



Universitat Autònoma de Barcelona

**ADVERTIMENT.** L'accés als continguts d'aquesta tesi queda condicionat a l'acceptació de les condicions d'ús establertes per la següent llicència Creative Commons:  [http://cat.creativecommons.org/?page\\_id=184](http://cat.creativecommons.org/?page_id=184)

**ADVERTENCIA.** El acceso a los contenidos de esta tesis queda condicionado a la aceptación de las condiciones de uso establecidas por la siguiente licencia Creative Commons:  <http://es.creativecommons.org/blog/licencias/>

**WARNING.** The access to the contents of this doctoral thesis it is limited to the acceptance of the use conditions set by the following Creative Commons license:  <https://creativecommons.org/licenses/?lang=en>



Universitat Autònoma de Barcelona

---

**La Pedra Rosetta de les Proteïnes  
Fluorescents: una Descripció  
Atomística**

---

TESI DOCTORAL

*Autor:*

Pau Armengol Torrella

*Supervisors:*

Dr. Ricard Gelabert

Dr. Miquel Moreno

Department de Química  
Unitat de Química Física

2017



# Declaració

Memòria presentada per aspirar al grau de Doctor per

Pau Armengol Torrella

vist-i-plau

Ricard Gelabert Peiri    Miquel Moreno Ferrer

Bellaterra, 2 de juny del 2017



*"I have a friend who's an artist and has sometimes taken a view which I don't agree with very well. He'll hold up a flower and say "look how beautiful it is", and I'll agree. Then he says "I as an artist can see how beautiful this is but you as a scientist take this all apart and it becomes a dull thing", and I think that he's kind of nutty. First of all, the beauty that he sees is available to other people and to me too, I believe. Although I may not be quite as refined aesthetically as he is ... I can appreciate the beauty of a flower. At the same time, I see much more about the flower than he sees. I could imagine the cells in there, the complicated actions inside, which also have a beauty. I mean it's not just beauty at this dimension, at one centimeter; there's also beauty at smaller dimensions, the inner structure, also the processes. The fact that the colors in the flower evolved in order to attract insects to pollinate it is interesting; it means that insects can see the color. It adds a question: does this aesthetic sense also exist in the lower forms? Why is it aesthetic? All kinds of interesting questions which the science knowledge only adds to the excitement, the mystery and the awe of a flower. It only adds. I don't understand how it subtracts."*

Richard P. Feynman



## Agraïments

El procés de creació d'una tesi doctoral, amb la recerca que comporta, és un procés llarg i a vegades extenuant, però sobretot és un procés que, en el meu cas, hauria estat impossible sense els inestimables suports rebuts. Ja d'antuvi em disculpo per la gent que em pugui deixar...

Jo vaig entrar en el món de la recerca científica de la mà del Ricard Gelabert, que després ha esdevingut el meu director de la tesi doctoral que teniu a les mans, juntament amb el Miquel Moreno. El Ricard havia estat el meu professor de Química Quàntica i de Programació i, sense dubte, ell va ser la persona que em va despertar l'interès pel món que he explorat els darrers anys. L'estiu posterior a tenir el Ricard de professor vaig començar a col·laborar amb el grup de recerca "Dinàmica i Mecanismes de les Reaccions Químiques i Bioquímiques", on finalment he desenvolupat la meva tesi doctoral, i a ell li haig d'agrair, entre altres coses, que em guiés en els meus primers passos de la meva vida com a investigador, en aquell calorós estiu. Durant el doctorat ha estat, sense dubte, un punt de suport intel·lectual i anímic i les hores al despatx han donat lloc a les més variades converses: des de resolucions de dubtes de programació, a classes improvisades de teoria, però també expansives dissertacions sobre filosofia de la ciència o, perquè no, història de la humanitat. També vull agrair especialment al Miquel (que ha compartit la direcció de la tesi amb el Ricard) i al Josep Maria Lluch les nombroses discussions sobre proteïnes fluorescents i tot el què les acompanya, així com el consell i recolzament rebut. També a la resta de membres sèniors del grup, la Laura Masgrau, la Mireia Garcia i especialment a l'Àngels Gonzàlez. Durant aquests anys també he pogut comptar amb la companyonia dels altres membres junior del grup, passats i presents, amb els quals he tingut el gust de coincidir: el Carlos, la Patricia, la Maria Fernanda, l'Anna, l'Alejandro, el Jordi i, especialment, el Marc, que era el doctorand amb més experiència quan jo vaig començar el meu camí al grup, i que em va ajudar en els durs inicis de la química computacional.

Vull agrair als meus amics i família el recolzament que m'han donat en tot moment. Sense cap mena de dubte, sense el seu suport el camí hauria estat molt més complicat.

També haig d'agrair sincerament l'ajuda prestada durant el desenvolupament de la recerca que a desembocat en aquesta tesi a un bon nombre de persones. Als tècnics: el Carles, la Irene, l'Elisabeth i finalment el Sergi, per haver estat sempre disposats a solucionar tots els problemes informàtics que han sorgit. Al Walter Thiel li vull agrair que m'acceptés al seu



grup per fer una estada que, encara que curta, ha resultat molt profitosa i al Lasse Spörkel li agraeixo l'acolliment i la companyonia que em va brindar en terres germàniques. També vull agrair a la Neus Sánchez la seva implicació en la correcció ortogràfica d'aquest document.

Pel què fa al suport insitucional i econòmic, haig d'agrair a la Universitat Autònoma de Barcelona la beca de Personal Investigador en Formació (PIF) que he gaudit dels anys 2013 a 2017. A l'Agència de Gestió d'Ajuts Universitaris i de Recerca (AGAUR) del departament d'Empresa i Coneixement de la Generalitat de Catalunya, la concessió de la Beca per a les Activitats Acadèmiques Dirigides al Suport al Professorat (AAD) durant el període 2014-2016. Al Max-Planck-Istitut für Kohlenforschung a Mülheim an der Ruhr per l'acolliment durant la meva estada doctoral. Finalment cal agrair el finançament rebut durant el projecte al Ministerio de Economía y Competitividad mitjançant els projectes CTQ2011-24292 i PK612960 i a la Generalitat de Catalunya a través del projecte 2009SGR409.

# Índex

<b>Declaració</b>	<b>iii</b>
<b>I Exordium</b>	<b>1</b>
<b>1 Introducció</b>	<b>3</b>
1.1 Pròleg: Avís per a navegants . . . . .	3
1.2 Proteïnes Fluorescents . . . . .	6
1.2.1 Brevíssima Historia de les Proteïnes Fluorescents . . .	7
1.2.2 Mecanisme de Fluorescència de la GFP . . . . .	9
1.2.3 La Paleta de Colors . . . . .	11
L'Embolcall . . . . .	12
1.2.4 La Pedra Rosetta: la Mutant S65T/H148D-GFP . . . .	15
<b>2 Marc Teòric</b>	<b>19</b>
2.1 Càlcul de l'Energia Electrònica . . . . .	20
La Separació Espai-temporal . . . . .	20
L'Aproximació de Born-Oppenheimer . . . . .	22
2.1.1 La Teoria del Funcional de la Densitat . . . . .	23
La Teoria del Funcional de la Densitat Dependent del Temps . . . . .	24
2.1.2 Els Mètodes Semiempírics . . . . .	26
2.1.3 Mecànica Molecular . . . . .	27
2.1.4 Els Mètodes Híbrids QM/MM . . . . .	29
El Codi ChemShell . . . . .	31
2.2 Dinàmica . . . . .	32
2.2.1 Dinàmica Clàssica . . . . .	33
2.2.2 Dinàmica Semi-clàssica . . . . .	34
2.2.3 Dinàmica Quàntica . . . . .	38
<b>3 Objectius</b>	<b>41</b>
<b>II Narratio et Confirmatio</b>	<b>43</b>
<b>4 Task-Farming</b>	<b>45</b>
4.1 El Problema de la Funció de Partició Canònica . . . . .	45
4.2 <i>Tous pour Un, Un pour Tous</i> : el Mètode <i>Task-farming</i> . . . . .	48
<b>5 Propietats Espectroscòpies de les Proteïnes Fluorescents</b>	<b>51</b>

5.1	Proteïnes Verdes Fluorescents . . . . .	51
5.1.1	Metodologia . . . . .	51
	Sistemes Model . . . . .	52
	Simulacions de Dinàmica Molecular . . . . .	53
	Càlculs Estàtics . . . . .	54
5.1.2	Resultats . . . . .	55
	Resultats Estructurals . . . . .	55
	Espectres d'Absorció . . . . .	56
5.2	Proteïnes Vermelles Fluorescents . . . . .	60
5.2.1	Metodologia . . . . .	60
	Sistemes Model . . . . .	61
	Simulacions de Dinàmica Molecular . . . . .	63
	Càlculs Estàtics . . . . .	63
5.2.2	Resultats . . . . .	63
	Espectres d'Absorció . . . . .	63
	Naturalesa dels Estats Excitats . . . . .	65
	Una Nova Distribució de Càrregues . . . . .	69
	Anàlisi de les Interaccions del Cromòfor . . . . .	70
5.3	<i>Red Hot Chili Proteins</i> : Proposta d'una Nova RFP . . . . .	74
5.3.1	Selecció de Candidates . . . . .	74
	Metodologia . . . . .	77
5.3.2	Resultats . . . . .	77
	El Mètode OM2 . . . . .	79
5.4	Conclusions Parcial . . . . .	81
	Conclusions Metodològiques . . . . .	81
	Conclusions Químiques . . . . .	82
<b>6</b>	<b>Desxifrant la Pedra Rosetta</b> . . . . .	<b>85</b>
6.1	Metodologia . . . . .	85
	Càlculs Estàtics . . . . .	85
	Càlculs Dinàmics . . . . .	85
6.2	La Factibilitat de la ESPT . . . . .	87
6.2.1	Perfils Estàtics . . . . .	87
6.2.2	L'Acidesa del Cromòfor, a Prova . . . . .	89
6.2.3	Dinàmica QM/MM . . . . .	92
	L'Energètica de la Transferència Protònica . . . . .	93
	Perfils Dinàmics . . . . .	96
	La Inclusió de l'Energia Cinètica del Protó . . . . .	98
6.2.4	Una Nova Mutant . . . . .	100
	L'Efecte Papallona en una FP: Resultats Estructurals . . . . .	101
	L'Energètica de la Transferència Protònica . . . . .	103
6.3	Química Ultrarràpida a Càmera Lenta: la Vida de la Doble Mutant a l'Estat Fotoactiu . . . . .	105
6.3.1	Canals de Relaxació no Radiativa . . . . .	106
6.3.2	Espectres de Fluorescència . . . . .	112
6.3.3	La Transferència del Protó . . . . .	115
6.3.4	La Coordenada Complexa: la Relaxació de l'Entorn . . . . .	116

6.3.5	El Cau del Conill: el Destí Ocult de l'Espècie A*	118
6.4	Conclusions Parcial	120
6.4.1	Conclusions Metodològiques	120
6.4.2	Conclusions Químiques	121
<b>III</b>	<b>Peroratio</b>	<b>123</b>
<b>7</b>	<b>Conclusions</b>	<b>125</b>
<b>IV</b>	<b>Bibliografia</b>	<b>127</b>
<b>V</b>	<b>Apèndix</b>	<b>141</b>



# I

## EXORDIUM

Io fin qui t'ho condotto, or più non lice  
Teco venir, ch'amara legge il vieta,  
Legge iscritta col ferro in duro sasso  
De l'ima reggia in su l'orribil soglia,  
Che in queste note il fiero senso esprime:  
L'Lasciate ogni speranza, voi ch'entrate".

Speranza - L'Orfeo (C. Monteverdi)



# 1 Introducció

## 1.1 Pròleg: Avís per a navegants

Teniu a les mans una tesi doctoral. Idealment, la tesi, com a document, és la sublimació del procés de creació de coneixement. És l'eina per la qual el coneixement es pot construir doncs engloba el què entenc que és l'únic procés possible per crear-lo: descriu unes observacions, presenta uns resultats i conclou un graó més de coneixement a partir d'aquests. I el més important és que ho fa d'una manera traçable i reproduïble. Si mai algú discrepa de les conclusions presentades o de la metodologia utilitzada per la obtenció dels resultats que han desembocat en aquestes conclusions, pot seguir el procediment invers, remenant les entranyes d'aquest nou coneixement creat per, si cal, invalidar-lo o substituir-lo per un de més adequat. Així doncs, la presentació d'una tesi no deixa de ser un acte d'exhibicionisme científic, una manera de despullar-se per defensar una nova parcel·la d'univers colonitzada per la ciència.

Crec que un dels reptes que té la ciència avui dia és fer entendre al públic extra-científic quin tipus de coneixement és el coneixement científic, que vol dir fer ciència i què se'n pot esperar de la mateixa. Vivim en un món superinformat on costa molt discernir entre el què és coneixement i el què no ho és. David Roberts va fer servir el terme "post-veritat" fa pocs anys en un article on parlava de la discussió política al voltant del canvi climàtic per descriure la situació en què hi ha polítics que neguen aquest fet malgrat tota l'evidència científica que tenim avui dia per confirmar-lo, tan sols invocant arguments de tipus emocional.<sup>1</sup> Em permeto recordar que el que és el president de l'estat espanyol en el moment de redacció d'aquestes línies va gosar dir que no creia que el canvi climàtic fos un gran problema mundial"perquè així li ho va dir el seu cosí, o que el també actual president dels Estats Units es declara un negacionista del canvi climàtic de primera línia. L'any 2016, Oxford Dictionaries va erigir el terme *post-truth* (post-veritat en anglès) com a paraula de l'any<sup>2</sup> i és que, encara que inicialment va ser un concepte aplicat a la política, crec que és absolutament adaptable a qualsevol discussió on hi intervinguin les idees de coneixement o veritat. Sembla ser que actualment, "els fets"han passat a un segon terme: sovint ens estimem més descripcions senzilles i poc acurades del complex món que ens envolta, mentre no xoquin gaire amb les nostres idees preestablertes o gustos emocionals. Hi ha un problema generalitzat de pensament crític: sovint tendim a creure'ns allò que ens estimula o ens plau emocionalment enlloc d'allò que es pot considerar coneixement, allò



que compleix els requisits pertinents per poder ser elevat a tal categoria. Així doncs, adoptem una relació heurística amb la realitat. Només cal donar un cop d'ull a l'auge de les medicines pseudocientífiques per entendre que tenim un problema global de concepció de la realitat. Sovint es busquen respostes dogmàtiques a preguntes massa complexes, unes respostes que, malauradament, la ciència de vegades no pot proporcionar. Les pseudociències sempre tenen respostes per a qualsevol pregunta, però no són respostes basades en el coneixement ja que no són traçables ni falsables. Així doncs aquestes respostes, com que no són veritat, podem afirmar que són mentides.

Per tant, com hem d'entendre la ciència per tal de tenir una visió acurada i crítica de la realitat? La ciència no és més que una eina que hem creat per poder estructurar i descriure la visió que tenim del món que ens envolta. Vull remarcar que tan sols podem arribar a descriure *la visió* que tenim del món que ens envolta, no el propi món. Nosaltres, com a éssers humans, però abans com a mamífers, animals, organismes vius, experimentem un univers al nostre voltant i ens en fem imatges mentals. Amb la ciència, intentem descriure aquesta imatge. En cap cas serem capaços de descriure res que sigui fora de la imatge mental que ens fem del món, ja que fora de la nostra percepció, no hi ha res que puguem experimentar. De fet, ni tan sols podem estar segurs que hi hagi res fora del nostre pensament. El Geni Maligne<sup>3</sup> de Descartes és una opció que no es pot "descartar": podríem estar immersos en un món virtual sense que ho poguéssim notar, o prendre per lògiques algunes conclusions que en realitat no ho són, tot per obra d'aquest Geni Maligne. Com diu el propi Descartes en la seva cèlebre frase "*cogito ergo sum*", de l'única cosa que podem estar segurs és de la pròpia existència. Que metafísicament no puguem estar segurs de res del que ens proporcionen els nostres sentits o la nostra raó, no vol dir que no sigui convenient adoptar models mentals que ens permetin viure en aquest món, real o imaginari, i que a més ens permetin fer-ne prediccions. En aquest sentit, qualsevol model mental que puguem concebre que permeti explicar les nostres observacions seria un model de realitat vàlid. Stephen Hawking i Leonard Mlodinow explicaven aquesta concepció de realitat amb el següent exemple:<sup>4</sup> posem el cas d'un peix que viu en una peixera esfèrica. Ell observarà una realitat distorsionada respecte la nostra: distorsionada no vol dir menys real, tan sols diferent, doncs la nostra visió també està condicionada a altres circumstàncies. Amb la seva visió, podria observar una trajectòria corbada des de dins la peixera per a un moviment extern a la mateixa, que nosaltres interpretaríem com a rectilini. El peix podria formular lleis científiques que descrivissin aquest moviment i que sempre es complirien des de dins la peixera i, per tant, podria fer previsions sobre el comportament de l'univers "exterior" (exterior a la peixera, és clar). Les lleis del peix serien diferents a les nostres, però hauríem d'acceptar que tindrien la mateixa validesa. És evident que nosaltres també vivim dins d'una peixera, que és el nostre cervell, el nostre cos, o els nostres sentits. Per tant hem d'acceptar que qualsevol model de realitat que concordi amb les nostres observacions i ens permeti fer



FIGURA 1.1: En l'exemple de Hawking i Mlodinov, les possibles lleis científiques que desenvolupés el peix des de dins la peixera, que serien diferents a les nostres, serien igualment vàlides per al seu model de realitat. Extret de Hawking *et al.*<sup>4</sup>

prediccions, tot això de manera reproducible, serà acceptable. Hawking i Mlodinov anomenen aquesta concepció de realitat “realisme dependent del model”.

Crec que aquesta idea és principal i s'ha de tenir present en un treball com el que teniu a les mans. Anant una mica més enllà en aquesta descripció, em podria atrevir a dir que no existeix aquest llibre (com a objecte, vull dir), doncs no hi ha res fora del nostre cap que respongui a la idea de llibre. Algun observador extern que interaccionés amb el seu entorn d'alguna manera exòtica, podria no experimentar aquest llibre com a unitat i per tant es faria un model de realitat on aquest llibre no existeix com a tal, per tant seria possible un model on aquest llibre, com a tal, no existís. A nosaltres, ens és molt convenient adoptar un model de realitat on tenim objectes discrets (llibres inclosos), i per això parlem d'objectes amb naturalitat, però hem de tenir present que no és l'única manera possible d'entendre el món. És una qüestió de comoditat: el fet que acceptem que la Terra gira al voltant del Sol i no al contrari és senzillament perquè, amb les matemàtiques que hem creat, ens és molt més convenient explicar el moviment de la resta de planetes del Sistema Solar prenent un model heliocèntric però, com deia, fora del nostre pensament, ningú dona voltes al voltant de ningú ja que no existeix un sistema de referència absolut. Repeteixo, és una qüestió de conveniència del model. Per tant, seguint la *lex parsimoniae* d'Ockham, prendrem com a vàlid aquell model que minimitzi la quantitat d'assumpcions que hem de prendre per a entendre'l. Per exemple, imaginem que estem observant la presència d'una taula en una habitació. Si abandonem l'habitació per tornar-hi a entrar més tard i tornar-nos a trobar amb la taula, podríem pensar que aquesta ha desaparegut just quan hem tancat la porta per tornar a aparèixer quan hem entrat de nou a l'habitació així que podríem formular lleis físiques que expliquessin l'univers en

aquest context. Ara bé, triem fer-nos un altre model mental: la taula sempre és allà, la veiem o no. Generalment, es pren aquest últim model com a vàlid perquè és el que simplifica més clarament l'explicació de l'univers, i per tant, és la que ens és més còmode.

El tema que aquí tan sols enceto és prou extens per fer-ne unes quantes tesis doctorals addicionals, només l'he exposat fins on crec que és estrictament necessari pel seguiment de qualsevol tesi científica, però especialment aquesta. En treballs de caire teòric, com és el cas que us ocupa, haurem d'utilitzar visions diferents de la realitat per adaptar la descripció del món que observem a les nostres necessitats. A vegades podrem fer servir teories clàssiques per descriure algunes de les situacions que ens trobarem, d'altres vegades farem servir teories quàntiques, o fins hi tot aproximacions clàssiques a comportaments quàntics, entre d'altres... I sempre estarem descrivint el mateix univers. Navegant, esteu avisats.

## 1.2 Proteïnes Fluorescents

L'avanç de la tècnica sempre ha anat acompanyat del de la ciència. Sovint, la ciència es veu acotada per les limitacions tècniques a les que l'ésser humà pot accedir. Hi ha teories desenvolupades fa desenes d'anys que tot just recentment han pogut ser testades (per exemple, amb el descobriment del bosó de Higgs<sup>5,6</sup> l'any 2012, teoritzat als anys seixanta<sup>7,8</sup>). És habitual que la ciència resti a l'espera d'algun avanç tècnic (no pas teòric) que la permeti progressar. Quelcom similar ha passat amb la Química Teòrica i Computacional al llarg dels anys. L'arribada dels superordinadors ha obert un món de possibilitats per a aquesta disciplina. Càlculs que abans eren del tot inimaginables, avui dia es poden realitzar d'una manera relativament senzilla. Amb tot, la recerca és, per definició, expansiva: ningú que faci recerca bàsica està interessat en fer allò que ja es pot fer, si no en fer coses noves per tal d'estudiar casos que han restat inaccessibles fins al moment. En recerca, sovint es juga al límit de les pròpies possibilitats tècniques i per això ens cal, encara, adoptar visions aproximades de la realitat per a poder obtenir resultats concrets de sistemes complexos. Em remeto a la cèlebre cita de Dirac:

The underlying physical laws necessary for the mathematical theory of a large part of physics and the whole of chemistry are thus completely known, and the difficulty is only that the exact application of these laws leads to equations much too complicated to be soluble. It therefore becomes desirable that approximate practical methods of applying quantum mechanics should be developed, which can lead to an explanation of the main features of complex atomic systems without too much computation.

- Paul M. Dirac

L'estudi de sistemes fotobiològics de gran interès per a la comunitat científica, com ho són les Proteïnes Fluorescents o FPs (Fluorescent Proteins),<sup>9-12</sup> ha esdevingut un repte engrescador però assumible en els últims anys, doncs l'avanç de les teories de la Química així com de la possibilitat tècnica de les màquines actuals fa que, per fi, puguem estudiar i racionalitzar sistemes complexos a nivell multiescala. Amb tot, per poder entendre el gran interès científic que desperten les proteïnes fluorescents, cal contextualitzar el coneixement que es té sobre les mateixes.

### 1.2.1 Brevíssima Historia de les Proteïnes Fluorescents

A mitjans dels anys 50 del segle passat, es va observar el fenomen de la bioluminescència en la medusa *Aequorea victoria*, una medusa pròpia de les aigües del pacífic que emetia llum verda.<sup>13</sup> Aleshores, es pensava que totes les reaccions de bioluminescència es produïen per la reacció luciferina-luciferasa, i aquesta fou la sospita de la comunitat científica. En termes de bioluminescència, els complexos luciferina-luciferasa són parelles de biocompostos susceptibles de reaccionar entre si i emetre llum. Sovint, les luciferines són molècules relativament petites i actuen de substrat de les luciferases, enzims que oxiden les primeres per donar una reacció bioluminiscent. A partir d'estudis *in vitro* del complex luciferina-luciferasa trobat en l'*Aequorea victoria*, aequorina-coelenterazina, es va aconseguir l'emissió de llum, però de llum blava.<sup>14,15</sup> Faltava, encara, una peça del puzzle per entendre l'emissió de llum verda d'aquell organisme. L'any 1962 Osamu Shimomura va aconseguir aïllar aquesta peça,<sup>16</sup> la Proteïna Verda Fluorescent o GFP (Green Fluorescent Protein), d'aquí en endavant. Va extreure aquesta proteïna de la pròpia medusa i va observar que emetia llum verda al ser irradiada amb llum blava, tancant així el cicle: la reacció de l'aequorina (luciferina) amb la colenetraxina (luciferasa) produeix llum blava, que és rebuda per la GFP per mitjà d'una transferència d'energia no radiativa, qui finalment emet la característica llum verda.<sup>17,18</sup> La GFP presentà el seu màxim d'absorció als 398 nm, emetent als 508 nm.<sup>19,20</sup> La importància d'aquest descobriment rau en el fet que la GFP és "autofluorescent", és a dir que no requereix cap cofactor extern a ella per emetre fluorescència, si bé de forma natural participa d'un cicle bioluminiscent més complex.

L'any 1979 Shimomura va aconseguir deduir l'estructura del cromòfor dins la proteïna.<sup>21</sup> L'estructura 4-(*p*-hydroxybenzildè)imidazolid-5-ona del cromòfor (HBDI) va ser confirmada amb posterioritat<sup>22</sup> i es va proposar que es forma per la condensació i posterior oxidació dels residus Ser65, Tyr66 i Gly67.<sup>23</sup> La GFP havia estat descoberta, però era un gegant adormit. Com passa sovint en ciència, especialment en aquella que és multidisciplinària, va haver d'esperar que altres disciplines de la ciència i tècniques experimentals es posessin a l'altura del que el cas demanava. Amb tot, li va arribar el seu torn amb el boom de l'expressió genètica. L'any 1992 es va determinar per primera vegada la seqüència genètica de la GFP<sup>24</sup> i l'any 1994 va poder ser expressada,<sup>25</sup> tot permetent confirmar

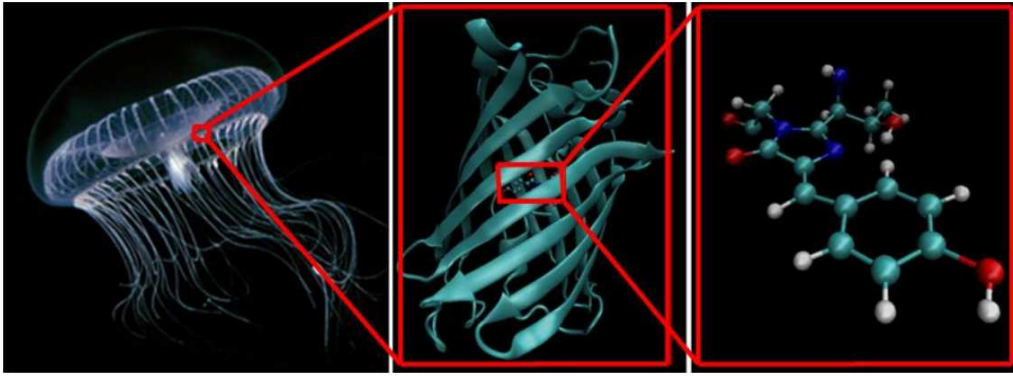


FIGURA 1.2: D'esquerra a dreta: la medusa *Aequorea victoria*, l'estructura de barril  $\beta$  de la GFP i l'estructura molecular del cromòfor

que el cromòfor es forma en una reacció autocatalítica dels residus anteriorment esmentats. L'elucidació de l'estructura de la GFP per difracció de rajos X va permetre observar com està formada d'onze fulles  $\beta$  que formen un barril, atravesat per una hèlix  $\alpha$  que conté el cromòfor.<sup>26</sup> Aquesta disposició interior del cromòfor, com veurem, és el que li dona a la GFP les seves característiques úniques (Figura 1.2).

L'interès en la GFP es va disparar un cop va aparèixer la possibilitat d'expressar-la unida a altres proteïnes. La GFP es va convertir en una opció de primer ordre pel marcatge de proteïnes *in vivo*, donades les seves característiques: és autofluorescent,<sup>17</sup> té un alt rendiment quàntic (QY) de 0,8<sup>20</sup> (en comparació amb el rendiment quàntic del cromòfor aïllat en solució, que és de  $10^{-3}$ ),<sup>27</sup> és molt resistent a una ampli rang de temperatura, pH, sals i proteases,<sup>28</sup> i donat el seu petit tamany (només té 238 aminoàcids),<sup>24</sup> es pot unir a altres proteïnes sense afectar la funció d'aquestes, a més de tenir una baixa toxicitat.<sup>25</sup> Així doncs, s'havia obert la capsula de Pandora, un marcador de proteïnes que permetia fer seguiments no invasius de processos biològics. Marcant altres proteïnes o teixits amb GFP i d'altres proteïnes fluorescents desenvolupades posteriorment, es va poder observar la reactivitat dels primers, el seu moviment, comportament, etc.<sup>29-33</sup> Potser l'aplicació de les proteïnes fluorescents, a vegades anecdòtica, però que més ha trascendit al públic no científic, és la possibilitat d'unir-les a teixits de mamífers vius creant, així, mamífers fluorescents<sup>34-36</sup> (Figura 1.3). Avui en dia, les proteïnes fluorescents són cabdals en els estudis de bioimatge i representen una eina indispensable en la biologia molecular moderna.<sup>37-39</sup>

Les utilitats que van sorgir de la possibilitat d'expressar la proteïna, amb aquest nou món d'aplicacions que s'obria, va centrar l'atenció de molts científics de camps diferents. Així doncs, l'interès per les aplicacions de la proteïna, però també per conèixer el mecanisme de la fluorescència, la relació estructura-propietats, entre d'altres, va començar a créixer. Tota aquesta producció científica que va comportar un gran avenç ja no només per l'estudi de les aplicacions de la GFP si no també per l'assentament de

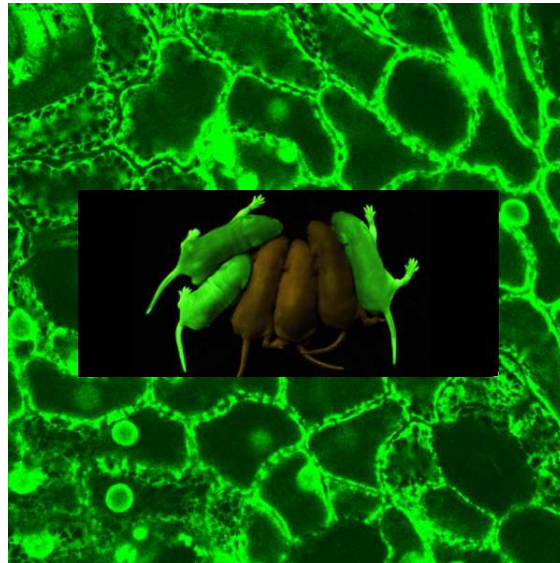


FIGURA 1.3: Fons: Imatge microscòpica de la GFP expressada en teixit vegetal. Primer pla: Ratolins en els quals ha estat expressada la GFP al costat de ratolins no modificats, tots il·luminats amb llum ultraviolada.

la GFP com a model de processos fotobiològics, va portar a que se li atorgués el Nobel de l'any 2008 a Shimomura, Chalfie i Tsien per "*the discovery and development of the green fluorescent protein, GFP*". A més, el premi Nobel de l'any 2014, atorgat a Betzig, Hell i Moerner per "*the development of super-resolved fluorescence microscopy*" està també estretament lligat a les proteïnes fluorescents ja que sense el desenvolupament de les mateixes i el gran impacte que han tingut en les tècniques de microscòpia modernes, tampoc no hauria estat possible.

### 1.2.2 Mecanisme de Fluorescència de la GFP

El fotocicle de la GFP ha estat intensament estudiat a nivell experimental i teòric, i hi ha consens sobre les generalitats del seu funcionament, encara que moltes de les subtileses del mateix no han estat encara explicades. A temperatura ambient, l'espectre d'absorció de la GFP presenta dos pics, el més intens és l'anomenat A (398 nm) i el menys intens, l'anomenat B (471 nm),<sup>19,20</sup> que s'associen a la forma neutra i ionitzada del cromòfor, respectivament.<sup>27,40,41</sup> El cromòfor, com he dit anteriorment, es forma espontàniament dins de la proteïna després de l'expressió de la mateixa per una reacció de ciclació del tripèptid Ser65-Tyr66-Gly67, seguida de la 1,2-deshidrogenació de la Tyr66.<sup>19,42</sup> El protó fenòlic, provinent originàriament de la Tyr66, i degut a la nova deslocalització del sistema  $\pi$ , guanya propietats àcides, propietats que es veuran incrementades a l'estat excitat fotoactiu. Després de l'excitació de la banda A, la GFP emet radiació a 508 nm en una escala de temps de piconsegons.<sup>43</sup> L'espècie emetent d'aquesta radiació es coneix per I (o I\* si es troba a l'estat excitat) i també correspon,

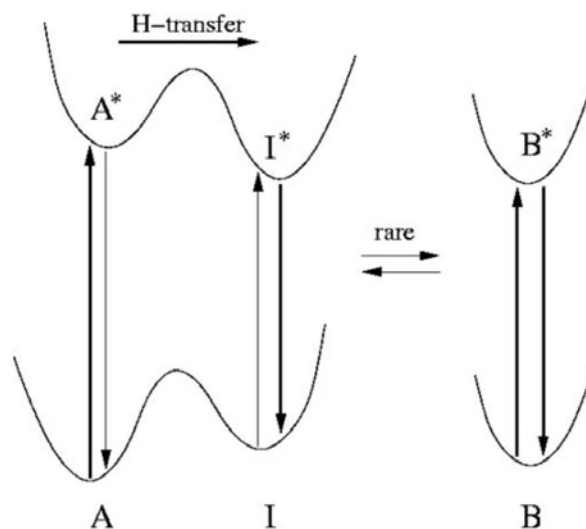


FIGURA 1.4: Representació esquemàtica del mecanisme de fluorescència de la GFP amb els possibles estats A, I o B implicats. L'excitació de l'espècie A porta a  $A^*$ , que evoluciona fins a  $I^*$  per mitja de l'ESPT. L'espècie  $I^*$  és l'emissora de la fluorescència ja que la transformació de  $I^*$  a  $B^*$  és un procés lent. Extret de Lill *et al.*<sup>46</sup>

com B, a una forma aniònica del cromòfor, essent la diferència entre I i B que en la primera, la proteïna encara no ha adaptat la seva conformació a la nova distribució electrònica. La transformació de  $I^*$  en  $B^*$  és un procés lent, ja que comporta el moviment de tota la proteïna, especialment l'entorn del cromòfor, i la de B en A, ja a l'estat fonamental, pot trigar hores<sup>44,45</sup> (Figura 1.4). S'accepta que el mecanisme de transformació de  $A^*$  a  $I^*$  implica una transferència protònica a l'estat excitat, ESPT (Excited State Proton Transfer) d'aquí en endavant, i això es recolza en el considerable efecte cinètic d'isòtop H/D ( $\sim 5$ ) trobat en l'aparició de la fluorescència de l'espècie  $I^*$ . A partir de les dades estructurals obtingudes per la difracció de rajos X en estructures cristal·lines de la GFP salvatge o *wt*-GFP<sup>26</sup> (*wild-type*-GFP) d'aquí en endavant i de la mutant S65T,<sup>47</sup> es proposa la transferència neta d'un protó a través d'un cable de protons format pel propi cromòfor, una molècula d'aigua i els residus Ser205 i Glu222, que resulta ser l'acceptor final del protó, tal i com han demostrat experiments de IR resolts en el temps,<sup>48,49</sup> i s'ha confirmat per estudis teòrics<sup>50,51</sup> (Figura 1.5). S'assumeix que la força motriu de la transferència protònica és el gran guany d'acidesa del cromòfor a l'estat excitat. Per mitjà de càlculs teòrics, s'ha estimat que la caiguda del  $pK_a$  del cromòfor en solució després de l'excitació és d'unes 8 unitats (és a dir, 100.000.000 cops més àcid).<sup>53</sup>

El fet que el rendiment quàntic de la GFP sigui de 0,8 (a comparar amb el valor de  $10^{-3}$  pel cromòfor lliure en solució), s'explica pel confinament que pateix aquest a dins la proteïna: quan es troba en solució, i després de l'excitació, el cromòfor trenca la conjugació entre l'anell fenòlic (l'anell de

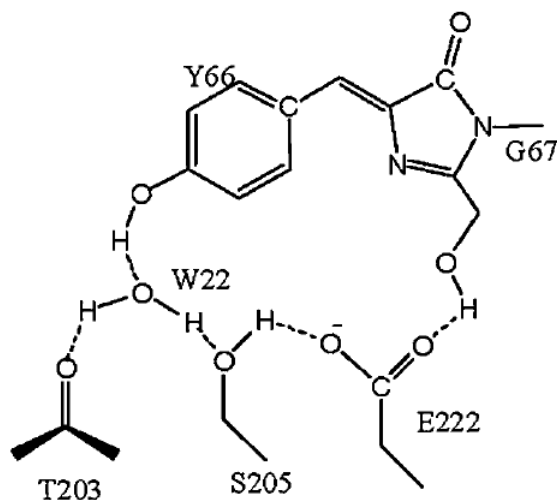


FIGURA 1.5: Representació de l'estructura del cromòfor juntament amb els altres actors del cable de protons que fa possible la ESPT: la molècula d'aigua, etiquetada com a W22 per raons de nomenclatura, la Treonina 203, la Serina 205 i el Glutamat 222, que serà l'acceptor final del protó. Extret de Di Donato *et al.*<sup>52</sup>

sis carbonis, vegeu la Figura 1.5) i l'imidazolinònic (l'anell de cinc membres, entre els quals 2 nitrògens, vegeu la Figura 1.5) permetent, així, una desactivació no radiativa en el camí de la pròpia isomerització cis-trans. La situació dins la proteïna és ben diferent, doncs la pressió que exerceixen la resta de residus sobre el cromòfor no el permeten rotar lliurement tot evitant, en gran mesura, l'accés a aquesta via de desactivació no radiativa i per tant forçant el cromòfor a tornar a l'estat fonamental per la via radiativa. A més, l'entorn proteic protegeix el cromòfor del possible *quenching* per xocs amb molècules d'aigua.<sup>54-56</sup>

### 1.2.3 La Paleta de Colors

El boom dels anys 90 i la possibilitat d'expressar les proteïnes fluorescentes, va portar ràpidament a explorar les possibilitats mutagèniques que presentava la GFP. Tsien *et al.* van sintetitzar un reguitzell de proteïnes de diferents colors, creant, així, un ventall de proteïnes fluorescentes.<sup>57,58</sup> Amb això, no es va aconseguir sintetitzar cap proteïna vermella, tot i l'interès que despertava aquest color. El vermell (i fins hi tot zones properes a l'IR) era un objectiu per tot aquell que es dedicava al *bioimaging* en biomedicina. La raó principal és que el teixit de mamífer és només transparent a radiació d'entre uns 600 nm i 1000 nm, ja que l'hemoglobina i l'aigua absorbeixen fortament per sota i per sobre d'aquests límits, respectivament.<sup>59,60</sup> A més, la dispersió de Rayleigh és menys intensa per a longituds d'ona grans, ja



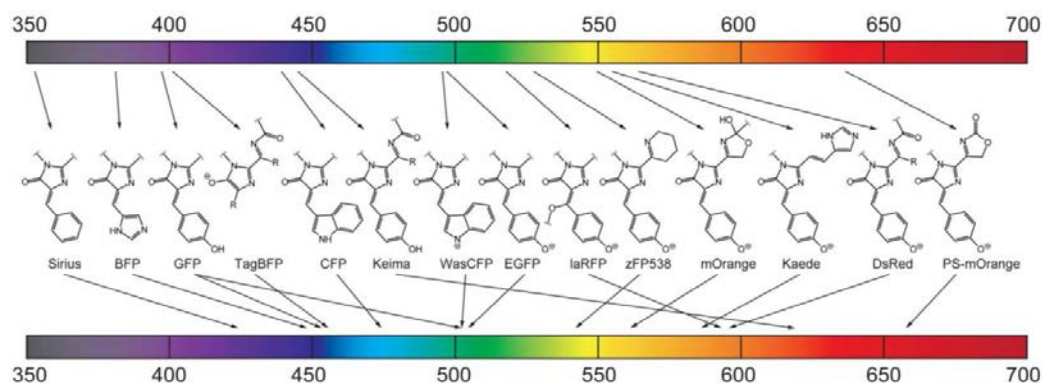


FIGURA 1.6: Les diferents estructures dels cromòfors comporten diferents longituds d'ona d'absorció i d'emissió. Es mostren els cromòfors de les principals famílies de proteïnes amb la zona de l'espectre on absorbeixen (barra superior) i on emeten (barra inferior). Extret de Acharya *et al.*<sup>12</sup>

que obeeix la relació  $\lambda^{-4}$ , i en ser una radiació menys energètica, redueix el risc de danyar els teixits vius. Per totes aquestes raons, la recerca de proteïnes que emetessin en aquesta finestra d'entre 600 nm i 1000 nm es va intensificar. Lukyanov i els seus col·laboradors van publicar el descobriment i caracterització de la que fou la primera proteïna fluorescent vermella: la DsRed,<sup>61,62</sup> que s'extregué d'algunes espècies de corall (*Discosoma sp.*). Tsien va determinar l'estructura del cromòfor d'aquesta nova proteïna i va elucidar que la conjugació del mateix està ampliada per un doble enllaç C=N respecte al cromòfor de la GFP<sup>63</sup> (vegeu la Figura 1.6), fet que fa disminuir l'energia d'excitació tot augmentant la longitud d'ona d'absorció.<sup>64</sup> A partir de la DsRed, moltes d'altres proteïnes van ser desenvolupades, obtenint un altre cop una paleta de colors, aquest cop en la zona del vermell.<sup>65-67</sup> Gran part de les mutacions emprades per a l'obtenció de proteïnes de diferents colors basen la seva acció sobre els residus que posteriorment formaran el cromòfor obtenint, així, una extensa paleta de colors basada en la varietat de cromòfors obtinguts (Figura 1.6). Cal remarcar que sovint, en les proteïnes vermelles, el cromòfor es troba ja en la seva forma aniònica a l'estat fonamental, així que el desplaçament d'Stokes observat, es deu essencialment a la reorganització de l'entorn proteic del cromòfor.

## L'Embolcall

Les modificacions explícites dels cromòfors de les FPs, que naturalment comporten canvis en les propietats espectroscòpiques de les mateixes, tenen un recorregut limitat ja que no tots els aminoàcids són capaços de formar un cromòfor un cop expressats. En aquesta direcció, la modificació dels residus veïns al cromòfor i que interaccionen amb ell, ha obert una porta a la possibilitat de modular les propietats espectroscòpiques del mateix, encara ajustant-les més a les idealment desitjades.<sup>68</sup>

Una de les mutants "estrella" de la GFP és la S65T<sup>47</sup> (això és, substituir la Serina 65 per una treonina) ja que va ser una mutant indispensable per a l'elucidació del mecanisme de fluorescència de la *wt*-GFP. El subtil canvi que suposa la substitució d'una serina per una treonina (només representa la inclusió d'un grup metil), fa canviar dràsticament les propietats espectroscòpiques de la proteïna: l'espectre d'absorció només presenta un pic a 489 nm (que s'atribueix a la forma ionitzada del cromòfor), i el rendiment quàntic de la fluorescència es veu dràsticament reduït fins a 0,02.<sup>69</sup> Aquesta pèrdua de fluorescència s'atribueix al trencament del cable de protons que suposa la inclusió de la treonina. Al contrari que la GFP, la mutant S65T presenta una forta dependència de l'espectre d'absorció respecte el pH, essent el seu  $pK_a=6$ .<sup>70</sup> Segons Remington *et al.*,<sup>71</sup> aquesta propietat es podria fer servir per obtenir medidors intracelulars de pH *in vivo* tot evitant tècniques invasives.<sup>72,73</sup> En aquesta direcció, Elsliger *et al.*<sup>74</sup> introduïren la mutació H148D a la mutant S65T per observar l'efecte d'una càrrega negativa prop del cromòfor, doncs la mutació en qüestió suposa la substitució de la histidina frontal al cromòfor per un aspartat (presumiblement aniònic a pH neutre). La doble mutant S65T/H148D presenta, altre cop, dues bandes d'absorció anomenades A (415 nm), que es troba desplaçada uns 20 nm cap al vermell respecte la *wt*-GFP i B (487 nm).<sup>75</sup> A part de la recuperació de la forma neutra del cromòfor, el que va atraure la major atenció científica cap a aquesta mutant va ser la reaparició de la fluorescència a 510 nm amb un rendiment quàntic de 0,21. A més, la distància cristal·logràfica determinada entre el cromòfor i l'aspartat fou unusualment curta (2,32 Å),<sup>75</sup> fet que va portar a pensar en una forta deslocalització del protó del cromòfor i en la recuperació del mecanisme d'ESPT per la fluorescència de la proteïna. Les característiques mecanístiques i estructurals d'aquesta mutant seran presentades i discutides més endavant. Continuant amb les propietats espectroscòpiques d'equilibri, i en la línia de la doble mutant S65T/H148D, d'altres variants amb la mutació H148D van ser estudiades, concloent que el desplaçament cap al vermell de la banda A possiblement està causat per la proximitat de l'aspartat al cromòfor. Remington ho explicà no en base a l'efecte electrostàtic que pugui crear la càrrega negativa de l'aspartat sobre el cromòfor, si no per la presumible deslocalització del protó entre l'oxigen del cromòfor i el de l'aspartat, que podria induir una certa conjugació extra del sistema  $\pi$  del cromòfor cap al residu D148.<sup>76</sup>

Com ja s'ha explicat, un dels interessos principals de l'actual recerca en proteïnes fluorescents és l'obtenció de proteïnes que emetin en l'anomenada finestra de transparència del teixit de mamífer (entre 600 nm i 1000 nm), ja que així es poden maximitzar les possibles aplicacions en biomedicina.<sup>77-82</sup> Així doncs, partir d'un cromòfor que emeti a una zona més propera a la finestra de transparència sembla una bona aproximació. És el cas de la ja anomenada DsRed, però també la resta de proteïnes Vermelles Fluorescents o RFPs d'aquí en endavant (Red Fluorescent Proteins) que, com he dit, contenen un cromòfor amb una conjugació més estesa respecte al cromòfor de les GFPs, donada la inclusió d'un doble enllaç C=N. Una

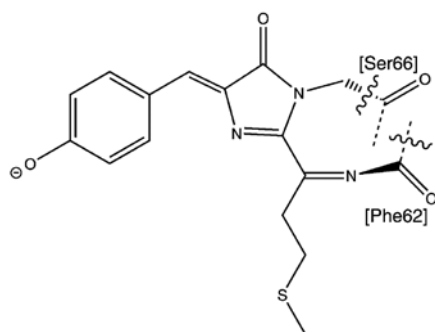


FIGURA 1.7: Representació de l'estructura del cromòfor de les RFPs. S'hi pot apreciar l'extensió de la conjugació pel grup C=N acilimínic. També s'hi representa el residu Phe62, que també influeix a la posició del màxim de la banda d'absorció.

de les primeres RFPs que va ser estudiada fou la mKate, que presenta un màxim d'absorció a 588 nm, molt proper a la desitjada regió de transparència.<sup>83,84</sup> En els últims anys, emperò, una mutació de la mKate s'ha erigit com a candidata molt prometedora: la mNeptune, desenvolupada per Lin *et al.*<sup>77</sup> La mNeptune (mKate Met41Gly Ser61Cys Ser158Cys Tyr197Phe Met146Thr) mostra un espectre d'absorció i emissió desplaçat uns 15 nm cap al vermell respecte la mKate, absorbint a 600 nm i emetent a 650 nm.

