V,W,R,M,C,G,T,A,F,Q,D
	ARG B26	13.6	54	9	H,R
2J7A-1	PHE A111			9	A,F,M,Y
	ARG A133	11.3	100	9	S,P,R
	LYS A295	9.4	100	9	A,F,H,Q,K,Y
2OT4-1	PHE A109			9	F
	ARG A131	11.1	100	9	R
	LYS A358	9.6	100	9	K
2VLR-1	PHE E121			4	S,H,F,Q,I,L,K,E,R,T,N
	ARG E187	11.7	34	1	L,V,F,S,N,T,R,K
	ARG E227	11.3	15	6	N,G,R,T,E,W,S,D
3C5Z-1	PHE B117			5	A,S,F,T,N,P,K,E,V,H,Q,R,I,L
	ARG B183	11.6	33	6	S,T,N,K,Y,V,H,Q,M,R,L
	ARG B223	11.4	14	7	S,W,N,R
3CES-1	PHE A2			1	F,T,W,N,K,E,Y,H,Q,D,R,I,G,L
	LYS A145	9.6	51	1	S,F,T,N,K,Y,E,V,Q,M,C,L,A,P,H,D,R,I
	ARG A147	12.0	23	1	S,F,T,N,K,E,Y,Q,M,C,L,A,W,P,H,D,I,R,G
3HUI-1	PHE F124			6	A,S,F,T,N,P,K,E,V,H,Q,R,I,L
	ARG F190	12.0	22	7	S,T,N,K,Y,V,H,Q,M,R,L
	ARG F230	11.3	17	7	S,W,N,R,E

Table S8: SCS-MP2/CBS interaction energies (kcal/mol) predicted for the cation- π -cation interactions in vacuum.

ID	ΔE_{XAr}	ΔE_{ArY}	ΔE_{XY}	ΔE_{XArY}	ΔE_{3-body}
X-Trp-Y					
1A22-1	-5.5	-8.2	39.7	28.2	2.1
1CN4-1	-9.1	-8.7	43.9	32.1	5.9
1EGJ-1	-11.5	-8.5	41.9	26.3	4.4
1EGJ-2	-9.8	-4.2	40.5	30.4	4.0
1FG9-1	-6.6	-7.1	43.9	32.4	2.2
1GXS-1	-12.0	-7.1	42.4	27.0	3.6
1GXS-2	-6.9	-11.3	42.1	27.4	3.5
1IWB-2	-8.7	-4.5	45.8	33.8	1.2
1KXV-1	-11.9	-16.8	48.1	25.9	6.5
1MDX-1	-11.2	-6.9	42.5	28.0	3.6
1SQV-2	-8.3	-5.4	45.9	33.8	1.6
1UA4-1	-10.9	-4.1	48.0	34.2	1.2
1VFF-1	-11.0	-15.0	43.1	22.5	5.3
2CQQ-1	-12.0	-15.1	43.3	21.5	5.3
2HCS-1	-10.0	-8.8	38.5	23.3	3.7
2PI0-1	-4.9	-11.7	42.4	28.7	2.9
2QOW-1	-6.4	-10.5	42.7	29.4	3.6
2ZTG-1	-10.3	-8.6	42.3	27.4	4.0
3DGC-1	-3.8	-8.0	41.5	32.5	2.8
3FQD-1	-8.2	-9.6	41.0	26.5	3.4
3G9V-1	-8.7	-9.8	44.4	29.8	3.8
3G9V-2	-8.0	-6.0	41.5	30.2	2.6
3GVP-1	-8.2	-11.0	42.1	26.2	3.3
X-Tyr-Y					
1A22-2	-5.0	-4.7	34.0	27.0	2.7
1TJ0-1	-8.2	-5.5	40.7	30.6	3.5
1TZ9-2	-7.0	-6.0	44.1	35.4	4.4
2GRE-1	-7.2	-13.1	41.2	25.6	4.7
2J47-1	-8.5	-7.7	41.1	29.0	4.2
2J47-2	-7.2	-6.2	42.4	31.2	2.2
2J47-4	-5.9	-8.1	39.8	27.8	2.0
2JKV-1	-9.9	-8.9	39.9	26.5	5.4
2JKV-2	-4.7	-12.4	37.2	23.6	3.5
2PAM-1	-11.9	-7.2	39.4	25.0	4.7
2PAM-2	-7.9	-12.6	43.5	26.6	3.6
2QWN-1	-8.1	-8.1	38.9	26.9	4.1
2YU9-1	-7.6	-8.8	40.3	28.1	4.2
3E0M-1	-7.5	-7.6	42.4	31.0	3.7
X-Phe-Y					
1A22-3	-6.7	-5.0	39.5	30.4	2.6

1B35-1	-7.2	-9.5	41.7	28.5	3.6
1GQK-2	-7.9	-6.5	41.0	30.3	3.7
1T72-1	-6.7	-8.1	41.4	30.4	3.8
1T72-2	-7.6	-8.4	40.3	28.5	4.2
2AK4-1	-7.0	-6.3	41.1	30.9	3.1
2CVF-1	-12.5	-11.1	45.7	29.5	7.4
2CYC-1	-8.4	-5.3	39.5	29.5	3.9
2H39-1	-6.3	-6.9	39.5	29.4	3.1
2H39-2	-6.4	-6.7	38.8	28.8	3.2
2J7A-1	-6.4	-7.4	38.5	27.7	3.0
2OT4-1	-6.0	-11.9	40.4	27.2	4.8
2VLR-1	-6.9	-6.0	39.5	29.5	3.0
3C5Z-1	-8.9	-6.6	43.0	32.2	4.7
3CES-1	-7.2	-7.7	40.6	29.4	3.8
3HUI-1	-7.0	-6.9	41.8	31.2	3.2

Table S9: Interaction energies (kcal/mol) predicted for the cation– π –cation interactions in vacuum, in a polar (water) and apolar (*n*-octanol) environments, and in the protein environment. Vacuum interaction energies were computed at the SCS-MP2/CBS level of theory, octanol and water solvation energies by MST-IEFPCM B3LYP/6-31G(d) calculations, and water to protein transfer free energies by classical MM-GBSA calculations.

ID	ΔE_{vac}	$\Delta\Delta G_{solv}^{oct}$	$\Delta\Delta G_{solv}^{wat}$	$\Delta G_{wat \rightarrow prot}$	ΔE_{oct}	ΔE_{wat}	ΔE_{prot}
X-Trp-Y							
1A22-1	28.2	-21.5	-26.6	-18.3	6.7	1.5	-16.7
1CN4-1	32.1	-20.2	-19.6	-2.8	11.9	12.5	9.7
1EGJ-1	26.3	-22.0	-22.8	1.3	4.3	3.6	4.8
1EGJ-2	30.4	-24.4	-27.2	30.1	6.1	3.2	33.3
1FG9-1	32.4	-24.1	-29.6	-10.5	8.2	2.7	-7.8
1GXS-1	27.0	-24.0	-25.3	-26.2	3.0	1.7	-24.5
1GXS-2	27.4	-24.0	-26.1	-26.2	3.3	1.2	-25.0
1IWB-2	33.8	-22.7	-28.8	-8.2	11.1	5.0	-3.3
1KXV-1	25.9	-19.3	-19.1	4.6	6.6	6.8	11.4
1MDX-1	28.0	-24.3	-26.2	-16.9	3.7	1.8	-15.2
1SQV-2	33.8	-25.4	-29.2	-3.9	8.4	4.6	0.7
1UA4-1	34.2	-25.6	-29.7	-10.7	8.6	4.5	-6.2
1VFF-1	22.5	-19.1	-20.0	-7.0	3.4	2.5	-4.5
2CQQ-1	21.5	-14.9	-15.5	13.9	6.6	6.1	20.0
2HCS-1	23.3	-21.7	-23.9	-1.0	1.7	-0.6	-1.6
2PI0-1	28.7	-18.7	-23.6	-12.2	9.9	5.0	-7.2
2QOW-1	29.4	-25.9	-28.1	-44.6	3.5	1.3	-43.3
2ZTG-1	27.4	-25.3	-27.2	-13.2	2.0	0.2	-13.0
3DGC-1	32.5	-23.2	-28.0	-1.0	9.3	4.5	3.5
3FQD-1	26.5	-24.5	-26.5	0.0	2.1	0.0	0.0
3G9V-1	29.8	-22.9	-25.5	9.0	6.9	4.3	13.2
3G9V-2	30.2	-26.6	-29.9	-3.8	3.5	0.3	-3.5
3GVP-1	26.2	-20.4	-22.2	-8.3	5.8	4.0	-4.3
X-Tyr-Y							
1A22-2	27.0	-20.8	-26.1	-1.9	6.1	0.9	-1.0
1TJ0-1	30.6	-27.1	-30.1	3.5	3.5	0.5	4.0
1TZ9-2	35.4	-26.9	-28.0	-11.3	8.6	7.5	-3.8
2GRE-1	25.6	-20.1	-21.7	-22.2	5.5	4.0	-18.3
2J47-1	29.0	-22.8	-27.6	-9.1	6.2	1.4	-7.7
2J47-2	31.2	-23.8	-27.8	-7.5	7.3	3.3	-4.1
2J47-4	27.8	-22.0	-26.2	0.7	5.8	1.6	2.3
2JKV-1	26.5	-17.3	-20.2	11.7	9.3	6.3	18.0
2JKV-2	23.6	-18.6	-20.4	-9.6	5.0	3.2	-6.4
2PAM-1	25.0	-14.5	-16.3	-2.9	10.5	8.7	5.8
2PAM-2	26.6	-17.0	-20.0	-4.5	9.5	6.6	2.0

2QWN-1	26.9	-20.3	-22.9	-4.6	6.6	4.0	-0.6
2YU9-1	28.1	-21.2	-24.8	-4.4	6.9	3.3	-1.0
3E0M-1	31.0	-28.3	-30.3	-27.0	2.7	0.7	-26.3
<hr/>							
X-Phe-Y							
1A22-3	30.4	-25.8	-29.0	-9.2	4.6	1.4	-7.8
1B35-1	28.5	-22.5	-24.7	-1.7	6.0	3.9	2.2
1GQK-2	30.3	-28.0	-30.5	-22.3	2.3	-0.3	-22.6
1T72-1	30.4	-26.2	-28.6	-10.7	4.2	1.8	-8.9
1T72-2	28.5	-25.6	-27.8	-6.1	2.9	0.7	-5.4
2AK4-1	30.9	-28.4	-31.2	-10.5	2.5	-0.3	-10.8
2CVF-1	29.5	-19.3	-18.7	8.5	10.2	10.8	19.3
2CYC-1	29.5	-23.1	-27.0	12.0	6.4	2.5	14.5
2H39-1	29.4	-26.9	-29.7	-22.0	2.5	-0.3	-22.2
2H39-2	28.8	-25.7	-28.0	-23.0	3.1	0.8	-22.2
2J7A-1	27.7	-22.0	-26.2	6.3	5.7	1.5	7.7
2OT4-1	27.2	-21.6	-23.1	-7.4	5.6	4.1	-3.3
2VLR-1	29.5	-27.1	-30.8	-4.3	2.4	-1.3	-5.6
3C5Z-1	32.2	-26.7	-28.6	-8.0	5.6	3.6	-4.4
3CES-1	29.4	-23.5	-27.2	9.1	5.9	2.2	11.3
3HUI-1	31.2	-28.5	-31.4	-9.2	2.7	-0.2	-9.4