L'anàlisi de l'estructura cristal·lina de la mNeptune revela alguns dels factors que possiblement afecten a la desplaçada posició del màxim d'absorció. En particular, es va observar que la mutació Met41Gly causava la creació d'un buit just darrere del cromòfor. En l'estructura cristal·lina es pot observar com aquest buit està ocupat per una molècula d'aigua que pot interaccionar per mitjà d'un pont d'hidrogen amb el residu Phe62, limitrof amb el cromòfor. El grup C=O del residu Phe62 està a una distància raonable de la conjugació del sistema  $\pi$  del cromòfor, així que es proposa que la interacció de l'aigua amb l'anomenat residu podria estabilitzar l'estat excitat, tot reduint l'energia d'excitació.

Les característiques presentades per la mNeptune han estat més profundament explorades amb el desenvolupament d'una nova candidata, la mCardinal<sup>79</sup> (mNeptune Ala104Val Ile121Leu Ile171His Ser28Thr Gly41Gln Ser143Thr), que presenta un espectre d'absorció i emissió amb màxims a 604 nm i 659 nm, respectivament. De totes les mutacions de la mCardinal respecte la mNeptune, els autors només en resalten una: la Gly41Gln. L'estructura cristal·lina revela una interacció de pont d'hidrogen entre la Gln41 i el grup C=O del residu Phe62, sense intervenció de molècules d'aigua. Basant-nos en la distància cristal·logràfica de la interacció Gln41-Phe62 (2,5 Å) amb la de l'interacció Aigua-Phe62 trobada en la mNeptune (2,8 Å), sembla evident que la primera és més forta. Això donaria suport a la hipòtesi que una interacció de pont d'hidrogen amb el residu Phe62 fa baixar l'energia d'excitació per als cromòfors tipus RFPs (Figura 1.7).

Tot i la gran quantitat d'informació experimental disponible, sembla evident que cal racionalitzar aquest coneixement i dotar-lo d'un context on es pugui modificar, de forma sistemàtica, les característiques dels cromòfors i dels residus que l'envolten per a modular les propietats dels mateixos i, en últim terme, de les proteïnes fluorescents que els contenen.

#### 1.2.4 La Pedra Rosetta: la Mutant S65T/H148D-GFP

La Pedra Rosetta és un fragment d'un antic monument egipci on hi ha escrit un decret publicat a Menfis l'any 196 a. C. en nom del faraó Ptolomeu V. El text està reproduït en tres escriptures diferents: el text superior en jeroglífics egipcis, la part intermèdia en escriptura demòtica i la part inferior en grec antic. Com que el text reproduït amb les tres escriptures és el mateix, va ser una eina clau per la comprensió dels jeroglífics egipcis antics.

En ciència, és una pràctica comuna estudiar casos senzills, entendre'ls a un nivell profund, per després poder generalitzar aquest coneixement per a casos més complexos dins d'un marc ja creat, doncs l'abordatge directe d'un cas complex, moltes vegades es fa inviable. Naturalment la ciència es construeix des de baix cap amunt, així que necessitem obtenir coneixement ferm de casos senzills per poder estudiar com cal sistemes complexos. Hom coneix la molècula d'hidrogen com el principal banc de proves de la història de la química i el primer sistema model per a qualsevol nova teoria molecular. El comportament d'aquesta molècula ha servit per generalitzar el comportament de les molècules diatòmiques, i aquestes per generalitzar el comportament de les poliatòmiques. Seguint aquesta filosofia, per crear un coneixement profund sobre les proteïnes fluorescents, vam decidir centrar-nos en un cas paradigmàticament "senzill" (permeteu-me les cometes), per estudiar-lo a fons i, a partir d'aquí, poder generalitzar el comportament a casos més complicats. El banc de proves que vam triar per a les proteïnes fluorescents és la mutant S65T/H148D-GFP, anteriorment anomenada en aquest treball, i en la que la transferència protònica que porta de l'estat  $A^*$  a  $I^*$  és directament entre el donador (cromòfor) i l'acceptor final (l'Aspartat 148), que es troba davant del mateix, essent una transferència protònica simple, l'objecte d'estudi. Aquesta proteïna, a més, ha estat estudiada experimentalment a un nivell molt fi, fet que ens permet validar constantment les nostres explicacions atomístiques del funcionament global d'aquesta proteïna fluorescent arquetípica. En aquest treball, doncs, s'utilitzarà aquesta mutant com a "Pedra Rosetta" de les proteïnes fluorescents, doncs el coneixement profund que es pugui obtenir sobre la mateixa serà de gran ajuda per a desxifrar el funcionament general d'aquests complicats foto-biosistemes.

La doble mutant S65T/H148D-GFP, recordem, va ser obtinguda en un intent (realitzat amb èxit) de modificar el  $pK_a$  de la mutant S65T-GFP, que podia ser utilitzada com a marcador *in vivo* de proteïnes intracel·lulars.

La inclusió de la mutació H148D va portar a la recuperació de la forma A (perduda en la mutant S65T) i també a la recuperació de la fluorescència a 510 nm (espècie I\*) amb un rendiment quàntic considerable de 0,21. Així doncs, l'espectre d'absorció d'aquesta proteïna presenta els pics corresponents a la forma neutra i ionitzada del cromòfor, A (415 nm) i B (487 nm), desplaçats cap al vermell uns 15 nm respecte la *wt*-GFP.<sup>74</sup> Remington i els seus col·laboradors van publicar una sèrie de treballs on analitzaven les propietats espectroscòpiques resoltes en el temps de la doble mutant. Un dels descobriments més remarcables va ser la pujada ultrarràpida de la fluorescència a 510 nm, que estava per sota del límit de detecció del seu instrumental (<170 fs) i que, a més, no presentava efecte cinètic d'isòtop (KIE) H/D mesurable. Aquest comportament és del tot contrastant amb el de la *wt*-GFP, on la fluorescència apareix en l'ordre dels pico-segons i és molt sensible a la substitució per deuteri. El fet que ni a baixes temperatures es pogués observar el component de pujada de la fluorescència ni l'efecte cinètic d'isòtop féu pensar en una ESPT sense barrera. Un altre dels resultats de l'estudi de Remington va ser que no es va trobar una correspondència absoluta entre la baixada de població de l'espècie A\* amb la pujada de l'espècie I\*, fet que indicava una possible via de desactivació prèvia a la transferència.

L'estructura cristal·lina de la doble mutant mostra un contacte entre un dels oxígens carboxílics de l'Asp148 i l'oxigen fenòlic del cromòfor inusualment curt (2,32 Å), fet que evidencia que l'aspartat és l'acceptor del protó del cromòfor, després de l'excitació. Aquesta distància tant curta, també seria la causant del desplaçament cap al roig de la banda A d'absorció respecte la *wt*-GFP, ja que podria contribuir a una deslocalització del sistema  $\pi$  del cromòfor sobre el carboxilat de l'aspartat. Remington va explorar la possibilitat que, donada aquesta curta distància donador/acceptor, es pogués descriure una situació de pont d'hidrogen de barrera baixa o LBHB (Low Barrier Hydrogen Bond) d'aquí en endavant: si el protó es trobés en tal situació, seria possible una transferència gairebé instantània després de l'excitació. Per fer tal cosa, es va analitzar l'espectre de NMR de <sup>1</sup>H a la zona de camp baix però no es va trobar cap senyal corresponent a un protó especialment desapantallat, característica típica dels LBHB. Reacis a abandonar la idea d'un protó deslocalitzat, van proposar que aquest protó podria estar subjecte a una ràpida interconversió amb el solvent, a més de recordar altres casos on un protó en un pont d'hidrogen molt curt no presenta un desplaçament químic tant gran com es podria esperar.<sup>85</sup>

Per tal d'explorar més intensament el procés a l'estat excitat, Stoner-Ma *et al.* van enregistrar l'espectre d'infraroig resolt en el temps i no van ser capaços d'observar cap canvi significatiu en els modes de vibració del C=O de l'aspartat, fet que va fer concloure que no hi havia un canvi d'estat de protonació clar del mateix.<sup>87</sup> Altre cop es va explicar aquest fenomen en base a la deslocalització del protó ja a l'estat fonamental en un possible LBHB: si el protó es troba prou deslocalitzat entre donador i acceptor a l'estat fonamental, no es possible observar una transferència *completa* a

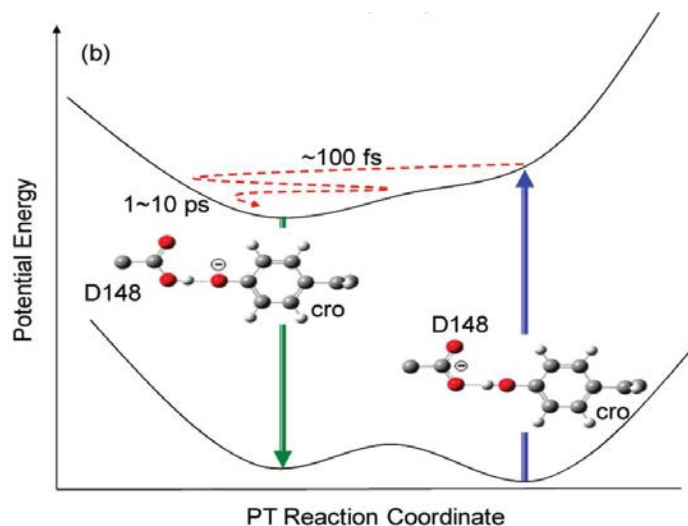


FIGURA 1.8: Representació esquemàtica del procés proposat per Kondo *et al.*: després d'absorbir la radiació de 411 nm es dona una transferència ultrarràpida (100 fs) i sense barrera, seguida d'una relaxació que ocorre en l'ordre dels picosegons. Extret de Kondo *et al.*<sup>86</sup>

L'estat excitat, després de l'excitació, essent només una aproximació del paquet protònic el que està tenint lloc. En el mateix estudi van realitzar els mateixos experiments amb la mutant E222Q/H148D i la *wt*-GFP: per a la primera, els resultats obtinguts foren gairebé els mateixos que per la doble mutant S65T/H148D-GFP, doncs l'acceptor del protó es presumia, altre cop, l'Asp148. Aquests resultats contrastaren amb els obtinguts per a la *wt*-GFP ja que, en aquest cas, es pogué observar clarament el canvi d'estat de protonació del carboxilat del Glutamat 222.

L'any 2010, i ja amb una resolució temporal molt més acurada (50 fs), Kondo *et al.* van enregistrar l'espectre d'emissió resolt en el temps<sup>86</sup> però tampoc van ser capaços de resoldre la pujada de l'espècie I\*, tot i que van poder determinar dos components en la caiguda de la fluorescència. Van assignar el primer component a la transferència ultrarràpida del protó que tindria lloc en l'ordre dels 100 fs i el segon component a una relaxació de l'entorn on s'acomodaria aquesta nova situació electrònica, que ocorreria en l'ordre dels pico-segons. Altre cop es va invocar el suposat gran guany d'acidesa del cromòfor a l'estat excitat i a l'absència de barrera per a la transferència per justificar aquest comportament (Figura 1.8).

Quedà establert, així, un dibuix raonable del funcionament d'aquesta proteïna però encara mancaven explicacions atomístiques a les hipòtesis mecanístiques plantejades pels estudis experimentals. Parafrasejant a Stoner-Ma *et al.*, aquest dibuix és compatible amb les evidències experimentals però s'ha fet referència a conceptes com els ponts d'hidrògen de barrera baixa o a certes característiques dels estats electrònics fonamental i excitats que "... need to be supported by detailed high-level quantum chemical calculations. Even if such surfaces become available (in itself, a challenging objective),

*the subsequent calculation of the ultrafast rate coefficients will require nonequilibrium dynamics simulations.*"<sup>87</sup> Així doncs, es fa evident la necessitat d'un punt de vista teòric i computacional al problema, que pugui aportar el detall atòmic que es necessita per escatir els detalls del funcionament d'aquesta excepcional proteïna.

## 2 Marc Teòric

El marc teòric emprat és, juntament amb el tema tractat, el factor que més influencia en la personalitat d'un treball de recerca. I m'atreveria a dir que ho fa gairebé en més mesura que el tema en qüestió. Defineix una estructura sobre la qual podem construir el nou coneixement obtingut amb les eines emprades i limita l'aplicabilitat del mateix. En la present tesi, de caire exclusivament teòric i computacional, cal donar un marc teòric general per comprendre el tipus de coneixement i conclusions que es poden extreure de la mateixa.

Les FPs, com a objecte d'estudi, presenten tots els reptes als que es pot enfrontar un treball de recerca en química computacional. Per una banda, són sistemes biològics de tamany considerable. Això fa que càlculs que serien factibles per a molècules relativament petites es tornin del tot inviables, havent d'emprar aproximacions addicionals a les ja utilitzades pels casos més senzills per a poder obtenir resultats numèrics en un temps raonable. D'altra banda, al tractar el fenomen de la fluorescència, ens hem d'enfrontar amb el càlcul d'estats electrònics excitats de proteïnes fluorescentes, fet que comporta dificultats addicionals. A més, hi ha consens en que el càlcul de propietats fisicoquímiques de proteïnes (i de sistemes amb un gran nombre de graus de llibertat en general) a partir de l'exploració i caracterització d'alguns punts concrets de la superfície d'energia potencial pot ser una aproximació pobre al problema que plantegen aquests sistemes tant complexos, així que sovint ens veurem forçats a utilitzar mètodes dinàmics per explorar l'espai configuracional del sistema.<sup>88</sup> Tanmateix, els processos descrits en la introducció i susceptibles d'ésser estudiats són prou complexos com perquè s'hagi d'emprar una aproximació multiescala del problema.

Durant el desenvolupament d'aquesta tesi, sempre s'ha pensat en la millor manera possible de resoldre els problemes químics que han presentat els casos d'estudi, no pas en quins casos d'estudi podien complir els requisits per a ser estudiats amb les tècniques i mètodes prèviament coneguts. Això ha fet que s'hagi utilitzat una gran varietat de mètodes de caires molt diferents (des de càlculs de mecànica molecular fins a càlculs de dinàmica quàntica, passant per càlculs d'energies electròniques a nivells semiempírics, DFT o post-Hartree-Fock), i a vegades ha suposat el desenvolupament i implementació de metodologia concreta en nou programari. Com que per raó d'extensió seria impossible explicar a un nivell fi tots i cadascun dels mètodes emprats, em centraré només en aquells que crec que han tingut més rellevància en l'obtenció de resultats.



## 2.1 Càlcul de l'Energia Electrònica

En el context de la química teòrica, es parla sovint d'energia electrònica. És un concepte acceptat per la comunitat i es maneja amb lleugeresa. Una gran part (per no dir la pràctica totalitat) de la química teòrica actual basa la seva recerca en nous mètodes per calcular-la, aplicacions d'aquests mètodes, càlculs realitzats a partir de l'obtenció de la mateixa, etc. És per això que em sembla de vital importància, abans de continuar, entendre encara que sigui de manera descriptiva quin grau d'aproximació estem assumint ja només pel fet de parlar d'energia electrònica.

### La Separació Espai-temporal

En el context de la mecànica quàntica, per conèixer l'evolució temporal de qualsevol estat quàntic, cal resoldre l'equació d'Schödinger dependent del temps, l'equivalent microscòpic de la segona llei de Newton:

$$i\hbar \frac{\partial}{\partial t} |\psi(t)\rangle = \hat{H} |\psi(t)\rangle \quad (2.1)$$

on  $|\psi(t)\rangle$  és l'estat quàntic a un temps  $t$  i  $\hat{H}$  és l'operador Hamiltonià.

Per a la majoria de problemes químics, ens interessarà obtenir una funció que depengui de les coordenades espacials del sistema i que encara contingui tota la informació del mateix. És per això que es projecta l'estat  $|\psi(t)\rangle$  sobre l'espai de coordenades  $\vec{q}$  per a obtenir l'eina per excel·lència de la química teòrica, la funció d'ona:

$$\langle \vec{q} | \psi(t) \rangle = \psi(\vec{q}, t) \quad (2.2)$$

La funció d'ona és una funció matemàtica, en general en un espai complex ( $\psi(\vec{q}, t) \in \mathbb{C}$ ), que depèn de les coordenades del sistema i que conté tota la informació del mateix. Amb això, com a tal, no té cap significat físic directe. El mòdul quadrat de la funció d'ona,  $|\psi(\vec{q}, t)|^2 = \psi(\vec{q}, t)^* \psi(\vec{q}, t)$  s'interpreta com a la densitat de probabilitat de trobar el sistema a unes coordenades  $\vec{q}$  a un temps  $t$ . Així doncs, podem reescriure l'equació 2.1 en:

$$i\hbar \frac{\partial}{\partial t} \psi(\vec{q}, t) = \hat{H} \psi(\vec{q}, t) \quad (2.3)$$

Aquesta equació és de caire general perquè podem calcular l'evolució de qualsevol sistema quàntic a partir d'una funció d'ona i un operador Hamiltonià. Si tenim en compte  $\vec{q}$  com a el conjunt de coordenades del sistema que representen les posicions d'electrons i nuclis, després de negligir els possibles efectes relativistes en el sistema, el Hamiltonià per un sistema molecular format per electrons i nuclis (això és, considerem que els nuclis

són una sola identitat i no tenen estructura interna), esdevé (prenent unitats atòmiques):

$$\begin{aligned}\hat{H} &= \hat{T}_N + \hat{T}_e + \hat{V}_{NN} + \hat{V}_{ee} + \hat{V}_{Ne} \\ &= -\sum_{i=1}^N \frac{\nabla_i^2}{2} - \sum_{\alpha=1}^M \frac{\nabla_\alpha^2}{2m_\alpha} + \sum_{\alpha>\beta}^M \frac{Z_\alpha Z_\beta}{|\vec{R}_\alpha - \vec{R}_\beta|} \\ &\quad + \sum_{i>j}^N \frac{1}{|\vec{r}_i - \vec{r}_j|} + \sum_{i=1}^N \sum_{\alpha=1}^M \frac{-Z_\alpha}{|\vec{r}_i - \vec{R}_\alpha|}\end{aligned}\quad (2.4)$$

on els dos primers termes corresponen a l'energia cinètica d'electrons i nuclis respectivament i la resta als termes potencials nucli-nucli, electró-electró i electró-nucli, respectivament. Això es desenvolupa per a els  $N$  electrons i  $M$  nuclis que tenen posicions  $\vec{r}_i$  i  $\vec{R}_\alpha$ , respectivament.

Segons l'equació 2.3, la dependència de  $\psi$  amb  $\vec{q}$  i  $t$  no hauria de ser *a priori* separable, però si considerem que el Hamiltonià no té cap tipus de dependència temporal i assumim que la funció d'ona és separable:

$$\psi(\vec{q}, t) = \Psi(\vec{q})\Phi(t) \quad (2.5)$$

llavors l'equació 2.3 esdevé:

$$i\hbar \frac{\partial}{\partial t} [\Psi(\vec{q})\Phi(t)] = \hat{H}[\Psi(\vec{q})\Phi(t)] \Rightarrow i\hbar \frac{1}{\Phi(t)} \frac{\partial \Phi(t)}{\partial t} = \frac{\hat{H}\Psi(\vec{q})}{\Psi(\vec{q})} \quad (2.6)$$

L'única manera que es compleixi aquesta igualtat és que l'equació 2.6 no depengui ni de  $\vec{q}$  ni de  $t$ , pel que ha de ser constant. Així doncs, tenint en compte que el Hamiltonià té correspondència amb l'energia total del sistema, del terme dret de l'equació 2.6 obtenim:

$$\frac{\hat{H}\Psi(\vec{q})}{\Psi(\vec{q})} = E \Rightarrow \hat{H}\Psi(\vec{q}) = E\Psi(\vec{q}) \quad (2.7)$$

coneguda com a equació d'Schrödinger independent del temps. La resolució d'aquesta equació, que és el principal objectiu dels mètodes de química quàntica, ens permet obtenir l'energia total d'un sistema electrònic en un estat estacionari. Del terme esquerre de l'equació 2.6 obtenim:

$$i\hbar \frac{1}{\Phi(t)} \frac{\partial \Phi(t)}{\partial t} = E \Rightarrow \Phi(t) = e^{-iEt/\hbar} \quad (2.8)$$

I si substituïm a 2.5:

$$\psi(\vec{q}, t) = \Psi(\vec{q})e^{-iEt/\hbar} \quad (2.9)$$

És a dir que la funció d'ona és separable en una part espacial i una de temporal. D'altra banda, la funció d'ona és efectivament depenent del temps ja que conté l'exponent que li dona la fase. Amb això, considerant el Hamiltonià independent del temps, si fem el mòdul quadrat de l'equació 2.9, podem veure fàcilment com la densitat de probabilitat del sistema no canvia amb el temps i que, per tant, qualsevol observable física associada al mateix tampoc ho fa.

### L'Aproximació de Born-Oppenheimer

Amb la simplificació aconseguida amb la separació espai-temporal, el problema s'ha reduït a un problema de funcions i valors propis, havent de resoldre l'equació 2.7 on el Hamiltonià molecular inclou termes cinètics de nuclis i electrons, a més de termes d'interacció nucli-nucli, electró-electró i electró-nucli. Per poder resoldre l'equació cal emprar l'aproximació de Born-Oppenheimer, o BO d'aquí en endavant.<sup>89</sup> Sense entrar en els detalls del seu desenvolupament, que es poden trobar diversos textos acadèmics de referència,<sup>90</sup> l'aproximació BO separa el moviment nuclear de l'electrònic assumint que els electrons es mouen infinitament més ràpid que els nuclis. Si tenim en compte la massa relativa de nuclis i electrons ja en l'àtom menys pesat ( $m_H > 1.800m_{el}$ ), aquesta és una aproximació assumible. Un altre punt de l'aproximació BO que s'acostuma a passar pel alt és que cal aproximar que els estats electrònics no estan en absolut acoblats i per què això passi, les diferències energètiques entre els estats ha de ser gran. Tenint en compte aquests dos punts, podem separar el moviment electrònic del nuclear i reescriure l'equació d'Schrödinger independent del temps per un sistema molecular com a:

$$[\hat{T}_N(\vec{Q}) + \hat{V}_{NN}(\vec{Q}) + E_e(\vec{Q})] \Psi_N(\vec{Q}) = E \Psi_N(\vec{Q}) \quad (2.10)$$

on l'energia electrònica,  $E_e(\vec{Q})$  s'obté de:

$$\hat{H}_{el} \Psi_e(\vec{r}; \vec{Q}) = E_e(\vec{Q}) \Psi_e(\vec{r}; \vec{Q}) \quad (2.11)$$

que és l'equació d'Schrödinger electrònica independent del temps i on el Hamiltonià electrònic conté totes les contribucions electròniques a l'energia,  $\hat{H}_{el} = \hat{T}_e + \hat{V}_{ee} + \hat{V}_{Ne}$ .

Si definim  $U(\vec{Q}) = V_{NN}(\vec{Q}) + E_e(\vec{Q})$  podem finalment reescriure l'equació 2.10 com a:

$$[\hat{T}_N(\vec{Q}) + U(\vec{Q})] \Psi_N(\vec{Q}) = E \Psi_N(\vec{Q}) \quad (2.12)$$

Aquesta és l'equació d'Schrödinger pel moviment nuclear un cop s'ha aplicat la aproximació BO. En aquesta, apareix el potencial  $U(\vec{Q})$  i que anomenem superfície d'energia potencial o PES (Potential Energy Surface). Des d'un punt de vista interpretatiu, l'aproximació BO comporta que

els nuclis es moguin en un potencial creat pels electrons, que s'adaptin instantàniament a cada nova configuració nuclear.

La resolució de l'equació electrònica (equació 2.11) hauria de ser tansols el primer pas per a la resolució del moviment nuclear, ja que és aquest el que comporta reaccions químiques. Els mètodes de la química quàntica, basen el seu estudi en la resolució d'aquesta equació electrònica, però encara han d'aplicar una altra aproximació per arribar a tal fi: l'aproximació orbital. La necessitat d'aquesta aproximació neix de la impossibilitat de resoldre un sistema de 3 o més partícules interaccionants o, el que és el mateix, una molècula de més de 2 electrons en un context BO. En l'aproximació orbital, es considera que la funció d'ona electrònica és separable en un producte de funcions monoelèctriques, anomenades spin-orbitals. Aquesta funció, a més, i per tal de complir el principi d'antisimetria que ha de regir en qualsevol sistema de fermions, haurà de ser antisimètrica respecte el bescanvi de dos electrons.<sup>91</sup> Per a assegurar aquestes propietats matemàtiques s'utilitzen els determinants d'Slater, que a més, compleixen el principi d'exclusió de Pauli. Així doncs, de forma general, s'expressa la funció d'ona electrònica com a combinació lineal d'aquests determinants d'Slater. La cèlebre aproximació de Hartree-Fock, de la qual parteixen tots els mètodes *ab initio*, expressa la funció d'ona com a un sol determinant d'Slater. Els mètodes clàssics de la química quàntica tracten els problemes derivats d'aquesta aproximació (la manca de correlació electrònica en la funció d'ona d'un sol determinant) i són els anomenats *post-Hartree-Fock*.

Així doncs, de manera qualitativa, s'ha arribat al punt d'on parteixen els mètodes de la química quàntica pel càlcul d'estructura electrònica. És primordial entendre que, més enllà de les aproximacions que introduiran els mètodes que seran presentats a continuació, hi ha una sèrie d'aproximacions "çongènites" que hem d'assumir per a tots ells i que han estat tot just explicades.

### 2.1.1 La Teoria del Funcional de la Densitat

Els mètodes *post-Hartree-Fock* proporcionen una manera sistemàtica de millorar la pobre descripció Hartree-Fock de la funció d'ona però sovint esdevenen massa cars computacionalment per a ser aplicats a molècules mitjanes. La Teoria del Funcional de la Densitat o DFT (Density Functional Theory) proporciona una manera de trobar l'energia i densitat electrònica de l'estat fonamental d'un sistema molecular, considerant que l'energia és un funcional de la densitat. Això vol dir que l'energia és una funció que depèn d'una altra funció, la densitat electrònica del sistema. D'aquesta manera, s'introdueix de forma natural la correlació electrònica.

El DFT es basa en dos teoremes desenvolupats per Hohenberg i Kohn l'any 1964.<sup>92</sup> El primer teorema presenta que qualsevol observable física d'un sistema en un estat fonamental no degenerat es pot expressar com a funcional de la densitat electrònica. És a dir que l'energia electrònica, com a

observable del sistema, és un funcional de la densitat electrònica:  $E[\rho]$ . El segon teorema diu que es pot calcular exactament la densitat electrònica del sistema tot determinant la densitat que minimitza l'energia. Aquest és l'equivalent del teorema variacional:

$$E[\rho \neq \rho_{exacte}] > E[\rho_{exacte}] \quad (2.13)$$

L'energia com a funcional de la densitat es pot separar en els següents termes:

$$E[\rho] = T[\rho] + V_{Ne}[\rho] + V_{ee}[\rho] + V_{NN} \quad (2.14)$$

on  $T[\rho]$  correspon a l'energia cinètica electrònica i  $V_{Ne}[\rho]$  i  $V_{ee}[\rho]$  als termes potencials electró-nucli i electró-electró, respectivament.

Els termes  $T[\rho]$  i  $V_{ee}[\rho]$  no són coneguts però l'any 1965 Kohn i Sham van proposar un mètode per calcular el terme d'energia cinètica.<sup>93</sup> En el seu tractament es proposava un sistema fictici on els electrons, que descriu amb funcions d'ona monoelectròniques, no interaccionen entre si però presenten la mateixa densitat electrònica que el sistema real. Així doncs, l'expressió de l'energia es pot reescriure:

$$E_v[\rho] = T_s[\rho] + \int v_n(\vec{r})\rho(\vec{r})d\vec{r} + J[\rho] + E_{xc}[\rho] \quad (2.15)$$

on  $T_s[\rho]$  correspon al terme cinètic del sistema no interaccionant,  $v_n(\vec{r})$  és un potencial extern introduït en el tractament i  $J[\rho]$  és l'energia Coulòmbica. El problema d'aquest tractament és que fa aparèixer un terme, el funcional de correlació-intercanvi  $E_{xc}[\rho]$ , del qual no se'n coneix la forma. Diferents intents per a l'expressió de la forma aproximada d'aquest funcional han donat lloc a la gran varietat de funcionals disponibles pel mètode DFT.<sup>94</sup>

## La Teoria del Funcional de la Densitat Dependent del Temps

Tot i que el DFT va ser pensat per al tractament dels estats fonamentals dels sistemes electrònics, existeix una versió d'aquesta teoria que permet el càlcul d'estats electrònics excitats. La Teoria del Funcional de la Densitat Dependent del Temps o TD-DFT (Time-Dependent Density Functional Theory),<sup>94</sup> desenvolupada a mitjans dels anys 90 del segle passat per Casida *et al.*,<sup>95-97</sup> parteix de la idea d'aplicar una pertorbació dependent del temps a un sistema electrònic en un context DFT. El mètode es basa en el principi d'acció estacionària i el teorema de Runge-Gross,<sup>98</sup> l'equivalent dependent del temps dels teoremes de Hohenberg-Kohn, i que manté la mateixa filosofia: correlaciona inequívocament l'estat inicial del sistema i la pertorbació dependent del temps que se li aplica amb la densitat electrònica de l'estat final. D'una manera similar al formalisme DFT, en que es considera un sistema fictici d'electrons que no interaccionen però que descriuen la mateixa densitat que el sistema real, es poden derivar les

equacions Kohn-Sham depenent del temps:

$$i\hbar\frac{\partial}{\partial t}\Psi(\vec{r}, t) = \left[ -\frac{1}{2}\nabla^2 + v_{ks}(\vec{r}, t) \right] \Psi(\vec{r}, t) \quad (2.16)$$

on el potencial efectiu depenent del temps,  $v_{ks}(\vec{r}, t)$ , es descompon en un primer terme corresponent al potencial extern, un segon terme que té en compte la interacció Coulòmbica i un últim terme de correlació-intercanvi:

$$v_{ks}(\vec{r}, t) = v_{ext}(\vec{r}, t) + v_c(\vec{r}, t) + v_{xc}(\vec{r}, t) \quad (2.17)$$

Runge i Gross van aproximar la forma del potencial de correlació intercanvi a l'integral d'acció de Dirac:

$$A[\Psi]_{xc} = \int \langle \Psi(t) | \hat{H} - i\hbar\frac{\partial}{\partial t} | \Psi(t) \rangle dt \quad (2.18)$$

Assumida la correspondència entre la funció d'ona i la densitat electrònica, Runge i Gross tractaren l'integral d'acció de Dirac com a funcional de la densitat. Donada la impossibilitat de trobar una expressió pel funcional d'acció, aplicaren l'aproximació ALDA (Adiabatic Local Density Approximation):

$$\frac{\delta A_{xc}[\rho]}{\delta \rho(\vec{r}, t)} \approx \frac{\delta E_{xc}[\rho]}{\delta \rho(\vec{r})} \quad (2.19)$$

Aquesta aproximació imposa que el potencial extern és independent del temps. Això pot semblar una incongruència però és una aproximació acceptable si el potencial en qüestió presenta una dependència temporal suau però incorpora grans avantatges ja que implica que podem utilitzar els funcionals de correlació-intercanvi de l'estat fonamental per al càlcul d'estats excitats.

A partir d'aquí, per evitar l'integració temporal de les equacions dependents del temps, s'aplica l'aproximació de resposta lineal, en que s'assumeix que la pertorbació és prou petita com per poder expressar l'estat final del sistema a partir de l'estat fonamental DFT, això és, en la base d'orbitals Kohn-Sham optimitzats. Així doncs, es poden obtenir energies d'excitació i forces d'oscil·lador per a sistemes moleculars de mida considerable amb un relatiu esforç computacional.

El TD-DFT presenta grans avantatges davant del seu equivalent en mètodes basats en funció d'ona<sup>94</sup> (el CIS) ja que incorpora la correlació electrònica a un cost raonable. Amb tot, pateix del mateix mal que el seu germà petit, el DFT: no existeix manera sistemàtica de millorar els resultats, que sempre estaran condicionats pel funcional triat.

En aquesta línia, per a descriure l'excitació qualitativament i quantitativa, s'ha utilitzat el funcional CAM-B3LYP.<sup>99</sup> Aquest funcional, desenvolupat

L'any 2004, combina les qualitats del popular B3LYP<sup>100</sup> amb les correccions a llarga distància proposades per Tanawa *et al.*<sup>101</sup> El funcional CAM-B3LYP dona bones descripcions d'excitacions de transferència de càrrega, punt on fallen estrepitosament els funcionals sense correcció de llarga distància. Aquesta correcció fa variar la proporció Hartree-Fock del funcional en funció de la distància dels orbitals que participen en l'excitació. El TD-DFT amb el funcional CAM-B3LYP s'ha posicionat com un mètode fiable per a la descripció d'estats excitats de sistemes de mida considerable i és el mètode de referència utilitzat en el càlcul de propietats espectroscòpiques dels Capítols 5 i 6.

### 2.1.2 Els Mètodes Semiempírics

Tot i la gran utilitat dels mètodes basats en el DFT, de vegades cal fer una aproximació teòrica diferent per a poder realitzar càlculs en temps raonables. Més enllà de càlculs puntuals o de tipus estàtic, pels quals s'ha utilitzat el DFT/TD-DFT en aquest treball, per a poder assumir càlculs de dinàmica sobre PESs, ens cal utilitzar els mètodes semiempírics, que presenten grans avantatges pel que fa la relació cost/fiabilitat.<sup>102</sup>

Els mètodes semiempírics són mètodes de la química quàntica on es busca una màxima simplificació del càlcul de l'estructura electrònica, ja sigui per la parametrització o la directa elusió de certes integrals que apareixerien en el càlcul purament *ab initio*. La filosofia dels mètodes semiempírics es basa en el formalisme dels mètodes *ab initio*, on s'hi introdueixen modificacions dràstiques per tal d'accelerar el càlcul, normalment basades en negligir el càlcul dels termes menys importants en les equacions subjacents. A fi i efecte de compensar els possibles errors que puguin aparèixer, s'afegeixen paràmetres empírics al càlcul, per tal d'ajustar els resultats a una sèrie de dades contrastades. Per tant, si el mètode semiempíric en qüestió conté la física essencial del problema, aquest serà prou bo si contempla els casos d'aplicació en la seva parametrització.

Existeixen una gran varietat de mètodes semiempírics que han anat sorgint de les necessitats computacionals i possibilitats tècniques de cada moment. Els primers mètodes que tractaven la totalitat dels electrons de valència van ser proposats per Pople:<sup>103</sup> basaven la seva parametrització amb la comparació amb resultats Hartree-Fock i introduïen diferents aproximacions en el càlcul d'integrals. Foren el CNDO (*complete neglect of differential overlap*), l'INDO (*intermediate neglect of differential overlap*) i el NDDO (*neglect of diatomic differential overlap*).

Dewar va prendre la aproximació NDDO i la va ajustar, aquest cop contra dades experimentals, per acabar creant el mètode MINDO, que va desembocar en el MNDO<sup>104,105</sup> i l'AM1.<sup>106</sup> Una posterior parametrització del mètode MNDO va donar lloc al PM3.<sup>107,108</sup> Formalment l'AM1 i el PM3 es diferencien de l'MNDO tansols en la tria de la funció de repulsió nuclear.

Posteriors reparametritzacions han donat lloc a mètodes com l'RM1,<sup>109</sup> el PM6<sup>110</sup> o el PM7.<sup>111</sup>

Anant encara més enllà en el model MNDO, W. Thiel va proposar els mètodes OMx (OM1,<sup>112</sup> OM2<sup>113</sup> i OM3<sup>114</sup>), que inclouen correccions d'ortogonalització als termes monoelèctrics de la matriu de Fock NDDO per tenir en compte els efectes de correlació de Pauli. El formalisme NDDO, com la resta de mètodes semiempírics basats en la resolució autoconsistent d'orbitals moleculars, negligeix totes les integrals bielectròniques de 3 i 4 centres. El NDDO s'aplica a totes les integrals que impliquen interaccions Coulòmbiques i també a les integrals de solapament que apareixen a les equacions de Hartree-Fock. Les correccions a aquest mètode que proposa Thiel amb els OMx, tot incloent efectes d'ortogonalització, s'apliquen a les integrals monoelèctriques d'un centre (OM1) o també a les bielectròniques de dos centres (OM2 i OM3). El mètode OM3 negligeix algunes de les correccions de l'OM2 sense grans conseqüències en la precisió del mètode. Es poden trobar els detalls de les aproximacions dels mètodes NDDO i OMx a la bibliografia especialitzada.<sup>112,113</sup>

Els mètodes OMx s'han erigit en els darrers anys com una alternativa prou fiable per al càlcul de propietats termodinàmiques i dinàmiques de sistemes a l'estat fonamental<sup>115-117</sup> però a més, degut a que la correcció d'ortogonalització fa pujar adequadament els orbitals virtuals que s'obtenen del mètode autoconsistent, també es poden aplicar per estudis d'estats excitats.<sup>118</sup> Especialment el mètode OM2 ha demostrat ser capaç de reproduir PES d'estats electrònics excitats per a molècules petites, essent útil en la integració de dinàmiques en estats fonamental i excitats.<sup>119-121</sup>

En aquest treball, s'han utilitzat els mètodes OMx per a la propagació de les dinàmiques QM/MM tant a l'estat fonamental com a estats excitats presentades en els Capítols 5 i 6 amb un èxit considerable, tenint en compte el grau d'aproximació assumit.

### 2.1.3 Mecànica Molecular

En ocasions, tot i l'elevat nombre d'aproximacions emprades en els mètodes semiempírics, que encara es troben dins del context de la química quàntica, es fa computacionalment impossible la seva aplicació a sistemes de gran volum, com ho pot ser una proteïna solvatada en aigua. Per aquesta raó, es van desenvolupar els mètodes de mecànica molecular o MM d'aquí en endavant, pel càlcul de l'energia. Aquests mètodes basen el càlcul



de l'energia (i del gradient, si és necessari) en una expressió parametritzada per a cada cas. Les expressions dels càlculs MM tenen la forma general:

$$E_{MM} = \sum_{\text{enllaços}} k_r (r - r_0)^2 + \sum_{\text{angles}} k_\theta (\theta - \theta_0)^2 + \sum_{\text{díedres}} k_\phi [1 + \cos(n\phi + \delta)] + \sum_{A,B} \left\{ \left[ \varepsilon_{A,B} \left( \frac{\sigma_{A,B}}{r_{A,B}} \right)^{12} - \left( \frac{\sigma_{A,B}}{r_{A,B}} \right)^6 \right] + \frac{1}{4\pi\epsilon_0} \frac{q_A q_B}{r_{A,B}} \right\} \quad (2.20)$$

on  $r$ ,  $\theta$  i  $\phi$  corresponen a les distàncies d'enllaç, angles i torsions, respectivament, essent  $r_0$  i  $\phi_0$ , les distàncies i angles d'equilibri. Les constants  $k_r$ ,  $k_\theta$  i  $k_\phi$  defineixen la força dels enllaços, angles i díedres, respectivament i els paràmetres  $n$  i  $\delta$  defineixen la multiplicitat i fase de les torsions. La suma sobre A i B corre només per a parelles d'àtoms no enllaçats entre si i els termes  $\varepsilon_{A,B}$  i  $\sigma_{A,B}$  són els paràmetres de Lennard-Jones, essent  $r_{A,B}$  la distància entre tals parelles. Finalment, el terme  $q_i$  representa la càrrega de l'àtom en qüestió. Des d'un punt de vista interpretatiu, el primer terme de l'equació correspon a la força de cada enllaç, el segon a la força dels angles i el tercer a la dels díedres. Les constants per a cada terme donen una idea de la força d'aquesta interacció. L'últim terme de l'equació representa les interaccions no enllaçants: les interaccions de Van der Waals i les interaccions coulòmbiques càrrega-càrrega. Com he dit, aquesta és la versió general i bàsica de qualsevol mètode MM però hi ha versions que inclouen altres termes per tenir en compte d'altres possibles interaccions com poden ser la interacció  $\pi\pi$ -stacking o l'estabilització de l'esquelet de les proteïnes degut a la deslocalització de l'enllaç peptídic.

A banda de l'inclusió de termes addicionals pel càlcul de l'energia, el que diferencia els mètodes MM entre ells és el conjunt de constants i paràmetres que defineixen l'energia. Em refereixo a les constants  $k_r$ ,  $k_\theta$ ,  $k_\phi$ , les distàncies i angles d'equilibri, els paràmetres de Lennard-Jones, les càrregues atòmiques, etc. Cadascun d'aquest conjunt de paràmetres es deriven de manera consistent tot ajustant-los a càlculs de referència de caràcter *ab initio* i s'anomenen "campos de forces". En el present treball, quan ha estat necessari, s'ha utilitzat el camp de forces CHARMM27,<sup>122,123</sup> que va ser parametritzat especialment per a la correcta descripció de sistemes biològics i que es troba implementat en el programari CHARMM-35.<sup>124,125</sup> Cal anomenar, també, d'altres camps de força d'ús comú com l'AMBER<sup>126,127</sup> o el GROMACS<sup>128</sup> que també han demostrat descriure sistemes biològics amb èxit.

El càlcul de l'energia per mitjà de l'equació 2.20 té evidents avantatges computacionals i s'utilitza generalment per determinar propietats d'equilibri o dinàmiques de sistemes de gran volum, com poden ser les proteïnes. Amb tot, presenta unes limitacions intrínseques molt clares: una d'elles és que no descriu bé l'enllaç químic si aquest es troba prou desplaçat de la configuració d'equilibri. El terme que descriu l'enllaç en qualsevol mètode MM és un terme quadràtic, que tendeix a infinit si  $r$  es fa gran.

L'enllaç químic, en general, segueix un potencial de tipus Morse, on el potencial s'estabilitza a un valor en concret a partir d'una certa distància  $r$  respecte  $r_0$ . Això fa que aquests mètodes siguin del tot inadequats per descriure qualsevol procés químic on es trenquin o formin enllaços, o que ocorri fora de les distàncies d'equilibri. Per tractar de solucionar aquest problema, s'utilitzen càlculs híbrids de mecànica quàntica (QM) i mecànica molecular (MM), els anomenats QM/MM. En aquest context, es tracta una part petita del sistema a un nivell quàntic, ja sigui purament *ab initio*, DFT o semiempíric, i la resta a nivell de mecànica molecular. La regió QM ha d'incloure tots aquells àtoms que formen o trenquen enllaços durant el procés a estudiar, a més d'incorporar els àtoms que participen de l'energia amb notoris efectes quàntics.

### 2.1.4 Els Mètodes Híbrids QM/MM

Els mètodes híbrids QM/MM han marcat un abans i un després en l'estudi computacional de sistemes complexos des que van ser introduïts per Warshel i Karplus l'any 1972<sup>129</sup> i utilitzats per primer cop en l'estudi d'una reacció enzimàtica l'any 1976,<sup>130</sup> essent avui en dia una eina principal en la elucidació teòrica del funcionament dels sistemes químics de gran mida.<sup>131</sup> Tant és així que l'any 2013 es va atorgar el premi Nobel a M. Karplus, M. Levitt i A. Warshel per "*the development of multiscale models for complex chemical systems*". Com s'ha dit anteriorment, la idea dels mètodes QM/MM és dividir el sistema en una part QM, on ocorre el trencament o formació d'enllaços, i la resta, la part MM, que afecta de manera no enllaçant a la primera. La tria de la part QM és un punt crític pel bon funcionament del mètode, doncs ha d'incloure tots els àtoms que contribueixin amb efectes quàntics de manera substancial. Naturalment, per tal de mantenir a ratlla el cost computacional, s'intenta sempre triar una part QM el més compacta possible. En aquest treball, típicament, s'han fet servir sistemes on la part QM ha estat entre els 50 i 100 àtoms, deixant per la part MM entre 20.000 i 35.000 àtoms, és a dir, la resta del sistema.