Table S10: Comparison of SCS-MP2 and CCSD interaction energies (kcal/mol) predicted for the multiple cation– π –cation interactions identified in the human growth hormone receptor (PDB entry 1A22) in vacuum.

ID	Method	ΔE_{XAr}	ΔE_{ArY}	ΔE_{XY}	ΔE_{XArY}	ΔE_{3-body}
1A22-1						
	SCS-MP2/aug-cc-pVDZ	-5.4	-7.6	39.7	28.8	2.2
	SCS-MP2/aug-cc-pVTZ	-5.4	-7.9	39.7	28.5	2.1
	SCS-MP2/CBS	-5.5	-8.2	39.7	28.2	2.1
	CCSD/aug-cc-pVDZ	-5.4	-7.4	39.7	29.0	2.2
1A22-2						
	SCS-MP2/aug-cc-pVDZ	-4.8	-4.6	34.0	27.3	2.7
	SCS-MP2/aug-cc-pVTZ	-4.9	-4.6	34.0	27.1	2.7
	SCS-MP2/CBS	-5.0	-4.7	34.0	27.0	2.7
	CCSD/aug-cc-pVDZ	-4.6	-4.6	34.0	27.4	2.7
1A22-3						
	SCS-MP2/aug-cc-pVDZ	-6.1	-4.7	39.5	31.3	2.6
	SCS-MP2/aug-cc-pVTZ	-6.5	-4.8	39.5	30.8	2.6
	SCS-MP2/CBS	-6.7	-5.0	39.5	30.4	2.6
	CCSD/aug-cc-pVDZ	-5.9	-4.5	39.5	31.6	2.6

Apéndice II. Información suplementaria: Can Förster theory describe stereoselective energy transfer dynamics in a protein-ligand complex?

**Can Förster theory describe stereoselective energy transfer
dynamics in a protein-ligand complex?**

Journal of Physical Chemistry B, **2017**, 121 (10), 2265–2278

Silvana Pinheiro¹ and Carles Curutchet¹

¹Departament de Farmàcia i Tecnologia Farmacèutica i Fisicoquímica and Institut de Biomedicina (IBUB), Facultat de Farmàcia i Ciències de l'Alimentació, Universitat de Barcelona, Barcelona Spain

APÉNDICE II

Supporting Information for
Can Förster Theory Describe Stereoselective Energy Transfer
Dynamics in a Protein-Ligand Complex?

Silvana Pinheiro¹ and Carles Curutchet^{1,*}

¹Departament de Farmàcia i Tecnologia Farmacèutica i Fisicoquímica and Institut de Biomedicina (IBUB), Facultat de Farmàcia i Ciències de l'Alimentació, Universitat de Barcelona, Barcelona Spain

*Corresponding author: CC (carles.curutchet@ub.edu)