Com passa sempre en física quan s'intenta utilitzar dues teories de naturalesa diferent en el mateix context, més enllà de les limitacions i problemes de cada teoria per separat, apareixen problemes relacionats amb la conjunció de les mateixes. En aquest sentit, els mètodes QM/MM pateixen de la descripció de la unió entre les parts QM i MM, la zona frontera. Com he dit, al ser zones tractades amb teories de naturalesa diferent, no hi ha manera totalment coherent de tractar aquesta discontinuïtat, havent d'utilitzar, encara, aproximacions addicionals.

A no ser que els àtoms de la part QM no estiguin enllaçats amb cap altre de la part MM, fet extraordinàriament excepcional, la definició de la part QM fa que, inevitablement, s'hagi de tallar algun enllaç químic i definir un àtom QM al costat d'un MM. Per abordar la problemàtica de tallar un enllaç covalent, existeixen dues aproximacions:

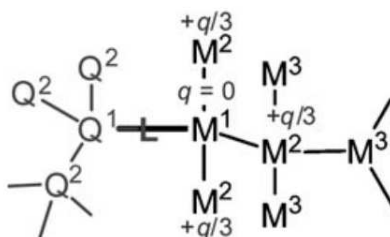


FIGURA 2.1: En el tractament *charge-shift*, la càrrega del primer àtom de la part MM ( $M^1$ ) que fa frontera amb la part QM ( $Q^1$ ), es reparteix amb els àtoms MM contigus ( $M^2$ ). Parelles de càrregues addicionals es situen a prop dels àtoms  $M^2$  per a restaurar els dipols  $M^1$ - $M^2$ .

Extret de Senn *et al.*<sup>131</sup>

- **Orbitals localitzats:**<sup>132-135</sup> En aquest context, es localitzen orbitals híbrids en els àtoms QM fronterers que saturen electrònicament la zona QM. Aquests mètodes són el *Frozen Orbital Method*, el *Generalized Hybrid Orbital*, etc.
- **Link atoms:**<sup>136,137</sup> En aquesta aproximació, es saturen els enllaços trencats de la part QM amb àtoms addicionals, normalment hidrogens. D'aquesta manera s'aconsegueix igualment saturar electrònicament la part QM, evitant les complicacions que suposa la inclusió d'orbitals localitzats.

Encara que des d'un punt de vista formal els mètodes d'orbitals híbrids localitzats són més correctes, l'ús dels *link atoms* està àmpliament acceptat donada la simplicitat del model i la qualitat dels resultats que proporcionen.<sup>138</sup> Amb això, l'addició de nous àtoms al sistema porta alguns problemes, per exemple la proximitat d'aquests a la zona MM, que pot donar lloc a una sobrepolarització de la densitat electrònica atès que les càrregues MM quedaràn molt a prop de la densitat electrònica QM. Per solucionar aquest problema, tot i que existeixen diferents aproximacions, s'ha utilitzat el mètode anomenat *charge shift*, introduït per Sherwood i col·laboradors, que tracta de repartir la càrrega de l'àtom MM fronterer amb els àtoms MM contigus<sup>139,140</sup> (vegeu la Figura 2.1). La combinació del *link atom* amb el *charge shift*, esquema que s'ha utilitzat en els càlculs de tipus QM/MM realitzats en aquesta tesi, és una metodologia conceptualment senzilla però que dóna bons resultats, a més de ser de fàcil implementació i aplicació.<sup>131</sup>

Pel què fa al càlcul de l'energia, aquesta s'obté com a suma de l'energia de les parts QM i MM per separat, més un terme d'interacció, que essencialment té en compte la interacció de la densitat electrònica de la part QM amb les càrregues de la part MM:

$$E_{QM/MM}(S) = E_{MM}(\odot) + E_{QM}(I + L) + E_{QM-MM}(I, \odot) \quad (2.21)$$

on les lletres entre parèntesi representen els subespais sobre els quals es realitza el càlcul:  $\mathbb{S}$  representa l'espai de tot el sistema,  $\mathbb{O}$  el de les càrregues MM,  $\mathbb{I}$  el dels àtoms QM i  $\mathbb{L}$  el dels *link-atoms*.

L'energia MM s'obté directament amb l'expansió de l'equació 2.20 però sovint s'aplica l'anomenat *cutoff*, que és una distància a partir de la qual no es calculen les interaccions no enllaçants. Cal considerar que, al tractar-se de sistemes en 3 dimensions, el nombre d'interaccions no enllaçants per a un àtom creix amb  $r^3$ , així que realment es pot fer molt més eficient el càlcul negligint les interaccions a partir d'una distància considerable, distància a partir de la qual la interacció serà mínima. Ara bé, com s'introdueix l'interacció de les càrregues MM amb els electrons de la part QM? Tot i que primerament, es van plantejar mètodes menys sofisticats com l'anomenat *mechanical embedding* (ME), en el desenvolupament d'aquest treball s'ha utilitzat, quan ha calgut, el tractament anomenat *electrostatic embedding*<sup>141,142</sup> (EE). En aquest context, s'inclouen les càrregues MM en el Hamiltonià electrònic, així que es resol el mètode autoconsistent en aquest "bany" de càrregues puntuals. Per tant hi ha una polarització directa de la densitat electrònica de la part QM per causa de les càrregues MM, creada per aquest hamiltonià d'interacció QM/MM:

$$\hat{H}_{\text{int.QM/MM}}^{\text{el}} = - \sum_i^N \sum_{j \in \mathbb{O}}^P \frac{q_j}{|\vec{r}_i - \vec{R}_j|} + \sum_{\alpha \in \mathbb{I} + \mathbb{L}}^M \sum_{j \in \mathbb{O}}^P \frac{q_j Q_\alpha}{|\vec{R}_\alpha - \vec{R}_j|} \quad (2.22)$$

on  $q_j$  són les càrregues MM que es troben a  $\vec{R}_j$ ,  $Q_\alpha$  són les càrregues nuclears de la part QM localitzades a  $\vec{R}_\alpha$  i  $r_i$  correspon a la posició dels electrons. Els índex  $i$ ,  $j$  i  $\alpha$  corren pels  $N$  electrons,  $L$  càrregues MM i  $M$  nuclis QM.

Tot i això, cal notar que la part MM no es polaritza per l'acció de la densitat QM, fet que pot portar a alguns problemes a nivell de la descripció energètica i estructural del sistema.

## El Codi ChemShell

Per tal de poder dur a terme els càlculs de tipus QM/MM en aquest treball, els quals són la majoria dels realitzats, s'ha utilitzat la plataforma ChemShell.<sup>143</sup> El ChemShell, és una interfície que permet combinar l'acció de diferents programes per la part QM i per la part MM, afegint els càlculs corresponents a la interacció QM/MM. El programa és capaç de connectar càlculs QM d'altres programes externs com ara Gaussian, Turbomole, MNDO, etc, amb d'altres MM com ara Amber, Gromacs o Charmm. Aquesta propietat el fa increïblement versàtil per a qualsevol aplicació que necessiti de càlculs QM/MM. Incorpora mòduls per realitzar dinàmiques, minimitzacions, càlculs d'energia, etc., i si s'és coneixedor del codi font, permet la incorporació de noves funcionalitats que permetin realitzar nous tipus de càlculs, amb combinacions de mòduls QM i MM encara

per desenvolupar. En aquest treball, s'ha utilitzat la versió estàndard del Chemshell per als càlculs QM/MM dels Capítols 5 i 6 i s'ha desenvolupat un mòdul nou en una versió de proves en col·laboració amb el Max-Planck Institut für Kohlenforschung i el prof. W. Thiel per realitzar els càlculs de dinàmica no adiabàtica presentats al Capítol 6.

## 2.2 Dinàmica

En molts aspectes de la química, no només teòrica, es duen a terme discussions basades en l'anàlisi topogràfica de la superfície d'energia potencial. Efectivament, la superfície d'energia potencial (PES) és una gran eina per entendre la química d'un sistema però la descripció de processos complexos basada tan sols en la determinació de punts concrets i matemàticament característics de la mateixa (essencialment mínims i punts de sella), a vegades pot resultar insuficient. La PES és una hipersuperfície de  $3N-6$  dimensions (per a molècules no lineals) però pot ser relativament clara en el seu anàlisi per a molècules o sistemes de petita dimensió. Ara bé, per a sistemes biològics, la quantitat de punts estacionaris que presenta pot ser molt gran i, per tant, la importància relativa que tenen entre ells es veu qüestionada. Quin sentit té la consideració d'un mínim trobat en una proteïna quan sabem que prop del mateix és molt probable que m'hi hagi molts d'altres? Sovint es farà la incorrecta interpretació que el sistema es troba en aquest mínim i que, per tant, la geometria del mateix és representativa de la situació real a una certa temperatura. Amb tot, podem obtenir una informació molt més fina i representativa si, per mitjà de simulacions de Dinàmica Molecular o MD (*molecular dynamics*) d'aquí en endavant, calculem com es mou el sistema al voltant d'aquest mínim (i d'altres) a una temperatura  $T$  i amb quina freqüència visita cada configuració. Les propietats macroscòpiques dels sistemes químics són el resultat de l'observació promig de les propietats de cada sistema molecular que forma la mostra a observar així que la manera més correcta de reproduir dades experimentals és promitjar les propietats moleculars d'un col·lectiu que representi correctament la població de l'espai de fases, col·lectiu que podem obtenir de les simulacions dinàmiques.<sup>144,145</sup>

A part de la complexitat que presenten les PES dels sistemes biològics, hi ha un problema conceptual en el sol anàlisi de les mateixes a l'hora de parlar de reactivitat: no es té en compte l'energia cinètica que, efectivament, té el sistema i que és el motor de les reaccions químiques. Les posicions i velocitats dels àtoms d'un sistema molecular (en un context clàssic) poden determinar camins de reacció que siguin més exòtics que els previstos pel camí de mínima energia. Per fer una analogia macroscòpica, posem el cas que volem preveure per on passarà un avió que creui una serralada: la visió clàssica i sovint estaticista que malauradament encara contamina les discussions en química teòrica, és que passarà per la vall de menys

alçada i a més, seguint la llera del riu (que posem pel cas que és el camí de menys alçada possible). Això no és una visió realista del procés, doncs tothom convindrà en que, a priori, no podem saber per on passarà l'avió a partir del sol anàlisi de la topografia de la serralada. La manera de saber per on passa és, efectivament, calcular per on passa. Hauríem de calcular a quina alçada vola i, per tant, per quines zones pot passar. A partir d'aquí, coneixent les seves condicions inicials i les lleis que regeixen el seu moviment (les lleis de Newton en aquest cas) podríem, finalment, determinar la seva trajectòria. Es podria donar el cas que volés a suficient alçada com per creuar la serralada just passant per sobre del pic més alt, sense que els passatgers puguin, ni de bon tros, intuir la llera del riu, conclusió totalment contraposada a la obtinguda de l'anàlisi de la topografia de la serralada. De la mateixa manera, per escatir de manera fina com transcorre una reacció química a nivell molecular, no n'hi ha prou amb analitzar uns punts de la PES (típicament 3 per una reacció senzilla: el mínim de reactius, el de productes i el punt de sella que els connecta), si no que hem de calcular les trajectòries dels successos que tenen lloc. En aquest sentit, les simulacions de Dinàmica Molecular són necessàries per a la correcta i completa descripció del processos químics que ocorren en un sistema molecular complex.

### 2.2.1 Dinàmica Clàssica

La denominació de "clàssica" pel què fa a la dinàmica, fa referència al tractament que es dona als nuclis en la mateixa. En el context de la mecànica clàssica, els nuclis segueixen la segona llei de Newton per al seu moviment:

$$-\frac{\partial V}{\partial q_i} = \frac{dp_i}{dt} \quad (2.23)$$

on  $V$  és l'energia potencial,  $q_i$  les coordenades de la partícula  $i$  i  $p_i$  el seu moment lineal. Si considerem que la massa de cada partícula és constant en el temps, podem escriure:

$$-\frac{\partial V}{\partial q_i} = m_i \frac{d^2 q_i}{dt^2} \quad (2.24)$$

on  $m_i$  és la massa de la partícula  $i$ .

La resolució de l'equació del moviment de Newton requereix mètodes numèrics per a l'integració d'equacions diferencials. Un dels mètodes inicialment utilitzats fou l'integrador de Verlet,<sup>146</sup> basat en una doble expansió de Taylor. Actualment s'utilitzen d'altres mètodes més avançats d'ordre superior, com el Velocity-Verlet<sup>147</sup> o el Leap-Frog,<sup>148</sup> que és el que s'ha utilitzat en aquest treball. Tanmateix, la integració numèrica de les equacions fa que haguem de transformar el  $dt$  en un  $\Delta t$ , fent passos temporals no infinitesimals. Naturalment, com més petit sigui el pas d'integració, més

analíticament correcta serà la trajectòria obtinguda, però més computacionalment cara resultarà, doncs per a cada pas de dinàmica cal calcular l'energia i el gradient. Sovint, s'accepta un pas d'integració que sigui algun ordre de magnitud més petit que el procés que es vol observar. Típicament, per a dinàmiques clàssiques, el pas d'integració és de l'orde de  $\Delta t = 1$  fs.

Vull remarcar que, al tractar-se de dinàmica clàssica, qualsevol efecte quàntic relacionat amb la inherent naturalesa quàntica dels nuclis, quedarà anul·lat en tota simulació de MD. Així doncs, fenòmens com la deslocalització nuclear o l'efecte túnel no podran ser representats. Com sempre, si aquests fenòmens no són importants pel cas d'estudi, un tractament clàssic dels nuclis no hauria de donar resultats significativament diferents a un tractament quàntic.

Tot i que, com he dit, la dinàmica és purament clàssica pel moviment dels nuclis, el potencial que noten els mateixos es pot obtenir a nivell de mecànica molecular (MM-MD) a nivell de mecànica quàntica (QM-MD) o a un nivell híbrid (QM/MM-MD). En els casos QM o QM/MM, també es pot obtenir una superfície analítica sobre la qual es propaguin les equacions de moviment (tal i com es fa en les MM-MD), però davant l'extrema dificultat de l'exploració de la PES i el següent ajust d'una funció analítica que la descriu, sovint s'utilitza l'anomenat mètode *on-the-fly* Born-Oppenheimer MD. En aquest context es calcula l'energia i el gradient per cada instant i es propaguen les posicions fins al següent punt de l'espai de fases. Aquest tractament es pot fer gràcies a que la teoria clàssica (al contrari que la quàntica) és una teoria local: les partícules només noten la força del potencial en el punt on es troben, per tant no ens cal tenir una funció de potencial global. En altres paraules, no hem de conèixer la forma analítica de la PES. La superfície sobre la qual es propaguen les equacions en el cas QM o QM/MM és la PES, però sempre obtingudes en un context adiabàtic. S'assumeix que l'aproximació BO té validesa en el rang de la dinàmica i per tant es pot obtenir una representació adiabàtica del moviment nuclear. Això vol dir que els nuclis es mouren en una sola PES, associada a un estat electrònic definit. Per a la majoria de casos que s'estudien, la representació adiabàtica és totalment vàlida perquè els estats electrònics romanen desacoblats durant la propagació de la dinàmica però, què passa quan la dinàmica és susceptible de visitar zones de l'espai de fases on l'aproximació de Born-Oppenheimer deixa de tenir validesa?

### 2.2.2 Dinàmica Semi-clàssica

Anteriorment s'ha descrit qualitativament les bases de l'aproximació de BO, però es fa evident que les condicions necessàries per l'aplicabilitat de la mateixa no sempre es compliran. En aquests casos, una aproximació purament clàssica del moviment nuclear es fa insuficient. Davant la impossibilitat de realitzar un tractament quàntic dels mateixos en sistemes de

mida considerable, s'opta per una aproximació semi-clàssica on, a partir d'un tractament clàssic dels nuclis, s'intenta descriure un comportament quàntic.

En el desenvolupament complert de l'aproximació de Born-Oppenheimer apareixen dues magnituds que són l'acoblament escalar i l'acoblament derivatiu,  $G_{ij}$  i  $\vec{F}_{ij}$  respectivament, i que es defineixen:

$$G_{ij} = \sum_{\alpha} \frac{1}{2m_{\alpha}} \langle \varphi_j | \hat{\nabla}_{\alpha}^2 | \varphi_i \rangle \quad (2.25)$$

$$\vec{F}_{ij} = \sum_{\alpha} \frac{1}{2m_{\alpha}} \langle \varphi_j | \hat{\nabla}_{\alpha} | \varphi_i \rangle \quad (2.26)$$

on  $m_{\alpha}$  representa les masses nuclears i els estats  $\varphi_i$  i  $\varphi_j$  són els estats electrònics adiabàtics. Per tal que l'aproximació de BO es compleixi, com s'ha dit, els acoblaments entre els estats, és a dir les magnituds  $G_{ij}$  i  $\vec{F}_{ij}$ , han de ser 0. Com que la massa dels nuclis és molt més gran que la dels electrons, això serà així mentre els termes d'acoblament entre estats es mantinguin petits. Es pot demostrar la següent relació:

$$\langle \varphi_j | \hat{\nabla}_{\alpha}^2 | \varphi_i \rangle = \frac{\langle \varphi_j | (\hat{\nabla}_{\alpha}^2 \hat{H}_{el}) | \varphi_i \rangle}{E_i(\vec{R}) - E_j(\vec{R})} \quad (2.27)$$

Això vol dir que en situacions on els estats electrònics adiabàtics s'acosten en energia, l'aproximació de BO deixa de ser aplicable ja que el terme de l'Equació 2.27 es fa massa gran. De fet, en punts on hi ha degeneració dels estats electrònics, hi ha una singularitat. Aquests punts s'anomenen interseccions còniques (CI) i tenen una implicació rellevant en la fotoquímica. Un sistema molecular pot canviar el seu estat electrònic tot passant per una CI, ja que en aquest punt, l'estat electrònic no està definit com a tal. Les dimensió d'una intersecció cònica és, en general, la mateixa que la de la PES, menys dos graus de llibertat ( $3N - 8$  per a molècules no lineals). Això és degut a que només hi ha dues direccions on es trenca la degeneració, que són les que defineixen el vector de diferència de gradients i el vector d'acoblament entre estats,  $\vec{g}$  i  $\vec{h}$  respectivament, que s'expressen com a:

$$\vec{g} = \frac{\partial(E_i - E_j)}{\partial q} \quad (2.28)$$

$$\vec{h} = \left\langle \varphi_j \left| \frac{\partial \hat{H}}{\partial q} \right| \varphi_i \right\rangle \quad (2.29)$$

Aquests dos vectors, que són els únics que trenquen la degeneració, defineixen el que s'anomena *branching space* de la CI en qüestió. L'espai complementari del *branching space*, és a dir, aquell que manté la degeneració



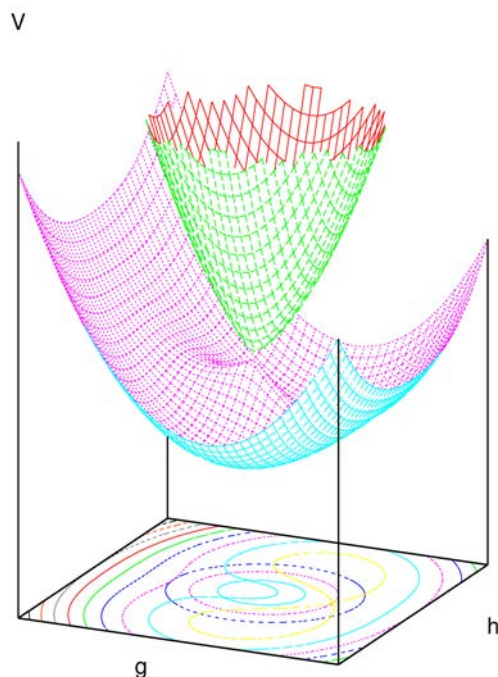


FIGURA 2.2: Representació model de dues PES en funció dels vectors  $\vec{g}$  i  $\vec{h}$ . La CI es localitza en el punt on es troben i es produeix la degeneració.

i és ortogonal a les direccions  $\vec{g}$  i  $\vec{h}$ , s'anomena *seam*. Les interseccions còniques, doncs, formen un *seam* (un espai de  $3N - 8$  dimensions per a molècules no lineals) per on un sistema molecular pot canviar d'estat electrònic. Si es representen gràficament l'energia potencial dels estats electrònics implicats en una CI en funció de les direccions  $\vec{g}$  i  $\vec{h}$ , la CI quedarà definida en un punt, però cal tenir en compte que en realitat es tracta d'un "hiperpunt" de dimensió  $3N - 8$  (Figura 2.2).

Les simulacions dinàmiques en estats excitats són susceptibles d'explorar zones on l'aproximació BO no sigui aplicable, ja sigui perquè el terme de l'Equació 2.27 es faci massa gran o perquè directament apareguin interseccions còniques. Així doncs, com es desenvolupa una dinàmica fora dels límits d'aplicabilitat de l'aproximació BO? S'ha desenvolupat un gran nombre de mètodes per al tractament d'efectes no adiabàtics en dinàmiques moleculars. Entre ells, un dels que gaudeix de més acceptació és l'anomenat *surface hopping*, en que els nuclis evolucionen en una sola PES però es permet el transit entre estats electrònics per mitjà d'un sorteig que es fa en base al pes de l'estat electrònic en qüestió.<sup>149</sup> Cal notar que l'aproximació BO és tant robusta que fins i tot es fa servir en els casos on no és aplicable per a calcular els diferents estats adiabàtics sobre els quals propagar una dinàmica no adiabàtica.

En el context del mètode *surface hopping*, es pren un eixam de trajectòries on cadascuna d'elles evolucionarà en una sola PES. A cada pas de dinàmica, es calcula la probabilitat que una trajectòria faci un salt i canviï de PES.

Aquesta probabilitat està controlada per un algoritme que garanteix que la fracció de trajectòries en un estat electrònic determinat és aproximadament igual a la població d'aquest estat.<sup>150</sup>

En el *surface hopping* els nuclis es mouen clàssicament en un trajectòria  $\vec{R}(t)$ . El Hamiltonià pel moviment electrònic és:

$$\hat{H}(\vec{r}, \vec{R}) = -\frac{\hbar^2}{2} \sum_{\alpha} \frac{1}{m_{\alpha}} \nabla_{\vec{r}_{\alpha}}^2 + \hat{V}_{rR}(\vec{r}, \vec{R}) \quad (2.30)$$

on  $\alpha$  correspon als graus de llibertat electrònics i el potencial  $V_{rR}$  inclou els potencials electró-electró i electró-nucli. Com que el Hamiltonià electrònic és dependent del temps a través de  $\vec{R}(t)$ , la funció d'ona electrònica  $\Phi(\vec{r}, \vec{R}, t)$  associada evoluciona segons l'equació d'Schrödinger dependent del temps:

$$i\hbar \frac{\partial}{\partial t} \Phi(\vec{r}, \vec{R}, t) = \hat{H} \Phi(\vec{r}, \vec{R}, t) \quad (2.31)$$

La funció d'ona electrònica es pot expressar com a expansió d'una base orthogonal coneguda  $\phi_i(\vec{r}, \vec{R})$  per la qual prendrem les funcions d'ona de la representació adiabàtica BO, és a dir les solucions de l'equació d'Schrödinger electrònica independent del temps:

$$\Phi(\vec{r}, \vec{R}, t) = \sum_i c_i(t) \phi_i(\vec{r}, \vec{R}) \quad (2.32)$$

La substitució de l'equació 2.32 a l'equació d'Schrödinger (2.31), després de la multiplicació per l'esquerra de  $\phi_j^*(\vec{r}, \vec{R})$  i la integració sobre  $\vec{r}$  porta a:

$$i\hbar \frac{d}{dt} c_j(t) = \sum_i c_i(t) [H_{ij} - i\hbar \dot{\vec{R}} \cdot \vec{d}_{ij}] \quad (2.33)$$

on  $H_{ij}$  i  $\vec{d}_{ij}$  són:

$$H_{ij}(\vec{R}) = \langle \phi_j(\vec{r}, \vec{R}) | \hat{H}(\vec{r}, \vec{R}) | \phi_i(\vec{r}, \vec{R}) \rangle \quad (2.34)$$

$$d_{ij}(\vec{R}) = \langle \phi_j(\vec{r}, \vec{R}) | \nabla_{\vec{R}} | \phi_i(\vec{r}, \vec{R}) \rangle \quad (2.35)$$

i on s'ha utilitzat la regla de la cadena:

$$\left\langle \phi_j(\vec{r}, \vec{R}) \left| \frac{\partial \phi_i(\vec{r}, \vec{R})}{\partial t} \right. \right\rangle = \dot{\vec{R}} \cdot \vec{d}_{ij} \quad (2.36)$$

En la representació adiabàtica, la matriu Hamiltoniana (2.34) és diagonal així que l'equació 2.33 queda simplificada:

$$i\hbar \frac{d}{dt} c_j(t) = c_j(t) \epsilon_j - i\hbar \sum_i c_i(t) \dot{\vec{R}} \cdot \vec{d}_{ij} \quad (2.37)$$

Així doncs, l'amplitud dels estats quàntics a cada temps ( $c_j(t)$ ) es calcula integrant l'equació 2.37 durant la trajectòria  $\vec{R}(t)$  i amb aquesta informació es pot fer el sorteig estocàstic que decidirà si hi ha o no hi ha canvi d'estat electrònic.

Existeixen un gran nombre d'algoritmes que determinen la probabilitat de saltar d'estat electrònic. Un dels utilitzats més habitualment és el Fewest Switches Algorithm (FSA). Aquest algoritme calcula una probabilitat de transició per a un pas d'integració no infinitesimal ( $\Delta t$ ) a partir dels coeficients  $c_j(t)$  a cada temps, llavors es tria un nombre aleatori entre 0 i 1 i es decideix si hi ha transició. Si hi ha canvi de PES, cal reajustar les velocitats per garantir la conservació de l'energia. Es poden trobar detalls sobre l'algoritme FSA i el reajust de velocitats a la bibliografia especialitzada.<sup>149–151</sup>

En aquest tractament, altre cop, es fa servir una aproximació clàssica a un comportament quàntic, així que naturalment no es tracta d'un mètode lliure de certa problemàtica. Amb tot, el *surface hopping* generalment es comporta bé per a sistemes moleculars i permet incloure efectes no adiabàtics que són bàsics per explicar processos en fotoquímica, especialment transicions no radiatives<sup>120, 121, 152–155</sup> i reaccions de transferència de càrrega.<sup>156</sup>

En el transcurs d'aquesta tesi, s'ha desenvolupat una implementació d'aquest mètode fent servir el codi ChemShell juntament amb el grup del prof. W. Thiel del Max-Planck-Institut für Kohlenforschung i s'ha utilitzat pels resultats presentats en el Capítol 6.

### 2.2.3 Dinàmica Quàntica

Tot el descrit fins ara, fins hi tot l'aproximació semiclàssica del *surface hopping*, fa servir la teoria clàssica per a propagar el moviment nuclear. Això fa que els efectes purament quàntics com ara la ressonància, la deslocalització o l'efecte túnel no puguin aparèixer de forma natural. Eliminar la naturalesa quàntica dels nuclis en pro de la simplificació de les equacions és un requisit indispensable per a poder fer càlculs sobre sistemes químics i a més és una aproximació raonable per a nuclis pesats que no presentaran grans efectes quàntics. Ara bé, recordem que la reacció responsable de la fluorescència en les proteïnes verdes fluorescentes és una transferència protònica. El protó és una partícula que, donat la seva reduïda massa, pot presentar efectes quàntics importants, sobretot en casos on es pugui donar una certa deslocalització de la seva funció d'ona associada o si es troba

amb una barrera. Per aquesta raó, hem de presentar alguna metodologia basada en la dinàmica quàntica (QD) que permeti el tractament purament quàntic del moviment nuclear, per quan sigui necessària la seva aplicació. En aquest treball s'ha utilitzat dos tipus de simulacions de QD: la determinació d'estats estacionaris vibracionals (funcions pròpies i valors propis vibracionals) i la propagació del paquet d'ones nuclear a l'estat excitat.

Per a la determinació d'estats vibracionals, s'ha pres alguns perfils d'energia potencial per a la transferència protònica i s'ha resolt la següent equació d'Schrödinger independent del temps:

$$\hat{H}\psi_i = [\hat{T} + \hat{V}]\psi_i = E_i\psi_i \quad (2.38)$$

On  $\hat{V}$  és el potencial calculat per a la transferència protònica i  $\hat{T}$  l'operador d'energia cinètica per al moviment del protó (o deuteró, quan ha calgut). S'ha utilitzat la representació *sync-DVR* de Colbert i Miller<sup>157</sup> per a la construcció de la matriu Hamiltoniana corresponent, utilitzant 50 punts com a base. Els vectors propis ( $\{\psi_i\}$ ) i valors propis ( $\{E_i\}$ ) s'han obtingut per mitjà de la diagonalització de tal matriu.

Per al càlcul de l'evolució temporal del paquet d'ones nuclear s'ha assumit que a temperatura ambient només hi ha població significativa a l'estat fonamental vibracional de l'estat fonamental electrònic ( $|\psi_0^{S_0}\rangle$ ), així que aquest s'ha promocionat a l'estat electrònic excitat fotoactiu ( $\psi^{S_1}$ ). Com que aquest estat  $|\psi_0^{S_0}\rangle$  no és a priori un estat propi del Hamiltonià de l'estat excitat, s'ha determinat la seva evolució temporal seguint l'equació d'Schrödinger depenent del temps. Com que el Hamiltonià sí que és independent del temps, podem escriure la funció d'ona en funció de la forma espectral de l'operador d'evolució temporal<sup>158</sup>

$$\Psi(\vec{q}, t) = \langle \vec{q} | e^{-i\hat{H}t/\hbar} | \psi_0^{S_0} \rangle = \sum_i \langle \vec{q} | \psi_i^{S_1} \rangle \langle \psi_i^{S_1} | \psi_0^{S_0} \rangle e^{-iE_i t/\hbar} \quad (2.39)$$

en la que el sumatori corre per tots els estats vibracionals de l'estat excitat  $S_1$  i on  $E_i$  són les energies de tals estats. La integració de l'equació 2.39, ens permet el seguir l'evolució del paquet d'ones a l'estat excitat.



### 3 Objectius

En el present treball, com en qualsevol tesi, l'objectiu global és crear un coneixement ferm i estable sobre el tema tractat, en aquest cas, el comportament espectroscòpic de les proteïnes fluorescents, la relació estructura-propietats i la fotoquímica d'aquests sistemes. A partir de casos particulars estudiats, es pretén generalitzar un coneixement aplicable a la resta de casos podent fer, així, certes prediccions del funcionament de sistemes que encara no han estat testats experimentalment.

En el Capítol 5 de la part de resultats es tractarà amb el càlcul i anàlisi dels espectres d'absorció de proteïnes fluorescents vermelles i verdes, tot relacionant les propietats espectroscòpiques d'equilibri de les mateixes amb les seves característiques estructurals en solució aquosa. Els objectius d'aquest capítol són:

- Reproduir qualitativament i quantitativa els espectres d'absorció de les proteïnes estudiades.
- Ser capaços de descriure les diferències espectroscòpiques més importants entre els principals grups de proteïnes objecte d'estudi, les GFPs i les RFPs, així com les diferències a dins de cada grup.
- Caracteritzar els estats electrònics excitats fotoactius dins del context de la Teoria del Funcional de la Densitat Dependent del Temps.
- Concloure raons el més clares i sistemàtiques possible per a la posició dels màxims de les bandes d'absorció de les proteïnes estudiades.
- Racionalitzar la relació estructura-espectre, descrivint les interaccions del cromòfor amb els residus veïns.
- Amb el coneixement adquirit, ser capaços de proposar noves mutacions que permetin la síntesi de proteïnes amb propietats modulades *ad hoc*.

En el Capítol 6 de l'apartat de resultats, es tractarà de donar descripcions atomístiques a tot el procés físic que ocorre en la doble mutant S65T/H148D-GFP, la Pedra Rosetta. En aquest sentit, s'intentarà mostrar una visió completa del procés per entendre com funciona i, altra vegada, ser capaços de modular-lo a plaer. Els objectius d'aquest capítol són:

- Confirmar que l'estructura del cromòfor i l'apartat amb relació donador-acceptor és estable en solució aquosa, més enllà de l'estructura cristal·lina.

- Contrastar el coneixement assentat que la força motriu de la transferència protònica a l'estat excitat és el guany d'acidesa del cromòfor després de rebre la radiació.
- Elucidar si hi ha components estructurals de l'entorn proteic que influeixen en el procés que té lloc a l'estat excitat fotoactiu.
- Explorar els graus de llibertat de la superfície d'energia potencial involucrats en la transferència protònica i determinar si hi ha, o no, barrera associada a aquest procés
- Explicar, mitjançant un tractament quàntic del protó transferit, l'inusual comportament observat pels estudis experimentals: la transferència ultrarràpida i la presumible deslocalització a l'estat fonamental.
- Donar una explicació atomística del procés que té lloc a l'estat excitat després de la transferència protònica, i que s'associa amb la relaxació de l'entorn proteic.
- Caracteritzar possibles vies de desactivació no radiativa del cromòfor.
- Amb tot l'anterior, ser capaços de reproduir la major part possible de les dades experimentals disponibles per a aquesta doble mutant, especialment les dades de fluorescència resoltes en el temps.
- Amb tot el coneixement adquirit, proposar mutacions que modifiquin el comportament de la proteïna per que adquireixi les propietats desitjades.

Aquests objectius, comporten una sèrie de reptes de caire metodològic. En aquest sentit, es planteja el desenvolupament de les tècniques i mètodes necessaris per a l'estudi de cada cas, com a un dels objectius principals d'aquest treball. Les aportacions metodològiques es presenten al Capítol 4 i al Capítol 6.

# II

## NARRATIO ET CONFIRMATIO

Ecco il mondo - vuoto e tondo  
S'alza, scende - balza, splende  
Fa carole - intorno al sole,  
Trema, rugge - dà e disturgge.  
Ora sterile, or fecondo  
Ecco il mondo.

Mefistofele - Mefistofele (A. Boito)





## 4 Task-Farming

El primer (i present) capítol de resultats que es troba en aquest treball presenta un mètode per a l'acceleració de dinàmiques moleculars que ha estat del tot indispensable per a l'obtenció dels resultats que es presentaran als Capítols 5 i 6. Aquest mètode, al què ens referirem com a *task-farming*, es podria dir que és una obra d'enginyeria, ja que no incorpora nou coneixement o teoria a la ja existent, només aplica alguns conceptes físics per obtenir una optimització de la despesa computacional que comporten els costosos càlculs que es presentaran en capítols posteriors. Amb tot, per ser un mètode desenvolupat durant el transcurs del doctorat que ha donat lloc a aquesta tesi, es presenta a continuació.

### 4.1 El Problema de la Funció de Partició Canònica

Sovint, el químic computacional busca la relació de magnituds macroscòpiques mesurables experimentalment amb la descripció molecular que se'n desprèn. Per a tal fi, s'utilitza la formulació de la termodinàmica estadística, que permet establir relacions entre els estats de les partícules que formen un sistema molecular i les propietats macroscòpiques del mateix sistema. Prenent un col·lectiu canònic (on la temperatura, número de mols i volum es consideren constants), la probabilitat de població dels estats segueix una distribució de Boltzman:

$$p_i = \frac{e^{-E_i/k_bT}}{Q} \quad (4.1)$$

on  $E_i$  és l'energia de l'estat  $i$  en qüestió,  $k_b$  és la constant de Boltzmann i  $Q$  és la funció de partició canònica:

$$Q = \sum_j e^{-E_j/k_bT} \quad (4.2)$$

La funció de partició, que actua com a funció normalitzadora de la probabilitat, es pot utilitzar per a calcular les propietats macroscòpiques del sistema. A partir del càlcul dels estats energètics possibles, podem calcular quina serà la població mitjana en aquests estats. El valor mig de la

propietat a observar entre els estats poblats, serà aquell valor observable a nivell macroscòpic.

Així doncs, per tal de calcular qualsevol propietat del sistema, en principi en tindriem prou amb calcular tots els estats possibles per al mateix, o en altres paraules, conèixer la funció de partició. Per exemple, es pot demostrar que l'energia mitjana del sistema en el col·lectiu canònic és:

$$\bar{E} = k_b T^2 \left( \frac{\partial \ln Q}{\partial T} \right)_{N,V} \quad (4.3)$$

o la pressió mitjana:

$$\bar{p} = k_b T \left( \frac{\partial \ln Q}{\partial V} \right)_{N,T} \quad (4.4)$$

El càlcul de la funció de partició d'una molècula és factible realitzant unes aproximacions addicionals. Si s'assumeix la separabilitat de l'energia de la molècula:

$$E_{total} = E_{elect} + E_{trans} + E_{rot} + E_{vib} \quad (4.5)$$

es pot separar la funció de partició:

$$Q_{total} = q_{elect} \cdot q_{trans} \cdot q_{rot} \cdot q_{vib} \quad (4.6)$$

Cal notar que aquí hem assumit que tota l'energia prové dels termes anteriorment expressats, sense cap altra contribució. Això és naturalment una altra aproximació.

A partir de l'equació 4.6, podem aproximar la forma de les funcions de partició mitjançant models mecanoquàntics per acabar obtenint expressions analítiques que permetin el càlcul de les variables desitjades. En aquests context, altre cop, cal assumir noves aproximacions (pròpies de cada model). Per exemple, per al moviment de translació es pren el model de la partícula lliure en una caixa tridimensional o pel moviment de vibració, es pren el model de l'oscil·lador harmònic. Aquests tipus de càlculs són els que fan els programes tipus Gaussian09 quan calculen propietats termodinàmiques a partir de càlculs de freqüències.

Aquestes aproximacions són relativament bones per a sistemes de mida petita i la utilització de les mateixes per a descriure propietats macroscòpiques de tals sistemes dona resultats acceptables. Ara bé, els sistemes biològics de gran mida, com ho són les FPs, no poden suportar aquestes aproximacions. La gran quantitat de mínims conformacionals que hi ha en les PES d'aquests sistemes fa que no puguem aplicar l'aproximació harmònica a la seva vibració, per exemple. És per això que ens cal canviar l'estratègia per a l'obtenció de propietats macroscòpiques dels biosistemes.

Per a solucionar aquest problema, assumim la hipòtesi ergòdica, que diu que el promig temporal a temps prou llarg d'una propietat d'un membre del col·lectiu coincideix amb el promig a un temps  $t$  de tal propietat pel col·lectiu, si el col·lectiu és prou gran (estrictament, "temps prou llarg i col·lectiu prou gran", vol dir  $t = \infty$  i  $N = \infty$ ). En altres paraules, la distribució de valors d'una propietat en un col·lectiu, serà la mateixa que el promig temporal d'aquesta propietat en un dels membres d'aquest col·lectiu. Enlloc d'obtenir els estats del sistema de manera analítica (i calcular el pes que tindran per mitjà de l'equació 4.1), els obtenim *in situ*, ja amb el pes que tenen, que serà el nombre de vegades que han estat visitats en el lapse temporal explorat. Això només es complirà si, efectivament, el sistema a tractar és ergòdic. Si el sistema no és ergòdic, vol dir que hi ha zones del seu espai de fases que són literalment inconnexes, és a dir que mai es podrà accedir a una zona si s'està a l'altra, ni a temps infinits. Per tal d'utilitzar les dinàmiques moleculars com a eina per a la predicció de propietats macroscòpiques, hem d'assumir la hipòtesi ergòdica, i així ho fem.

Així doncs, hem transformat el problema teòric de l'obtenció de la funció de partició analítica d'un sistema complex, en un problema pràctic, sobre l'exploració prou exhaustiva de l'espai de fases. En aquest punt, cal esmentar les limitacions computacionals que encara marquen les possibilitats de la recerca que es realitza avui en dia. La integració d'una trajectòria fins a temps molt llargs és, sovint, inviable, especialment si el programa utilitzat no paral·lelitzava l'esforç computacional de manera òptima. En els casos que es presentaran en aquesta tesi, ha calgut emprar el programari ChemShell per a poder propagar simulacions dinàmiques BO a un nivell QM/MM, utilitzant el programa MNDO99 per al càlcul de la part QM. Aquest últim programa no paral·lelitzava eficientment l'esforç computacional, així que l'obtenció de simulacions dinàmiques prou llargues seria del tot inviable.

Per tal de sobrevenir aquest problema, es proposa una metodologia que permetrà accelerar el càlcul de manera espectacular, fins a fer-lo viable. Es tracta de fer una paral·lelització "manual" del càlcul, ja que la pròpiament computacional no és possible.

Assumint que els sistemes tractats són ergòdics, enlloc de propagar una sola simulació dinàmica fins a un temps  $t$ , es propagarà  $n$  simulacions dinàmiques fins a un temps  $t/n$ . Per assegurar que s'escombra adequadament l'espai de fases, i tenint en compte que els sistemes tractats són deterministes (es regeixen per les lleis de Newton pel moviment nuclear), només ens cal iniciar les simulacions dinàmiques des de punts diferents de l'espai de fases. Així doncs, formalment, l'obtenció de microestats a partir d'un paquet de dinàmiques curtes ha de ser igual de vàlida que amb una de llarga. De fet, si explorem l'espai de fases amb un eixam de dinàmiques ens assegurem que la descripció del mateix serà molt més representativa, tenint en compte que integrem les dinàmiques fins a temps no infinits: una sola dinàmica té el risc d'entrar en una zona poc representativa de l'espai

de fases i quedar-hi atrapada, fet que faria que s'obtingués un mal resultat dels promitjos macroscòpics. Si es disposa de diverses simulacions, el pes que eventualment pugui tenir aquesta simulació esbiaixada, és molt menor.

Utilitzar aquesta estratègia té unes claríssimes avantatges computacionals. El programari que s'ha utilitzat en aquest treball només permet la utilització d'un processador per a cada càlcul de simulació dinàmica doncs, com s'ha dit, no paral·lelitzava eficientment. Així doncs, es pot multiplicar per  $n$  la velocitat d'exploració de l'espai de fases si es propaguen  $n$  simulacions independents al mateix temps. Això, tenint en compte els recursos computacionals emprats, ha permès multiplicar la velocitat del càlcul entre  $\sim 30$  i  $\sim 150$  vegades, en funció del cas.

## 4.2 *Tous pour Un, Un pour Tous: el Mètode Task-farming*

En aquesta secció es presenta la implementació que s'ha fet de la idea descrita anteriorment, basada en propagar diverses simulacions dinàmiques de temps més curt, per substituir la propagació d'una dinàmica de temps més llarg i aconseguir, així, una acceleració del càlcul, que s'ha anomenat mètode *task-farming*.

Partint d'un sistema equilibrat, s'ha propagat una sola dinàmica durant un temps  $t$ . Com que el sistema ja està equilibrat, s'assumeix que les estructures obtingudes d'aquesta primera simulació pertanyen, pròpiament, a l'espai de fases que es vol explorar. S'han pres  $n$  fotogrames equiespaiats i se n'ha reiniciat les velocitats. Si no es reiniciessin les velocitats, òbviament, totes les simulacions seguirien la mateixa trajectòria. Després d'un breu període d'equilibratge, s'ha propagat cadascuna de les dinàmiques fins a el temps desitjat. A la Figura 4.1 es representa esquemàticament el protocol clàssic per a les simulacions dinàmiques juntament amb el protocol *task-farming*.

Naturalment, aquest mètode presenta una limitació: en el cas que s'ha-gi d'esdevenir un succés a temps més llargs que els de cadascuna de les simulacions, aquest no apareixerà. Això no es un problema sempre que assumim que partim d'un sistema equilibrat, assumpció que s'ha de fer per a qualsevol simulació dinàmica que pretengui descriure situacions d'equilibri (també totes aquelles que basen els seus resultats en una sola trajectòria).

Aquesta metodologia s'ha utilitzat per a totes les simulacions dinàmiques d'equilibri que es presenten en aquest treball, amb variacions en el temps d'integració total i el nombre de dinàmiques per a cada cas. Cal apuntar que, sense l'aplicació d'aquest mètode, hauria estat del tot inviable obtenir els temps d'integració que s'ha obtingut, amb els recursos computacionals

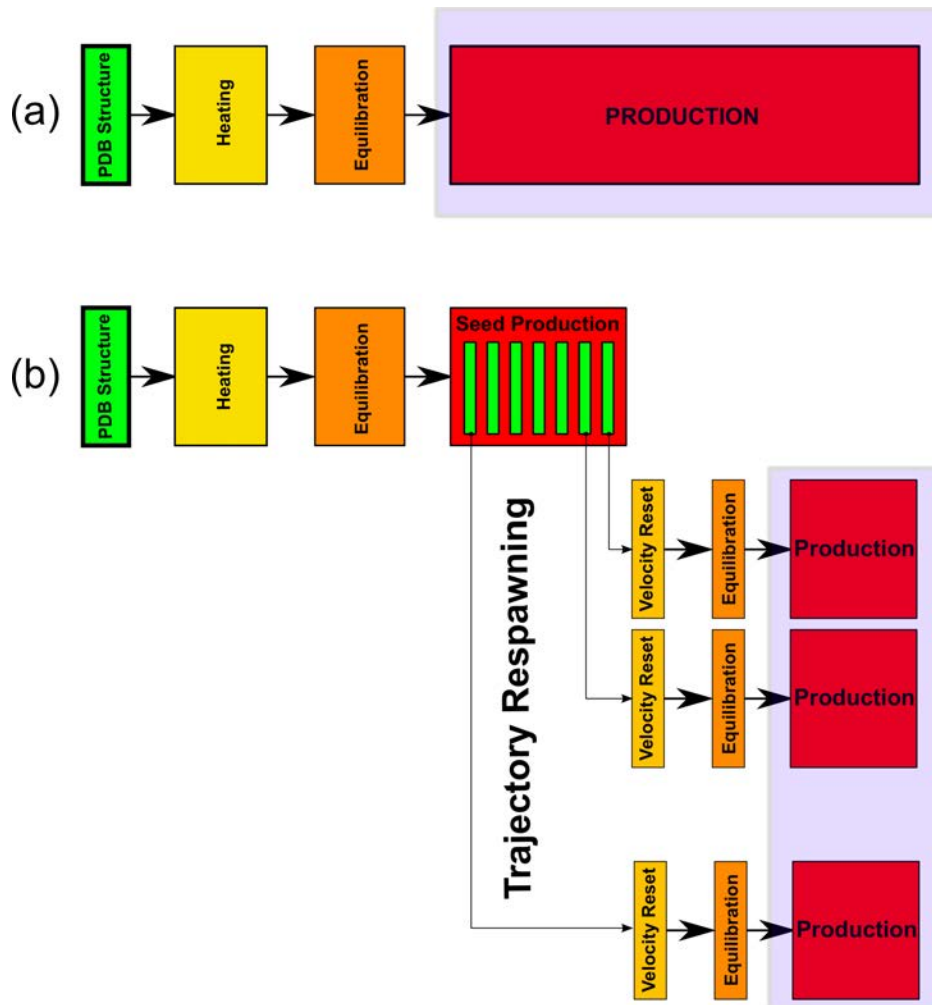


FIGURA 4.1: Representació esquemàtica de (a) un protocol estàndard per a la realització d'una simulació MD i de (b) el protocol *task-farming*

dels quals disposem (o disposàvem en el moment de la realització dels càlculs). A tall d'exemple, gràcies a l'aplicació del mètode *task-farming*, hem pogut simular fins a 100 ns de simulació MM (2016) o fins a 1,0 ns de simulació QM/MM (2014), que fou una de les simulacions QM/MM més voluminoses mai realitzada, en l'any que fou presentada.