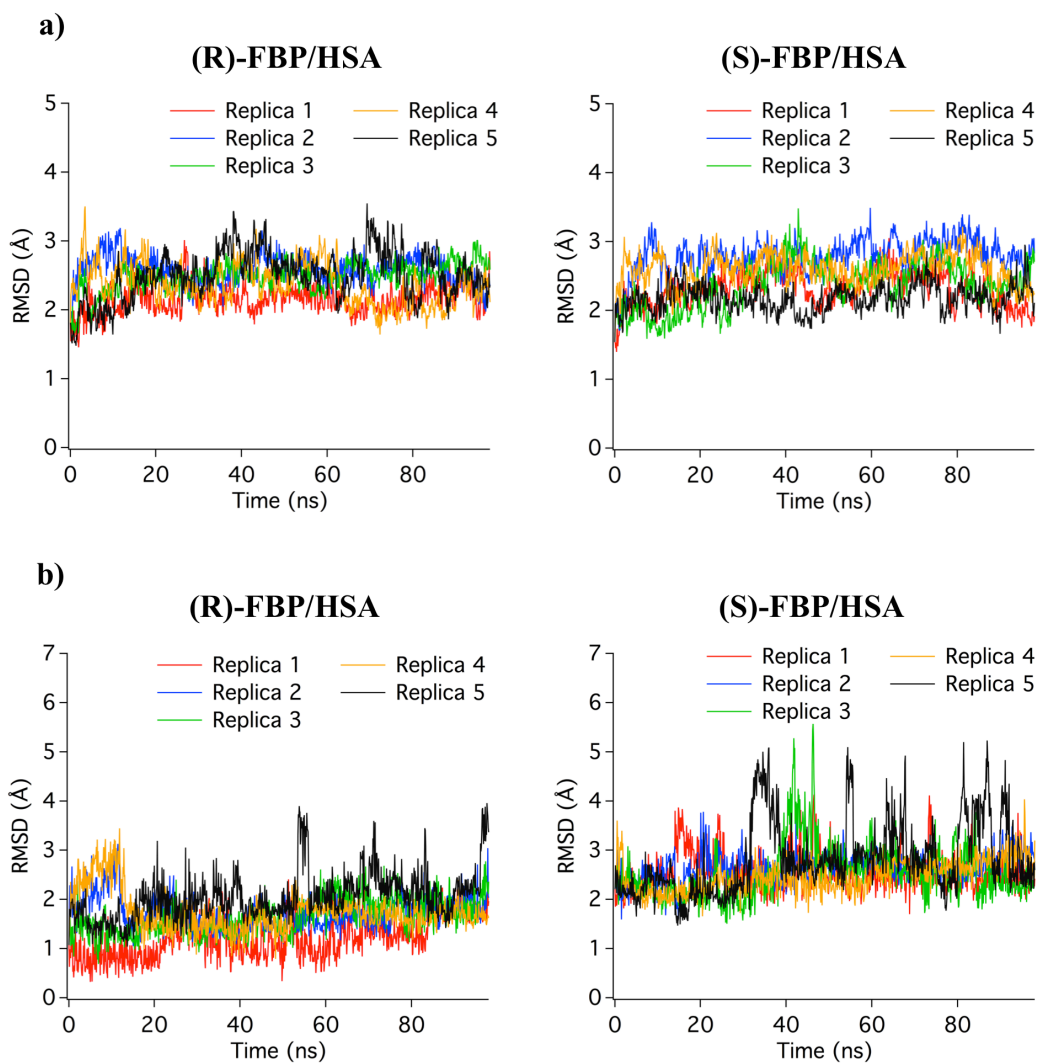


Figure S1. Root-mean-squared deviation of atomic positions calculated along each MD replica of (R)- and (S)-FBP/HSA complexes with respect to the initial structure: a) protein backbone and FBP ligand; b) FBP ligand. All structures have been aligned based on the protein backbone.

Table S1. Transition energies and transition dipole moments calculated for the (S,S)- and (R,S)-FBP–TrpMe model dyads at different QM/PCM levels of theory. Values are averaged over the Boltzmann populations of the conformers.

		FBP		Trp	
		$\langle\Delta E\rangle$ (eV)	$\langle\mu^T\rangle$ (a.u.)	$\langle\Delta E\rangle$ (eV)	$\langle\mu^T\rangle$ (a.u.)
ZINDO					
D-A	R	4.37	2.51	4.23	1.72
	S	4.39	2.34	4.23	1.72
D-B-A	R	4.24	2.58	4.12	1.72
	S	4.25	2.32	4.12	1.84
CIS					
D-A	R	5.59	2.01	5.67	1.38
	S	5.62	1.97	5.68	1.37
D-B-A	R	5.51	2.31	5.61	1.34
	S	5.54	2.28	5.61	1.35
TD-CAM-B3LYP					
D-A	R	5.15	1.68	5.17	0.95
	S	5.17	1.61	5.17	0.94
D-B-A	R	5.05	2.11	5.01	0.93
	S	5.05	2.03	5.01	0.90

Table S2. Interchromophoric distances, orientation factors, screening factors and electronic couplings calculated for the (S,S)- and (R,S)-FBP–TrpMe model dyads at different QM/PCM levels of theory. Values are averaged over the Boltzmann populations of the conformers.

		$\langle R \rangle$ (Å)	$\langle \kappa^2 \rangle$	$\langle V_{dip-dip}^2 \rangle^a$ (cm ⁻¹)	$\langle s^2 \rangle$	$\langle V_{QM/PCM}^2 \rangle$ (cm ⁻¹)
ZINDO						
D-A	R	4.14	0.10	7.68×10^4	0.42	3.95×10^3
	S	4.14	0.13	1.12×10^5	0.37	1.56×10^4
D-B-A	R	4.16	0.10	7.56×10^4	0.33	2.17×10^3
	S	4.16	0.08	7.20×10^4	0.36	7.26×10^3
CIS						
D-A	R	4.14	0.07	2.13×10^4	0.58	1.37×10^4
	S	4.14	0.13	5.53×10^4	0.55	3.75×10^4
D-B-A	R	4.16	0.06	2.05×10^4	0.51	2.20×10^4
	S	4.16	0.09	5.28×10^4	0.54	3.59×10^4
TD-CAM-B3LYP						
D-A	R	4.14	0.09	1.03×10^4	0.64	4.88×10^3
	S	4.14	0.18	2.33×10^4	0.55	9.29×10^3
D-B-A	R	4.16	0.07	1.57×10^4	0.49	3.11×10^4
	S	4.16	0.12	3.04×10^4	0.53	2.78×10^4

^aDipole-dipole couplings include Förster screening factor $s^2 = 1/n^4 = 0.31$ ($n^2 = 1.806$ in acetonitrile).

Table S3. Transition energies and transition dipole moments, as well as the corresponding standard deviations, calculated for the (S)- and (R)-FBP/HSA complexes at different QM/MMPol levels of theory. Values are averaged over the MD trajectories.

		FBP		Trp	
		$\langle\Delta E\rangle$ (eV)	$\langle\mu^T\rangle$ (a.u.)	$\langle\Delta E\rangle$ (eV)	$\langle\mu^T\rangle$ (a.u.)
ZINDO					
R		4.14	2.27	4.20	1.31
		± 0.20	± 0.35	± 0.16	± 0.24
S		4.15	2.26	4.17	1.34
		± 0.17	± 0.31	± 0.13	± 0.22
CIS					
R		5.48	1.95	5.59	1.04
		± 0.24	± 0.36	± 0.18	± 0.16
S		5.48	1.97	5.58	1.04
		± 0.23	± 0.34	± 0.12	± 0.16
TD-CAM-B3LYP					
R		5.08	2.04	5.07	0.90
		± 0.37	± 0.36	± 0.35	± 0.15
S		5.02	2.04	4.92	0.84
		± 0.28	± 0.34	± 0.22	± 0.10

Table S4. Interchromophoric distances, orientation factors, screening factors and electronic couplings, as well as the corresponding standard deviations, calculated for the (S)- and (R)-FBP/HSA complexes at different QM/MMPol levels of theory. Values are averaged over the MD trajectories.

	$\langle R \rangle$ (Å)	$\langle \kappa^2 \rangle$	$\langle V_{dip-dip}^2 \rangle^a$ (cm ⁻¹)	$\langle s^2 \rangle^b$	$\langle V_{QM/MMPol}^2 \rangle$ (cm ⁻¹)
ZINDO					
R	11.01	0.12	166	0.31	173
	± 0.56	± 0.14	± 241	± 0.09	± 180
S	10.52	0.27	800	0.30	324
	± 0.88	± 0.31	± 1412	± 0.10	± 331
CIS					
R	11.01	0.23	146	0.38	128
	± 0.56	± 0.20	± 148	± 0.12	± 311
S	10.52	0.26	311	0.35	229
	± 0.88	± 0.28	± 557	± 0.11	± 288
TD-CAM-B3LYP					
R	11.01	0.28	127	0.38	76
	± 0.56	± 0.24	± 112	± 0.10	± 123
S	10.52	0.26	176	0.35	119
	± 0.88	± 0.27	± 272	± 0.11	± 154

^aDipole-dipole couplings include Förster screening factor $s^2 = 1/n^4 = 0.25$ (n^2 is assumed to be 2.0 in a protein). ^bScreening factors averaged over structures with $V_{QM/MMPol} > 5$ cm⁻¹ to avoid numerical inaccuracies in the ratio of Eq. 9.

Table S5. Atom names, optimized geometry Cartesian coordinates, atom types and point charges derived for (R)-FBP.

Atom name	Coordinates (Å)			Atom type	Charge (a.u.)
C1	4.572	-1.179	-0.664	CA	-0.191881
C2	3.177	-1.210	-0.564	CA	-0.121912
C3	2.473	-0.148	0.036	CE	0.075782
C4	3.204	0.950	0.527	CA	-0.121912
C5	4.598	0.984	0.420	CA	-0.191881
C6	5.289	-0.083	-0.170	CA	-0.133100
H1	5.098	-2.007	-1.136	HA	0.138117
H2	2.616	-2.048	-0.972	HA	0.131652
H3	2.672	1.774	0.995	HA	0.131652
H4	5.147	1.840	0.807	HA	0.138117
H5	6.373	-0.055	-0.252	HA	0.125012
C7	1.004	-0.219	0.172	CE	-0.009636
C8	0.159	0.848	-0.168	CA	0.188723
C9	0.375	-1.385	0.651	CA	-0.243923
C10	-1.222	0.799	-0.050	CA	-0.205155
C11	-1.014	-1.457	0.785	CA	-0.172587
H6	0.999	-2.228	0.943	HA	0.141140
C12	-1.837	-0.368	0.443	CA	0.024484
H7	-1.842	1.626	-0.387	HA	0.197291
H8	-1.474	-2.368	1.163	HA	0.137158
F1	0.724	1.982	-0.669	F	-0.206646
C13	-3.330	-0.429	0.612	CT	0.076163
C14	-3.788	0.567	1.685	CT	-0.160624
H9	-3.596	-1.448	0.915	HC	-0.020891
H10	-4.870	0.482	1.839	HC	0.023584
H11	-3.578	1.587	1.352	HC	0.023584
H12	-3.279	0.384	2.641	HC	0.023584
C15	-4.046	-0.138	-0.764	C	0.793372
O1	-4.023	1.071	-1.116	O2	-0.794634
O2	-4.535	-1.142	-1.334	O2	-0.794634

Table S6. Atom names, optimized geometry Cartesian coordinates, atom types and point charges derived for (S)-FBP.

Atom name	Coordinates (Å)			Atom type	Charge (a.u.)
C1	-4.573	-1.178	-0.665	CA	-0.188793
C2	-3.178	-1.209	-0.565	CA	-0.116989
C3	-2.473	-0.148	0.035	CE	0.072636
C4	-3.203	0.950	0.528	CA	-0.116989
C5	-4.598	0.984	0.422	CA	-0.188793
C6	-5.290	-0.082	-0.169	CA	-0.141630
H1	-5.099	-2.004	-1.137	HA	0.136498
H2	-2.618	-2.047	-0.974	HA	0.130612
H3	-2.671	1.773	0.997	HA	0.130612
H4	-5.146	1.840	0.810	HA	0.136498
H5	-6.373	-0.053	-0.250	HA	0.127615
C7	-1.004	-0.220	0.171	CE	-0.011268
C8	-0.159	0.847	-0.168	CA	0.197490
C9	-0.376	-1.386	0.649	CA	-0.252411
C10	1.222	0.798	-0.050	CA	-0.213892
C11	1.013	-1.458	0.784	CA	-0.165662
H6	-1.000	-2.229	0.940	HA	0.143017
C12	1.837	-0.369	0.442	CA	0.023772
H7	1.842	1.626	-0.386	HA	0.199400
H8	1.474	-2.369	1.160	HA	0.134294
F1	-0.724	1.982	-0.669	F	-0.209020
C13	3.330	-0.430	0.611	CT	0.094819
C14	3.787	0.564	1.686	CT	-0.174648
H9	3.596	-1.449	0.914	HC	-0.025482
H10	4.869	0.480	1.841	HC	0.026139
H11	3.577	1.585	1.356	HC	0.026139
H12	3.278	0.379	2.642	HC	0.026139
C15	4.047	-0.136	-0.763	C	0.789755
O1	4.023	1.073	-1.114	O2	-0.794929
O2	4.537	-1.139	-1.335	O2	-0.794929

Table S7. Atom names, optimized geometry Cartesian coordinates, atom types and point charges derived for (R,S)-FBP–TrpMe.