Cal esmentar que de manera contemporània al mètode aquí explicat, es va publicar un treball de Kohlhoff *et al.* a la revista *Nature Chemistry*, on s'utilitzava la mateixa idea (naturalment implementada de manera diferent) per integrar dinàmiques MM en receptors lligats a proteïnes (GPCRs).<sup>159</sup>

## 5 Propietats Espectroscòpies de les Proteïnes Fluorescents

Tal i com s'ha descrit anteriorment en aquest treball, la modulació de la longitud d'ona d'absorció de les proteïnes fluorescents és un dels principals focus de la recerca científica bàsica que es fa al voltant d'aquests complexos foto-biosistemes. És evident que, per tal d'intentar evadir les tècniques mutagèniques basades en aproximacions aleatòries, ens cal crear un coneixement ferm de la naturalesa de les excitacions que ocorren en les FPs. En aquest sentit, es fa palesa la necessitat d'una descripció atòmica d'aquests sistemes, que sigui capaç de racionalitzar a un nivell fi les singularitats de les proteïnes fluorescents i dels processos foto-biofísics que experimenten.

### 5.1 Proteïnes Verdes Fluorescents

En aquest apartat es presenta la recerca realitzada respecte el comportament dels espectres d'absorció de les proteïnes verdes fluorescents. Per a poder construir un coneixement ferm s'ha estudiat un grup de 3 GFPs: la *wt*-GFP i les dobles mutants S65T/H148D-GFP i E222Q/H148D-GFP. La rellevància de les dues primeres ha estat explicada a la Introducció: la *wt*-GFP per ser la primera descoberta i més extensament estudiada i la S65T/H148D-GFP per representar un sistema model de les proteïnes fluorescents, donades les seves característiques dinàmiques tant concretes. La mutant E222Q/H148D és una proteïna que té un comportament molt similar a la S65T/H148D, encara que no existeixen dades espectroscòpiques resoltes en el temps per a la mateixa, ni tampoc estructura cristal·lina. Amb això, se li suposa, igual que a la S65T/H148D, una ESPT entre el cromòfor i l'Asp148. Vegeu la Figura 5.1, panell dret. Encara que la mutant E222Q/H148D-GFP no ha pres tanta rellevància científica com la seva germana, ens servirà per confirmar i afinar el coneixement que poguem extreure de la S65T/H148D.

#### 5.1.1 Metodologia

L'objectiu en aquest capítol és reproduir observables macroscòpiques per a poder obtenir descripcions atòmiques dels processos que representen



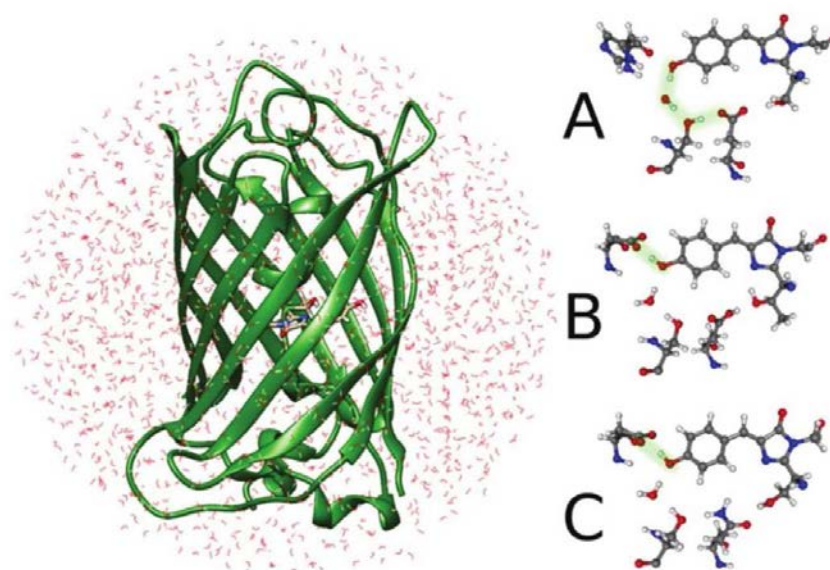


FIGURA 5.1: Esquerra: Representació gràfica del model utilitzat, que inclou tots els àtoms de la proteïna i la gota d'aigües de 60 Å de diàmetre. Dreta: Etiquetats per A, B i C, el cromòfor i els residus propers al mateix per a les proteïnes *wt*-GFP, S65T/H148D i E222Q/H148D, respectivament. El camí del cable de protons està ressaltat en verd.

així que, mancant-nos les funcions de partició de les proteïnes solvatades en aigua, hem de realitzar simulacions de dinàmica molecular per a generar microestats a partir dels quals promitjar les propietats moleculars o les característiques estructurals d'interès.

### Sistemes Model

Per a la *wt*-GFP i la doble mutant S65T/H148D, s'han extret les coordenades de l'estructura cristal·lina respectiva del Protein Data Bank (PDB ID: 1EMB<sup>47</sup> i 2DUF,<sup>75</sup> respectivament). Els àtoms d'hidrogen (que no estan resolts per difracció de rajos X) s'han afegit utilitzant el servidor PROPKA a pH 5,7.<sup>160-163</sup> Per tal de reproduir correctament l'efecte de la solvatació, s'ha situat cada proteïna al centre d'una esfera d'aigües explícites (tractades amb el model TIP3<sup>164</sup>) de 60 Å de diàmetre (Figura 5.1). Per evitar mals contactes entre molècules d'aigua i àtoms de la proteïna que hagin pogut aparèixer en el procés de solvatació, s'ha aplicat uns pocs passos de minimització. Tot el procés de creació del sistema model s'ha realitzat amb el programa CHARMM<sup>125</sup> amb el camp de forces CHARMM27<sup>123</sup> i els paràmetres desenvolupats per Thiel i col·laboradors per als àtoms del cromòfor.<sup>165</sup>

Per a la mutant E222Q/H148D no podem prendre l'estructura del PDB perquè no existeix cap estructura resolta disponible. És evident, doncs, que hem d'introduir les mutacions manualment però, a partir de quina

estructura? Com que les propietats espectroscòpiques conegudes de la mutant E222Q/H148D són molt similars a les de la S65T/H148D,<sup>87</sup> sembla adient prendre l'estructura d'aquesta última per a aplicar-li les mutacions T65S i E222Q. Amb això, s'assumeix que el plegament de la proteïna no es veurà afectat per aquest canvi però, tenint en compte les dades experimentals disponibles de la *wt*-GFP, la S65T, la S65T/H148D i la E222Q/H148D, no sembla una suposició molt agosarada. Després d'aplicar la mutació, s'ha realitzat el mateix procés de solvatació descrit anteriorment. Els sistemes complets inclouen al voltant de 18500 àtoms, depenent del cas.

### Simulacions de Dinàmica Molecular

Per tal d'explorar les diferències en els espectres d'absorció de les tres proteïnes verdes fluorescents hem de calcular tres simulacions dinàmiques diferents, una per cada cas. Pel cas de la *wt*-GFP, i donat el tipus d'interaccions que predominen en l'estructura de la mateixa, s'accepta que una descripció MM del seu moviment és satisfactòria.<sup>50,166-168</sup> El problema sorgeix amb les dues dobles muntants: a l'estructura cristal·lina de la mutant S65T/H148D trobem una interacció de pont d'hidrogen extraordinàriament curta (2,32 Å) entre el cromòfor i l'aspartat, i se suposa que aquesta interacció resulta en la presència d'un protó altament deslocalitzat entre donador i acceptor.<sup>75</sup> Aquest tipus d'interacció, tant diferent del que és una típica interacció covalent, no pot ser descrita per un camp de forces MM. De fet, s'ha realitzat proves preliminars amb la mutant S65T/H148D on s'ha intentat descriure la seva estructura per mitjà de dinàmiques MM. En tots els casos, en qüestió de 500 ps, es trenca la interacció entre el cromòfor i l'aspartat, quan aquest últim es torça per acabar apuntant cap a l'exterior de la proteïna en una situació molt menys constreta que la inicial. Sabem que aquesta interacció cromòfor-aspartat és estable en el temps, així que per tal de descriure-la, ens farà falta una aproximació QM/MM al problema, que només serà utilitzada per a les dues dobles mutants. El nivell QM/MM utilitzat ha estat el semiempíric OM3<sup>114</sup> amb el camp de forces CHARMM27.

Aquest ús combinat de dinàmiques MM i QM/MM pot semblar controvertit però no hauria de comportar problemes si assumim que cadascun dels mètodes per separat és prou bo per ser aplicat al cas que li pertoca i que, per tant, és capaç d'explorar adequadament l'espai de configuracions en cada cas, cosa que naturalment faríem sense problema, si no comparéssim els resultats d'una i altra metodologia. Per confirmar que una descripció MM és suficient pel cas de la *wt*-GFP s'ha seleccionat a l'atzar 5 estructures de la dinàmica MM d'aquesta proteïna i s'han optimitzat a nivell MM (CHARMM27) i a nivell QM/MM (OM3/CHARMM72). Els resultats indiquen que l'aproximació és assumible ja que l'RMS mitjà per les parelles d'estructures optimitzades és només de 0,88 Å i la diferència mitjana entre les energies d'excitació TD-DFT és de 0,11 eV, que es troba dins l'error del mètode.

Els càlculs purament MM s'han realitzat amb el programari CHARMM i els QM/MM amb el codi ChemShell.<sup>143</sup> La part QM per a les dues dobles mutants comprèn el cromòfor, l'Asp148, i la part més propera dels residus limítrofs als mateixos, essent un total de 58 àtoms QM per a la doble mutant S65T/H148D i 55 àtoms QM per a la mutant E222Q/H148D (la diferència de 3 àtoms es deu a la mutació S65T, que afecta al cromòfor). La resta del sistema s'ha considerat MM. La zona fronterera s'ha tractat amb l'aproximació *link-atom* amb el tractament *charge-shift*, i la interacció QM/MM s'ha tractat a nivell EE. S'ha definit com a part activa tota la proteïna i les molècules d'aigua més properes, deixant una capa esfèrica de molècules d'aigua congelada a la part més externa de la gota. Aquesta part activa és d'un total de 7050 àtoms.

Pel que fa a la MM-MD de la *wt*-GFP, s'ha escalfat el sistema tot fent increments de 100 K cada 50 ps per un total de 150 ps, tot mantenint l'estructura del cable de protons fixada per mitjà de restriccions harmòniques (ja que volem reproduir la forma A). A continuació s'ha realitzat l'equilibratge del sistema en què s'ha eliminat gradualment la restricció pel cable de protons durant 200 ps, per finalment propagar 150 ps de dinàmica sense cap tipus de restricció. A partir d'aquí, s'ha calculat 1 ns de MM-MD.

Pel què fa a les simulacions QM/MM, s'ha utilitzat el cap de forces CHARMM27 implementat en el mòdul DL-POLY de l'aplicació ChemShell per a la part MM i el mètode OM3 per a la part QM, implementat en el programa MNDO99.<sup>169</sup> Per tal d'acumular un temps suficientment llarg de simulació, s'ha utilitzat el mètode de paral·lelització de dinàmiques descrit en el Capítol 4 per a obtenir 1 ns de simulació QM/MM per a cadascuna de les dobles mutants. Pel cas de la mutant E222Q/H148D, s'ha aplicat una fase d'equilibratge de 40 ps prèvia a l'escalfament per assegurar l'acomodació de les mutacions realitzades.

### **Càlculs Estàtics**

Els càlculs estàtics realitzats sobre les estructures obtingudes de les simulacions MD, també s'han realitzat a un nivell QM/MM. La part QM presa ha estat la mateixa que en les simulacions QM/MM i, en el cas de la *wt*-GFP, ha inclòs només el cromòfor (un total de 40 àtoms). En aquest cas, encara utilitzant el *link-atom* amb el tractament *charge-shift* i la interacció EE, s'ha emprat el programa Gaussian09<sup>170</sup> pel càlcul de la part QM. Les energies de l'estat fonamental i dels estats excitats s'han extret per mitjà de càlculs DFT i TD-DFT respectivament, utilitzant el funcional CAM-B3LYP i la base 6-31G(d,p).

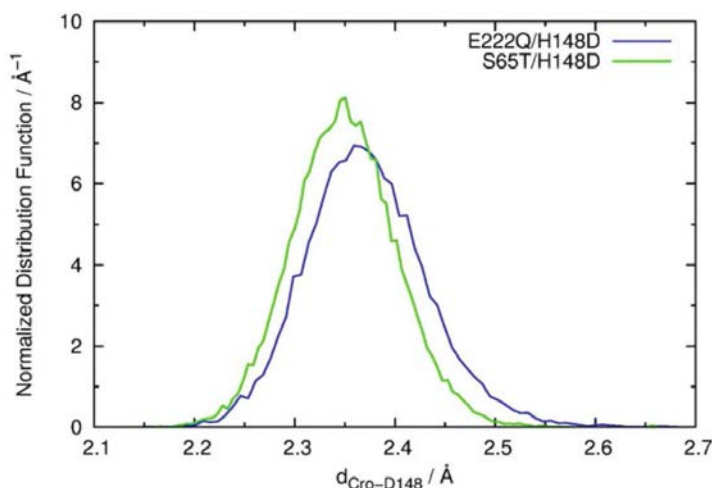


FIGURA 5.2: Distribució de la distància entre els oxígens del cromòfor i l'Asp148 al llarg del nanosegon de simulació MD per les mutants E222Q/H148D (blau) i S65T/H148D (verd).

## 5.1.2 Resultats

### Resultats Estructurals

Les simulacions de MD, a més de ser la base sobre la qual calcular els espectres d'absorció de les proteïnes estudiades, ens proporcionen informació estructural per elles mateixes. En aquest sentit, hem analitzat algunes propietats estructurals de la mutant E222Q/H148D (per la qual no s'ha resolt l'estructura cristal·lina) i les hem comparat amb les de la mutant S65T/H148D per tal de predir algunes de les característiques de la primera.

Segurament, el paràmetre estructural més important en aquestes dues mutants és la distància entre l'Asp148 i el cromòfor. Aquesta distància, que és extremadament curta en la mutant S65T/H148D, és crítica en la fotoquímica de les dues mutants doncs permet la transferència ultraràpida i sense barrera del protó, si es manté prou constreta. Com es pot observar en la Figura 5.2, la distància O(Asp148)-O(Cro) és estable i presenta una distribució gaussiana en ambdós casos.

La distància mitjana entre el cromòfor i l'Asp148 és de 2,35 Å i de 2,37 Å en les mutants S65T/H148D i E222Q/H148D, respectivament. Com que aquest valor concorda amb l'experimental pel cas de la primera mutant, podem predir que també en el cas de la mutant E222Q/H148D podem trobar aquesta interacció de pont d'hidrogen curt. Així doncs, aquests resultats recolzen els indicis experimentals, que apuntaven a aquesta interacció.<sup>87</sup> Un altre paràmetre estructural diferencial trobat en l'estructura de rajos X de la doble mutant S65T/H148D és l'angle díedre entre els dos anells del cromòfor, que està força allunyat de la planaritat (9,5°), mentre que en la *wt*-GFP és de 0,0°. Com es mostra a la Figura 5.3, l'angle díedre

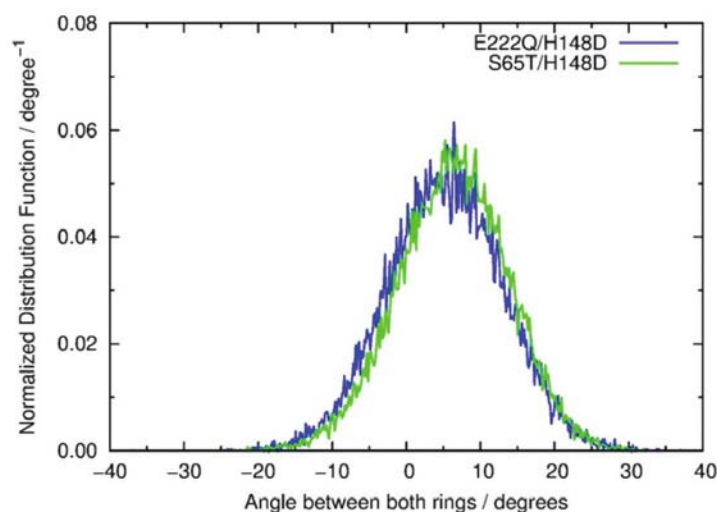


FIGURA 5.3: Distribució de la l'angle dièdre entre els dos anells del cromòfor al llarg del nanosegon de simulació MD per les mutants E222Q/H148D (blau) i S65T/H148D (verd).

entre els dos anells del cromòfor en les simulacions que s'han dut a terme compleixen una distribució gaussiana amb màxims a  $6,3^\circ$  i  $5,2^\circ$  per a les mutants S65T/H148D i E222Q/H148D, respectivament. La concordança entre els resultats experimentals i teòrics per a la primera mutant, altre cop ens permeten fer una predicció del valor que tindrà aquest paràmetre en la segona, que estarà al voltant de  $5^\circ$ .

### Espectres d'Absorció

El càlcul d'espectres d'absorció de proteïnes a partir de simulacions MD és una metodologia d'ús relativament comú i que dóna resultats fiables.<sup>81,171-174</sup> Pel aquest cas, s'ha pres 500 estructures temporalment equiespaiades (una estructura cada 2 ps) de les simulacions MD i s'ha realitzat els càlculs d'estats excitats TD-DFT a nivell QM/MM sobre les mateixes. Llavors s'ha sumat el component de cada transició electrònica trobada ponderada per la seva força d'oscil·lador, que és una mesura de com d'intensa serà la transició en qüestió.

Com es pot observar en la Figura 5.4, els espectres calculats mostren 2 pics principals (3 per a la *wt*-GFP). El més intens correspon a l'excitació localitzada en el cromòfor i representa el que es denomina com a banda A. Els màxims d'aquests pics es troben a 333 nm per a la *wt*-GFP, 365 nm per a la mutant E222Q/H148D i 370 nm per a la mutant S65T/H148D. Les bandes menys intenses situades a  $\sim 225$  nm, es deuen a excitacions localitzades en els enllaços peptídics. En els espectres experimentals, aquests darrers són els més intensos ja que de manera natural tots els enllaços peptídics de la proteïna participen en l'excitació mentre que, en el nostre model QM, només una mínima part dels enllaços peptídics han estat inclosos.

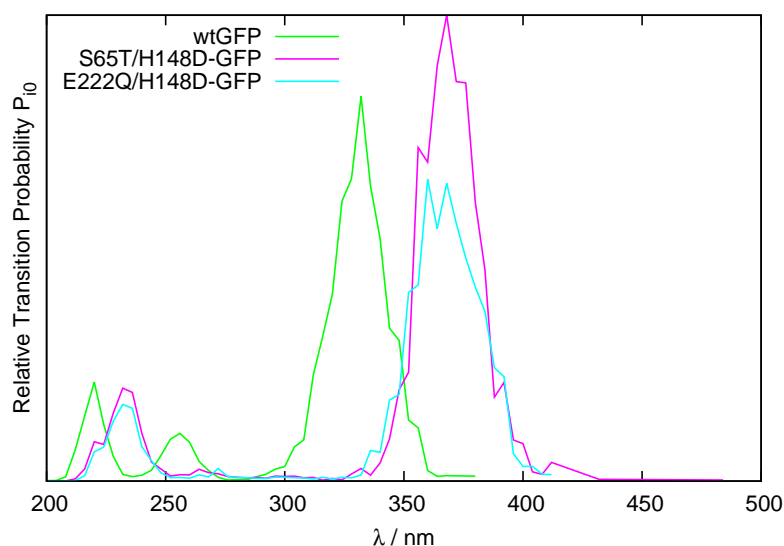


FIGURA 5.4: Espectres d'absorció calculats per a la *wt*-GFP (verd), E222Q/H148D (blau) i S65T/H148D (lila).

En termes relatius, els resultats que aquí es presenten mostren una bona concordança amb els experimentals.<sup>75</sup> Com es pot observar a la Figura 5.4, les dues dobles mutants tenen un espectre d'absorció molt similar, tal i com van publicar Stoner-Ma *et al.*<sup>87</sup> i, el què és més important, els màxims d'absorció es troben desplaçats cap al vermell 32,5 nm i 36,9 nm respecte el màxim d'absorció de la *wt*-GFP. Tal i com es pot observar qualitativament a la Figura 5.5 i quantificar per mitjà del càlcul del coeficient de correlació de Pearson, hi ha una correlació negativa entre la distància O(Asp148)-O(Cro) i la longitud d'ona d'absorció. El coeficient de correlació de Pearson és de -0,32 per la mutant E222Q/H148D i -0,37 per la mutant S65T/H148D indicant que, com més curta és la distància O(Asp148)-O(Cro), més desplaçada cap al vermell apareix la banda d'absorció. Experimentalment, la diferència entre les bandes d'absorció de les dues mutants és de 5 nm i el valor extret del càlcul que aquí es presenta és de 4,4 nm. El fet que el desplaçament cap al vermell de la mutant E222Q/H148D sigui lleugerament inferior que el de la mutant S65T/H148D està en línia amb el coeficient de Pearson calculat ja que, recordem, la distància mitjana O(Asp148)-O(Cro) per la primera mutant és lleugerament superior a la de la mutant S65T/H148D. A més, cal notar que el desplaçament cap al vermell obtingut experimentalment és lleugerament inferior al calculat en aquest treball, però això es pot justificar per la tria del mètode QM utilitzat en la propagació de la dinàmica. És ben sabut que l'OM3, tot i descriure correctament les interaccions presentades, tendeix a sobreestimar la força de les interaccions de pont d'hidrogen,<sup>115</sup> així que probablement les estructures obtingudes de les simulacions MD tenen una distància O(Asp148)-O(Cro) lleugerament inferior a la real. Invocant altre cop el valor negatiu del coeficient de correlació de Pearson, podem predir que, si el mètode QM no sobreestimés la força de la interacció entre el cromòfor i l'aspartat, encara obtindríem una millor descripció quantitativa del desplaçament cap al roig de les dues dobles

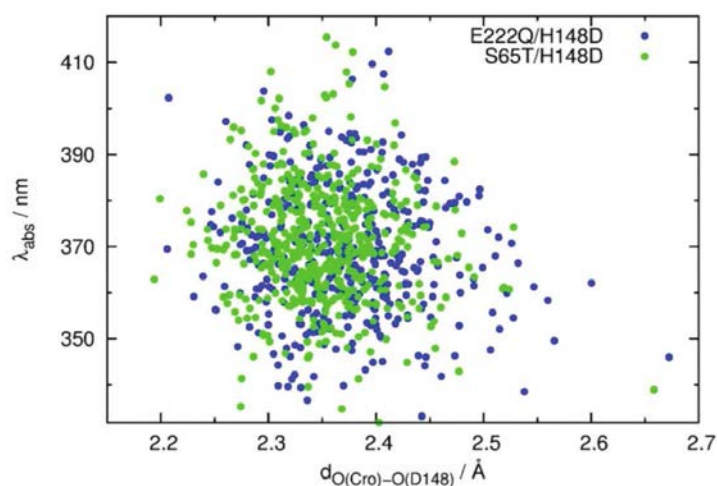


FIGURA 5.5: Representació de la longitud d'ona d'absorció de l'estat fotoactiu en funció de la distància entre cromòfor i Asp148. Es pot observar una correlació negativa.

mutants.

Amb tot, no és molt agosarat afirmar que els resultats són prou bons, tenint en compte la complexitat dels sistemes estudiats i dels mètodes i càlculs duts a terme. Tot i la bona descripció en termes relatius, hi ha una desviació sistemàtica de les energies d'excitació que ronda els 0,35 eV. Aquesta sobreestimació és un error sistemàtic conegut pel TD-DFT/CAM-B3LYP per aquest tipus de sistema, com ja s'ha publicat anteriorment.<sup>175-177</sup>

Per tal de racionalitzar perquè una càrrega negativa prop de l'anell fenòlic del cromòfor causa un desplaçament cap al roig, hem d'analitzar els orbitals que participen en l'excitació TD-DFT. L'excitació corresponent a la banda A és l'esperada HOMO→LUMO  $\pi\pi^*$ , similar a la descrita anteriorment per al cromòfor de les GFPs,<sup>178,179</sup> i que, com es pot observar qualitativament a la Figura 5.6 i quantitativa a la Figura 5.7, implica un cert caràcter de transferència de càrrega de l'anell fenòlic al carboni metínic. Amb aquestes dades davant, és fàcil veure perquè una càrrega negativa prop de l'anell fenòlic redueix l'energia de l'excitació, tot incrementant la longitud d'ona d'absorció: l'excitació implica una pèrdua de densitat electrònica de l'anell fenòlic a favor del pont metínic, així que és raonable pensar que l'acostament d'una càrrega negativa al primer afavorirà aquesta pèrdua de densitat electrònica. Dit d'altra manera, l'acostament d'una càrrega negativa a una zona on la densitat electrònica és més alta a l'estat fonamental que a l'excitat, desestabilitza el primer respecte el segon, fet que redueix l'energia d'excitació tot produint el desplaçament cap al vermell.

El que s'ha exposat fins ara pel què fa la descripció de l'excitació està en conflicte amb el raonament fet per Remington i els seus col·laboradors que explicava l'efecte del residu Asp148 sobre el corriment cap al roig de la banda A d'absorció de la doble mutant S65T/H148D. Per mitjà d'estudis



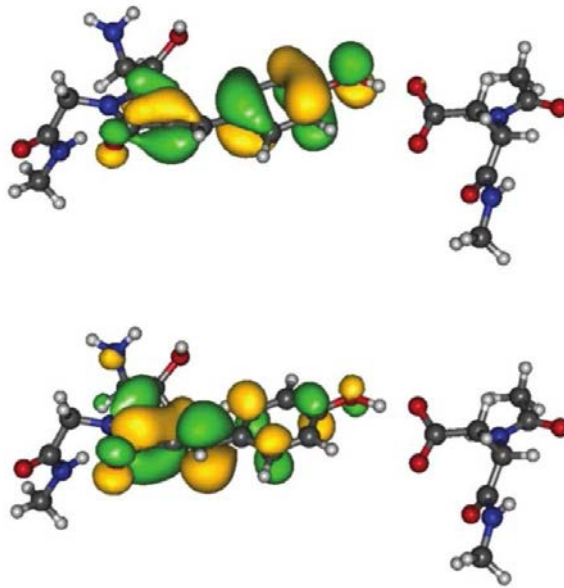


FIGURA 5.6: Orbitals HOMO (superior) i LUMO (inferior) involucrats en la excitació  $\pi\pi^*$  de la mutant S65T/H148D, descrita a nivell TD-DFT. Els orbitals per a la mutant E222Q/H148D són essencialment iguals i no es mostren per evitar redundància.

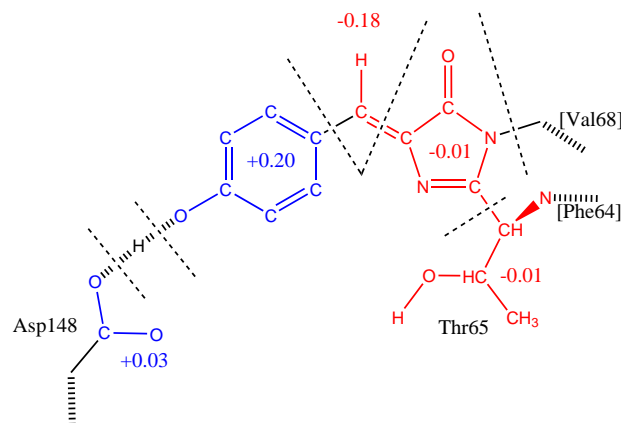


FIGURA 5.7: Variació de la càrrega dels diferents fragments del cromòfor durant el procés de fotoexcitació calculada a partir de les càrregues NBO de l'estat fonamental i excitat fotoactiu. En blau es mostren els fragments que perden càrrega electrònica i en vermell els fragments que en guanyen.



de mutagènesi, van concloure que "*a presumed negative charge on Asp148 cannot be responsible for the red shift in the S65T/H148D mutation*".<sup>75</sup> Van atribuir aquest desplaçament cap al vermell a la possible deslocalització electrònica del sistema  $\pi$  del cromòfor sobre el grup carboxilat de l'Asp148. Van basar la seva argumentació en el fet que la mutant S65T/H148E no presenta desplaçament cap al vermell: si la presència d'un glutamat (carregat negativament) no causa el corriment cap al roig, la càrrega negativa de l'Asp148 no pot ser la causant de tal desplaçament en la doble mutant S65T/H148D. Aquí s'ha demostrat qualitativament i quantitativa que, efectivament, és la proximitat de la càrrega negativa la que causa aquest desplaçament. A més, també s'ha demostrat que l'explicació proposada per Remington respecte la deslocalització addicional del sistema  $\pi$  és incorrecta ja que els orbitals involucrats en la transició estan clarament localitzats en el cromòfor. D'altra banda, aquí es proposa la següent explicació per a l'absència de desplaçament cap al vermell de la mutant S65T/H148E, en termes estructurals: com que no existeix una estructura cristal·lina resolta per aquesta última mutant, no es pot assegurar la proximitat de la càrrega negativa del glutamat a l'anell fenòlic del cromòfor. És més, es proposa que aquest glutamat estarà apuntant, probablement, fora de la cavitat de la proteïna, però en cap cas, enfrontant-se al cromòfor. Això és absolutament imaginable si tenim en compte que, ja en la mutant S65T/H148D, hi ha molta pressió entre l'Asp148 i el cromòfor, així que probablement no hi hagi espai per acomodar el grup  $\text{CH}_2$  extra que suposa la mutació H148E.

## 5.2 Proteïnes Vermelles Fluorescents

En la línia de maximitzar l'aplicabilitat de les proteïnes fluorescents en biomedicina, cal explorar les possibilitats mutagèniques de les RFPs, properes a la "finestra de transparència" del teixit humà. Les proteïnes mNeptune i mCardinal s'han situat en els últims anys com a candidates molt prometedores per a aquest fi. En aquest apartat es presenta l'estudi realitzat sobre les propietats espectroscòpiques d'equilibri d'aquestes dues mutants, que ens permetrà escatir els detalls del funcionament de les mateixes, en concret, posar nom i cognoms a les interaccions clau que determinen el corriment cap al roig de les RFPs estudiades.

### 5.2.1 Metodologia

De la mateixa manera que ocorre amb les proteïnes verdes, per tal d'analitzar les causes estructurals que indueixen a canvis en les propietats espectroscòpiques de les RFPs, necessitem realitzar simulacions MD per poder promitjar les propietats moleculars de les estructures obtingudes.

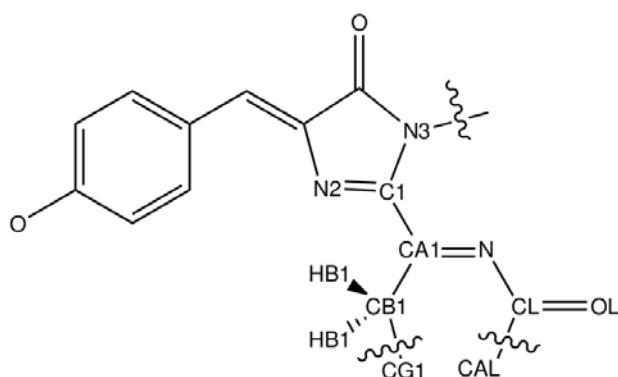


FIGURA 5.8: Etiquetatge dels àtoms en els cromòfors de la mNeptune i mCardinal. Els àtoms sense etiqueta (amb els seus paràmetres associats) es poden consultar a la bibliografia.<sup>165</sup>

### Sistemes Model

Les estructures inicials s'han pres del PDB: codis 3IP2<sup>77</sup> (mNeptune) i 4OQW<sup>79</sup> (mCardinal). L'estructura 4OQW conté un total de 8 cadenes de les quals s'ha pres la primera. A més, aquesta estructura correspon a una mutació de la mCardinal on s'ha metilat una lisina per a facilitar la cristal·lització. Com que les dades espectroscòpiques van ser extretes de la mCardinal sense tal mutació, s'ha eliminat aquesta de l'estructura 4OQW. S'ha afegit els àtoms d'hidrogen altre cop amb el servidor PROPKA a pH 7,0. A continuació s'ha solvatat cada proteïna amb una caixa cúbica d'arestes de 71 Å (mNeptune) i 73 Å (mCardinal) de molècules explícites d'aigua (TIP3). El recompte total d'àtoms és de 38108 per la mNeptune i de 41408 per la mCardinal. S'ha fet correr 400 passos d'optimització sobre cada estructura per a eliminar possibles contactes que puguin aparèixer durant el procés de solvatació. El procés s'ha realitzat amb el programari CHARMM i el camp de forces CHARMM27. Per als paràmetres MM dels cromòfors s'ha utilitzat una versió dels ja publicats i utilitzats amb èxit pel nostre grup anteriorment<sup>81,172</sup> i que es basen en els paràmetres del Prof. W. Thiel pel cromòfor de la GFP.<sup>165</sup> La nova versió dels paràmetres utilitzada en aquesta ocasió incorpora millores en la descripció del doble enllaç addicional C=N. Aquest refinament dels paràmetres s'ha calculat ajustant càlculs QM sobre la superfície d'energia potencial a nivell MP2. A la Taula 5.1 es pot trobar el valor d'aquests nous paràmetres on l'etiquetatge dels àtoms segueix el de la Figura 5.8. La resta de paràmetres (que no han estat modificats) es poden trobar en la publicació original de W. Thiel i col·laboradors.<sup>165</sup>

Tipus d'àtom	Càrrega / au	$\epsilon$ / kcal mol <sup>-1</sup>	$R_{\min}/2$ / Å
N	-0,8651	-0,09649	1,9352
CA1	0,1339	-0,00043	2,0835

Enllaç	$k$ / kcal mol <sup>-1</sup> Å <sup>-2</sup>	$r_e$ / Å
CA1-C1	376,10	1,4215
CA1-CB1	250,54	1,5610
CA1-N	539,04	1,3538
N-CL	414,93	1,3907

Angle	$k_\theta$ / kcal mol <sup>-1</sup> rad <sup>-2</sup>	$\theta_e$ / graus
C1-CA1-CB1	126,02	114,822
C1-CA1-N	124,88	120,320
CA1-CB1-HB1	73,56	106,870
CA1-CB1-CG1	89,64	111,240
N2-C1-CA1	304,06	121,330
N3-C1-CA1	456,96	122,670
N-CL-CAL	126,67	117,690
N-CL-OL	124,29	122,011
CA1-N-CL	47,10	146,309
CB1-CA1-N	112,31	122,005

Díedre	$k_\chi$ / kcal mol <sup>-1</sup> rad <sup>-2</sup>	n	$\delta$ / graus
X-CA1-C1-X	4,50	2	180,000

Torsió Impròpia	$k_\phi$ / kcal mol <sup>-1</sup> rad <sup>-2</sup>	$\phi_0$ / graus
C1-CA1-N2-N3	60,37	-2,97
CA1-N-CB1-C1	103,33	-2,41

TAULA 5.1: Paràmetres refinats per al cromòfor de les RFPs mNeptune i mCardinal. La resta de paràmetres, prèviament publicats, es poden trobar a la bibliografia.<sup>165</sup>

L'etiquetatge dels àtoms segueix el de la Figura 5.8.

## Simulacions de Dinàmica Molecular

Les proteïnes vermelles presenten un cromòfor aniònic a l'estat fonamental així que, al contrari que en el cas de les mutants dobles de la GFP estudiades en la secció anterior, no s'espera haver de reproduir cap situació d'un protó deslocalitzat. Per aquesta raó, una aproximació MM al seu moviment serà suficient. Les simulacions s'han realitzat amb condicions periòdiques de contorn (PBC) i prenent un temps d'integració d'1 fs. A partir de les estructures minimitzades després de la solvatació, s'ha escalat el sistema fins a 300 K, seguint un escalat de 5 K per picosegon. Durant aquest procés, s'ha mantingut una restricció harmònica al moviment de l'esquelet de la proteïna, que s'ha alliberat posteriorment i de forma gradual, durant 80 ps. Finalment, s'ha equilibrat el sistema, ja lliure de tota restricció, durant 920 ps. Per a la fase de producció, altre cop s'ha utilitzat el mètode presentat al Capítol 4 que, en un context MM, ens ha permès acumular 100 ns de trajectòria per a cadascuna de les proteïnes.

## Càlculs Estàtics

Per al càlcul d'estats excitats a partir de les estructures obtingudes de les simulacions MM-MD, s'ha utilitzat un context QM/MM. La part QM ha estat el cromòfor juntament amb alguns dels àtoms dels residus limítrofs al mateix, essent la resta àtoms de la part MM. Per al càlcul de l'energia de la part QM s'ha utilitzat el mètode TD-DFT amb el funcional CAM-B3LYP i la base 6-31G(d,p) implementat en el programa Gaussian09. Per a la part MM s'ha utilitzat el mòdul DL-POLY del programari ChemShell amb el camp de forces CHARMM27. La interacció QM/MM s'ha tractat en el context EE, amb l'aproximació *link-atom* i la correcció *charge shift*.

Per contrastar la correcció del mètode TD-DFT en aquests sistemes s'han fet càlculs de referència en un fotograma seleccionat utilitzant el mètode SOS-CIS(D)<sup>180</sup> amb la base auxiliar RIMP2-cc-pVDZ, utilitzant el programari Qchem 4.3<sup>181</sup>

## 5.2.2 Resultats

### Espectres d'Absorció

Per al càlcul dels espectres d'absorció de les RFPs s'ha seguit una metodologia similar a la de la secció anterior. S'ha pres una estructura cada 100 ps de dinàmica (un total de 1000 estructures) i s'ha realitzat un càlcul d'estats excitats nivell QM/MM sobre les mateixes. Llavors s'ha sumat la contribució de cada excitació en base a la probabilitat de transició a aquell estat. La histogramització d'aquest sumatori ha portat a l'obtenció dels espectres d'absorció de la mNeptune i mCardinal.

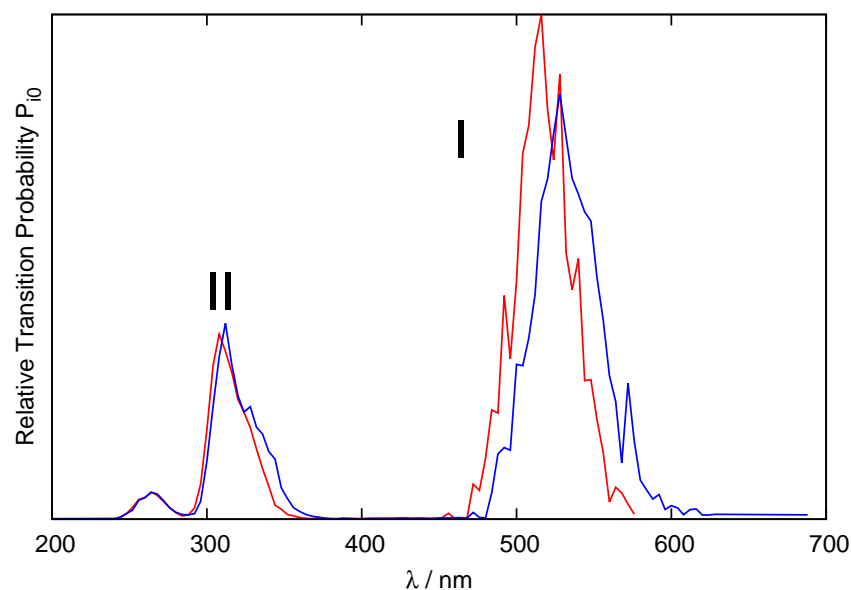


FIGURA 5.9: Espectre d'absorció calculat a nivell QM/MM a partir dels 1000 fotogrames extrets dels 100 ns de la simulació MM-MD per a les proteïnes mNeptune (vermell) i mCardinal (blau). Les etiquetes I i II identifiquen els diferents pics.

Els espectres d'absorció de la mNeptune i mCardinal es mostren a la Figura 5.9 i presenten dos pics. El pic més intens (Banda I) es troba a longituds d'ona més llargues, amb màxims a 516 nm per la mNeptune i 528 nm per la mCardinal. Els valors experimentals són 600 nm i 604 nm però en aquesta zona de l'espectre, aquesta diferència de  $\sim 75$  nm es tradueix en uns  $\sim 0,3$  eV, que és l'error típic trobat pel TD-DFT/CAM-B3LYP en cromòfors de proteïnes fluorescents a la bibliografia especialitzada i que ja s'ha acusat a la secció anterior. La banda menys intensa (Banda II) presenta el seu màxim gairebé a la mateixa longitud d'ona per a les dues proteïnes (308 nm per a la mNeptune i 312 nm per a la mCardinal) i és de l'ordre de 2,5 vegades menys intensa la primera. L'espectre d'absorció experimental<sup>79</sup> presenta dos pics de baixa intensitat a uns  $\sim 350$  nm i  $\sim 375$  nm, sense diferència entre les 2 proteïnes. Aquests dos pics podrien correspondre a la Banda II, encara que aquesta és més intensa i definida en termes relatius que la trobada experimentalment. Encara que la descripció de la Banda II no és totalment satisfactòria, l'objectiu d'aquest apartat és l'anàlisi de l'excitació associada a la Banda I, que efectivament queda ben descrita amb el mètode emprat. En aquest sentit s'ha de remarcar que la posició de la Banda I per les dues proteïnes estudiades, tenint en compte les limitacions del mètode, és prou bona com per poder confiar en la qualitat de les simulacions. Cal remarcar, particularment, la posició relativa de la Banda I per les dues proteïnes: s'ha reproduït una diferència de 12 nm mentre la diferència experimental és de 4 nm. Tenint en compte la complexitat del sistema estudiat, es tracta d'un molt bon resultat qualitatiu que ens permetrà fer anàlisis comparatius entre les dues proteïnes.

### Naturalitat dels Estats Excitats

Per entendre la naturalesa de les excitacions que donen lloc a les dues bandes dels espectres d'absorció les mNeptune i mCardinal, ens cal realitzar un anàlisi fi dels orbitals implicats. L'explicació que segueix és vàlida per a les dues mutants ja que contenen el mateix cromòfor, i les excitacions trobades per una i altra proteïna són molt similars. Un cop d'ull general a les excitacions que tenen lloc revela que existeixen 3 estats excitats amb forces d'oscil·lador considerables i que, per tant, contribueixen més o menys notablement a l'espectre d'absorció. La transició més intensa (amb força d'oscil·lador  $f \sim 0,8$ ) és essencialment una HOMO→LUMO, doncs el determinant que descriu aquesta excitació té un coeficient de  $\sim 0,7$ .

Les altres dues transicions són menys intenses (amb forces d'oscil·lador  $f \sim 0,3$  i  $f \sim 0,1$  respectivament). La primera, que també és la més energètica (308-312 nm), correspon a una excitació HOMO→LUMO-1 amb coeficient  $\sim 0,5$  i la segona ( $\sim 325$  nm) correspon a una excitació HOMO-3→LUMO amb un coeficient de  $\sim 0,6$ . La resta d'estats excitats calculats (7 dels 10 totals) tenen forces d'oscil·lador inferiors a 0,05 i, tot i que han estat sumats a l'espectre d'absorció per correcció formal, hi han contribuït molt pobrament així que seran considerats "estats foscos" i no seran discutits aquí.

És conegut que el TD-DFT pot convergir resultats erronis com a conseqüència de l'anomenat "error d'auto-interacció" (SIE).<sup>182</sup> Quan aquest fenomen ocorre, convergeixen arrels espúries normalment associades a estats de transferència de càrrega, que acostumen a tenir una força d'oscil·lador menyspreable. Encara que donada la baixa força d'oscil·lador d'aquests estats, el SIE no embrutaria el càlcul de l'espectre d'absorció, podria dur a malinterpretacions en l'anàlisi de les excitacions en termes d'orbitals moleculars. Per assegurar que no ens trobem amb aquesta casuística, s'ha seleccionat un fotograma representatiu de la simulació MD de la mNeptune i s'ha comparat els resultats TD-DFT amb els resultats d'un altre mètode on el SIE no pugui ocórrer.

Per a tal fi, s'ha triat el mètode *scaled-opposite-spin* CIS(D) (SOS-CIS(D)), que és un mètode referencial basat en funció d'ona però encara és assequible pel càlcul en molècules de mida mitjana.<sup>180</sup> Els resultats de tal comparació es mostren a la Taula 5.2.

Els estats trobats a nivell TD-DFT i SOS-CIS(D) tenen una bona correlació tant a nivell del caràcter de l'excitació com a nivell de la força d'oscil·lador. Les discrepàncies principalment es troben en les energies d'excitació, encara que es conserva una bona correlació en la majoria dels casos. Aquest resultat ens assegura que els estats trobats a nivell TD-DFT no pateixen del SIE i que per tant poden ser tinguts en compte en els raonaments que segueixen.

A fi i efecte de clarificar l'anàlisi dels orbitals que participen en els estats excitats, s'ha utilitzat els Orbitals Naturals de Transició o *Natural Transition*

TD-DFT/CAM-B3LYP			SOS-CIS(D)		
Estat	Energia d'Excitació / eV	<i>f</i>	Estat CIS	Energia d'Excitació / eV	<i>f</i>
1	2,3027	0,7847 *	1	1,6193	1,2855
2	2,7987	0,0002	6	2,6589	0,0007
3	2,9377	0,0028	2	2,5704	0,0229
4	3,5040	0,0063	mescla 7&8		
5	3,5886	0,0198	3	3,5862	0,0475
6	3,7101	0,2858 *	4	3,5491	0,2904
7	3,8862	0,1371 *	5	3,6554	0,1977
8	4,0994	0,0005	7	4,5759	0,0112
9	4,6953	0,0249	9	4,8445	0,0039
10	4,7183	0,0477	10	4,6085	0,1573

TAULA 5.2: Relació dels estats TD-DFT amb els SOS-CIS(D). Per a fer tal relació s'ha analitzat els pesos i participacions de cadascuna de les excitacions pels dos mètodes, que no es mostren aquí. Les excitacions marcades amb un asterisc són aquelles que més contribueixen a les bandes de l'espectre calculat.

*Orbitals* (NTOs). La transformació que implica la obtenció d'aquests orbitals simplifica la interpretació de les excitacions quan aquestes tenen força components si s'expressen en funció dels orbitals canònics ja que resulta en l'obtenció d'una sola (o unes poques) transició(ns) des d'un orbital ocupat a un altre de virtual (que seran combinació lineal dels orbitals moleculars canònics).<sup>183</sup> S'ha obtingut els NTOs pel mateix fotograma seleccionat de la mNeptune i aquests es mostren a la Figura 5.10.

Com es pot observar, els NTOs que descriuen l'excitació corresponent a la Banda I estan deslocalitzats per tot el cromòfor i mostren un desplaçament de càrrega des de l'anell fenòlic cap al fragment acilimínic del cromòfor (que conté el doble enllaç C=N, característic del cromòfor de les RFPs). Aquest comportament és similar a l'observat en el cromòfor de les GFPs, on la densitat electrònica es transferia de l'anell fenòlic al carboni metínic. Això suggereix que, a banda de rebaixar l'energia d'excitació situant un grup donador d'electrons prop de l'anell fenòlic (o una càrrega negativa, com en el cas de les GFPs), es podria aconseguir un desplaçament cap al vermell tot situant un grup atraient d'electrons a prop del fragment acilimínic.

Hi ha dues excitacions corresponents a la Banda II, també mostrades en base als NTOs calculats a la Figura 5.10. Les dues excitacions són de longitud d'ona similars i intensitats comparables. Per tant, sota el què s'ha etiquetat com a Banda II, probablement hi hagi en realitat dues bandes que se solapen (es pot intuir un colze a la Banda II), i que cadascuna d'elles correspongui a cadascuna de les excitacions mostrades. Òbviament, és temptador especular amb la possibilitat que aquestes dues excitacions

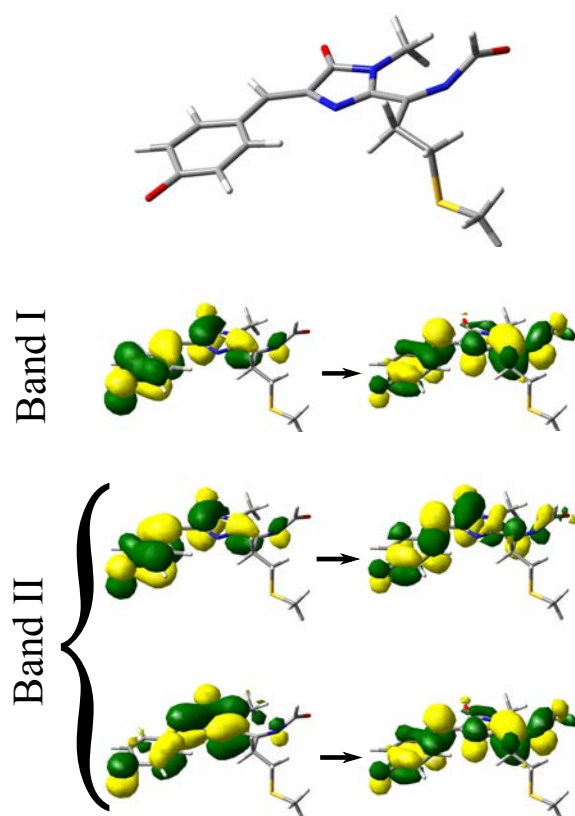


FIGURA 5.10: Representació de la part QM presa en els càlculs QM/MM i NTOs associats a les 3 transicions amb forces d'oscil·lador considerables, amb la seva correspondència a les bandes de l'espectre d'absorció de la mNeptune.



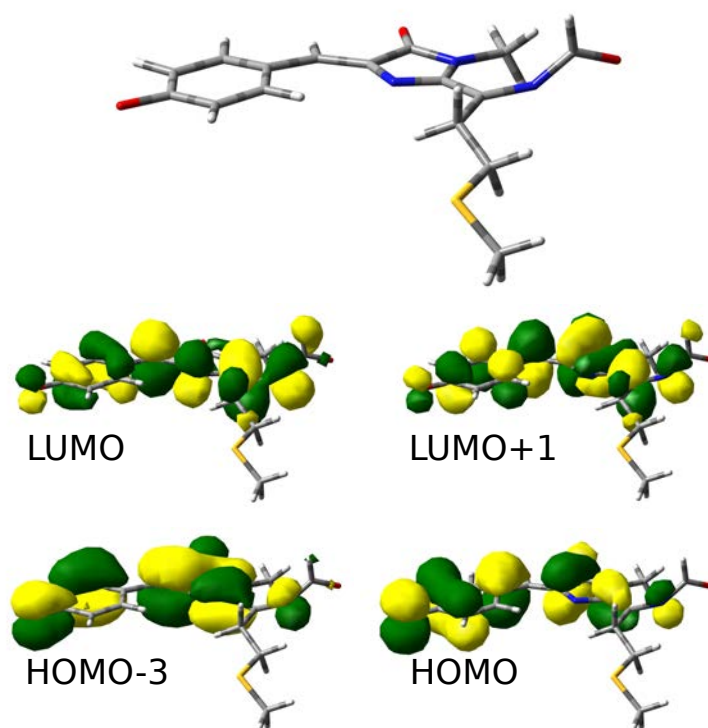


FIGURA 5.11: Representació de la part QM presa en els càlculs QM/MM i dels principals orbitals moleculars canònics involucrats en les excitacions de la Banda I i II de la mNeptune.

corresponguin a les dues bandes de baixa intensitat trobades experimentalment, ja mencionades anteriorment.

Sovint s'ha suggerit des d'estudis experimentals que la inclusió del grup acilimínic que comporta l'extensió de la conjugació del cromòfor de les RFPs en un enllaç C=N és la raó per la qual apareix el desplaçament cap al vermell de la banda d'absorció, doncs comporta la desestabilització de l'HOMO i l'estabilització del LUMO. De nou, una descripció teòrica d'alt nivell pot ajudar a la comprensió d'aquest fenomen: en la secció anterior s'ha descrit l'excitació fotoactiva en el cromòfor de les GFPs com una HOMO→LUMO, igual que en el cas que ens ocupa, de les RFPs. Ara bé, si comparem la forma dels orbitals frontera de les GFPs de la Figura 5.6 amb els de les RFPs, mostrats a la Figura 5.11, podem veure com l'orbital LUMO de les GFPs té una clara correspondència amb el LUMO+1 de les RFPs, mentre l'orbital HOMO es manté essencialment igual. Per tant, l'excitació HOMO→LUMO de les GFPs correspondria a l'excitació HOMO→LUMO+1 de les RFPs, és a dir, a la principal excitació present a la Banda II. En altres paraules, sembla que la inclusió del grup C=N en el cromòfor de les RFPs no desplaça l'energia d'excitació de la banda d'absorció de les GFPs substancialment, sinó que fa aparèixer una banda nova, corresponent a un altre estat excitat, i que és la banda d'interès en les

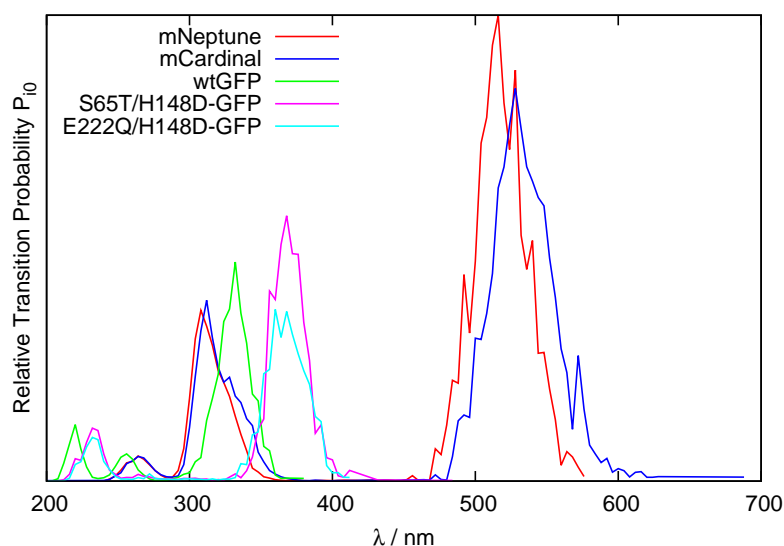


FIGURA 5.12: Espectres d'absorció de les 5 proteïnes estudiades en aquest capítol: la *wt*-GFP (verd), es dues dobles mutants S65T/H148D i E222Q/H148D (blau clar i lila respectivament) i les RFPs mNeptune i mCardinal (vermell i blau, respectivament).

RFPs. A més, a la Figura 5.12, que mostra els espectres de les GFPs i de les RFPs (i que és una combinació de les Figures 5.4 i 5.9), es pot comprovar com la posició i intensitat de la Banda II de les RFPs calculades correspon aproximadament amb la banda d'absorció de la *wt*-GFP.

### Una Nova Distribució de Càrregues

L'excitació electrònica, que naturalment implica una nova redistribució dels electrons, comporta canvis en els enllaços químics i, el què és més interessant pel cas que ens ocupa, una nova distribució de càrregues. Aquesta nova distribució de càrregues pot tenir efectes importants en la manera en com interacciona el cromòfor amb els residus propers. A fi i efecte de categoritzar els canvis en la distribució de la densitat electrònica en el cromòfor de les RFPs després de l'excitació, s'ha realitzat un anàlisi de càrregues NBO<sup>184</sup> a partir d'un fotograma agafat a l'atzar però que correspongui a una estructura típica de les obtingudes mitjançant la dinàmica MM. Els resultats es mostren a la Taula 5.3.

En termes de la redistribució de càrregues en cada secció del cromòfor (l'anàlisi d'àtom a àtom s'ha obviat per ser poc clarificador), es pot observar com la excitació implica un drenatge de càrrega de l'anell fenòlic que és rebut no per l'anell imidazolinònic, sinó pel grup acilimínic (que conté el doble enllaç C=N), característic del cromòfor de les RFPs. Aquest anàlisi porta a pensar, aquest cop des d'un punt de vista quantitatiu, que la interacció del grup acilimínic amb un grup atraient d'electrons farà baixar

Grup/Àtom	$Q(S_0)^a$	$Q(S_1)^a$	$\Delta Q^a$
Fenol	-0,61	-0,47	+0,15
Pont metí	+0,20	+0,14	-0,06
Imidazolinona	-0,45	-0,38	+0,08
Acilimina	-0,20	-0,35	-0,14
Cua de metionina	+0,07	+0,04	-0,02
O (fenol)	-0,831	-0,771	+0,060
O (imidazolinona)	-0,731	-0,698	+0,033
O (acilimina)	-0,688	-0,703	-0,015

<sup>a</sup> Les càrregues es mostren en a.u.

TAULA 5.3: Càrregues NBO en el cromòfor de les RFPs per l'estat electrònic fonamental i excitat corresponent a la Banda I.

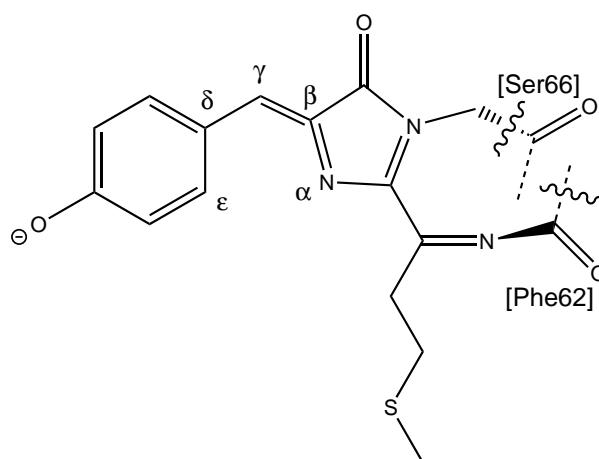


FIGURA 5.13: Representació gràfica del cromòfor de les RFPs on s'ha etiquetat els àtoms a partir dels quals es defineix l'angle *tilt*  $\tau$  (angle diedre format pels àtoms  $\alpha-\beta-\gamma-\delta$ ) i *twist*  $\phi$  (angle diedre format pels àtoms  $\beta-\gamma-\delta-\epsilon$ )

l'energia d'excitació. També confirma que l'acostament de grups electrò-donadors a prop de l'anell fenòlic tindrà el mateix efecte.

### Anàlisi de les Interaccions del Cromòfor

Un dels paràmetres estructurals que modula la posició de la longitud d'ona d'absorció de les proteïnes fluorescents és el grau de planaritat del cromòfor. La pèrdua de planaritat del cromòfor desestabilitza l'estat fonamental respecte l'excitat, tot disminuint l'energia d'excitació. En la bibliografia especialitzada es defineixen dos angles, *tilt* ( $\tau$ ) i *twist* ( $\phi$ ), per descriure quantitativament la planaritat del cromòfor. Per a una definició gràfica dels angles anomenats, referiu-vos a la Figura 5.13. L'anàlisi d'aquests dos angles denota que, en els dos casos estudiats, el cromòfor

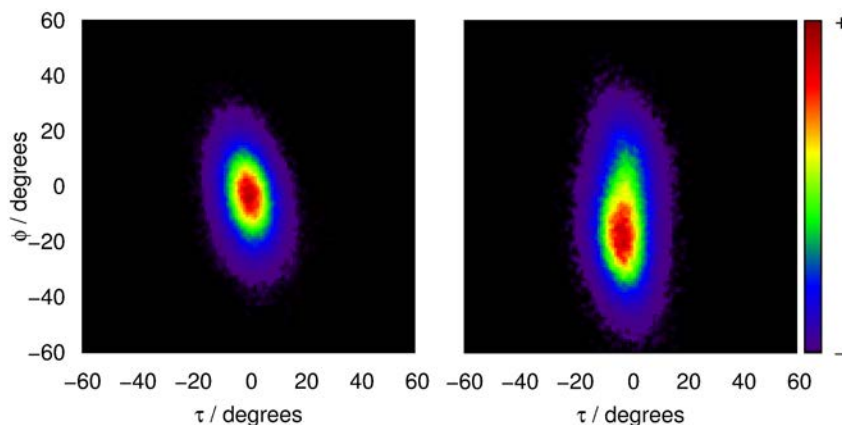


FIGURA 5.14: Anàlisi de la planaritat del cromòfor en les simulacions de MD en termes de la distribució dels angles  $\tau$  i  $\phi$  per a la mNeptune (panell esquerra) i mCardinal (panell dret).

manté la planaritat al amb valors voltant dels  $0^\circ$  (Figura 5.14). El valor mitjà per l'angle  $\tau$  és de  $0,57^\circ$  per la mNeptune (estructura cristal·lina:  $5,3^\circ$ ) i  $-1,9^\circ$  per la mCardinal (estructura cristal·lina:  $-7,8^\circ$ ). L'angle  $\phi$ , que presenta una major dispersió, té un valor mitjà de  $-2,8^\circ$  per la mNeptune (estructura cristal·lina:  $-9,4^\circ$ ) i de  $-11,0^\circ$  per la mCardinal (estructura cristal·lina:  $-11,0^\circ$ ). Així doncs, pel que fa a la planaritat del cromòfor, es pot afirmar que el cromòfor de la mNeptune és més pla que el de la mCardinal, i això pot tenir un efecte sobre la longitud d'ona d'absorció. La pèrdua de planaritat del cromòfor en la mCardinal fa que l'estat fonamental es desestabilitzi respecte l'excitat, tot fent augmentar la longitud d'ona d'absorció.

D'altra banda, com indiquen els estudis experimentals de mutagènesi, la mutació que comporta més diferències entre la mNeptune i la mCardinal és la Gly41Gln.<sup>79</sup> La mKate original té una metionina a la posició 41, que és substituïda per una glicina a la mNeptune. La gran diferència de volum entre els dos residus explica l'existència d'un forat a la mNeptune, que és ocupat per una molècula d'aigua en l'estructura cristal·lina resolta per rajos X. Se suposa que aquesta aigua interacciona simultàniament amb l'oxigen pertanyent a l'enllaç peptídic de la Phe62 (el justament anterior al cromòfor) i la Ser28, que es troba en una de les fulles- $\beta$  de la proteïna. Aquesta molècula d'aigua és considerada la responsable del corriment de 12 nm cap al vermell de la mNeptune respecte la mKate doncs la interacció de pont d'hidrogen que es dona a la mNeptune entre l'aigua i la Phe62, impossible en la mKate, estabilitza l'excés de càrrega negativa que rep el grup acilimínic a l'estat excitat fotoactiu. En la mCardinal la Gly41 es substitueix per una glutamina que interacciona alhora amb la Phe62 i la Thr28 (vegeu la Figura 5.15, panell dret).

Sembla ser que l'oxigen de la Phe62 captura les possibles interaccions dels grups atractants d'electrons que poden ajudar a estabilitzar l'estat excitat

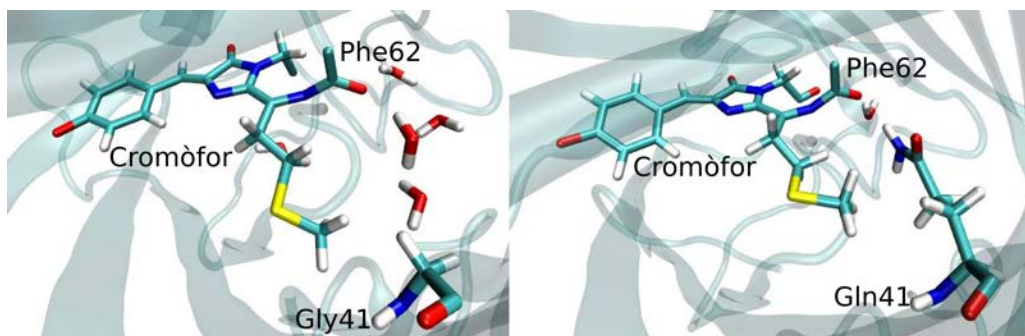


FIGURA 5.15: Panell esquerre: dibuix d'un fotograma representatiu de la simulació MD de la mNeptune. S'hi mostra el cromòfor, l'oxigen de la Phe62, la Gly41 i les aigües que ocupen el forat que deixa aquesta última. Panell dret: dibuix d'un fotograma representatiu de la simulació MD de la mCardinal. S'hi mostra el cromòfor, la Gln41, l'oxigen de la Phe62 i l'aigua més propera al mateix.

fotoactiu que presenta un excés de càrrega al grup acilimínic, situat just al costat del propi oxigen. S'ha analitzat la distribució de la distància entre l'oxigen de la Phe62 i la Gly41 per la mNeptune i entre l'oxigen de la Phe62 i la Gln41 per la mCardinal (Figura 5.16). Com es pot observar, en el cas de la mCardinal apareix una distribució bimodal de distàncies amb un pic que integra aproximadament el 50% de l'àrea total sobre els  $\sim 3 \text{ \AA}$  i l'altre sobre els  $\sim 4,5 \text{ \AA}$ , fent pensar que en ocasions hi ha quelcom entre la Gln41 i l'oxigen de la Phe62. Aquest quelcom és aigua. La distribució de distàncies de la Figura 5.16 per la mNeptune no mostra una interacció real, però denota que hi ha un espai buit considerable entre la Gly41 i l'oxigen de la Phe62. Com es pot veure a la Figura 5.15, aquest espai està ocupat per molècules d'aigua. En definitiva, per la mNeptune, l'oxigen de la Phe62 interacciona amb molècules d'aigua i, per la mCardinal, aquest mateix oxigen interacciona o bé amb la Gly41 (aproximadament la meitat del temps) o bé amb molècules d'aigua.

Ara bé, com d'important és la interacció amb l'aigua pels dos casos que ens ocupen? Per poder respondre aquesta pregunta s'ha analitzat les esferes de solvatació de l'oxigen de la Phe62 tot calculant la funció de distribució radial d'aquest àtom per les molècules d'aigua més properes (Figura 5.17). Mentre que la distribució de la primera i segona aigua més propera és molt similar en les dues proteïnes, la diferència principal que s'observa és en la distribució de la tercera aigua. En el cas de la mCardinal queda clar com la tercera aigua gairebé no interacciona amb l'oxigen de la Phe62, però en el cas de la mNeptune, es veu una certa interacció efectiva. La integració del les corbes fins a la primera esfera de solvatació, dóna un total de 2 aigües per a la mCardinal i 2,2 aigües per a la mNeptune. Això vol dir que la mCardinal té, de mitjana, dues aigües interaccionant amb l'oxigen de la Phe62 de manera estable mentre la mNeptune en té dues interaccionant de manera estable i una tercera que interacciona amb aquest àtom un 20% del

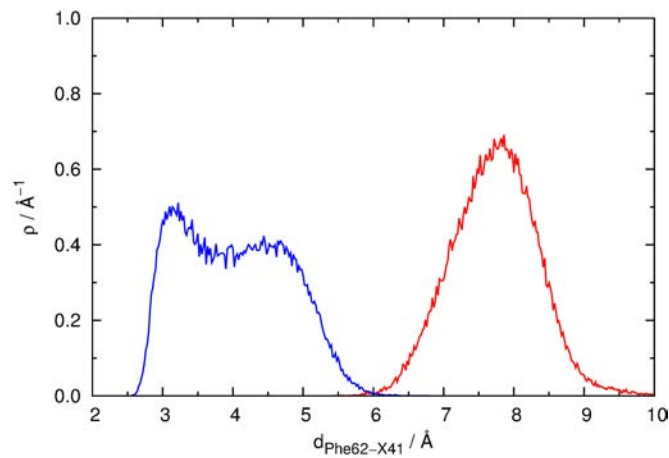


FIGURA 5.16: Distribució de distàncies al llarg dels 100 ns de simulació MD. En blau es mostra la distància entre l'oxigen de la Phe62 i el nitrogen de la Gly41 per la mCardinal i en vermell la distància entre el mateix oxigen i el  $C_{\alpha}$  per la mNeptune.

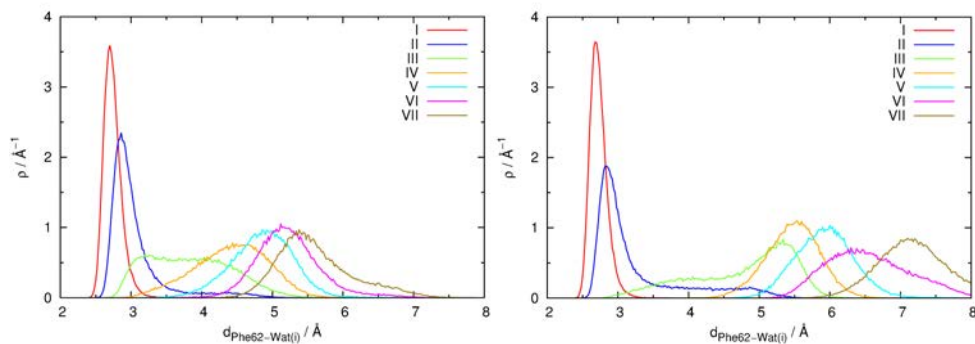


FIGURA 5.17: Estructura de solvatació de l'oxigen de la Phe62 per a la mNeptune (panell esquerre) i la mCardinal (panell dret). Es mostren les funcions de distribució radial per a les aigües més properes a l'oxigen de la Phe62 en cada cas.

temps. Recordem, emperò, que la mCardinal, a més de la interacció amb les aigües, té una interacció (present un 50% del temps) amb la Gln41, així que, de forma global, podem afirmar que l'excés de càrrega rebut després de l'excitació en la mCardinal queda més estabilitzat que en la mNeptune, fet que sembla causar el desitjat corriment cap al vermell.

En definitiva, s'ha identificat que l'excitació corresponent a la Banda I de l'espectre d'absorció comporta un desplaçament de la densitat electrònica des de l'anell fenòlic del cromòfor cap al grup acilimínic, així que s'ha aconseguit explicar la diferència entre les dues RFPs estudiades en base a d'interacció d'aquest grup amb els residus o aigües properes, sempre a través de l'oxigen del residu Phe62.

### **5.3 *Red Hot Chili Proteins: Proposta d'una Nova RFP***

El coneixement desenvolupat fins ara sobre les característiques de l'excitació en les GFPs i les RFPs ens ha de permetre, finalment, comprendre com funcionen aquests biosistemes per tal que puguem dissenyar noves mutants amb característiques obtingudes *ad hoc*. En aquest sentit, cal remarcar que, sovint, la concepció de les tècniques mutagèniques experimentals tenen un cert component aleatori. Es fan bancs de proves de proteïnes mutades i es mesuren les seves propietats. En el nostre cas, gràcies al coneixement adquirit, podem reduir les possibles mutacions a provar tot guiant-nos per la coneixença de les característiques espectroscòpiques del sistema, sabent que caldrà fer un estudi seriós sobre les mutants obtingudes per a poder concloure si s'ha aconseguit l'efecte volgut. Com s'ha anat repetint al llarg del treball, l'interès en les RFPs, des del punt de vista de les aplicacions biomèdiques, és aconseguir desplaçaments cap al roig de les bandes d'absorció i emissió. Desplaçar l'absorció cap al roig vol dir reduir l'energia entre els dos estats electrònics i per aconseguir-ho, podem desestabilitzar l'estat fonamental respecte l'excitat o estabilitzar l'estat excitat respecte el fonamental. Aquest comportament queda esquematitzat, que no totalment representat, a la Figura 5.18.

#### **5.3.1 Selecció de Candidates**

La primera part d'aquest capítol parlava de les GFPs i l'efecte d'una càrrega negativa sobre la densitat de l'anell fenòlic. Es conclouïa que la proximitat d'aquesta càrrega feia augmentar la longitud d'ona d'absorció ja que es desestabilitzava l'estat fonamental respecte l'excitat. En el cas de les RFPs, es donava l'altre opció, l'estabilització de l'estat excitat respecte el fonamental. Com que l'excitació transfereix densitat de càrrega cap al grup imidazolinònic, la proximitat a aquest d'un grup electró-atraient



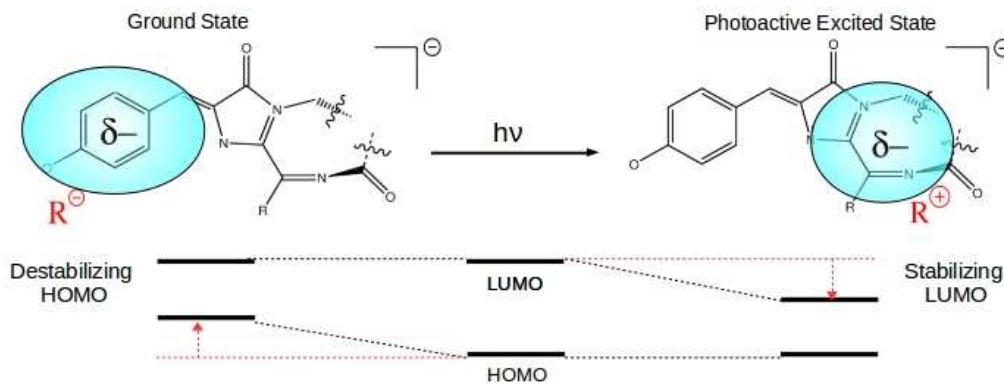


FIGURA 5.18: Esquematzació de les possibles estratègies a seguir per aconseguir un desplaçament cap al roig de la longitud d'ona d'absorció en les FPs. Es mostra el cromòfor de les RFPs, però qualitativament el cromòfor de les GFPs és equivalent.

(o directament carregat positivament), feia augmentar la longitud d'ona d'absorció (Figura 5.18).

La mCardinal és, de les proteïnes estudiades, la que està més desplaçada cap al roig així que la prendrem com a candidata sobre la qual proposarem noves mutacions per encara aconseguir un desplaçament major. Tenint en compte el coneixement adquirit i l'entorn dels dos fragments objectius de les mutacions (l'anell fenòlic i el grup acilimínic), que es mostren a la Taula 5.4, podem fer el següent anàlisi sobre els mateixos:

- **Interaccions amb el fragment acilimínic:** Els residus candidats són la Leu13, la Gln41 i la Ser64. La idea serà substituir els residus esmentats per d'altres que incrementin la càrrega de l'àtom interaccionant per així estabilitzar l'estat excitat i disminuir l'energia d'excitació.
  - Leu13: Al ser un residu de caràcter alifàtic, la leucina no interacciona efectivament amb l'oxigen de la Phe62, però ocupa una posició propera al mateix. Així doncs, la mutació de la Leu13 per un aminoàcid que pugui interaccionar amb la Phe62 segurament tindrà algun efecte en les propietats espectroscòpiques de la proteïna. Es proposa la substitució de la Leu13 per una asparagina (mutació **L13N**), que és gairebé isostèrica i a més conté un grup  $\text{NH}_2$ , susceptible d'interaccionar amb l'oxigen de la Phe62 i també la inclusió de una lisina (mutació **L13K**) de la qual esperarem una interacció encara major ja que conté un grup carregat positivament (encara que és més voluminosa).
  - Gln41: A la secció anterior s'ha demostrat com la interacció entre el nitrògen de la Gln41 i l'oxigen de la Phe62 afecta notablement a la longitud d'ona d'absorció. Es proposa augmentar el



Anell fenòlic		Grup acilimínic	
Residu : Àtom	Distància a l'oxigen fenòlic (Å)	Residu : Àtom	Distància a l'oxigen acilimínic (Å)
Thr60 : OG1	7,5	Leu13 : CD1	3,6
Glu139 : CD	10,1 (fora)	Gln39 : NE2	5,3
Ala140 : CB	6,1 (fora)	Gln40 : NE2	3,0
Thr141 : OG1	2,8	Cys61 : SG	8,0 (fora)
Met158 : CE	3,7	Ser64 : OG	4,1
Arg195 : CZ	8,0	Gln104 : NE2	5,8
Arg196 : CZ	9,7 (fora)	Tyr115 : OH	5,2
Leu197 : CG	5,7		

TAULA 5.4: Relació de residus més propers als fragments fenòlic i acilimínic del cromòfor de la mCardinal amb les seves distàncies, preses de l'estructura cristal·lina resolta per rajos X. L'etiquetatge dels àtoms és el propi del camp de forces. La indicació "fora" significa que el residu en qüestió apunta en direcció oposada al cromòfor.

caràcter de càrrega positiva del nitrògen interaccionant incloent una lisina (que conté un grup  $-NH_3^+$ ) per fer la mutació **Q41K**. Cal apuntar que la lisina és un aminoàcid més voluminós que la glutamina, fet que pot tenir conseqüències en l'orientació de la primera en la mutant **Q41K**.

- Ser64: Encara que aquest aminoàcid seria susceptible d'interaccionar amb el cromòfor, la seva orientació no és adequada per a tal fi. Els aminoàcids amb grups electró-atraients o amb càrrega positiva són voluminosos, així que la substitució de la Ser64 per algun d'aquests aminoàcids segurament no resultaria en cap interacció efectiva.
- **Interaccions amb el fragment fenòlic:** Els residus candidats són la Thr141, la Met158 i la Thr60. La idea serà substituir els residus esmentats per d'altres que trenquin la interacció estabilitzant que fan aquests residus amb el cromòfor a l'estat fonamental per així desestabilitzar-lo i disminuir l'energia d'excitació.
  - Thr141: Aquest residu forma un pont d'hidrogen amb l'oxigen fenòlic del cromòfor així la substitució de la treonina per una valina (mutació **T141V**), que és isostèrica, però no pot formar pont d'hidrogen amb el fenolat del cromòfor, podria causar un efecte de desestabilització de l'estat fonamental.
  - Met158: Aquest aminoàcid ja és hidrofòbic així que només la substitució del mateix per algun que tingui una càrrega negativa podria desestabilitzar l'estat fonamental. Un candidat seria el glutamat (mutació **M158E**).
  - Thr60: La Thr60 reposa sobre el cromòfor, així que la mutació de

la mateixa per un residu carregat negativament podria desestabilitzar l'estat fonamental per la interacció de tal càrrega negativa amb la densitat electrònica de l'anell fenòlic. La substitució d'aquest residu per un aspartat per a crear la mutant **T60D** podria ser vàlida per a tal fi.

## Metodologia

Per tal de comprovar l'efecte que causarien les mutacions anteriorment proposades en l'espectre d'absorció de la proteïna, cal calcular els espectres corresponents a les mateixes en un context similar al seguit en aquest capítol.

Així doncs, s'ha realitzat les mutacions proposades sobre l'estructura cristal·logràfica resolta per rajos X de la mCardinal utilitzant el programa Chimera i, sobre totes elles, s'ha aplicat el mateix protocol de solvatació, escalfament i equilibratge seguit en l'apartat de metodologia de la secció anterior (secció 5.2.1).

Pel què fa a la producció, i donada la gran quantitat de casos a tractar i els recursos computacionals disponibles, s'ha propagat 10 ns de simulació MM per a cada cas. Un eventual estudi en profunditat d'alguna de les mutants comportarà una major exploració de l'espai conformacional d'algunes proteïnes, tot utilitzant el mètode *task farming* a partir d'aquesta simulació dinàmica primigènia, per a aquelles mutants més prometedores.

A l'igual que en el casos tractats en la secció anterior, per al càlcul d'estats excitats a partir de les estructures obtingudes de les simulacions MM-MD, s'ha utilitzat un context QM/MM. La part QM ha estat el cromòfor juntament amb alguns dels àtoms dels residus limítrofs al mateix, essent la resta, àtoms de la part MM. Per al càlcul de l'energia de la part QM s'ha utilitzat el mètode TD-DFT amb el funcional CAM-B3LYP i la base 6-31G(d,p) implementat en el programa Gaussian09. Per a la part MM s'ha utilitzat el mòdul DL-POLY del programari Chemhell amb el camp de forces CHARMM27. La interacció QM/MM s'ha tractat en el context EE, amb l'aproximació *link-atom* i la correcció *charge shift*.

A fi i efecte d'explorar el rendiment dels mètodes semiempírics en aquests sistemes, també s'ha repetit els càlculs fent servir el mètode OM2 per a la part QM. En aquest context s'ha recalculat els espectres a nivell CIS-OM2, CISD-OM2 i CISDT-OM2.

### 5.3.2 Resultats

Els espectres d'absorció calculats per a les mutants proposades es mostren a la Figura 5.19. Els màxims dels pics de la Banda I corresponent a cada mutant es poden trobar a la Taula 5.5. Aquests resultats gràfics i numèrics denoten com, de totes les mutants proposades, les úniques que presenten

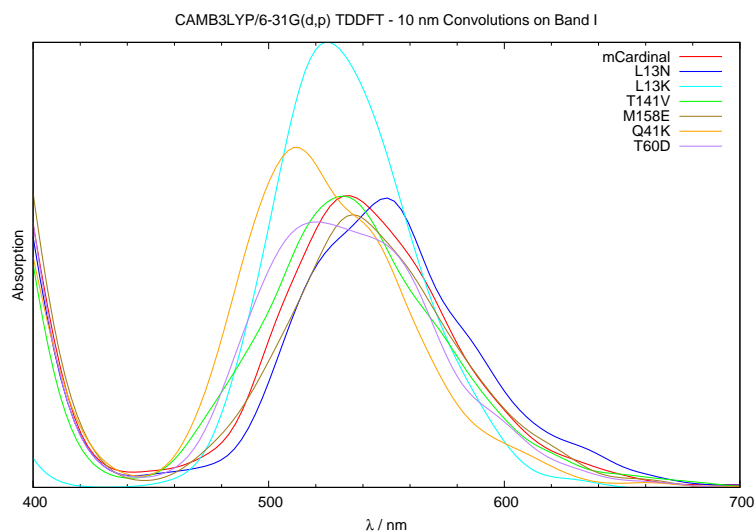


FIGURA 5.19: Espectre d'absorció calculat per a les mutants proposades. S'ha utilitzat una convolució amb funcions gaussianes per reduir el soroll estadístic amb una  $\omega=10$  nm.

Proteïna	CAM-B3LYP/TD-DFT	CIS-OM2	CISD-OM2	CISDT-OM2
mCardinal	516,4	419,9	458,2	525,3
L13N	527,8	424,4	465,3	527,0
L13K	509,9	433,7	471,0	526,3
T141V	512,7	423,5	449,2	514,0
M158E	519,7	418,4	455,6	506,2
Q41K	495,5	416,6	451,5	510,7

Longituds d'ona d'absorció mostrades en nm

TAULA 5.5: Màxims de les longituds d'ona d'absorció per a cadascuna de les mutants estudiades i per a cada mètode de càlcul emprat.

desplaçament cap al vermell són la M158E i la L13N, essent aquesta última la millor candidata (doncs la mutant M158E presenta un desplaçament molt suau).

A aquestes alçades del treball, el lector ja haurà pogut intuir que les proteïnes fluorescents, són foto-biosistemes extremadament complicats. L'anàlisi en profunditat del comportament trobat a les GFPs (secció 5.1) i les RFPs (secció 5.2), s'ha realitzat per a poder crear (amb un èxit notable, cal afegir) un coneixement ferm sobre els factors que influencien en la longitud d'ona d'absorció les proteïnes fluorescents. En la secció que ens ocupa, emperò, l'objectiu és proposar una o diverses mutants que presentin un desplaçament cap al roig addicional al que mostra la mCardinal. És evident que hi ha raons estructurals rere el fet que la majoria de mutants no presentin desplaçament cap al roig, però s'evadeix aquesta discussió per no ser l'objectiu del present apartat. A més, un anàlisi detallat de totes i

cadascuna de les mutants, dut a terme al nivell del realitzat per la mNeptune i mCardinal es fa del tot inviable i està fora de lloc, doncs només estariem parlant d'estructures hipotètiques, de les quals no hi ha dades experimentals. Amb això, es pot afirmar que molt probablement, les mutacions proposades no hagin tingut l'efecte esperat, o n'hagin comportat d'addicionals que hagin anul·lat l'efecte conegut i racionalitzat en l'apartat anterior, sense perjudici del mateix. Així doncs, a continuació només es discutiran els casos que han resultat positius: les mutants M158E i L13N.

La mutació M158E, que pretenia desestabilitzar l'estat fonamental respecte l'excitat tot situant una càrrega negativa a prop de l'anell fenòlic del cromòfor ha resultat en un tímid efecte de 3 nm. Encara que aquest desplaçament és molt petit i s'entén que pot no ser significatiu per estar dins l'error del mètode, cal presentar la mutant M158E com a possible candidata. La raó del poc efecte aconseguit sembla ser que la càrrega negativa del Glu158 es situa a una distància mitjana de  $\sim 6 \text{ \AA}$  de l'anell fenòlic del cromòfor.

La mutació més prometedora en base a l'efecte aconseguit respecte l'espectre d'absorció de la proteïna (que en aquest cas comporta un desplaçament cap al roig d'11 nm) és la L13N. Recordem que la mutació L13N s'havia proposat a fi i efecte d'afegir una interacció addicional entre l'oxigen de la Phe62 (que és un ancoratge per estabilitzar l'excés de densitat electrònica rebuda a l'estat excitat fotoactiu pel grup imidazolinònic) i un grup electró-atraient, a part de la ja existent interacció entre tal oxigen i la Gln41. Efectivament, durant el procés d'equilibratge, el nitrogen de l'Asn13 forma una interacció efectiva amb l'oxigen de la Phe62 que tanmateix no afecta a la interacció del mateix oxigen amb la Gln41. Així doncs, com que s'introdueix una nova interacció estabilitzant de l'estat excitat i que, a més, no trenca les interaccions prèviament existents, s'aconsegueix una nova mutant encara més desplaçada cap al roig. Val a dir que, per al lector no especialitzat, desplaçaments de 3 nm o 11 nm poden semblar irrisoris però cal tenir en compte que en la comunitat científica especialitzada en la recerca en proteïnes fluorescentes, aquests són resultats molt prometedors (especialment el corresponent a la L13N): el sonat descobriment de la mNeptune suposava un corriment cap al vermell de només 15 nm respecte la mKate i la mCardinal, que també va suposar un nou avenç important, presenta un desplaçament addicional de només 4 nm respecte la mNeptune. Així doncs, possibles desplaçaments de 3 nm i 11 nm respecte la mCardinal poden suposar, altre cop, avenços importants.

## **El Mètode OM2**

Com a afegit al treball presentat i per tal de comprovar com de bé poden descriure els espectres d'absorció els mètodes semiempírics, s'ha recalculat els espectres a nivell QM/MM on la part QM s'ha calculat amb els mètodes CIS-OM2, CISD-OM2 i CISDT-OM2. El càlcul CISDTQ-OM2 ha estat inviable doncs les màquines utilitzades no disposen de la memòria

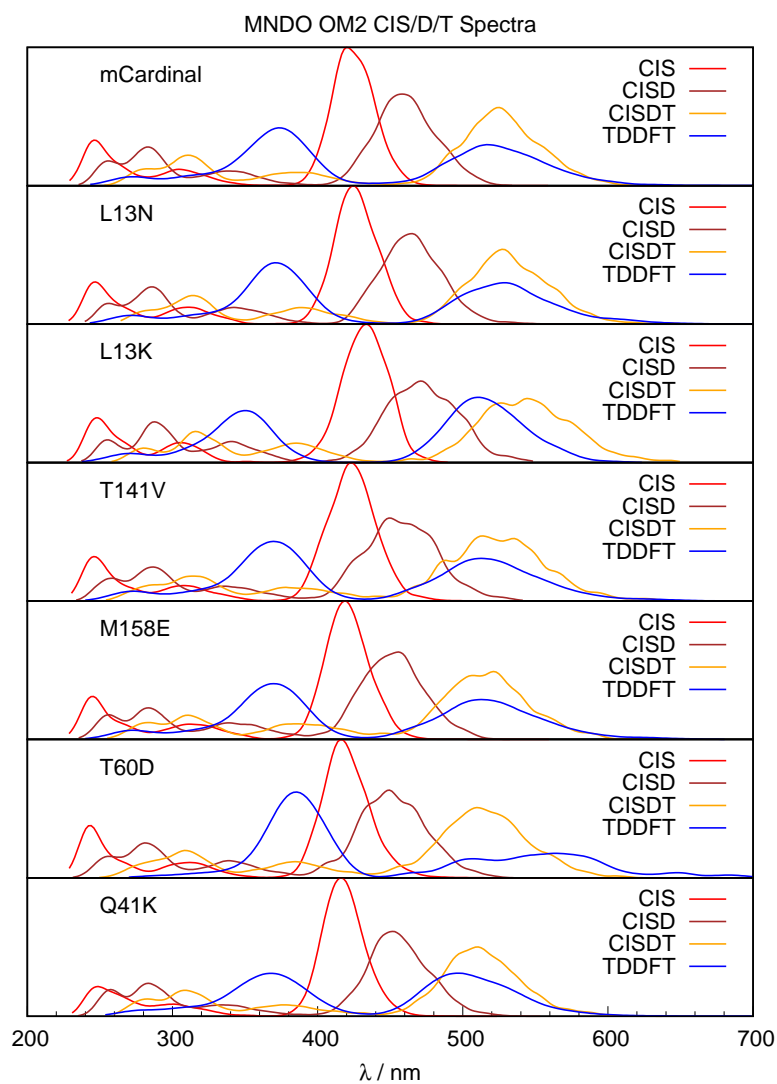


FIGURA 5.20: Espectre d'absorció calculat per a les mutants proposades amb els mètodes OM2 (CIS, CISD i CISDT), on també s'hi inclouen els espectres calculats a nivell TD-DFT. S'ha utilitzat una convolució amb funcions gaussianes per reduir el soroll estadístic amb una  $\omega=10$  nm.

necessària per tenir en compte les excitacions quàdruples amb l'espai actiu triat, que ha estat de 20 electrons en 20 orbitals (10 d'ocupats i 10 de virtuals).