Atom name	Coordinates (Å)			Atom type	Charge (a.u.)
N1	2.292	-0.692	0.392	N	-0.388481
H1	2.122	0.191	0.851	H	0.218423
C1	1.535	-1.776	0.718	C	0.498187
O1	1.660	-2.852	0.136	O	-0.540981
C2	0.511	-1.558	1.845	CT	0.043296
C3	0.537	-2.755	2.813	CT	-0.202119
H2	1.527	-2.862	3.267	HC	0.066906
H3	-0.200	-2.612	3.609	HC	0.066906
H4	0.308	-3.682	2.282	HC	0.066906
H5	0.796	-0.656	2.400	HC	0.050306
C4	-0.874	-1.315	1.256	CA	-0.020674
C5	-1.571	-0.140	1.559	CA	-0.178638
C6	-1.492	-2.255	0.416	CA	-0.131444
C7	-2.843	0.064	1.038	CA	0.194477
C8	-2.767	-2.021	-0.092	CA	-0.206623
H6	-0.967	-3.168	0.156	HA	0.155866
C9	-3.493	-0.850	0.202	CE	-0.009502
H7	-3.233	-2.773	-0.722	HA	0.149447
H8	-1.139	0.630	2.191	HA	0.147983
C10	-4.849	-0.620	-0.363	CE	0.062540
C11	-5.117	-0.952	-1.703	CA	-0.115400
C12	-5.897	-0.101	0.417	CA	-0.115400
C13	-6.392	-0.778	-2.245	CA	-0.194714
H9	-4.316	-1.334	-2.329	HA	0.134513
C14	-7.171	0.073	-0.125	CA	-0.194714
H10	-5.716	0.159	1.454	HA	0.134513
C15	-7.425	-0.264	-1.457	CA	-0.089130
H11	-6.574	-1.037	-3.284	HA	0.146935
H12	-7.968	0.470	0.498	HA	0.146935
H13	-8.417	-0.127	-1.877	HA	0.128621
F1	-3.460	1.239	1.355	F	-0.194280
C16	3.230	-0.737	-0.713	CT	-0.040931
C17	3.846	0.664	-0.975	CT	-0.043731
H14	4.508	0.916	-0.138	HC	0.055506
H15	4.482	0.577	-1.862	HC	0.055506
C18	2.828	1.755	-1.159	C*	-0.099685
C19	2.122	2.052	-2.301	CW	-0.142682
C20	2.360	2.671	-0.144	CB	0.058739
H16	2.175	1.599	-3.282	H4	0.200039
N2	1.244	3.092	-2.058	NA	-0.392419
C21	1.362	3.492	-0.742	CN	0.152332
C22	2.690	2.883	1.210	CA	-0.143474

H17	0.608	3.484	-2.733	H	0.362481
C23	0.690	4.492	-0.031	CA	-0.231607
C24	2.023	3.876	1.921	CA	-0.211486
H18	3.463	2.291	1.694	HA	0.142787
H19	-0.072	5.107	-0.502	HA	0.161148
C25	1.031	4.671	1.306	CA	-0.141957
H20	2.271	4.048	2.964	HA	0.150175
H21	0.527	5.437	1.887	HA	0.150106
C26	4.363	-1.728	-0.421	C	0.803941
O2	4.868	-1.912	0.665	O	-0.578543
O3	4.781	-2.315	-1.560	OS	-0.425515
C27	5.868	-3.251	-1.423	CT	-0.025497
H22	6.751	-2.754	-1.015	H1	0.088967
H23	5.575	-4.070	-0.762	H1	0.088967
H24	6.062	-3.619	-2.430	H1	0.088967
H25	2.712	-1.083	-1.613	H1	0.087208

Table S8. Atom names, optimized geometry Cartesian coordinates, atom types and point charges derived for (S,S)-FBP–TrpMe.