La motivació d'aquests càlculs, i amb la previsió futura d'utilitzar la mutagènesi *in silico* de forma més extensiva, és obtenir un mètode més barat que el TD-DFT però comparable a nivell dels resultats obtinguts i que, per tant, permeti una prospecció molt més ràpida, abans d'escollir les mutants prometedores sobre les quals realitzar els càlculs d'alt nivell TD-DFT. Només per tenir una idea del què pot suposar això a nivell del cost computacional, cal tenir en compte que un càlcul QM/MM a nivell TD-DFT per al cromòfor pres triga de l'ordre de 10 hores de CPU, mentre el mateix càlcul CIS-OM2 (amb l'espai actiu especificat) trigarà de l'ordre de 3 segons. Els casos intermedis, el CISD-OM2 i el CISDT-OM2, trigaran de l'ordre de 8 segons o 10 minuts, respectivament.

A la Figura 5.20 i a la Taula 5.5 es poden observar els resultats. Es fa evident que els mètodes CIS-OM2 i CISD-OM2 infravaloren la longitud d'ona d'absorció de manera flagrant. Amb això, el mètode CISDT-OM2 resulta donar una correspondència acceptable amb els resultats TD-DFT tant a nivell de la forma com de la posició de la banda i, per tant, es postula com a candidat per a ser utilitzat en exploracions mutagèniques futures ja que, recordem, és de l'ordre de  $\sim 60$  vegades més ràpid que el TD-DFT.

## 5.4 Conclusions Parcial

Arribats a aquest punt del treball, és convenient recapitular el coneixement presentat per a poder extreure unes conclusions parcials dels resultats obtinguts, que es classificaràn en conclusions metodològiques i conclusions químiques.

### Conclusions Metodològiques

- El mètode *task-farming* presentat al Capítol 4 ha permès l'obtenció de dinàmiques tant a nivell QM/MM com a nivell MM de longitud considerable en un temps raonable que ens han permès reproduir espectres d'absorció. Naturalment, no permet extreure conclusions sobre l'evolució temporal del sistema més enllà del temps d'integració de cadascuna de les dinàmiques que formen l'eixam, però això no té cap rellevància si es considera que el sistema està equilibrat quan comença la fase de producció (suposició assumida en qualsevol simulació MD).
- El mètode OM3 descriu satisfactòriament l'estat fonamental de les proteïnes fluorescents i, tot i que tendeix a sobreestimar la força de

la interacció de pont d'hidrogen, ha permès la integració de dinàmiques amb interaccions de pont d'hidrogen curts, reproduint-ne l'estructura satisfactòriament.

- Les simulacions dinàmiques presentades han reproduït satisfactòriament les dades experimentals, tan a nivell qualitatiu com quantitatiu, fet que ens ha permès extreure conclusions químiques. El fet que aquestes dinàmiques no estiguin condicionades per res que no sigui el propi mètode de càlcul i el sistema model triat, ens permet afirmar que aquests dos elements han aconseguit capturar la física del problema.
- Els paràmetres MM desenvolupats per al cromòfor de les RFPs han aconseguit donar una bona descripció del moviment del mateix en l'entorn proteic.
- El mètode TD-DFT amb el funcional CAM-B3LYP, tot i les conegudes limitacions que presenta, ha proporcionat una descripció satisfactòria de les longituds d'ona d'absorció i del tipus d'excitació trobades en temps de càlcul raonables, fet que el fa postular-se com a mètode de referència per al tipus de càlculs realitzats en aquest apartat.
- El mètode TD-DFT amb el funcional CAM-B3LYP no presenta SIE en els cromòfors de les RFPs estudiades
- El mètode TD-DFT amb el funcional CAM-B3LYP és sensible als petits canvis estructurals que es donen entre les diferents proteïnes, essent capaç de reproduir (fins hi tot quantitativament, en certa mesura) petites diferències experimentals trobades entre les mateixes.
- El mètode CISDT-OM2 per al càlcul d'estats excitats és un bon candidat per a ser utilitzat com a mètode exploratori de baix cost, ja que dona uns resultats comparables als del TD-DFT, en un temps de càlcul molt inferior.

### Conclusions Químiques

- L'excitació que porta a l'estat excitat fotoactiu en la *wt*-GFP i les seves mutants S65T/H148D i E222Q/H148D (i previsiblement a totes aquelles que comparteixin el mateix cromòfor) comporta el drenatge de densitat electrònica de l'anell fenòlic a favor del pont metínic.
- Es prediu que la doble mutant E222Q/H148D té un cromòfor que es troba lleugerament desviat de la planaritat i que aquest manté una interacció estable en el temps amb el residu Asp148. A més, es prediu que aquesta interacció serà lleugerament menys intensa que en el cas de la mutant S65T/H148D, encara que continuarà essent una interacció de pont d'hidrogen curt amb un protó deslocalitzat.

- Al contrari del què es pensava, la proximitat de la càrrega negativa de l'aspartat a l'anell fenòlic del cromòfor és la causant del desplaçament cap al roig de les mutants de la GFP estudiades. A més, hi ha una correlació directa entre la distància a la qual es situa aquesta càrrega i la longitud d'ona s'absorció (a menys distància, més desplaçament cap al roig).
- L'excitació que porta a l'estat excitat fotoactiu en les RFPs mNeptune i mCardinal (i previsiblement a totes aquelles que comparteixin el mateix cromòfor) comporta el drenatge de densitat electrònica de l'anell fenòlic a favor del grup acilimínic.
- La inclusió del grup acilimínic (amb el doble enllaç C=N) de les mutants RFPs no fa desplaçar substancialment la banda original de les GFPs si no que en fa aparèixer una altra de nova. S'ha demostrat com l'excitació de la Banda II de l'espectre calculat per a les RFPs es correspon qualitativament a l'excitació corresponent a la Banda A trobada en les GFPs.
- Un possible bon punt d'ancoratge per a capturar interaccions del grup acilimínic amb grups electró-donadors és l'oxigen de la Phe62.
- La diferència entre les longituds d'ona d'absorció entre la mNeptune i la mCardinal és que, en la segona, hi ha una interacció efectiva entre l'oxigen de la Phe62 i la Gln41, que estabilitza l'excés de densitat electrònica del grup acilimínic a l'estat excitat.
- L'estabilització de l'estat excitat per l'acostament de grups electró-atraients prop del grup acilimínic del cromòfor o la desestabilització de l'estat fonamental per l'acostament de grups electró-donadors a l'anell fenòlic del cromòfor, produeixen un desplaçament cap al roig (Figura 5.18).
- Es proposen les mutants M158E-mCardinal i L13N-mCardinal (especialment aquesta última) com a noves proteïnes vermelles fluorescentes capaces d'absorbir a la "finestra de transparència" encara en una zona més desplaçada cap al roig que la pròpia mCardinal.

Els resultats presentats en aquest capítol han donat lloc a dues publicacions, més una tercera que està en procés de redacció:

- P. Armengol, R. Gelabert, M. Moreno, J.M. Lluch "New Insights into the Structure-Spectrum relationship in S65T/H148D and E222Q/H148D Green Fluorescent Protein mutants. A theoretical assessment", *Org. Biomol. Chem.* 12(48):9845-9852, 2014.<sup>185</sup>
- P. Armengol, R. Gelabert, M. Moreno, J.M. Lluch "Chromophore interactions leading to different absorption spectra in mNeptune1 and mCardinal red fluorescent proteins", *Phys. Chem. Chem. Phys.* 18(25):16964-16976, 2016.<sup>186</sup>



- P. Armengol, R. Gelabert, M. Moreno, J.M. Lluch "Redder than Blood: Rational Design of a Red Fluorescent Protein Absorbing in the Mammalian Tissue Transparency Window- *en redacció*."

## 6 Desxifrant la Pedra Rosetta

En el present capítol s'explicaran els resultats obtinguts respecte la doble mutant S65T/H148D-GFP. Com s'ha explicat a la Introducció, es planteja un estudi complert i exhaustiu d'aquesta doble mutant, que representa un cas "senzilli paradigmàtic de les proteïnes fluorescents i que serà utilitzada com a clau per desxifrar el funcionament general de les proteïnes fluorescents.

### 6.1 Metodologia

Per tal d'estudiar la complexitat de la proteïna S65T/H148D, s'ha utilitzat una gran varietat de mètodes ja explicats amb anterioritat, prenent com a sistema model el mateix utilitzat en el Capítol 5, així que es remet al mateix per als detalls sobre el model utilitzat. En aquesta secció es donaran els detalls computacionals dels càlculs específics realitzats per aquest sistema.

#### Càlculs Estàtics

Els càlculs estàtics s'han realitzat en contextos QM i QM/MM, depenent del cas. Les energies i gradients dels estats fonamental i excitat per als càlculs QM s'han obtingut amb els mètodes DFT i TD-DFT, respectivament, utilitzant el funcional CAM-B3LYP i la base 6-31G(d,p), implementats en el programari Gaussian09. Els càlculs QM/MM s'han realitzat utilitzant la interfície ChemShell on la part QM s'ha calculat de la mateixa manera que la descrita anteriorment i la part MM s'ha calculat amb el programa DL-POLY, utilitzant el camp de forces CHARMM27. La part QM consta dels àtoms del cromòfor, l'Asp148 i algun dels àtoms dels residus més propers. La interacció QM/MM s'ha tractat amb l'aproximació *link-atom/charge-shift* en un context d'EE.

#### Càlculs Dinàmics

En aquest capítol, s'ha realitzat 3 classes de càlculs de dinàmica: simulacions QM/MM-MD a l'estat fonamental, simulacions QM/MM-MD en un context no adiabàtic (partint de l'estat excitat fotoactiu) i simulacions de QD. La simulació QM/MM-MD de l'estat fonamental utilitzada és la mateixa presentada en el Capítol 5, apartat 5.1, així que referiu-vos allà per

a més detalls. Per a les simulacions de QD s'ha utilitzat la representació *sync-DVR* de Colbert i Miller descrita a l'apartat 2.2.2, utilitzant 50 punts com a base i s'ha verificat la convergència del càlcul.

Per a la simulació no adiabàtica, que pretén simular els processos químics que tenen lloc després de l'excitació, cal un paquet d'estructures inicials que estiguin ja en equilibri a l'estat fonamental, doncs no tindria cap sentit equilibrar el sistema a l'estat excitat, ja que precisament el què es busca és simular els processos fora de l'equilibri que ocorren en el mateix. Per a fer tal cosa, s'ha emprat la simulació QM/MM de l'estat fonamental (OM3/CHARMM27) presentada al Capítol 5: per representar una mostra del sistema en una distribució d'equilibri a 300 K, s'ha pres aleatòriament un reguitzell de fotogrames d'aquesta trajectòria (un total de 350) i s'han propagat a l'estat excitat, prenent com a punt inicial per a la dinàmica clàssica les posicions i velocitats corresponents a cada fotograma. Per a la integració de la dinàmica no adiabàtica a l'estat excitat s'ha definit una zona activa de 20 Å de radi a partir del centre de la gota i la resta s'ha mantingut fix.

Per al càlcul d'energies i gradients de les simulacions no adiabàtiques s'ha utilitzat un context QM/MM on el mètode QM emprat ha estat una expansió multireferencial fins a dobles excitacions del mètode OM2 (MRCISD-OM2). Aquest mètode ha descrit amb èxit notable processos no adiabàtics en sistemes de mida reduïda, tant a nivell QM com QM/MM.<sup>118-121</sup> S'han pres 3 referències i han estat la corresponent a la capa tancada, la de la monoexcitació HOMO-LUMO i la de biexcitació HOMO-LUMO. Cal recordar que els orbitals HOMO i LUMO són els involucrats en l'excitació que porta a l'estat excitat fotoactiu. L'espai actiu que s'ha considerat per a les excitacions electròniques consta de 8 orbitals ocupats i 10 de virtuals, *ossia* 10 electrons en 18 orbitals. Per a la part MM s'ha utilitzat el camp de forces CHARMM27.

Les trajectòries a l'estat excitat s'han propagat en un context no adiabàtic fent servir el mètode *surface hopping* amb l'algoritme FSA, que s'ha activat només per a situacions on la diferència d'energia entre els estats d'interès ha estat inferior a 10 kcal/mol (suposant, així, que la probabilitat de transició entre estats quan aquests es troben a una separació major és menyspreable) i on també s'ha inclòs la correcció de decoherència proposada per Granucci *et al.*<sup>187</sup> S'ha utilitzat un temps d'integració d'1 fs pel moviment nuclear i de 0,0025 fs pel moviment electrònic. En una simulació dinàmica on el càlcul de l'energia electrònica està sotmès a la composició d'un espai actiu, i donat que durant la mateixa les configuracions nuclears poden canviar notablement, l'espai actiu en qüestió també pot patir modificacions dràstiques. Una discontinuïtat en l'espai actiu pot portar a canvis sobtats en l'energia i el gradient tot donant lloc a situacions no físiques. Per sobrevenir aquest problema, s'ha utilitzat el mètode de passos d'integració adaptatius desenvolupat per Spöerker *et al.* i que ha estat provat en el cromòfor de la GFP en solució aquosa.<sup>188</sup> Aquest mètode adapta el

pas d'integració en funció de la curvatura de la PES, tot reduint-lo al passar per zones conflictives on hi hagi canvis notables en la composició de l'espai actiu, fet que fa augmentar sensiblement l'efectivitat del mètode. Del total de 350 trajectòries a l'estat excitat, que han estat propagades durant 1,5 ps, només 22 han acabat en algun error degut a la convergència de la part QM, és a dir que hi ha hagut una supervivència del 93,7%. Tota aquesta metodologia ha estat implementada durant el desenvolupament d'aquesta tesi en una versió de desenvolupament del ChemShell, en estreta col·laboració amb el grup de recerca de Prof. W.Thiel del Max-Planck Institut für Kohlenforschung.

## 6.2 La Factibilitat de la ESPT

Les dades experimentals disponibles per a la doble mutant S65T/H148D exposades a la introducció d'aquesta tesi, poden ser explicades en base a les següents assumpcions: (1) que el procés de transferència protònica entre el cromòfor i l'Asp148 no és favorable a l'estat fonamental, (2) que la fotoexcitació del cromòfor resulta en un gran guany d'acidesa del mateix que (3) fa esdevenir favorable l'anomenada transferència que a més (4) és sense barrera. Aquestes assumpcions explicarien el sobtat creixement de la senyal de fluorescència (< 50 fs) de la banda I\* sense efecte cinètic d'isòtop mesurable. A més, la manca d'evolució del mode associat al grup carboxílic de l'Asp148 denotaria un protó altament deslocalitzat ja a l'estat fonamental. Amb les eines teòriques i computacionals de les quals disposem, podem comprovar la veracitat (o no) d'aquestes assumpcions.

### 6.2.1 Perfils Estàtics

Gran part de l'explicació proposada per a la fotoquímica de la S65T/H148D recau en el suposat gran guany d'acidesa del cromòfor a l'estat excitat. Aquest guany d'acidesa hauria d'estabilitzar el mínim d'energia potencial del fotoproducte respecte la forma neutre del cromòfor, que és la que s'associa a la banda A. Tenint en compte la redistribució electrònica que té lloc (recordem el drenatge de densitat electrònica de l'anell fenòlic), la visió qualitativa del guany d'acidesa del protó fenòlic del cromòfor és assumible però, en quina mesura? Per mitjà de càlculs teòrics, Scharnagl i Raupp-Kossmann van estimar que la caiguda de  $pK_a$  de l'HBDI (el cromòfor de la GFP lliure en solució) seria d'unes 8 unitats, anant des de 8,3 a l'estat fonamental a 0,1 a l'estat excitat.<sup>53</sup> Tenint en compte que el  $pK_a$  de l'aspartat en solució és de 3,7,<sup>189</sup> la transferència és imaginable.

S'ha calculat perfils per a la transferència protònica a nivell QM/MM partint de l'estructura cristal·lina del PDB i de l'estructura optimitzada a nivell DFT/CHARMM27 (vegeu Figura 6.1). Com es pot observar, ni a l'estructura cristal·lina ni a l'optimitzada a nivell DFT/CHARMM27 es

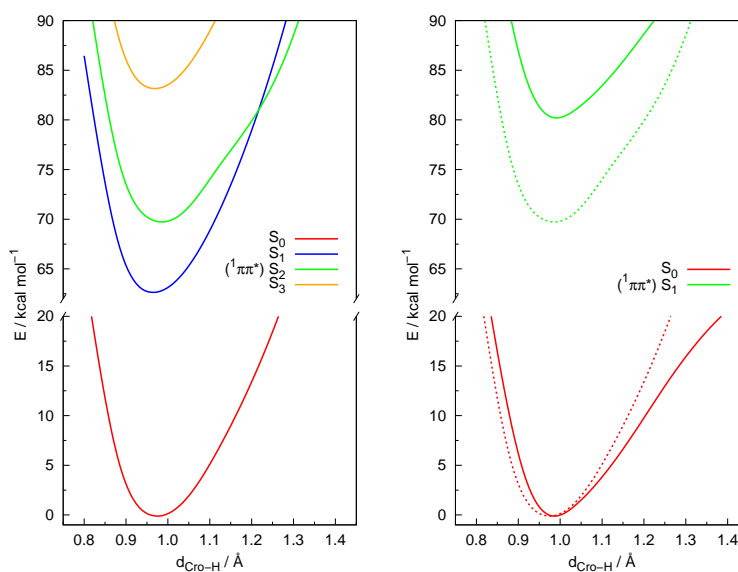


FIGURA 6.1: Perfils d'energia potencial per a la transferència protònica per a la mutant S65T/H148D-GFP per a diferents estats electrònics a nivell TD-DFT/CHARMM27. Esquerra: Perfil realitzat a partir de l'estructura cristal·lina del PDB però en el model de solvatació QM/MM explicat. S'hi mostra l'estat fonamental ( $S_0$ ) i els primers 3 estats excitats ( $S_1$ ,  $S_2$  i  $S_3$ , essent el segon l'estat fotoactiu (etiquetat com a  $\pi\pi^*$ ). Dreta: Perfil realitzat a partir de l'estructura optimitzada a l'estat fonamental a nivell QM/MM. S'hi mostra l'estat fonamental,  $S_0$ , i l'estat fotoactiu,  $S_1$  en aquest cas (etiquetat com a  $\pi\pi^*$ ). Per tal de facilitar la comparació, en línies puntejades, també s'ha representat els estats corresponents de l'estructura cristal·lina del PDB.

pot apreciar la formació d'un fotoproducte a l'estat excitat. A més, el mínim d'energia potencial de l'estat fonamental està ben localitzat a la forma neutra, fet que tampoc permetria explicar la suposada deslocalització del protó. Tot i això, es pot veure un cert guany de l'anharmonicitat en el perfil corresponent a l'estructura optimitzada, tot i que encara no seria suficient per explicar el comportament observat per Stoner-Ma *et al.* Així doncs, una descripció purament estàtica del problema no és suficient per explicar una transferència que sembla tenir lloc, com indiquen totes les dades experimentals. Per tant, aquests resultats indiquen que la suposició del gran guany d'acidesa del cromòfor a l'estat excitat que donaria una transferència intrínsecament lligada al mateix perd força, a favor d'una visió més completa on probablement l'entorn proteic hi jugui un paper important. Amb tot, abans de continuar, cal analitzar críticament la hipòtesis del gran guany d'acidesa, invocada tant sovint pels grups experimentals.

## 6.2.2 L'Acidesa del Cromòfor, a Prova

Scharnagl i Raupp-Kossmann van determinar el  $pK_a^*$  del cromòfor de la GFP lliure (HBDI) en solució per mitjà de càlculs teòrics de caire termodinàmic.<sup>53</sup> En aquest punt cal recordar el significat del  $pK_{a'}$  que es defineix com el logaritme en base 10 de la constant d'equilibri de la dissociació d'un àcid en aigua, canviat de signe. Així doncs, de manera estricta, parlar de  $pK_a$  implica parlar d'equilibri termodinàmic i sempre de l'espècie en qüestió amb l'aigua. Per tant, l'extrapolació d'un valor termodinàmic de reacció amb l'aigua per a intentar definir una propietat fora de l'equilibri entre un parell àcid/base (la transferència del protó entre el cromòfor i l'aspartat) ja és, en si, formalment incorrecta. Amb això, tot i assumir aquest fet, hi ha altres problemes amb l'assumpció que valor calculat per Scharnagl i Raupp-Kossmann respecte l'HBDI en solució sigui vàlid pel cromòfor dins de la proteïna. En el seu desenvolupament, Scharnagl i Raupp-Kossmann van derivar la següent expressió per a la diferència de  $pK_a$  entre l'estat fonamental i l'excitat:

$$pK_a^* - pK_a = \Delta pK_a^* + \Delta \Delta pK_a^S \quad (6.1)$$

on

$$\Delta pK_a^* = \frac{1}{RT \ln 10} [\Delta E_A^* - \Delta E_{AH}^*] \quad (6.2)$$

$$\Delta \Delta pK_a^S = -\frac{1}{RT \ln 10} \ln \frac{\langle e^{\beta \Delta \Delta E_A^*} \rangle_0}{\langle e^{\beta \Delta \Delta E_{AH}^*} \rangle_0} \quad (6.3)$$

Estat de Protonació	Experiment (aigua)	CISD/AM1 (aigua) <sup>a</sup>	TDDFT/CAM-B3LYP (aigua) <sup>b</sup> (aigua) <sup>c</sup> (prot.) <sup>d</sup>		
neutre	368 <sup>e,f</sup> , 370 <sup>g</sup>	323	332	388	408
aniònic	425 <sup>f</sup> , 426 <sup>g</sup> , 428 <sup>e</sup>	389	388	436	447

<sup>a</sup> Ref 53; <sup>b</sup> Estructura optimitzada a l'estat fonamental, dissolvent introduït a nivell PCM; <sup>c</sup> Estructura no relaxada (només del cromòfor) presa del PDB (ID: 2DUF), dissolvent introduït a nivell PCM; <sup>d</sup> Càlcul QM/MM sobre l'estructura del PDB (ID: 2DUF) amb aigües explícites; <sup>e</sup> Ref 190; <sup>f</sup> Ref 27; <sup>g</sup> Ref 191

TAULA 6.1: Longituds d'ona d'absorció (en nm) de la forma neutre i aniònica de l'HBDI en aigua i del cromòfor en la doble mutant S65T/H148D

L'equació 6.2 representa la diferència en les energies d'excitació pel cromòfor aniònic i neutre mesurada al seu mínim corresponent de l'estat fonamental i l'equació 6.3 quantifica les contribucions degudes a les diferències en les energies d'excitació del cromòfor aniònic i neutre en les seves zones configuracionals accessibles i essencialment aporta les correccions entròpiques al  $pK_a^*$ . Donada la gran quantitat de càlculs necessaris per obtenir els valors mitjos de l'equació 6.3, la metodologia emprada pel càlcul dels termes de l'equació 6.1 implica d'utilització del semiempíric AM1 per l'obtenció d'energies electròniques, tot introduint els efectes de solvatació per mitjà de mètodes de solvent continu (PCM). Les energies d'excitació foren determinades com a la diferència entre la primera i segona arrel de l'expansió de la matriu CI usant el mètode AM1.

L'ús del mètode AM1 per a calcular les energies d'excitació de les equacions 6.1 i 6.3 presenta certs problemes, ja que aquest mètode fou parame-tritzat per a l'estat fonamental, a més de no incorporar correccions d'ortogonalització, que milloren notablement la descripció energètica dels orbitals virtuals (i que sí que incorporen els mètodes OMx). Si bé és cert que en l'equació 6.3 aquestes energies apareixen com a part d'una fracció de valors mitjos (fet que podria comportar una certa cancel·lació d'errors sistemàtics), en l'equació 6.2 aquest errors poden tenir efectes desastrosos en el càlcul final. A fi i efecte de comprovar com de correcta és la descripció AM1 respecte la TD-DFT/CAM-B3LYP i a més refinar, si s'escau, els valors de les energies d'excitació susceptibles de resultar en errors greus, s'ha recalculat les energies d'excitació del cromòfor en un model PCM a nivell TD-DFT/CAM-B3LYP (Taula 6.1).

El resultat més sorprenent de la Taula 6.1 és la bona concordança que hi ha entre les energies d'excitació AM1 i TD-DFT/CAM-B3LYP per als casos comparables (és a dir, pel càlcul de l'energia d'excitació de l'estructura optimitzada del cromòfor en un entorn PCM). Això hauria de portar a pensar que l'ús del TD-DFT també hauria de donar un resultat similar (una gran caiguda del  $pK_a$  a l'estat excitat). Si assumim que els valors usats per Scharnagl i Raupp-Kossmann per a la magnitud  $\Delta\Delta pK_a^S$  es poden

utilitzar directament en el nostre càlcul (assumim que hi ha cancel·lació d'errors en aquest terme), podem recalcular el  $pK_a^* = 2,1$ . Aquest valor (tot i que no és tant exagerat com el valor AM1), encara representa una caiguda de més de 6 unitats respecte el cromòfor a l'estat fonamental. Això hauria de comportar que, donat que el  $pK_a$  de l'aspàrtic és de 3,7, la reacció hauria de ser espontània. Així doncs, perquè obtenim perfils endoèrgics a l'estat excitat fotoactiu?

Tot i que és obvi, l'HBDI en solució aquosa difereix essencialment en dos aspectes respecte l'HBDI en l'entorn proteic. Per una banda, hi ha l'entorn de càrregues determinat per l'estructura de la proteïna i, per l'altra, la distorsió que pateix el cromòfor dins la cavitat proteica: els dos anells del cromòfor, que són coplanars en l'estructura optimitzada en solució, no són a l'interior de la proteïna. Així doncs, si prenem l'estructura del cromòfor a l'interior de la proteïna i repetim el càlcul d'estats excitats sense l'optimització prèvia però en solució aquosa, obtindrem l'efecte que fa la distorsió geomètrica del cromòfor en l'energia d'excitació del mateix, sense tenir en compte l'efecte que puguin realitzar l'estructura de càrregues de la cavitat proteica. Aquest resultat es mostra a la cinquena columna de la Taula 6.1. Amb aquests nous valors, s'obté un  $pK_a^*$  de 5,3. Així doncs, en aquestes circumstàncies, des d'un punt de vista estàtic, la reacció ja no tindria lloc.

Si ara, a més, recalculem els valors de les energies d'excitació del cromòfor en l'entorn proteic (un càlcul de nivell QM/MM sobre l'estructura del PDB) podem afegir l'efecte que les càrregues tenen en la mateixa. Aquest resultat es mostra a l'última columna de la Taula 6.1. Prenent aquestes energies, el càlcul del  $pK_a^*$  dona un valor de 6,7 (a comparar amb el valor de 3,7 de l'aspàrtic). Així doncs, si tenim en compte la distorsió geomètrica que pateix el cromòfor i l'estructura de càrregues de la cavitat proteica, és fàcil entendre el perquè de l'obtenció de perfils endoèrgics per a la transferència protònica a l'estat excitat. Per tant, cal concloure que, més enllà de la validesa que puguin tenir els resultats obtinguts per Scharnagl i Raupp-Kossmann pel càlcul del  $pK_a^*$  de l'HBDI en solució aquosa, hi ha un error en l'extrapolació d'aquest valor al cromòfor a dins la proteïna, fet obviat gairebé de manera sistemàtica a gran part de la literatura. Paradoxalment, aquesta errònia extrapolació mai fou proposada per Scharnagl. Així doncs, cal tenir en compte que les propietats dels fragments moleculars en un entorn proteic poden ser notablement diferents dels mateixos en solució aquosa, o en fase gas, fet que aquí queda evidenciat.

Amb tot, si el cromòfor no té aquest suposat increment exagerat d'acidesa a l'estat excitat (tot i que, efectivament, esdevé més àcid), com es pot explicar la transferència protònica? Donat que ni l'estructura del PDB ni el mínim més proper poden explicar aquesta situació caldrà, altre vegada, assumir una explicació dinàmica del procés.



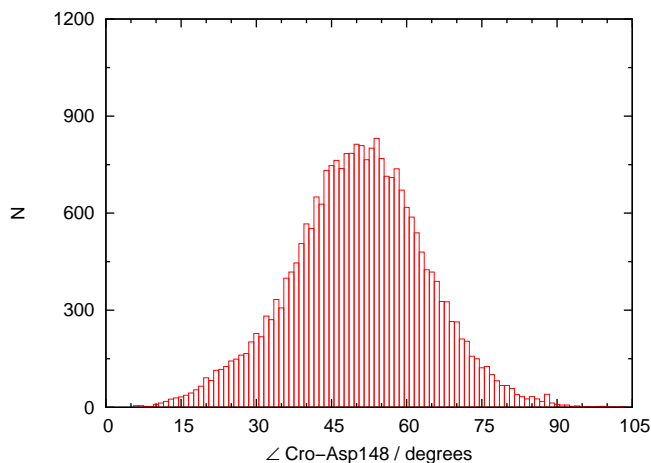


FIGURA 6.2: Distribució de l'angle que forma el grup carboxílic de l'Asp148 amb el pla del cromòfor durant el 1,0 ns de simulació QM/MM.

### 6.2.3 Dinàmica QM/MM

En aquest punt es recupera la simulació dinàmica a nivell QM/MM (OM3/CHARMM27) presentada en el Capítol 5 per a la doble mutant S65T/H148D. Recordem els valors ja discutits de la planaritat del cromòfor (el valor mig de l'angle díedre entre l'anell fenòlic i imidazolinònic del cromòfor és de  $6,3^\circ$ , vegeu la Figura 5.3) i de la distància entre donador i acceptor (el valor mig de la distància entre cromòfor i Asp148 és de  $2,35 \text{ \AA}$ , vegeu la Figura 5.2). Per a la discussió posterior, cal introduir l'anàlisi d'un altre paràmetre estructural, que és l'angle que formen el cromòfor i el grup carboxílic de l'Asp148 (Figura 6.2). Com es pot observar, també s'obté una distribució unimodal, aquest cop centrada a  $51,6^\circ$ .

Per tal d'analitzar si la descripció dinàmica aporta alguna novetat respecte a la descripció estàtica realitzada anteriorment a partir de l'estructura cristal·lina, i encara seguint amb la línia estadística que marca aquest treball, s'ha analitzat la factibilitat de la transferència protònica en les estructures obtingudes per mitjà de la simulació dinàmica, que és una representació molt més completa de la complexa realitat estructural de la proteïna. Per a tal fi, s'han pres 500 estructures a l'atzar (un fotograma cada 2 ps) de la simulació dinàmica QM/MM i, per cadascun d'ells, s'ha calculat el perfil de transferència protònica en els estats electrònics involucrats. Cadascun d'aquests perfils consta de 10 punts discrets, així que s'ha realitzat un total de 5000 càlculs puntuals per a tal anàlisi. A partir d'aquests perfils, s'ha realitzat l'anàlisi topogràfic pertinent per a determinar l'energètica de la transferència protònica a l'estat excitat. En aquest sentit, s'ha calculat l'ergicitat del procés de transferència protònica a l'estat fonamental ( $\Delta E$ ) com a la diferència d'energies entre l'estructura corresponent al cromòfor aniònic i l'estructura corresponent al cromòfor neutre, per a cada perfil. L'ergicitat de la transferència protònica a l'estat excitat ( $\Delta E^*$ ) s'ha calculat de manera anàloga.

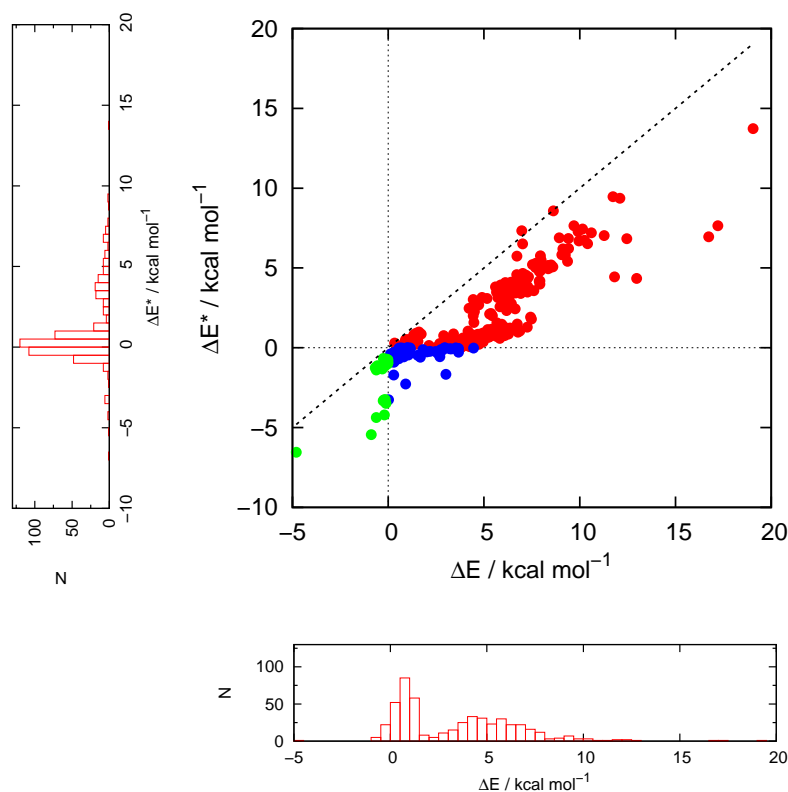


FIGURA 6.3: Anàlisi de l'energètica de la transferència protònica a l'estat fonamental i excitat fotoactiu, realitzada a partir dels 500 perfils per a la transferència entre el cromòfor i l'Asp148, obtinguts a nivell TD-DFT/CHARMM27. L'eix d'abscisses representa la exo- o endoergicitat del procés a l'estat fonamental ( $\Delta E$ ) per a cada perfil i el d'ordenades fa el mateix per a l'estat excitat ( $\Delta E^*$ ). Els punts vermells representen aquells perfils etiquetats com a "no reactiu" ( $\Delta E^* > 0$ ), els punts blaus aquells etiquetats com a "reactiu" ( $\Delta E^* < 0$ ) i els punts verds aquells en els quals, ja a l'estat fonamental, el protó es troba a l'Asp148 ( $\Delta E < 0$ ). Els histogrames que acompanyen l'eix d'abscisses i ordenades representen la distribució de cada valor ( $\Delta E$  i  $\Delta E^*$ ).

### L'Energètica de la Transferència Protònica

La Figura 6.3 mostra l'anàlisi realitzat per a l'energètica del procés de transferència protònica obtingut a partir de les 500 estructures provinents de la simulació dinàmica QM/MM. Es pot observar un comportament dràsticament diferent a l'obtingut amb l'estructura del PDB (Figura 6.1). L'histograma inferior de la Figura 6.3 representa la distribució de fotogrames respecte l'energètica de la transferència protònica a l'estat fonamental  $S_0$ ,  $\Delta E$ . Una gran quantitat de fotogrames mostren un procés gairebé isoèrgic, amb aproximadament la meitat d'estructures corresponent a situacions on  $\Delta E < 3$  kcal/mol. En aquest sentit, els efectes introduïts per la simulació dinàmica canvien totalment la visió de l'energètica del procés ja a l'estat fonamental: en l'estructura cristal·lina el protó estava fortament

l·ligat al cromòfor i hi havia una forta pujada d'energia per a la transferència, així que aquesta nova situació d'una certa isoergicitat ens pot portar a pensar que el protó estarà deslocalitzat entre donador i acceptor. Aquesta hipòtesi es veu reforçada pel fet que una petita (però considerable) part dels perfils tenen valors de  $\Delta E$  negatius (concretament, un 5,7 %). Això vol dir que en aquestes situacions el sistema es troba oscil·lant entre un cromòfor neutre i aniònic ja a l'estat fonamental. Aquests resultats són indicatius d'un canvi substancial respecte la situació en l'estructura cristal·lina i que, donat que la metodologia pel càlcul de l'energia QM/MM és la mateixa, aquest canvi ha de ser degut a que la simulació dinàmica ha portat al sistema a una zona de l'espai de configuracions significativament diferent al representat per l'estructura del PDB.

El dibuix a què porta l'estudi dinàmic mostra una situació de certa isoergicitat a l'estat fonamental. Amb diferències d'energia petites entre reactius i productes és possible que l'energia de punt 0 del protó ja quedi per sobre d'aquesta pujada d'energia i que, per tant, es pugui justificar el comportament observat per Stoner-Ma *et al.* respecte la impossibilitat de veure la protonació de l'Asp148 en els espectres de IR resolts en el temps, fet que es justificava, precisament, amb la proposta d'un protó deslocalitzat.<sup>87</sup> Els càlculs aquí presentats donen suport a aquesta hipòtesi. Amb això, posteriorment es discutirà el comportament quàntic del protó per a poder validar aquesta suposició.

També a l'estat fotoactiu, la situació és completament diferent de la trobada a l'estructura cristal·lina. Els càlculs mostrats anteriorment justificaven una baixada del  $pK_a$  del cromòfor a l'estat excitat, encara que una baixada molt més discreta de la que s'havia assumit fins ara a la literatura especialitzada. En aquest sentit, és esperable que trobem que la distribució de les energies  $\Delta E^*$  estigui desplaçada cap a valors més negatius respecte  $\Delta E$ . Això és exactament el que s'ha trobat a la Figura 6.3 (tots els punts es troben per sota de la diagonal). Els valors de  $\Delta E^*$ , que finalment determinaran la factibilitat de la transferència, estan distribuïts al voltant del 0, amb el  $\sim 70\%$  dels fotogrames trobant-se per sota de les 1,5 kcal/mol. Això vol dir que, la majoria del temps, el sistema es trobarà en una situació on, després de l'excitació, serà capaç de transferir el protó.

La qüestió que naturalment vindrà al cap del lector atent és: quins canvis estructurals ha comportat la simulació dinàmica QM/MM que hagin desembocat en aquesta situació energètica tant diferent? La resposta a aquesta pregunta està estretament lligada a l'increment de la factibilitat de la transferència protònica. De l'anàlisi de les dades estructurals provinents de la simulació MD, s'ha pogut identificar dos factors principals que influeixen en la PES per a la transferència protònica.

El primer canvi estructural important que té lloc està relacionat amb la Tirosina 145. Aquest residu, que apunta cap a l'interior del barril  $\beta$ , no sembla interaccionar amb cap dels actors principals de la transferència protònica en l'estructura cristal·lina, doncs es troba a 4,3 Å de l'oxigen

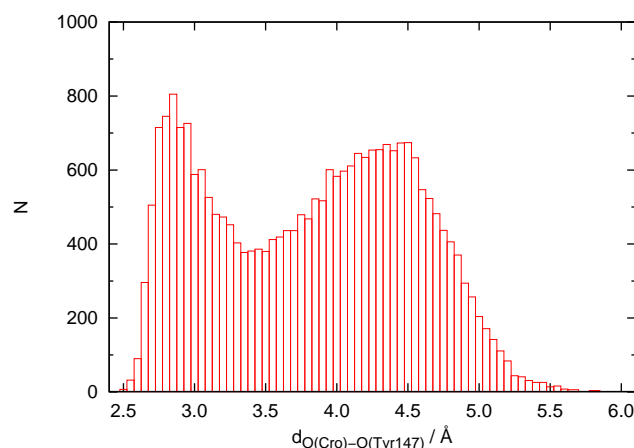


FIGURA 6.4: Histograma de la distribució de distàncies entre els oxígens fenòlics del cromòfor i de la Tyr145 durant el 1,0 ns de simulació dinàmica QM/MM.

fenòlic del cromòfor, que sosté el protó susceptible de ser transferit. Durant la simulació OM3/CHARMM22, emperò, aquest residu s'acosta i es manté a prop del cromòfor. La Figura 6.4 mostra la distribució de distàncies entre l'oxigen fenòlic del cromòfor i el de la Tyr145. Es pot observar una distribució bimodal on el pic que correspon a una clara interacció de pont d'hidrogen (amb màxim a  $\sim 2,8$  Å) integra aproximadament el 35% de l'àrea total del gràfic. És clar que aquesta interacció, que ja es dona amb el cromòfor neutre, establitzarà el fotoproducte, on l'oxigen fenòlic del cromòfor esdevindrà aniònic i que per tant és una interacció crítica per a la factibilitat del procés.

L'altre canvi estructural que s'ha identificat té a veure amb la solvatació del residu acceptor del protó, és a dir, de l'Aspartat 148. S'ha calculat la funció de distribució radial de les molècules d'aigua per als oxígens carboxílics de l'aspartat, a més dels oxígens fenòlics del cromòfor i de la Tyr145 (Figura 6.5). Com es pot observar, els oxígens de l'aspartat presenten un pic ben definit a  $r \sim 2,65$  Å, que indiquen una solvatació estable al llarg de la dinàmica per a aquest residu. Cal recordar que l'aspartat manté una orientació també estable al llarg de la dinàmica (Figura 6.2), fet que ajuda a la bona definició de la solvatació del mateix. Per a l'oxigen del cromòfor, s'observa una interacció menys intensa i per a la tirosina no es pot apreciar una solvatació estable clara al llarg del temps. La integració dels pics de la Figura 6.5 indiquen el nombre d'aigües interaccionant amb cada oxigen. Aquest càlcul dona per resultat:  $O_1(\text{Asp148}) = 1,26$  aigües,  $O_2(\text{Asp148}) = 1,10$  aigües,  $O(\text{Cromòfor}) = 0,92$  aigües i  $O(\text{Tyr145}) = 0,37$  aigües. En definitiva, un total de 2,36 aigües interaccionen amb l'Asp148 de mitjana. Aquesta solvatació contrasta amb la que es pot observar en l'estructura cristal·lina, on hi ha clarament 3 aigües interaccionant amb l'Asp148 a distància de pont d'hidrogen. Naturalment, que l'aspartat estigui menys solvatat durant la dinàmica que en l'estructura cristal·lina, el farà més bàsic en la primera situació que en la segona, tot fent possible la transferència

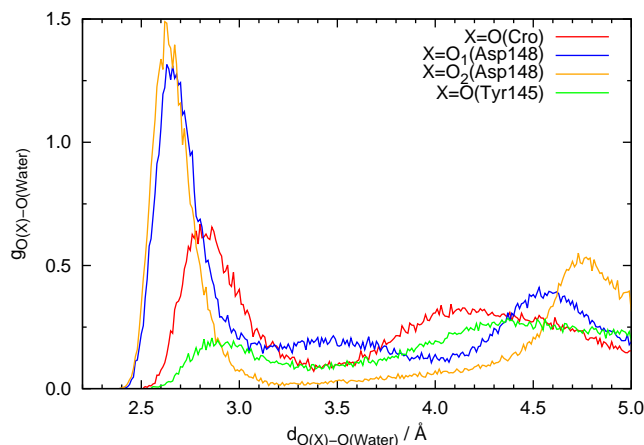


FIGURA 6.5: Funcions de distribució radial de les molècules d'aigua respecte els diferents oxígens indicats a la llegenda, obtingudes en la simulació QM/MM de la mutant S65T/H148D

protònica.

Així doncs, s'ha identificat dos factors estructurals que són crítics per a la transferència protònica: la interacció de la Tyr145 amb l'oxigen fenòlic del cromòfor, que estableix el fotoproducte i la solvatació de l'acceptor, el residu Asp148, que si és massa pronunciada, no permet la transferència per efecte de reduir la basicitat de l'esmentat residu.

### Perfils Dinàmics

S'ha realitzat l'anàlisi de la distribució de les energies QM/MM dels fotogrames obtinguts de la simulació dinàmica de la Figura 6.3. Com ja s'ha dit anteriorment, el 5,7% de les estructures descriuen un sistema amb un cromòfor aniònic ja a l'estat fonamental. El 64,8% dels perfils presenten una  $\Delta E^*$  positiva (alhora que petita) i per tant descriuen una situació on és més estable el cromòfor neutre encara a l'estat excitat. Tot i això, una gran proporció d'aquests perfils es troben a valors de  $\Delta E^*$  molt propers al 0, fet que indica que la endoergicitat del procés no és molt pronunciada. La resta de perfils, el 29,5 %, presenten una  $\Delta E^*$  negativa a l'estat excitat, i corresponen a situacions on la forma aniònica del cromòfor, des d'un punt de vista topogràfic, és més estable. Per tal de mostrar exemples concrets, a la Figura 6.6 es poden trobar 3 dels 500 perfils per a la transferència protònica calculats a nivell TDDFT/CHARMM27. S'hi mostra un perfil amb  $\Delta E^*$  positiva però propera al 0 ("no reactiu"), un amb  $\Delta E^*$  negativa i també propera al 0 ("reactiu"), i un amb  $\Delta E^*$  negativa i més pronunciada ("reactiu").

És remarcable que en els dos primers casos, els perfils són marcadament unimodals, amb un sol mínim que situa el protó a una zona més propera

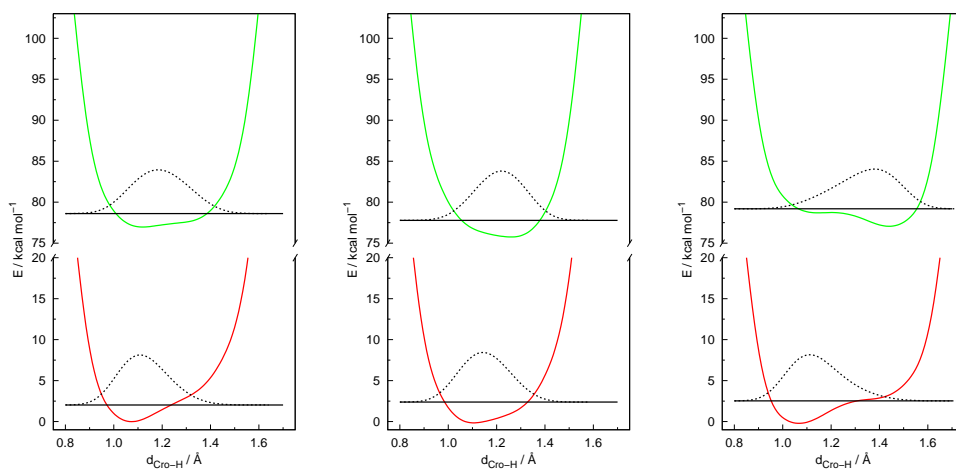


FIGURA 6.6: Perfils d'energia potencial per l'estat fonamental (vermell) i l'estat fotoactiu (verd) obtinguts a nivell DFT/CHARMM27 i TDDFT/CHARMM27, respectivament, a partir de 3 fotogrames seleccionats de la simulació dinàmica QM/MM. Al panell esquerra, es mostra un perfil típic "no reactiu" ( $\Delta E^* > 0$ ). Al panell central es mostra un cas típic "reactiu" ( $\Delta E^* < 0$ ). Al panell dret es mostra un perfil "reactiu" amb una  $\Delta E^*$  considerable. En vermell es mostren els perfils a l'estat fonamental i en verd el perfil a l'estat fotoactiu. Les línies horitzontals negres representen l'energia de punt 0 i les línies discontinues representen la densitat de probabilitat ( $|\psi|^2$ ) per a les funcions d'ona del protó.

al cromòfor que a l'aspartat a l'estat fonamental. Això indica que, efectivament, només existeix un mínim que descriu un protó certament deslocalitzat entre el cromòfor i l'Asp148. Això es pot corroborar quantitativament amb càlculs de QD. La resolució de l'equació d'Schrödinger independent del temps porta a l'obtenció de les energies de punt 0, que són de 2,0 kcal/mol pels dos primers casos i de 2,4 kcal/mol pel tercer. Aquests valors, juntament amb la funció de densitat de probabilitat  $|\psi|^2$ , també es troben representats a la Figura 6.6. Com es pot observar, la deslocalització del protó és evident a l'estat fonamental, fet que corrobora i explica el comportament observat per Stoner Ma. *et al.* respecte la manca d'evolució temporal dels modes de vibració de l'Asp148. Altre cop, una descripció teòrica de la situació permet una comprensió molt més complerta de la situació física. Cal remarcar que la inclusió del comportament quàntic del protó és primordial, també, per a la correcta descripció del fenomen que té lloc.

Per intentar escatir si hi ha algun factor estructural significatiu que afavoreixi els valors de  $\Delta E^*$  negatius, s'ha pres la població de tots els paràmetres estructurals en un radi partint del cromòfor pels perfils amb  $\Delta E^* > 0$  i  $\Delta E^* < 0$ . S'ha comparat estadísticament aquests paràmetres per veure si corresponen a poblacions diferents per mitjà d'un test *t* d'Student amb el 95% de confiança. El resultat d'aquests test és que no hi ha cap paràmetre que pertanyi a poblacions diferents entre els corresponents a  $\Delta E^*$  positives o negatives, per tant cal concloure que la diferenciació que es fa entre unes i altres situacions es deu a fluctuacions pròpies del sistema, accessibles per mitjà de la simulació dinàmica. Cal remarcar, altre cop, que la descripció dinàmica de la situació de la proteïna és absolutament imprescindible, ja que sense la mateixa no seria possible accedir a les fluctuacions tèrmiques que fan possible les situacions on la transferència ocorre.

### La Inclusió de l'Energia Cinètica del Protó

Un cop realitzat l'anàlisi topogràfic de la transferència (amb els perfils d'energia potencial), i haver introduït els efectes quàntics d'equilibri (amb la resolució de l'equació d'Schrödinger independent del temps) per a l'estat fonamental, manca quantificar el moviment del protó en la superfície d'energia potencial de l'estat excitat. Per fer tal cosa, s'ha realitzat càlculs de dinàmica quàntica seguint el procediment de simulació de fotoabsorció descrit a al Capítol 2, per als tres perfils presentats a la Figura 6.6. Primerament, havent trobat els vectors propis del Hamiltonià en els perfils de l'estat fonamental i excitat fotoactiu, s'ha simulat l'absorció Franck-Condon canviant l'estat electrònic mentre es manté la configuració nuclear (això és, canviat sobtadament la part potencial de l'operador Hamiltonià). Com que l'estat de vibració de l'estat electrònic fonamental ja no és vector propi del nou Hamiltonià, el primer evolucionarà d'acord amb l'equació d'Schrödinger depenent del temps. S'ha propagat el paquet d'ones del protó durant 150 fs. Per monitoritzar l'estat del sistema al llarg del temps,

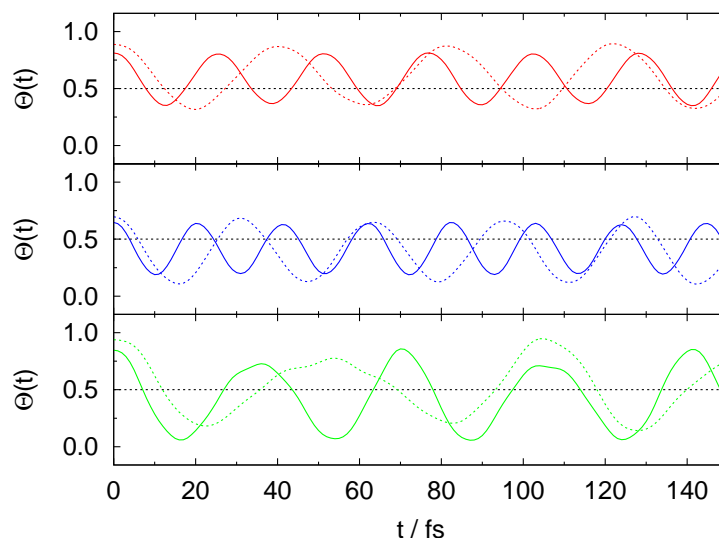


FIGURA 6.7: Representació de les funcions de probabilitat de supervivència en funció del temps, calculades sobre els mateixos perfils de la Figura 6.6. Les línies contínues representen la dinàmica del protó i les discontinues les del deuteró.

s'ha utilitzat la probabilitat de supervivència, que representa la fracció de la densitat de probabilitat que es troba a la zona de reactius (R), és a dir, la zona de l'espai de configuracions on el protó es troba més a prop del cromòfor que de l'Asp148:

$$\Theta(t) = \int_{q \in R} |\psi(q, t)|^2 dq \quad (6.4)$$

La funció de probabilitat de supervivència es pot interpretar com la probabilitat que, després de realitzar una mesura de la posició del protó, aquest es trobi en la zona corresponent a reactius.

Com es pot observar a la Figura 6.7, el comportament queda descrit com a un moviment periòdic de reactius a productes en el que es coneix com a oscil·lacions de Rabi.<sup>192</sup> El període d'aquestes oscil·lacions va de  $\sim 20$  a  $\sim 32$  fs, així que la meitat d'aquest temps és el que trigarà el protó a transferir-se. Cal notar que, fins i tot en els casos on la  $\Delta E^*$  és positiva (això és, una transferència "topogràficament inviable"), hi ha una transferència efectiva en menys de 20 fs, ja que més del 60% del paquet d'ones del protó es troba a la zona de productes. En els casos exoèrgics, naturalment, la transferència és molt més acusada, arribant a ser aquesta proporció del 80 o del 95%. Altre cop es fa palesa la necessitat d'una aproximació dinàmica per a la correcta descripció del procés: l'anàlisi del perfil endoèrgic faria pensar en una transferència inviable però la inclusió de l'energia cinètica juntament amb la simulació de la fotoexcitació ens porta a concloure que, en la gran majoria de casos estudiats, hi ha una transferència del protó efectiva a l'estat excitat. El temps de la transferència protònica està per



sota del límit de detecció experimental, així que concorda amb les mesures de Kondo *et al.*<sup>86</sup> A més, la substitució del protó per un deuteró en la simulació QD també explica l'absència d'efecte cinètic d'isòtop mesurable: encara que el deuteró es mou més lentament, es transfereix per sota del límit de detecció experimental, fet que no permet mesurar el KIE.

Cal comentar que la oscil·lació de Rabi predita pel càlcul QD no és en cap cas realista, doncs es deu a que el potencial és estable al llarg del temps (és a dir, que l'estructura es manté congelada). És evident que després de l'excitació el sistema evolucionarà per tal d'acomodar la nova distribució electrònica, fet que farà col·lapsar la funció d'ona del protó cap a la zona de productes. Les conclusions teòriques que es poden extreure del càlcul QD són que: (1) hi ha transferència protònica efectiva en la gran part dels perfils calculats (fins hi tot per aquells que presenten  $\Delta E^*$  positiva) i que (2) el temps de la transferència del protó i del deuteró està per sota del límit de detecció experimental, fet que ens permet donar una descripció atomística completa de la situació que porta a la transferència protònica.

## 6.2.4 Una Nova Mutant

En la secció anterior s'ha discutit la importància que té el residu Tyr145 en el procés de transferència protònica a l'estat excitat. La descripció positivista del procés assolida fins ara, emperò, ha de poder passar el test de la falsabilitat. Qualsevol coneixement científic ha de poder ser falsable, és a dir, ha de poder sobreviure a la contraposició del mateix a la proposició que sustenta. Precisament, Karl Popper va triar aquest criteri per separar allò que és ciència d'allò que no ho és.<sup>193</sup> Si un coneixement no pot ser falsejat, llavors no és coneixement científic. Que un coneixement o proposició sigui falsable vol dir que *pot* ser refutat per una nova observació, encara que si aquest coneixement és correcte, aquesta observació probablement no arribi mai. Per exemple, naturalment, englobar l'existència de Déu en un coneixement científic seria del tot erroni perquè no es pot provar la falsabilitat de l'assumpció d'aquesta existència. És impossible d'imaginar una observació que provés per ella mateixa que, eventualment, Déu no existeix. Per tant, l'existència de Déu no és falsable i no es coneixement científic. Com que en aquesta tesi es busca la creació de coneixement científic, cal que la hipòtesi plantejada en la secció anterior sobre la importància del residu Tyr145 sigui falsable, i ho és. Si es diu que la Tyr145 és un residu determinant per a la factibilitat de la transferència protònica a l'estat excitat, hom pot pensar que la mutació d'aquest residu hauria de tenir efectes sobre el comportament de la proteïna. Si el coneixement fins ara plantejat respecte la importància de la Tyr145 no és cert, la mutació Y145F no hauria de tenir més efectes. Com que la rellevància de la Tyr145 rau en la seva interacció de pont d'hidrogen amb el cromòfor, la substitució de la mateixa per una fenilalanina (que és isostèrica però no permet l'anomenada interacció), hauria de tenir conseqüències sobre el comportament de la proteïna.

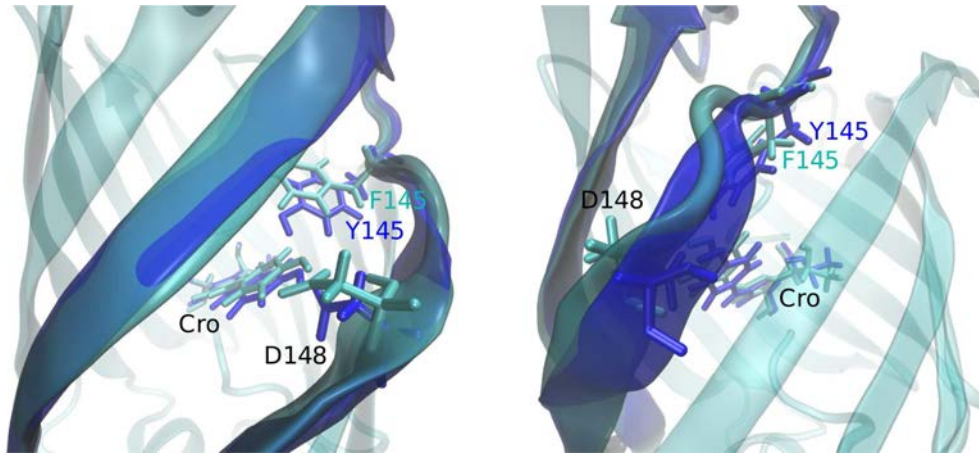


FIGURA 6.8: Aquí es mostra la superposició d'un fotograma representatiu de la simulació QM/MM per les dues mutants, S65T/H148D i S65T/H148D/Y145F, de color blau fosc i cian, respectivament. Al panell esquerre es pot observar com la Phe145 tendeix a estar més allunyada en la triple mutant que la Tyr145 en la doble. Això causa la distorsió de la fulla  $\beta$  que implica la coplanaritat de l'Asp148 amb el cromòfor. Com es pot observar, en el cas de la doble mutant, l'Asp148 roman essencialment perpendicular al mateix. Al panell dret, on es mostra el mateix fotograma però des d'un altre punt de vista, es pot apreciar l'anomenada distorsió de la fulla  $\beta$  causada per la pèrdua de la interacció de la Tyr145 amb el cromòfor.

Així doncs, a continuació es planteja un estudi realitzat sobre la triple mutant S65T/H148D/Y145F-GFP. En aquest estudi s'ha utilitzat una metodologia completament anàloga a la de la doble mutant S65T/H148D-GFP, sobre la qual s'ha aplicat la mutació Y145F.

### L'Efecte Papallona en una FP: Resultats Estructurals

La mutació Y145F comporta canvis notables en l'estructura de la proteïna. La distància entre cromòfor i l'Asp148 es manté estable en la triple mutant S65T/H148D/Y145F. Igual que les dobles mutants, manté una distribució unimodal, aquest cop centrada a 2,37 Å, valor lleugerament superior al trobat a la doble mutant S65T/H148D. Tot i això, aquest valor encara és prou petit com perquè sigui imaginable una transferència sense barrera, si aquesta és energèticament favorable.

El més sorprenent dels resultats de la simulació per a aquesta triple mutant no és en les similituds amb la doble mutant, si no en les diferències amb la mateixa. La més evident és que es perd la interacció de pont d'hidrogen entre la Tyr145 i el cromòfor present a la doble mutant, ja que la fenilalanina que substitueix aquest residu és hidrofòbica. Com que no hi ha possible interacció efectiva entre el cromòfor i la Tyr145 en la triple mutant, aquest residu tendeix a estar més allunyat de l'oxigen fenòlic.

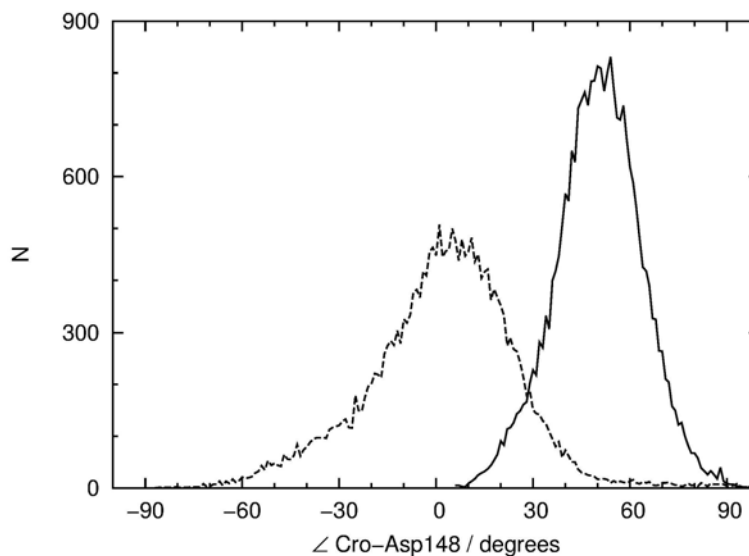


FIGURA 6.9: Distribució de l'angle diedre entre el pla del cromòfor i el grup carboxílic de l'Asp148 al llarg del 1,0 ns de simulació QM/MM per a les mutants S65T/H148D (línia contínua) i S65T/H148D/Y145F (línia discontinua).

La Tyr145 actua com a palanca sobre una de les fulles  $\beta$  de l'esquelet de la proteïna i, en adoptar aquesta nova posició, indueix una deformació de l'anomenada fulla  $\beta$ , com es pot veure al panell dret de la Figura 6.8. Aquest canvi de conformació de l'esquelet de la proteïna, al seu torn, produeix un canvi en la configuració espacial de l'Asp148 (l'acceptor del protó) que, encara que pugui semblar suau, tindrà conseqüències importants en l'energètica del procés.

Com es pot observar en la Figura 6.2, en la doble mutant, el grup carboxílic de l'Asp148 es troba clarament fora del pla del cromòfor. En el cas de la triple mutant, emperò, i degut a la deformació que produeix sobre l'esquelet de la proteïna l'allunyament de la fenilalanina respecte el cromòfor, el grup carboxílic de l'aspartat es situa en el mateix pla que el cromòfor. Això es pot observar numèricament a la Figura 6.9 i gràficament al panell esquerre de la Figura 6.8. Aquesta nova configuració de l'Asp148 comporta canvis en la solvatació del mateix, que recordem que era crítica per a la factibilitat de la transferència. S'ha calculat la funció de distribució radial de les molècules d'aigua respecte els dos oxígens carboxílics de l'Asp148 i també de l'oxigen fenòlic del cromòfor (Figura 6.10).

La integració d'aquests pics per a la primera esfera de solvatació mostra com, de mitjana, hi ha 1,0 molècules d'aigua interaccionant amb l'oxigen del cromòfor, 1,17 molècules interaccionant amb l'O1 i 1,93 amb l'O3. Això resulta en un total de 3,1 molècules d'aigua interaccionant amb l'aspartat respecte les 2,3 que interaccionaven amb el mateix en el cas de la doble mutant. A continuació es discuteix quin efecte tenen aquests canvis estructurals en l'energètica de la transferència protònica

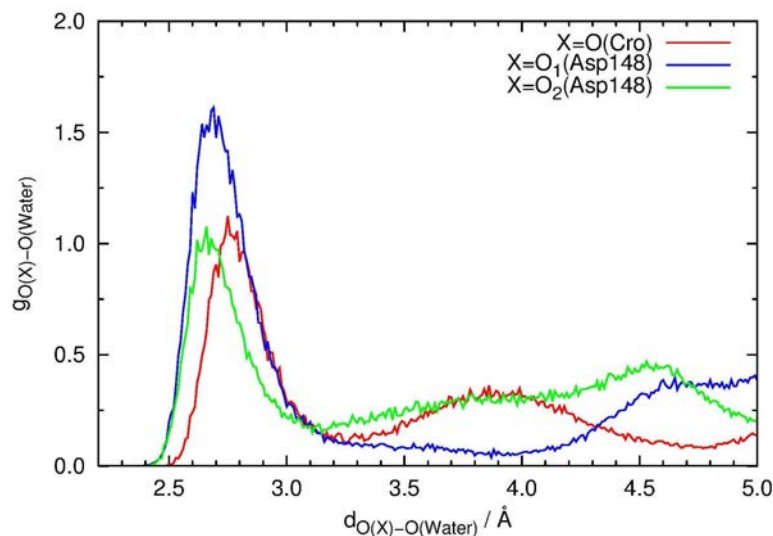


FIGURA 6.10: Funcions de distribució radial de les molècules d'aigua respecte els diferents oxígens indicats a la llegenda, obtingudes en la simulació QM/MM de la triple mutant.

### L'Energètica de la Transferència Protònica

L'increment de la solvatació de l'Asp148, combinat amb la pèrdua de d'interacció de pont d'hidrogen que es dona a la triple mutant, pot tenir efectes definitius en l'energètica del procés de transferència protònica. Per determinar aquest efecte quantitativament, s'ha repetit una estratègia idèntica a la utilitzada per a la doble mutant: s'ha pres 500 fotogrames del 1,0 ns de simulació QM/MM i s'ha calculat el perfil per a la transferència protònica a l'estat fonamental i excitat fotoactiu per cadascuna d'elles. Cada perfil s'ha obtingut a partir de 10 càlculs puntuals. Els resultats es mostren a la Figura 6.11.

Es poden apreciar canvis substancials respecte l'energètica de la doble mutant, que també es mostra a la Figura 6.11, per una millor comparació. Mentre a la doble mutant, recordem, al voltant del 30% de les estructures tenien  $\Delta E^* < 0$  i la gran majoria de perfils es trobaven en una situació d'una certa isoergicitat, en el cas actual de la mutant S65T/H148D/Y145F, només el 3% dels perfils tenen  $\Delta E^* < 0$  i la majoria dels mateixos són clarament endoèrgics. Altre cop, a tall il·lustratiu, s'ha seleccionat 3 dels 500 perfils calculats, que representen la situació típica (panell central), una d'especialment endoèrgica (panell esquerre) i una especialment exoèrgica (que resulta ser isoèrgica a l'estat excitat) i s'han representat a la Figura 6.12.

Com es pot observar, només és esperable una transferència efectiva en els casos més extrems d'isoergicitat per a la triple mutant. En el cas típic, la pujada d'energia és tant pronunciada que ni encara amb l'addició de l'energia cinètica al protó es pot assumir aquest increment. Cal recordar que aquests són perfils no relaxats, és a dir que no podem concloure d'aquestes

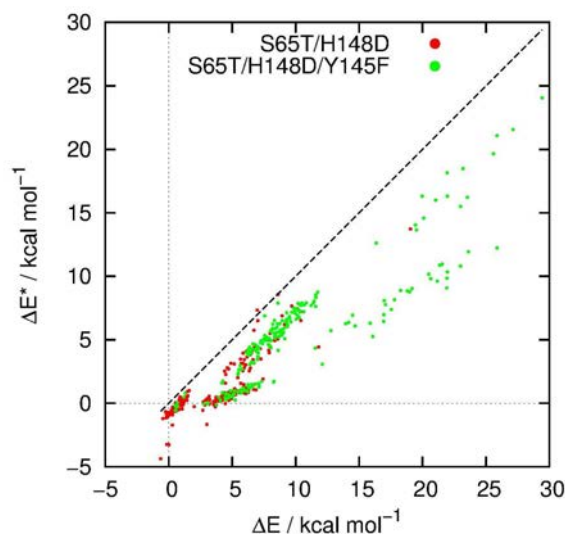


FIGURA 6.11: Anàlisi de l'energètica de la transferència protònica a l'estat fonamental i excitat fotoactiu, realitzada a partir dels 500 perfils per a la transferència entre el cromòfor i l'Asp148, obtinguts a nivell TD-DFT/CHARMM27, per a les mutants S65T/H148D (vermell) i S65T/H148D/Y145F (verd). L'eix d'abcises representa la exo- o endoergicitat del procés a l'estat fonamental ( $\Delta E$ ) per a cada perfil i el d'ordenades fa el mateix per a l'estat excitat ( $\Delta E^*$ ).

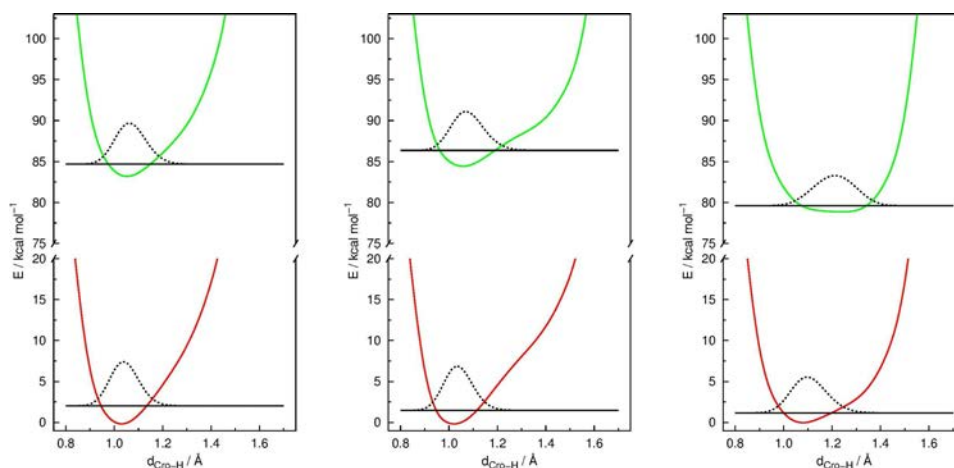


FIGURA 6.12: Perfils d'energia potencial obtinguts a nivell TDDFT/CHARMM27 a partir de 3 fotogrames seleccionats de la simulació dinàmica QM/MM per a la triple mutant. Al panell esquerre es mostra un perfil especialment endoèrgic, al panell central es mostra un cas típic, i al panell dret se'n mostra un d'especialment exoèrgic. Les línies horitzontals negres representen l'energia de punt 0 i les línies discontinues representen la densitat de probabilitat ( $|\Psi|^2$ ) per a les funcions d'ona del protó.

dades que no hi haurà transferència a l'estat excitat a temps prou llargs: podria ser que, després de la fotoexcitació, hi hagi un procés per acomodar la nova distribució electrònica que permeti la transferència. El que sí que es pot afirmar és que, segons aquestes dades, no és esperable una transferència ultrarràpida a l'estat excitat, com en el cas de la doble mutant. La triple mutant, un cop accedeixi a l'estat excitat, no podrà transferir el protó immediatament, així que l'eventual pujada de la fluorescència de l'espècie I\* hauria de ser mesurable, tot al contrari del que passa a la doble mutant. En el cas de la doble mutant, s'ha argumentat com una gran majoria de sistemes ja es troben en una situació susceptible de permetre la transferència protònica a l'estat excitat tot just després de la fotoexcitació, fet que no permet observar la pujada de l'espècie I\* amb el temps de resolució experimental (50 fs).<sup>86</sup>

D'altra banda, la resolució de l'equació d'Schrödinger independent del temps denota que en aquest cas, i contràriament a l'estudiat amb anterioritat, el protó es troba clarament localitzat a la zona de reactius, així que no s'espera el comportament observat en la doble mutant, propi d'un protó altament deslocalitzat. En el cas de la S65T/H148D/Y145F, si hi hagués transferència protònica a temps prou llarg gràcies al reordenament de l'entorn (conclusió inabastable amb l'estudi realitzat), es podria veure clarament un canvi en els modes de vibració associats al grup carboxílic de l'Asp148, ja que efectivament, el canvi en l'estat de protonació del mateix seria complet.

Així doncs, s'ha comprovat que el residu Tyr145 és necessari per a l'aparició ultrarràpida de la fluorescència corresponent a l'espècie I\*. Sense la interacció de pont d'hidrogen entre la Tyr145 i el cromòfor, i a causa de la deformació de l'esquelet de la proteïna, l'Asp148 canvia el seu estat de solvatació. Això, juntament amb la pèrdua de la pròpia interacció Tyr145-cromòfor, fa que l'energètica de la transferència ultrarràpida sigui molt menys favorable, tot desactivant-la.

Amb tot l'explicat fins ara, s'ha aconseguit donar una explicació atomística de les raons estructurals i electròniques que permeten que el protó es transfereixi del cromòfor a l'Asp148, tot explicant el fenomen en base dinàmica. Ara bé, què passa després de la transferència?

### 6.3 Química Ultrarràpida a Càmera Lenta: la Vida de la Doble Mutant a l'Estat Fotoactiu

A aquestes alçades, els únics detalls que queden per escatir respecte la proteïna fluorescent arquetípica són aquells relacionats amb la seva evolució temporal a l'estat excitat, després de la fotoexcitació. En el Capítol 5, i com a part d'un estudi més global que inclou altres proteïnes fluorescentes, s'ha caracteritzat l'excitació que porta la doble mutant S65T/H148D a l'estat fotoactiu. En aquest context, s'ha explicat la causa del corriment

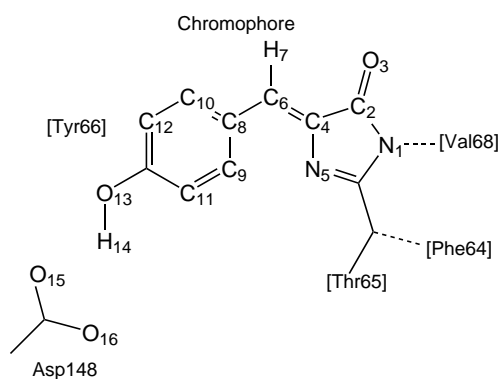


FIGURA 6.13: Etiquetatge dels àtoms que s'utilitzarà a la discussió que segueix.

cap al vermell de la banda d'absorció i s'ha donat detalls de la seva estructura d'equilibri en solució aquosa. En el present capítol s'ha explicat les raons estructurals i dinàmiques que expliquen la deslocalització del protó fenòlic del cromòfor així com la transferència ultrarràpida. S'ha mostrat les raons per les quals existeix una transferència gairebé instantània després de l'excitació que, en qualsevol cas, es troba per sota del límit de detecció experimental. Així, només queda per explicar quins són els processos físics i químics que ocorren després de la transferència protònica i que comporten la reorganització de l'entorn proteic al cromòfor o l'exploració de diferents vies no radiatives de desactivació. Per poder donar una explicació detallada d'aquests processos, s'ha propagat un paquet de 350 trajectòries QM/MM en un context no adiabàtic tal i com s'ha descrit en l'apartat de metodologia d'aquest capítol. Continuant amb la filosofia seguida fins ara en aquest treball, s'intentarà crear un model físic complet que doni un detall atomístic dels processos que tenen lloc. Si aquest model és, a més, capaç de reproduir les dades experimentals existents, podrem confiar en el mateix i en les conclusions que proporciona. Naturalment, les dades experimentals que permeten seguir l'evolució temporal del sistema i contra les quals s'ha contrastat els resultats obtinguts són els espectres de fluorescència resolts en el temps publicats per Kondo *et al.*<sup>86</sup> A fi i efecte de facilitar l'anàlisi estructural que es portarà a terme a continuació, a la Figura 6.13 es mostra l'etiquetatge d'àtoms que s'ha utilitzat en la discussió.

### 6.3.1 Canals de Relaxació no Radiativa

Abans d'analitzar l'espectre de fluorescència calculat a partir de les dinàmiques no adiabàtiques cal estudiar el comportament del sistema respecte l'eficiència de les vies de desactivació que ocorren pel trànsit a través d'una intersecció cònica (CI).

La Figura 6.14 mostra el nombre de sistemes que es transfereixen de manera no radiativa a l'estat fonamental i que, per tant, no són computades

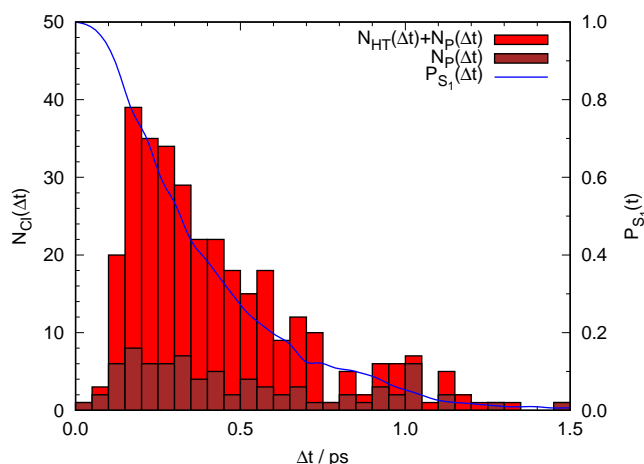


FIGURA 6.14: En vermell, representació del nombre total de trajectòries que es transfereixen de manera no adiabàtica des de l'estat excitat fotoactiu al fonamental via la HT-CI ( $N_{HT}$ ) o la P-CI ( $N_P$ ). Superposada, en marró, el nombre de trajectòries que es transferixen via la P-CI. La diferència entre les capses vermelles i marrons és la quantitat de sistemes que es transfereixen per mitjà de la HT-CI. En blau, es representa la funció fracció de població de l'estat excitat fotoactiu en funció del temps.

en el càlcul de l'espectre de fluorescència (a partir del moment de transició). Cal notar com, contràriament al comportament experimental (el rendiment quàntic de la proteïna és de 0,21), la totalitat dels sistemes es desactiven per via no radiativa durant els 1,5 ps de simulació. Naturalment, això denota que el sistema i mètode de càlcul emprats descriuen un accés a les vies de desactivació massa eficients. Aquest punt serà discutit amb posterioritat.

S'ha analitzat l'estructura de les proteïnes en el moment de la transició i s'ha pogut identificar dos tipus de CIs. En la Figura 6.15 es mostra un exemple representatiu de cadascuna d'elles. La primera (HT), que representa la via de desactivació majoritària, comporta una torsió al voltant del doble enllaç  $C_4C_6$ , que porta l'hidrogen metínic  $H_7$  fora del pla del cromòfor. Aquest tipus de desactivació ja ha estat prèviament estudiada per a cromòfors tipus GFP<sup>54-56</sup> i, en el cas que ens ocupa, comporta el 90% de les transicions no radiatives a l'estat fonamental.

Aquí és convenient recordar el treball realitzat per Zhang *et al.* sobre les vies de relaxació no radiatives del cromòfor de la S65T/H148D-GFP.<sup>194</sup> En aquell treball, els autors subdividien la coordenada en dues etapes: una primera on exploraven la formació del producte (la ionització del cromòfor per la transferència del protó a l'Asp148) i una segona on el cromòfor accedia a una CI que també comportava la rotació de l'enllaç  $C_4C_6$ . Aquest procés de torsió no tenia barrera si el càlcul es realitzava al buit, però presentava una barrera de 3,2 kcal/mol dins de l'entorn proteic. La similitud amb la CI trobada per mètodes dinàmics en aquest treball és evident, però



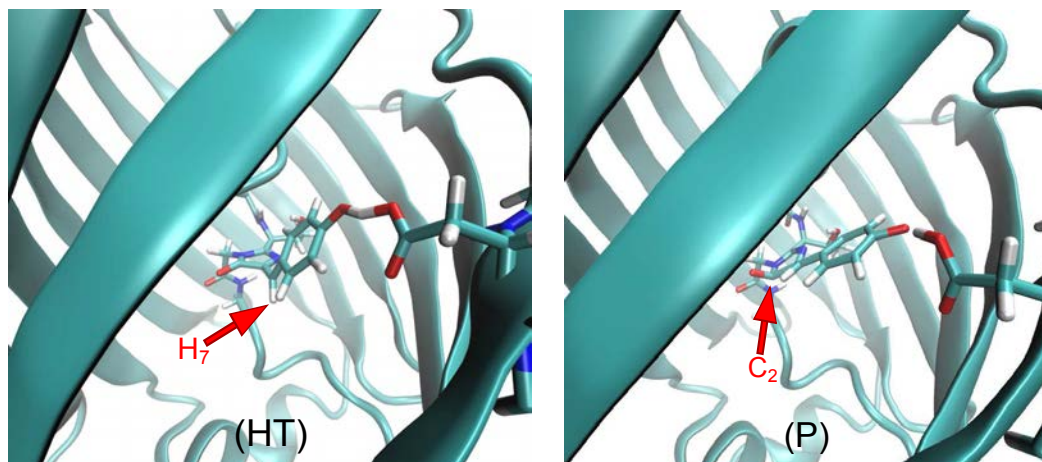


FIGURA 6.15: Representació gràfica de l'estructura de les dues CIs que porten a l'estat fonamental de manera no radiativa: (HT), que comporta la rotació de l'enllaç  $C_4C_6$  amb la torsió del  $H_7$  i (P), que comporta la piramidalització del  $C_2$ .

hi ha una diferència important a ressaltar: en el cas que presentava Zhang *et al.*, la CI era de tipus "producte" mentre que aquí, la CI és de tipus "reactiu". Aquesta diferenciació reactiu/producte fa referència a la definició clara del producte aniónic. En el cas presentat per Zhang *et al.*, la distància entre el cromòfor i l'Asp148 és significativament major, fet que permet una definició més clara del producte (recordem que a distàncies curtes entre donador i acceptor, s'accentua el comportament unimodal del perfil d'energia potencial per a la transferència protònica). Aquesta major definició del producte en el cas estudiat per Zhang *et al.* és natural ja que prèvia a l'exploració de la coordenada de la desactivació, com s'ha explicat, exploraven la coordenada que porta de reactius a productes. Ja que la CI trobada en aquest treball i la mencionada amb anterioritat només es diferencien en la definició del producte, es suggereix que la coordenada de transferència protònica i de relaxació de l'entorn són ortogonals a la que porta a la desactivació per la intersecció cònica que aquí es discuteix. Això voldria dir que la CI trobada per Zhang *et al.* i la trobada en aquest treball es troben en el mateix *seam*. La raó per la qual aquí no s'ha trobat cap CI amb una definició clara de productes és perquè, amb el temps propagat de dinàmica, no es pot aconseguir la relaxació estructural necessària per a la definició dels mateixos

El segon tipus de CI (P) representa al voltant del 10% de les transicions no radiatives a l'estat fonamental (vegeu la Figura 6.14). En aquesta CI, el carboni cetònic de l'anell d'imidazolinona apareix piramidalitzat, ja que l'oxigen que suporta surt del pla del cromòfor. No s'ha pogut trobar cap estudi que descrigui l'existència d'aquest tipus de còniques en els cromòfors tipus GFP en la literatura especialitzada.

La distribució d'aquestes dues CIs es pot monitoritzar durant la simulació MD a través dels angles díedres  $H_7C_6C_4N_5$  per a la tipus HT i  $C_4C_2O_3N_1$

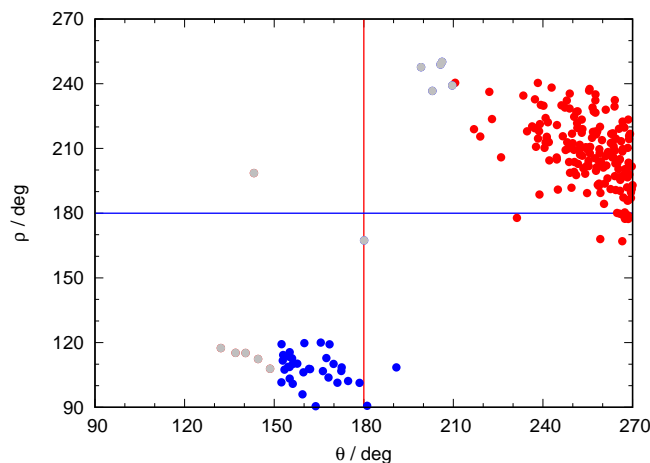


FIGURA 6.16: Valors dels angles  $\theta$  i  $\rho$  en el moment de la transició a l'estat fonamental. Els punts vermells corresponen a estructures tipus HT-CI i els punts blaus a estructures tipus P-CI. Els punts grisos corresponen a estructures que tenen un caràcter mixt.

per a la tipus P (angles  $\theta$  i  $\rho$ , respectivament). Mentre la trajectòria es manté lluny d'aquestes CIs, els valors de  $\theta$  i  $\rho$  es distribueixen al voltant dels  $180^\circ$ . Als voltants de la HT-CI s'hauria d'esperar  $\theta = 270^\circ$  mentre que al voltant de la P-CI caldria esperar que  $\rho = 120^\circ$ . A la Figura 6.16 es pot observar el valor d'aquests angles en el moment de la transició per a cadascuna de les trajectòries. Com es pot observar, en la gran majoria de casos, la distinció entre les CIs tipus HT o P és clara així que es suggereix que els camins que porten a tals vies de desactivació estan essencialment desacoblats.

Amb això, encara cal discutir l'excés d'efectivitat de les vies de desactivació no radiatives en la simulació que aquí es presenta. Un cop que el sistema arriba a una zona propera a la CI, i donat que aquesta causa un pou atractiu a l'estat excitat, la probabilitat que hi hagi una transició electrònica és elevada, si es deixa passar prou temps. Així doncs, la manera més intuïtiva d'explicar una menor eficiència de qualsevol CI, és en base a la presència d'una barrera energètica, que freni el procés. Per tant, el fet que en la simulació realitzada hi hagi una efectivitat exagerada dels canals de desactivació no radiatius, segurament és conseqüència d'una mala descripció de les barreres que porten a les CIs.

Per tal de comprovar si la metodologia utilitzada per propagar la dinàmica pateix d'aquesta limitació, s'ha realitzat càlculs de coordenades de reacció QM/MM que exploren la coordenada que defineix l'angle diedre  $\theta$ , que comprèn la rotació  $H_7C_6C_4N_5$ . S'ha realitzat la mateixa coordenada a nivell MRCISD-OM2/CHARMM27 i a nivell TD-DFT/CHARMM27, prenent aquest últim com a mètode de referència. Encara que el TD-DFT és un mètode monoreferencial que formalment no és capaç de localitzar CIs, s'ha demostrat que dona bones descripcions de les coordenades que

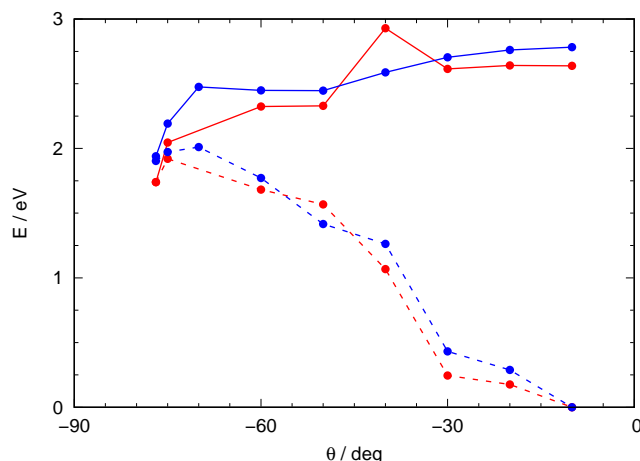


FIGURA 6.17: Coordenada de reacció a l'estat excitat fotoactiu calculada a nivell QM/MM per a la torsió del díedre  $H_7C_6C_4N_5$  ( $\theta$ ). Les línees contínues corresponen a l'estat excitat fotoactiu (on s'ha optimitzat les estructures) i les discontinuades a l'estat fonamental. Les línees blaves corresponen al càlcul MRCISD-OM2/CHARMM27 i les vermelles al càlcul TD-DFT(CAM-B3LYP)/CHARMM27.

porten a tals CI, ahora que en permet una localització aproximada prou bona de les mateixes, tot evadint els costos computacionals d'una aproximació més formalment correcta al problema.<sup>195-200</sup>

Els resultats es mostren a la Figura 6.17. Essencialment, es pot observar com a nivell MRCISD-OM2 no apareix cap barrera per a la rotació de l'angle  $\theta$  mentre que, a nivell TD-DFT es preveu una barrera d'unes 6 kcal/mol per a tal rotació. Cal apuntar que aquest valor és només significatiu a nivell qualitatiu. La coordenada que aquí es presenta és només un camí de mínima energia dels molts possibles. Cadascun dels mínims presents a l'espai de configuracions que tenen un pes significatiu en la representació de l'estructura a l'estat fonamental tindrà una coordenada diferent que el connecti amb la CI que aquí es tracta i, per tant, una barrera diferent. S'ha trobat un altre camí, que parteix d'una estructura diferent amb una barrera de 2,5 kcal/mol a nivell TD-DFT però altre cop, sense barrera a nivell MRCISD-OM2. La manera més formalment correcte de tractar aquest accés a la CI seria fer un eixam de coordenades i llavors realitzar un anàlisi estadístic sobre els resultats (semblant al procediment seguit en l'apartat 6.2.3 quan es parlava de la factibilitat de la transferència protònica), però aquest és un càlcul del tot inassumible a nivell TD-DFT/CHARMM27. Tot i que no es prendrà cap valor dels trobats (2,5 o 6 kcal/mol) com a valor de la barrera en termes absoluts, sí que es pot concloure que a nivell TD-DFT hi ha una barrera per a la rotació de l'angle  $\theta$  (que dona accés a la HT-CI) de l'ordre de les poques kcal/mol, mentre que a nivell MRCISD-OM2 aquesta barrera no existeix o és molt menor. Una exploració de la coordenada que porta a la P-CI ha resultat donar una barrera de 7 kcal/mol a nivell MRCISD-OM2/CHARMM27, fet que explica la menor eficiència

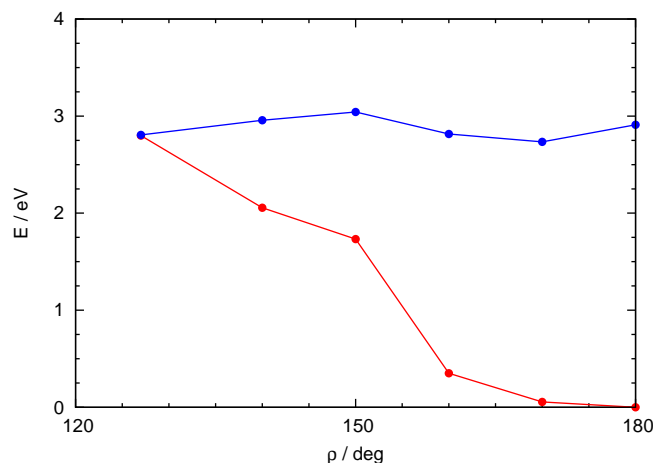


FIGURA 6.18: Coordenada de reacció a l'estat excitat fotoactiu calculada a nivell MRCISD-OM2/CHARMM27 per a la torsió del díedre  $C_4C_2O_3N_1$  ( $\rho$ ). La línia vermella correspon a l'energia de l'estat fonamental i la línia blava a l'energia de l'estat excitat fotoactiu

d'aquesta via de desactivació en les simulacions dinàmiques. Aquest resultat es mostra a la Figura 6.18. De fet reforça la idea que si el mètode MRCISD-OM2/CHARMM27 donés una barrera no menyspreable per a la HT-CI (tal i com ho fa per a la P-CI), la primera via de desactivació veuria molt minvada la seva eficiència.