Atom name	Coordinates (Å)			Atom type	Charge (a.u.)
N1	2.094	-1.441	-0.580	N	-0.238756
H1	1.192	-1.011	-0.423	H	0.128352
C1	-6.245	0.136	0.133	CA	-0.188015
C2	-5.128	-0.688	-0.014	CA	-0.091998
C3	-3.865	-0.273	0.446	CE	0.012142
C4	-3.757	0.993	1.050	CA	-0.091998
C5	-4.876	1.816	1.195	CA	-0.188015
C6	-6.125	1.391	0.737	CA	-0.106030
H2	-7.212	-0.206	-0.223	HA	0.148918
H3	-5.238	-1.662	-0.477	HA	0.128463
H4	-2.788	1.341	1.396	HA	0.128463
H5	-4.769	2.791	1.662	HA	0.148918
H6	-6.996	2.030	0.849	HA	0.138622
C7	-2.666	-1.143	0.334	CE	-0.061672
C8	-2.356	-1.887	-0.810	CA	0.306600
C9	-1.740	-1.252	1.390	CA	-0.112960
C10	-1.218	-2.674	-0.932	CA	-0.247274
C11	-0.591	-2.033	1.295	CA	-0.208829
H7	-1.942	-0.714	2.312	HA	0.137441
C12	-0.303	-2.757	0.126	CA	-0.004021
H8	-1.051	-3.206	-1.864	HA	0.147536
H9	0.089	-2.075	2.140	HA	0.164657
F1	-3.182	-1.822	-1.894	F	-0.209040
C13	0.963	-3.584	-0.025	CT	0.045144
C14	1.220	-4.539	1.157	CT	-0.185672
H10	0.845	-4.205	-0.923	HC	0.099232
H11	1.380	-3.992	2.091	HC	0.064016
H12	0.365	-5.208	1.295	HC	0.064016
H13	2.114	-5.132	0.960	HC	0.064016
C15	2.250	-2.769	-0.319	C	0.359308
O1	3.342	-3.331	-0.363	O	-0.528361
C16	3.211	-0.564	-0.858	CT	0.029517
H14	4.012	-1.184	-1.271	H1	0.105169
C17	2.828	0.508	-1.920	CT	-0.188404
C18	3.830	0.091	0.383	C	0.685562
H15	3.696	1.160	-2.053	HC	0.092959
H16	2.675	-0.024	-2.864	HC	0.092959
C19	1.602	1.319	-1.599	C*	-0.047190
O2	4.817	0.798	0.321	O	-0.576661
C20	0.322	1.053	-2.034	CW	-0.152710
C21	1.512	2.521	-0.797	CB	0.068073
H17	-0.034	0.257	-2.674	H4	0.177276

N2	-0.554	2.005	-1.550	NA	-0.318103
C22	0.146	2.920	-0.788	CN	0.078704
C23	2.452	3.309	-0.105	CA	-0.170968
H18	-1.548	2.020	-1.720	H	0.337217
C24	-0.300	4.059	-0.108	CA	-0.198014
C25	2.014	4.444	0.571	CA	-0.207906
H19	3.503	3.036	-0.097	HA	0.176741
H20	-1.347	4.349	-0.116	HA	0.139797
C26	0.651	4.814	0.572	CA	-0.153058
H21	2.732	5.060	1.104	HA	0.150275
H22	0.341	5.706	1.108	HA	0.148796
O3	3.185	-0.194	1.526	OS	-0.302478
C27	3.747	0.385	2.722	CT	-0.058394
H23	4.781	0.057	2.850	H1	0.089213
H24	3.715	1.476	2.663	H1	0.089213
H25	3.122	0.025	3.538	H1	0.089213

Table S9. Modified force field parameter file derived for (R)- and (S)-FBP (Amber frcmod file).

Parameters for (R)- and (S)-FBP

MASS

CE 12.01	0.360	sp2 C pure aromatic (benzene)		
----------	-------	-------------------------------	--	--

BOND

CA-CE	469.0	1.400	JCC,7,(1986),230; BENZENE,PHE,TRP,TYR	
CE-CE	469.0	1.400	JCC,7,(1986),230; BENZENE,PHE,TRP,TYR	

ANGLE

CA-CA-CE	63.0	120.00	changed from 85.0 bsd on C6H6 nmodes	
CA-CE-CA	63.0	120.00	changed from 85.0 bsd on C6H6 nmodes	
CA-CE-CE	63.0	120.00	changed from 85.0 bsd on C6H6 nmodes	
CE-CA-HA	50.0	120.00		
F-CA-CE	70.0	121.00	Junmei et al,99	
CA-CT-C	64.130	110.990	Taken from c -c3-ca (GAFF)	

DIHE

CA-CE-CE-CA	1	1.000	180.0	2.000	Parametrized torsion
X-CA-CE-X	4	14.50	180.0	2.000	intrapol.bsd.on C6H6

IMPROPER

CA-CA-CA-HA	1.1	180.0	2.0	General improper torsional angle (2 general atom types)	
CA-CE-CA-HA	1.1	180.0	2.0	General improper torsional angle (2 general atom types)	
CA-CE-CA-F	1.1	180.0	2.0	Using default value	
CT-O2-C -O2	10.5	180.0	2.0	General improper torsional angle (1 general atom type)	

NONBON

CE	1.9080	0.0860	Spellmeyer		
----	--------	--------	------------	--	--

Table S10. Modified force field parameter file derived for (R,S)- and (S,S)-FBP–TrpMe (Amber frcmod file).

Parameters for (R)- and (S)-FBPMASS

CE 12.01	0.360	sp2 C pure aromatic (benzene)		
----------	-------	-------------------------------	--	--

BOND

CA-CE	469.0	1.400	JCC,7,(1986),230; BENZENE,PHE,TRP,TYR	
CE-CE	469.0	1.400	JCC,7,(1986),230; BENZENE,PHE,TRP,TYR	

ANGLE

CA-CA-CE	63.0	120.00	changed from 85.0 bsd on C6H6 nmodes	
CA-CE-CA	63.0	120.00	changed from 85.0 bsd on C6H6 nmodes	
CA-CE-CE	63.0	120.00	changed from 85.0 bsd on C6H6 nmodes	
CE-CA-HA	50.0	120.00		
F -CA-CE	70.0	121.00	Junmei et al,99	
CA-CT-C	64.130	110.990	Taken from c -c3-ca (GAFF)	

DIHE

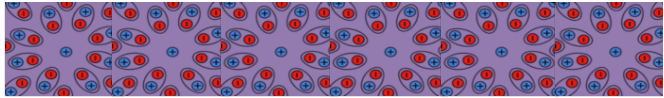
CA-CE-CE-CA	1	1.000	180.0	2.000	Parametrized torsion
X -CA-CE-X	4	14.50	180.0	2.000	intrpol.bsd.on C6H6

IMPROPER

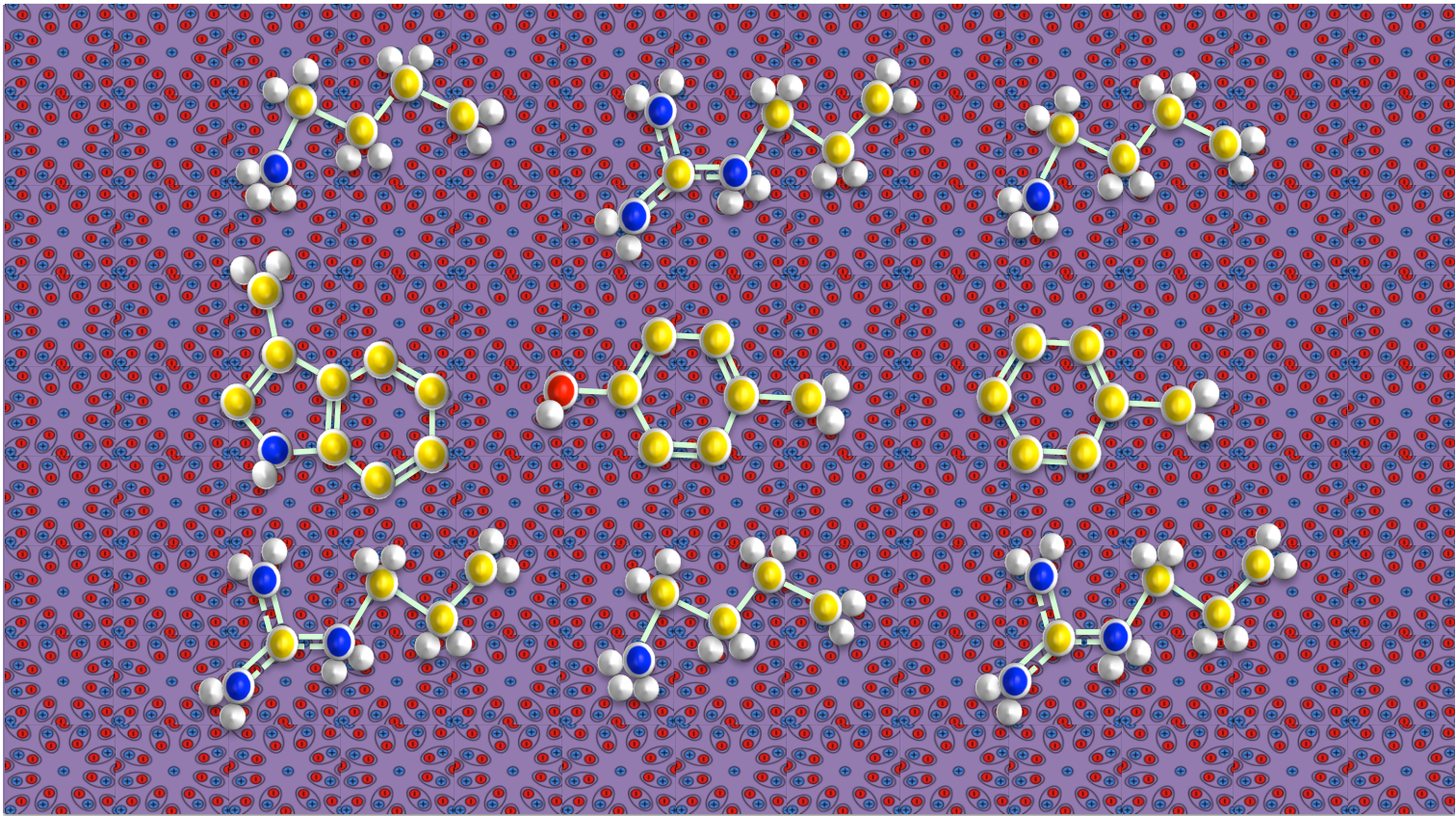
CA-CA-CA-HA	1.1	180.0	2.0	General improper torsional angle (2 general atom types)
CA-CE-CA-HA	1.1	180.0	2.0	General improper torsional angle (2 general atom types)
CA-CE-CA-F	1.1	180.0	2.0	Using default value
CT-N -C -O	10.5	180.0	2.0	General improper torsional angle (2 general atom types)
CT-O -C -OS	10.5	180.0	2.0	General improper torsional angle (2 general atom types)
C*-H4-CW-NA	1.1	180.0	2.0	General improper torsional angle (2 general atom types)
CN-CW-NA-H	1.0	180.0	2.0	General improper torsional angle (2 general atom types)
CA-CB-CN-NA	1.1	180.0	2.0	Using default value
CA-CB-CA-HA	1.1	180.0	2.0	General improper torsional angle (2 general atom types)
CA-CN-CA-HA	1.1	180.0	2.0	General improper torsional angle (2 general atom types)

NONBON

CE	1.9080	0.0860	Spellmeyer	
----	--------	--------	------------	--



UNIVERSITAT DE
BARCELONA



Conselho Nacional de Desenvolvimento
Científico e Tecnológico