Com s'ha dit, el mètode MRCISD-OM2 falla en la descripció d'una barrera que porta a la HT-CI, fet que fa exagerar l'eficiència d'aquesta via de desactivació no radiativa, però el tipus de dinàmiques no adiabàtiques realitzades a nivell QM/MM en sistemes biològics només són accessibles computacionalment a aquest nivell de càlcul. El fet que l'eficiència de la via de desactivació principal estigui exagerada emperò, no embruta els resultats que es presenten a continuació. Com s'ha explicat amb anterioritat, la coordenada de transferència protònica, la (complexa) coordenada de relaxació estructural i la coordenada que porta a la HT-CI són ortogonals (no des d'un punt de vista estricte, però si aproximat). Això vol dir que la descripció de la transferència protònica i de la relaxació estructural consegüent quedarà poc embrutada per l'alta eficiència de la HT-CI. El problema més greu que presenta aquesta alta eficiència és que no permet la propagació de simulacions a temps gaire llargs, però val a dir que tampoc és l'objectiu d'aquest treball, a més de ser un objectiu computacionalment inviable en l'actualitat. Per fer-ne una imatge gràfica, imaginem un camí de muntanya que passi per una angosta carena i que sigui estret i perillós, però que estigui limitat per un parell de baranes, a banda i banda del mateix. Volem explorar aquest camí i hi enviem excursionistes. Aquests excursionistes podran fer ús de les baranes i arribaran, probablement (sempre en pot relliscar algun), fins al final del camí o tant lluny com vulguin. Aquesta seria la situació amb la barrera correctament descrita. Ara imaginem que traiem les baranes del camí. Si fem això, el camí

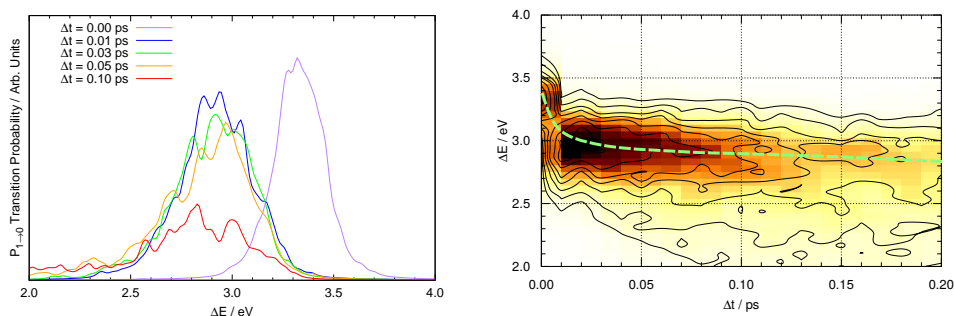


FIGURA 6.19: Panell esquerre: Espectres de fluorescència calculats a diferents temps indicats ( $\Delta t$ ) després de la fotoexcitació. Panell dret: Evolució temporal de la fluorescència on els punts més foscos representen més intensitat. La línia discontinua de color verd segueix aproximadament el màxim de fluorescència (la moda, no la mitjana, de la distribució).

a recórrer no queda modificat. L'únic que passarà és que, malauradament, molts dels excursionistes cauran per la carena i no podran explorar el camí tant com voldrien, encara que el camí que explorin serà el mateix que explorarien si hi hagués les baranes instal·lades. En el nostre cas, busquem explorar el primer tram del camí, especialment el que passa als primers centenars de femtosegons de dinàmica, i aquest és un tros de camí encara accessible sense barreres.

### 6.3.2 Espectres de Fluorescència

L'espectre de fluorescència resultant en el temps calculat per a la doble mutant S65T/H148D-GFP es mostra a la Figura 6.19. Cal remarcar, per evident que sigui, que la línia marcada com a  $\Delta t = 0$  ps equival a la banda A d'absorció, calculada amb anterioritat en aquest treball.

És evident en un primer cop d'ull com hi ha una clara evolució temporal de l'espectre d'absorció en dos temps. Primerament s'observa un desplaçament brusca de l'energia d'excitació seguit d'una evolució més suau, que sembla que fa eixamplar la banda d'absorció. Aquest comportament recorda a l'observat per Kondo *et al.*<sup>86</sup> Per analitzar la presència dels dos règims descrits per Kondo *et al.*, s'ha representat l'espectre d'absorció bidimensional per als primers 200 fs de simulació després de la fotoexcitació (panell dret de la Figura 6.19). El camí seguit pel màxim d'absorció és indicatiu: primerament hi ha una baixada forta de l'energia d'excitació (d'uns 0,3 eV en uns 30 fs) que continua disminuint a un ritme molt inferior (sobre uns 0,1 eV cada 100 fs), alhora que s'observa un eixamplament de la banda per la part de baixes energies. Ja que el comportament de dos règims observat per Kondo *et al.* queda ben descrit per la simulació que aquí es presenta, s'ha representat també l'evolució temporal de l'energia d'emissió mitjana. Aquest resultat es presenta a la Figura 6.20, i concorda qualitativament altre cop amb els resultats experimentals.

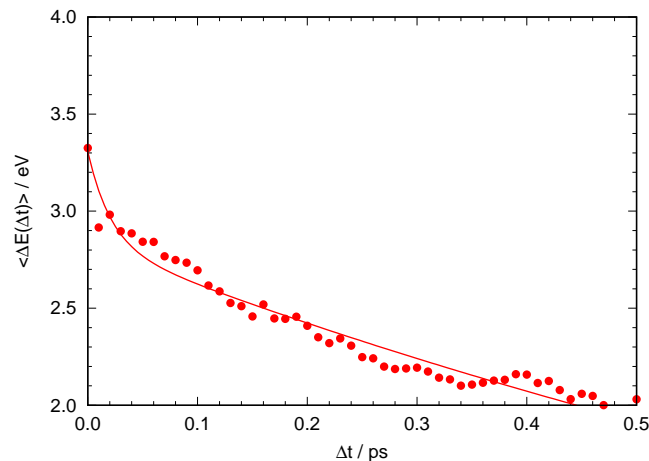


FIGURA 6.20: Evolució temporal de l'energia d'emissió mitjana. La línia contínua és una funció biexponencial ajustada.

Un altre resultat experimental que es presentava al treball de Kondo *et al.* és l'evolució temporal de la intensitat de fluorescència a determinades energies d'emissió. Kondo *et al.* excitaren la seva mostra a 415 nm i monitoritzaren la fluorescència a diferents longituds d'ona, des dels extrems més i menys energètics. En aquell treball s'explicava com la fluorescència a l'extrem vermell de la banda gairebé no mostrava cap decaïment, mentre en l'extrem blau es podia detectar un ràpid decaïment, en una escala de temps dels pocs centenars de fs. Aquest resultat es representa a la Figura 6.21. Per tractar de reproduir aquest gràfic, s'ha seleccionat quatre valors en el rang de l'a banda d'emissió calculada. S'han pres els valors de 2,6 eV, 3,0 eV, 3,1 eV i 3,2 eV, que son equivalents als valors presos experimentalment. Llavors s'ha representat la intensitat al llarg del temps per la emissió a aquestes energies. El resultat es pot veure a la Figura 6.22. Tal i com passa al cas experimental, la intensitat corresponent a l'extrem roig de l'espectre presenta poc decaïment, mentre a mesura que ens apropem a valors de l'extrem blau, el decaïment és més exagerat. Cal afegir que, donada la poca estadística que s'ha pogut acumular tot i el gran esforç computacional, hi ha una certa oscil·lació al voltant d'un valor mig, que s'interpreta com a soroll estadístic. Aquest resultat concorda prou bé altre cop, a nivell qualitatiu, amb el presentat experimentalment per Kondo *et al.*

En aquest punt es pot concloure que el model emprat en aquest treball aconsegueix reproduir satisfactòriament des d'un punt de vista qualitatiu els espectres resoltos en el temps obtinguts experimentalment per a la mutant S65T/H148D-GFP. Així doncs, com s'ha fet anteriorment, es pot concloure que, amb tota la precaució necessària i tenint en compte les pròpies limitacions, el model aquí utilitzat aconsegueix captar la física intrínseca del problema tractat. Tenint en compte això, es fa possible un anàlisi de les raons estructurals del fenomen observat, que al cap i a la fi, és l'objectiu d'un treball teòric (no la sola reproducció de les dades experimentals), ja que és la parcel·la de coneixement inaccessible, ara per ara, per mitjà de



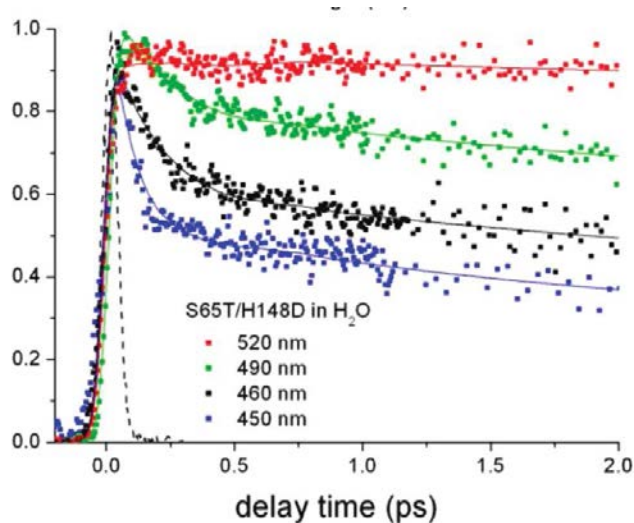


FIGURA 6.21: Intensitat de la flourescència resolta en el temps a diferents longituds d'ona seleccionades. Extret de Kondo *et al.*<sup>86</sup>

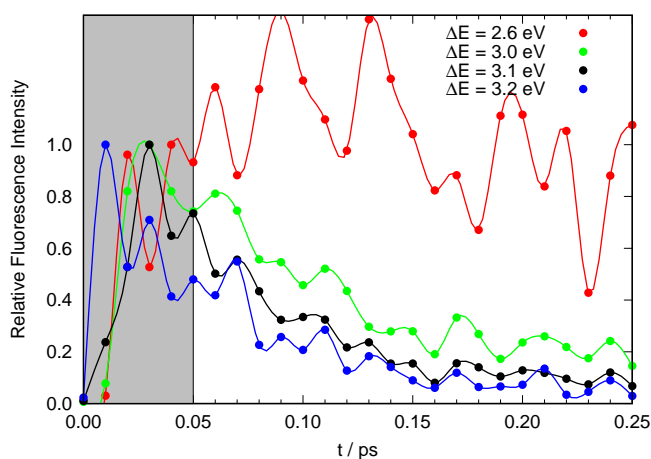


FIGURA 6.22: Simulació de la intensitat de fluorescència resolta en el temps publicada per Kondo *et al.*<sup>86</sup> El bloc gris representa la fracció de temps fora del límit de detecció experimental.

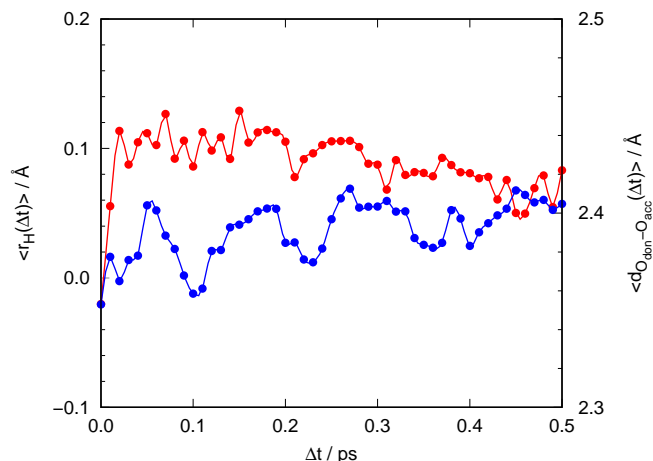


FIGURA 6.23: Evolució temporal de la distància del protó al centre de la distància donador-acceptor (Equació 6.5, vermell) i de la distància entre el cromòfor i l'Asp148 (blau)

mètodes experimentals.

### 6.3.3 La Transferència del Protó

Primerament es parlarà del primer procés químic que té lloc després de la fotoexcitació: la transferència protònica. Per tal d'analitzar la posició global del protó i la seva evolució al llarg del temps, es defineix la següent coordenada:

$$r_H = d_{O_{13}-H_{14}} - \frac{1}{2} \min(d_{O_{13}-O_{15}}, d_{O_{13}-O_{16}}) \quad (6.5)$$

que representa la distància entre el protó i el punt intermedi entre el donador i l'acceptor més propers (doncs l'Asp148 té dos possibles acceptors). La tria d'aquesta coordenada es fa perquè, independentment de la distància entre donador i acceptor, sempre es pot saber si l'estructura pertany a reactius ( $r_H < 0$ ) o productes ( $r_H > 0$ ).

La Figura 6.23 mostra l'evolució del valor mig d'aquesta coordenada al llarg del temps,  $\langle r_H(t) \rangle$ . S'ha d'apuntar que la magnitud  $\langle r_H(t) \rangle$  correspon al valor mig clàssic de la posició del protó. Encara que fer servir aquest valor per definir la posició del protó no és formalment correcte ja que no es té en compte la seva inherent naturalesa quàntica, com que els perfils de transferència protònica generalment no tenen barrera, no haurien d'aparèixer efectes d'interferència remarcables que poguessin afectar al valor mig de la posició. Com es pot observar, hi ha una transferència clàssica del protó que finalitza en aproximadament 20 fs (tal i com s'havia predit anteriorment amb càlculs de QD, vegeu la Figura 6.7). Després d'això el valor de  $\langle r_H(t) \rangle$  decreix a poc a poc. Això es pot interpretar com que el protó s'acosta al cromòfor (tot i encara mantenir-se transferit) o com que



la distància donador-acceptor augmenta. Aquesta distància també s'ha representat a la Figura 6.23 i es pot veure com aquest és el cas: després de la transferència, la distància entre cromòfor i l'Asp148 augmenta. Aquest és el moviment també predit per Zhang *et al.*<sup>194</sup> en l'exploració de la coordenada de reacció per a la transferència protònica prèvia a la desactivació i que aquí s'ha aconseguit reproduir de manera no condicionada, a partir d'una perspectiva dinàmica.

Així doncs, després de l'excitació, el moviment ultrarràpid del protó fa col·lapsar la funció d'ona electrònica a l'estat aniònic del cromòfor, fet que fa disminuir, també de manera notable, l'energia d'excitació. Per tant, s'associa directament la ultrarràpida disminució de l'energia d'excitació del cromòfor (per sota del límit de detecció experimental), a la també ultrarràpida transferència del protó des del cromòfor a l'Asp148. Naturalment, i donat que els perfils per a la transferència protònica no comporten baixades exagerades d'energia, el protó oscil·larà entre cromòfor i aspartat, com a mínim fins que la definició del fotoproducte no faci localitzar la funció d'ona del protó. Amb tot, parlant de posicions mitjanes, es pot parlar, efectivament, de transferència protònica ultrarràpida efectiva a l'estat excitat.

### 6.3.4 La Coordenada Complexa: la Relaxació de l'Entorn

Després del fort i sobtat decaïment de la l'energia d'excitació, hi ha una disminució d'aquest paràmetre que ocorre en un ritme molt més lent (uns 0,1 eV cada 100 fs). Els estudis experimentals parlen de *vibrational relaxation* però, quins són els moviments moleculars que hi ha darrera aquesta relaxació?

Fins ara ja s'ha exposat com la distància entre el cromòfor i l'aspartat creix monòtonament des del principi de la simulació, un moviment predit des d'estudis estàtics per Zhang *et al.*<sup>194</sup> Cal apuntar respecte la qualitat quantitativa de la simulació d'aquest moviment, i donades les limitacions del model emprat (una esfera d'aigües explícites amb una capa d'aigües congelades a l'exterior), que aquest allunyament quedarà segurament infravalorat respecte a una eventual simulació no tant constreta, com podria ser una PBC, que aquí no s'ha pogut assumir computacionalment.

Pel què fa a la resta de moviments que formen aquesta coordenada de relaxació, a la Figura 6.24 es presenten l'evolució temporal dels paràmetres  $\tau$  (que és l'angle díedre entre els àtoms C<sub>8</sub>C<sub>6</sub>C<sub>4</sub>C<sub>2</sub>) i  $\phi$  (que és l'angle díedre entre els àtoms C<sub>10</sub>C<sub>8</sub>C<sub>6</sub>C<sub>4</sub>), que representen la planaritat del cromòfor i  $\varphi$  (que és l'angle díedre entre els àtoms O<sub>16</sub>O<sub>15</sub>C<sub>11</sub>C<sub>12</sub>) que representa la coplanaritat entre el grup carboxílic de l'Asp148 i el cromòfor. Com es pot observar, mentre l'aspartat s'allunya del cromòfor, canvia lleugerament la seva orientació, respecte el mateix. Amb això, el canvi més important detectat té relació amb la coplanaritat dels dos anells del cromòfor. Mentre que l'angle  $\phi$  pateix un cert decreixement però es manté més o menys

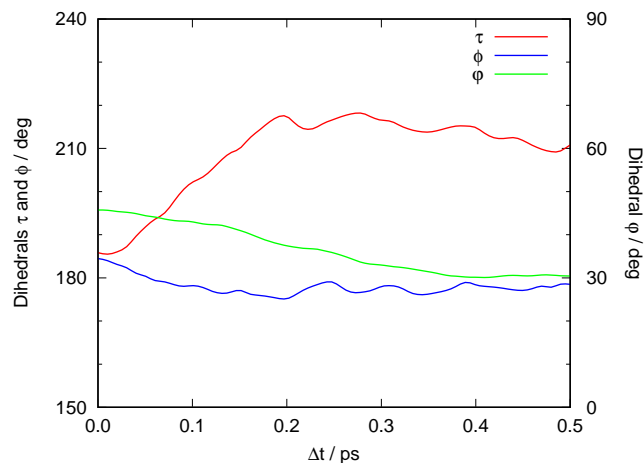


FIGURA 6.24: Evolució temporal dels angles  $\tau$  (díedre  $C_8C_6C_4C_2$ , vermell),  $\phi$  (díedre  $C_{10}C_8C_6C_4$ , blau) i  $\psi$  (díedre  $O_{16}O_{15}C_{11}C_{12}$ , verd).

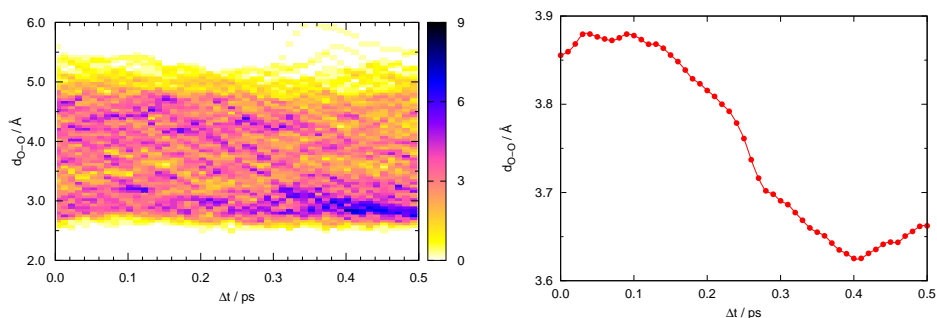


FIGURA 6.25: Panell esquerre: histogramització de l'evolució temporal de la distància entre l'oxigen fenòlic del cromòfor i de la Tyr145 al llarg de la dinàmica. Panell dret: evolució temporal de la mitjana d'aquest valor.

estable, l'angle  $\tau$  canvia bruscament en uns 200 fs fins a arribar a valors d'uns  $220^\circ$ , on s'estabilitza. Cal apuntar que s'ha suggerit que l'estat excitat del cromòfor de la GFP presenta una conformació menys plana, fet que facilitaria l'accés a una via de desactivació no radiativa.<sup>56</sup> Si s'analiza els orbitals moleculars involucrats en l'excitació, presentats a la Figura 5.6, es pot observar l'aparició d'un pla nodal del sistema  $\pi$  en l'enllaç  $C_6C_4$ , el que significa que, després de l'excitació, l'enllaç en qüestió queda sens dubte debilitat, fet que facilita la rotació del mateix. En termes energètics, aquest moviment desestabilitza més fortament l'estat fonamental (on implica la rotació d'un enllaç doble) que l'estat excitat (on implica la rotació d'un enllaç senzill), fet que fa reduir l'energia d'excitació.

En l'apartat anterior, s'ha discutit la importància de la interacció del cromòfor amb el residu Tyr145. Cal recordar que la Tyr145 es podia trobar en dues conformacions, una de les quals representava una interacció efectiva de pont d'hidrogen d'uns  $3,0 \text{ Å}$  (Figura 6.4). Al panell dret de la Figura 6.25 es mostra l'evolució temporal d'aquesta distància. Després d'uns

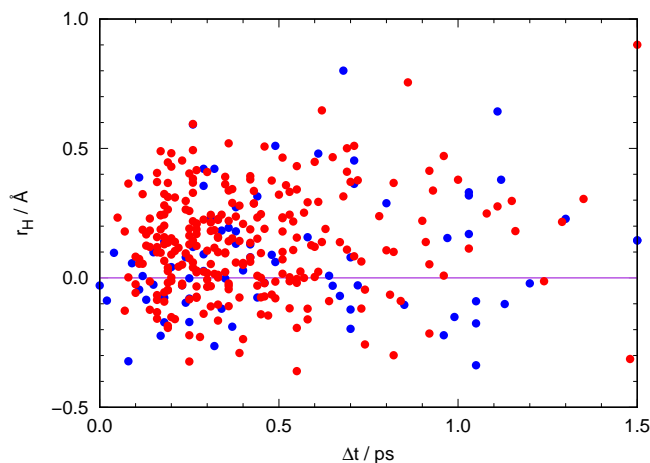


FIGURA 6.26: Posició del protó (usant l'Equació 6.5) en el moment del trànsit per alguna de les CIs, ja siguin del tipus HT-CI (vermell) o P-CI (blau).

100 fs, on sembla que no hi ha evolució en la posició de la Tyr145, es detecta una reducció brusca de la distància mitjana, que s'estabilitza cap als 3,65 Å als 400 fs. De l'anàlisi de l'histograma resolt en el temps mostrat al panell esquerre de la Figura 6.25, es pot comprovar com, en realitat, hi ha una transferència de població entre la conformació no interactuant cap a la interactuant, que passats els 400 fs de simulació, passa a ser molt majoritària. Sigui com sigui, la interacció entre la Tyr145 i el cromòfor es veu enaltida a l'estat excitat. Aquesta interacció establitzarà l'estat excitat preferentment sobre el fonamental, ja que a l'estat excitat la interacció de pont d'hidrogen és més intensa, donat que el cromòfor és aniònic. Aquesta establització de l'estat excitat respecte el fonamental, com es natural, comportarà una disminució de l'energia d'excitació.

En definitiva, s'ha aconseguit identificar tres moviments als quals associar l'anomenat *vibrational relaxation*, que ocorre després de la transferència protònica i que comporta una baixada suau de l'energia d'excitació en l'ordre dels picosegons: (1) la separació del cromòfor i l'Asp148, que ajuda a la definició del fotoproducte; (2) la pèrdua de planaritat dels anells del cromòfor, fet que desestablitzava l'estat fonamental respecte l'excitat (amb la conseqüent baixada de l'energia d'excitació) i (3) l'acostament de la Tyr145, que establitzava l'estat excitat respecte el fonamental, fet que fa baixar l'energia d'excitació.

### 6.3.5 El Cau del Conill: el Destí Ocult de l'Espècie A\*

Una dada experimental que encara no ha estat discutida i que ha de tenir a veure amb el procés dinàmic que ocorre a l'estat excitat fotoactiu és la concordança entre l'aparició de l'espècie I\* i la desaparició de l'espècie A\*, que van detectar Shi *et al.*<sup>76</sup>

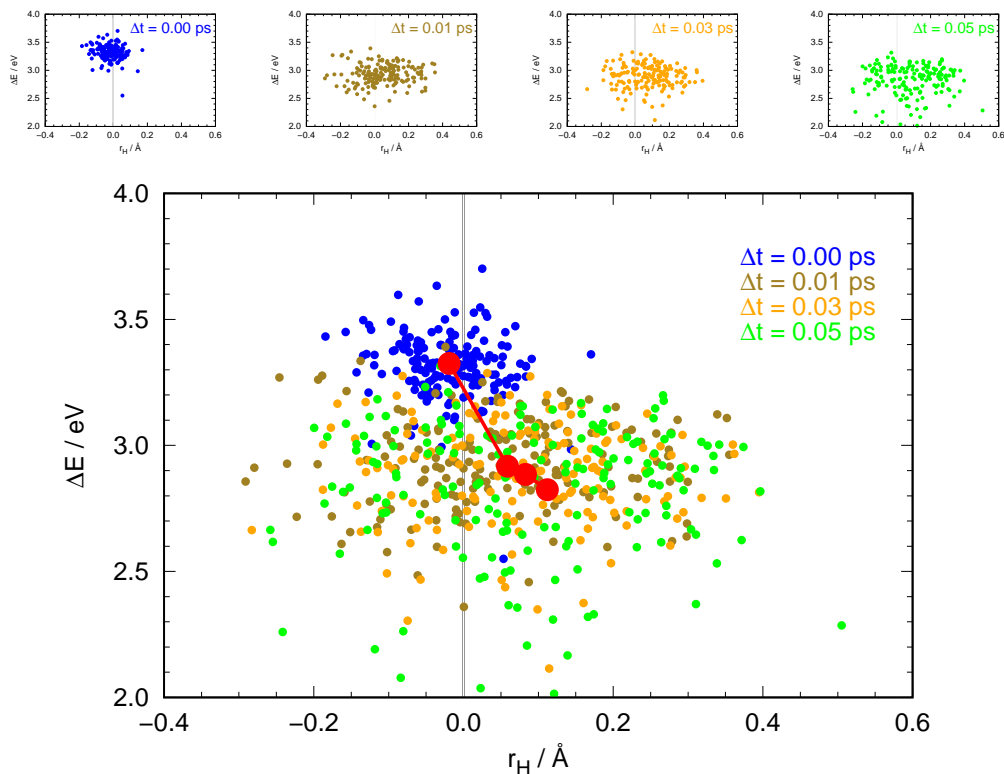


FIGURA 6.27: Relació entre la posició del protó (Equació 6.5) i l'energia d'emissió, en les etapes primerenques del procés. Els panells superiors descriuen el valor de l'energia d'emissió de cada estructura de la simulació MD respecte el paràmetre  $r_H$  pels temps indicats. En el panell inferior es representen tots els temps alhora en els seus colors corresponents, així com el camí seguit per l'energia d'excitació mitjana durant l'interval seleccionat, en color vermell i direcció descendent.

A la Figura 6.26 es representa la posició clàssica del protó en l'instant on té lloc la transició a l'estat fonamental. Cal notar com una quantitat no menyspreable de transicions ocorren amb el protó a la zona de reactius, és a dir amb  $r_H < 0$  (concretament, el 28,7 %). Si es pot establir que els cromòfors neutres tenen energies d'excitació més properes a les espècies A (més grans) que a les espècies I, es podria explicar la no concomitància en la mesura de les dues poblacions a l'estat excitat.

A la Figura 6.27 es mostra la relació entre la posició clàssica del protó i l'energia d'excitació, per a cada fotograma calculat. Es pot observar com el decreixement de l'energia d'excitació està correlacionada amb la posició del protó, encara que aquesta, naturalment, també depèn d'altres paràmetres del sistema, com s'acaba de demostrar, així que només cal esperar una certa correlació.

Així doncs, les trajectòries que pateixen una transició no radiativa a l'estat fonamental mentre són a la zona de "reactius", despoblen el nombre de sistemes que emeten a altes energies ( $A^*$ ) sense poblar els sistemes que

emetran a baixes energies ( $I^*$ ). Per tant, aquest comportament estaria d'acord amb l'observació realitzada per Shi *et al.*<sup>76</sup> En els seus experiments, van enregistrar l'espectre a temperatura ambient i a 140 K, veient que la pujada de l'espècie  $I^*$  era encara indetectable a 140 K. Això estaria d'acord amb la transferència protònica sense barrera. Ara bé, pel que fa a la desaparició no radiativa de  $A^*$ , aquesta sí que va ser frenada a baixes temperatures, fet que indica que el procés de desactivació segurament comporti una barrera. Això és el que s'ha trobat en aquest treball per a les CIs estudiades mitjançant coordenades de reacció a l'estat excitat que porten a la desactivació no radiativa de les espècies  $A^*$ .

## 6.4 Conclusions Parcial

A continuació es presenten les conclusions que s'extreuen dels resultats presentats en aquest capítol.

### 6.4.1 Conclusions Metodològiques

- La descripció estàtica que proporciona l'estructura cristal·lina és del tot insuficient per explicar la complexitat del fenomen físic que té lloc a l'estat excitat.
- Per tal d'obtenir una descripció completa del sistema, cal recórrer a un tractament estadístic basat en simulacions dinàmiques d'alt nivell. Només així es pot obtenir una població degudament ponderada d'estructures que expliquin un comportament global (per tant, macroscòpic) correcte.
- Tot i que la major part dels perfils obtinguts a l'estat excitat són (encara que lleugerament) endoèrgics, la inclusió de l'energia cinètica del protó mitjançant el procés de simulació de la fotoexcitació (QD), fa que la transferència es torni explicable. Per tant, cal introduir els efectes cinètics per entendre aquest tipus de sistemes.
- La mutagènesi *in silico*, aplicada en la mutant S65T/H148D/Y145F és una tècnica computacional vàlida ja que és sensible a les variacions proposades i permet fer prediccions sobre comportaments fins de noves mutants.
- S'ha aconseguit introduir modificacions i correccions imprescindibles per al càlcul de dinàmiques no adiabàtiques en sistemes biològics en el mòdul de desenvolupament NADYM del programari ChemShell.
- El mètode MRCISD-OM2/CHARMM27 ha permès la integració de dinàmiques no adiabàtiques amb èxit.

- El mètode MRCISD-OM2/CHARMM27 falla en la descripció de la barrera que porta a la HT-CI, i per tant limita l'aplicabilitat del mètode en el sistema triat.

### 6.4.2 Conclusions Químiques

- Encara que efectivament el cromòfor de la mutant S65T/H148D experimenta un guany d'acidesa a l'estat excitat, aquest no és ni de bon tros tant exagerat com s'havia extrapolat inicialment. Això és degut a que (1) hi ha una distorsió de la planaritat de l'estructura del cromòfor i a (2) l'estructura de càrregues de l'interior de la proteïna.
- Tant a l'estat fonamental com excitat fotoactiu, els perfils energètics per a la transferència protònica presenten una desviació suau respecte la isoergicitat (trobat només una fracció considerable de perfils exoèrgics a l'estat excitat). Així doncs, l'energètica del procés depèn d'oscil·lacions tèrmiques de l'entorn, que són les que determinen una transferència factible a l'estat excitat.
- En solució aquosa i a temperatura ambient, la configuració que adopta la proteïna permet explicar l'aparició instantània (per sota del límit de detecció experimental) de la radiació corresponent a l'espècie I\*.
- La interacció entre la Tyr145 i l'oxígen fenòlic del cromòfor, que aquí es descriu per primer cop, és crítica per a la factibilitat de la transferència protònica l'estat excitat
- En la mutant S65T/H148D/Y145F, i a causa de la pèrdua de la interacció entre el cromòfor i la Tyr145, que a més comporta una modificació en l'estat de protonació de l'Asp148, la transferència ultraràpida es fa inviable. En aquesta nova mutant, el protó es troba clarament localitzat al cromòfor a l'estat fonamental, així que s'espera un comportament diferent al de la doble mutant respecte a les seves propietats espectroscòpiques resoltes en el temps.
- L'aigua juga un paper crucial en la transferència protònica: la solvatació de l'Asp148, l'acceptor del protó, és determinant en la mateixa. Si l'Asp148 es troba massa solvatat, és incapaç d'acceptar el protó ja que perd la seva basicitat.
- La funció d'ona protònica del protó a l'estat electrònic fonamental, i donada la relativa isoergicitat del mateix, es troba altament deslocalitzada, fet que permet explicar els resultats observats per Stoner-Ma. *et al.*
- S'ha aconseguit reproduir el temps de la transferència del protó, que és d'uns  $\sim 15$  fs. El temps de transferència pel deuteró (uns  $\sim 20-25$  fs), que encara es troba per sota del límit de detecció experimental (50 fs), permet explicar l'absència d'efecte cinètic d'isòtop.

- S'ha aconseguit reproduir qualitativament els espectres d'emissió resolts en el temps, en els primers instants després de la fotoexcitació, fet que ha permès explicar el procés a l'estat excitat en dues etapes: una primera de transferència protònica ultrarràpida (en l'escala dels femtosegons) i una segona de relaxació vibracional, en l'ordre dels picosegons.
- El procés de relaxació vibracional s'ha aconeguit explicar en base a (1) l'allunyament de l'Asp148 del cromòfor, (2) la pèrdua de planaritat del segon i (3) l'acostament de la Tyr145 a l'oxigen fenòlic del mateix.
- S'ha identificat dues vies de relaxació no radiatives (CIs). La primera, HT-CI, comporta la rotació de l'enllaç C<sub>4</sub>C<sub>6</sub> amb la conseqüent torsió de l'H<sub>7</sub> fora del pla i pertany al mateix *seam* que la descrita per Zhang *et al.* Aquesta ha resultat ser la via de desactivació principal. La segona, P-CI, comporta la piramidalització del C<sub>2</sub> i representa una via de desactivació secundària.
- La no concordança entre la desaparició de l'espècie A\* i l'aparició de l'espècie I\* observada per Shi *et al.* es pot explicar en base a l'accés a les vies de desactivació no radiatives en la zona de reactius (això és, amb el protó al cromòfor). La desactivació de les espècies A\* per mitjà de les CIs trobades explica aquest comportament.

Els resultats presentats en aquest capítol han donat lloc a dues publicacions, més una tercera que està en procés de revisió:

- P. Armengol, R. Gelabert, M. Moreno, J.M. Lluch "Unveiling how an archetypal fluorescent protein operates: Theoretical perspective on the ultrafast excited-state dynamics of GFP variant S65T/H148D", *J. Phys. Chem. B* 119(6):2274-2291, 2015.<sup>201</sup>
- P. Armengol, R. Gelabert, M. Moreno, J.M. Lluch "Theoretical Computer-Aided Mutagenic Studies on the Triple Green Fluorescent Protein Mutant S65T/H148D/Y145F", *Chem. Phys. Chem.* 16(10):2134-2139, 2015.<sup>202</sup>
- P. Armengol, R. Gelabert, M. Moreno, J.M. Lluch "Ultrafast Action Chemistry in Slow Motion Atomistic Description of the Excitation and Fluorescence Processes in an Archetypal Fluorescent Protein- *en revisió.*

# III

## PERORATIO

Or che mi conoscete parlate voi.  
Deh parlate. Chi siete? Vi piaccia dir?

Rodolfo - La Bohème (G. Puccini)





## 7 Conclusions

En aquest treball he intentat, primerament, exposar una problemàtica científica, un tema d'interès per a la recerca actual com ho són les FPs. Dotant d'un context teòric i de pensament, he intentat donar explicacions clares, detallades i sistemàtiques sobre els processos associats als comportaments experimentals dels sistemes estudiats. Com en qualsevol treball de recerca científica actual, el nivell de detall assolit és molt profund, així que, per concloure el treball, crec que cal alçar-se sobre aquests resultats per donar una conclusió panoràmica dels resultats d'aquesta tesi. Apartar la vista després d'una descripció tan detallada crec que és un exercici necessari per comprendre l'impacte que pot tenir un treball de recerca actual en la ciència general. Ja no es tracta de que els arbres no et deixin veure el bosc, si no que en un treball com aquest, s'ha donat una descripció de les fulles dels arbres. En aquest sentit, cal retrocedir per tornar a veure els esmentats arbres i, finalment, el bosc.

Les conclusions concretes i detallades s'han donat als apartats finals dels Capítols 5 i 6, així que remeteu-vos als mateixos per a tal concreció. D'aquestes conclusions se'n poden extreure unes idees generals del funcionament dels foto-biosistemes, tant des del punt de vista de la seva interacció amb la llum com dels processos químics que tenen lloc en els mateixos en els seus diferents estats electrònics.

Ha quedat palesa la importància que tenen les descripcions atomístiques d'alt nivell en la comprensió dels fenòmens físic que tenen lloc en les FPs i, en general, als sistemes complexos. L'estudi detallat de l'excitació mitjançant els mètodes de la química quàntica ha permès assolir un nivell de coneixement fi sobre la relació estructura-espectre que ha permès, en últim terme, proposar mutacions que puguin donar lloc a noves mutants amb propietats espectroscòpiques modulades a plaer. En aquest sentit, cal remarcar la vital importància que té l'estructura de l'entorn proteic del cromòfor. Les interaccions que es tenen lloc, sovint tenen un caràcter dinàmic, a vegades participant en règims d'interacció plurimodals. La combinació de les tècniques de la química quàntica per determinar el moviment de la densitat electrònica a l'estat excitat amb les tècniques dinàmiques multi-escala, que permeten escatir d'una manera realista l'organització dels residus al voltant del cromòfor, han permès una profunda comprensió de les característiques que distingeixen les diferents FPs. D'altra banda, el paper que juga l'aigua, no com a dissolvent continu, si no com a actor discret dels processos químics que tenen lloc, és vital. Sovint es considera l'aigua

només com a dissolvent, però és evident que té un paper determinant en els processos estudiats.

Des d'un punt de vista del coneixement assolit respecte el mecanisme de fluorescència de les FPs, creat a partir de la proteïna fluorescent S65T/H148D-GFP, s'ha aconseguit donar un dibuix complert amb un alt grau de detall del procés que segueixen aquests sistemes després d'absorbir la radiació. S'ha aconseguit descriure, amb diferents tècniques, tot el procés que segueix una proteïna fluorescent amb la seva interacció amb la llum: des de la descripció de la seva situació d'equilibri a l'estat fonamental al procés que segueix a l'estat excitat un cop absorbeix la llum. Una de les conclusions importants que s'ha presentat és que les superfícies d'energia potencial que segueixen les transferències protòniques són essencialment planes, al contrari del què es pensava. Això vol dir que canvis subtils en l'entorn del cromòfor o en l'estructura del mateix poden decantar la balança cap a un o altre cantó. Altra vegada, s'ha demostrat com les molècules d'aigua, com a entitats discretes, són absolutament determinants en l'energètica dels processos estudiats. La metodologia desenvolupada i implementada en el transcurs d'aquesta tesi obre la porta a estudiar nous foto-biosistemes, que presentin processos complexos en estats electrònics excitats, processos que només poden ser visualitzats a nivell atòmic per mitjà de simulacions dinàmiques degudament implementades i executades.

Finalment, cal trencar una llança a favor de l'evident necessitat de descripcions dinàmiques dels processos químics. En aquest treball s'ha fet palesa la quantitat d'informació que pot aflorar d'una simulació dinàmica no condicionada, degudament realitzada. Els càlculs no basats en dinàmica tenen un significat limitat, doncs intenten descriure un procés essencialment dinàmic amb una visió i llenguatge que són estàtics. Les dinàmiques moleculars que segueixin de la manera més estricta possible les lleis físiques que regeixen el comportament de l'univers són una de les eines més potents que tenim per escatir la terrible complexitat que impera en els processos biològics. El procés d'integració d'una simulació dinàmica és equivalent a l'obtenció d'una petita part d'univers, de manera controlada. Les simulacions dinàmiques, en si mateixes, no aporten massa informació directa, però són contenidores d'una quantitat ingent de dades, ja que inclouen la física del problema, si han estat correctament realitzades. Així doncs, recau en l'habilitat del científic que s'hi enfronta, la responsabilitat de destriar el gra de la palla, i de poder obtenir informació útil de la mateixa, per a poder construir un coneixement a partir de la descripció atòmica que se'n desprèn. A més, de la capacitat del científic de presentar la informació de manera organitzada, clara i conscisa, en depèn que els seus descobriments passin la prova del temps i finalment visquin en l'inmaginari científic col·lectiu.

# IV

## BIBLIOGRAFIA

Madamina, il catalogo è questo  
Delle belle che amò il padron mio;  
un catalogo egli è che ho fatt'io;  
Osservate, leggete con me.

Leporello - Don Giovanni (W. A. Mozart)



## Bibliografia

- [1] Roberts, D. *Post-truth politics*; Grist, 2010  
<http://grist.org/article/2010-03-30-post-truth-politics>. Last accessed on: 2 de juny de 2017.
- [2] <https://www.oxforddictionaries.com/press/news/2016/12/11/WOTY-16>, Last accessed on: 2 de juny de 2017.
- [3] Descartes, R. *Meditationes de Prima philosophia*; : Paris, 1641.
- [4] Hawking, S.; Mlodinow, L. *The Grand Design*; Bentam Books: New York, 2010.
- [5] Chatrchyan, S. *et al. Phys. Lett. B* **2012**, 716, 30 – 61.
- [6] Aad, G. *et al. Phys. Lett. B* **2012**, 716, 1 – 29.
- [7] Higgs, P. W. **1964**, 12, 132 – 133.
- [8] Englert, F.; Brout, R. *Phys. Rev. Lett.* **1964**, 13, 321 – 324.
- [9] Zimmer, M. *Chem. Rev.* **2002**, 102, 759 – 781.
- [10] Meech, S. R. *Chem. Soc. Rev.* **2009**, 38, 2922–2934.
- [11] Dedecker, P.; De Schryver, F. C.; Hofkens, J. *J. Am. Chem. Soc.* **2013**, 135, 2387 – 2402.
- [12] Acharya, A.; Bogdanov, A. M.; Grigorenko, B. L.; Bravaya, K. B.; Nemukhin, A. V.; Lukyanov, K. A.; Krylov, A. I. *Chem. Rev.* **2016**, 117, 758–795.
- [13] Davenport, D.; Nicol, J. A. C. *Proc. R. Soc. London, Ser. B.* **1955**, 144, 399 – 411.
- [14] Shimomura, O.; Musicki, B.; Kishi, Y. *Biochem. J* **1988**, 251, 405 – 410.
- [15] Shimomura, O. *O. Biol. Bull.* **1995**, 189, 1 – 5.
- [16] Shimomura, O.; Johnson, F. H.; Saiga, Y. *J. Cell. Physio.* **1962**, 59, 223–239.
- [17] Morin, J. G.; Hastings, J. W. *J. Cell. Physiol.* **1971**, 77, 313 – 318.
- [18] Ward, W. W.; Cormier, M. J. *J. Biol. Chem.* **1979**, 254, 781 – 788.
- [19] Heim, R.; Prasher, D. C.; Tsien, R. Y. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, 91, 12501 – 12504.

- [20] Morise, H.; Shimomura, O.; Johnson, F. H.; Winant, J. *Biochemistry* **1974**, *13*, 2656 – 2662.
- [21] Shimomura, O. *FEBS Lett.* **1979**, *104*, 220 – 222.
- [22] Cody, C. W.; Prasher, D. C.; Westler, W. M.; Prendergast, F. G.; Ward, W. W. *Biochemistry* **1993**, *32*, 1212 – 1218.
- [23] Schwede, T. F.; Retey, J.; Schulz, G. E. *Biochemistry* **1999**, *38*, 5355 – 5361.
- [24] Prasher, D. C.; Eckenrode, V. K.; Ward, W. W.; Prendergast, F. G.; Cormier, M. J. *Gene* **1992**, *111*, 229 – 233.
- [25] Chalfie, M.; Tu, Y.; Euskirchen, G.; Ward, W. W.; Prasher, D. C. *Science* **1994**, *263*, 802 – 805.
- [26] Yang, F.; Moss, L. G.; Phillips Jr., G. N. *Nat. Biotechnol.* **1996**, *14*, 1246–1251.
- [27] Niwa, H.; Inouye, S.; Hirano, T.; Matsuno, T.; Kojima, S.; Kubota, M.; Ohashi, M.; Tsuji, F. I. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 13617 – 13622.
- [28] Ehrmann, M. A.; Scheyhing, C. H.; Vogel, R. F. *Letts. Appl. Microbiol.* **2001**, *32*, 230 – 234.
- [29] Lippincott Schwartz, J.; Snapp, E.; Kenworthy, A. *Nat. Rev. Mol. Cell Biol.* **2001**, *2*, 444 – 456.
- [30] Bastiaens, P. I. H.; Pepperkok, R. *Trends Biochem. Sci* **2000**, *25*, 631 – 637.
- [31] Billinton, N.; Knight, A. W. *Anal. Biochem.* **2001**, *291*, 175 – 197.
- [32] Misteli, T.; Spector, D. L. *Nat. Biotechnol.* **1997**, *15*, 961 – 964.
- [33] Wahlfors, J.; Loimas, S.; Pasanen, T.; Hakkarainen, T. *Histochem. Cell Biol.* **2001**, *115*, 59 – 65.
- [34] Rashidi, B.; Yang, M.; Jiang, P.; Baranov, E.; An, Z. L.; Wang, X.; Moossa, A. R.; Hoffman, R. M. *Clin. Exp. Metastasis* **2000**, *18*, 57 – 60.
- [35] Yang, M.; Baranov, E.; Jiang, P.; Sun, F. X.; Li, X. M.; Li, L. N.; Hasegawa, S.; Bouvet, M.; Al Tuwaijri, M.; Chishima, T.; Shimada, H.; Moossa, A. R.; Penman, S.; Hoffman, R. M. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 1206 – 1211.
- [36] Yang, M.; Baranov, E.; Li, X. M.; Wang, J. W.; Jiang, P.; Li, L.; Moossa, A. R.; Penman, S.; Hoffman, R. M. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 2616 – 2621.
- [37] Wouters, F. S.; Verveer, P. J.; Bastiaens, P. I. H. *Trends Cell Biol.* **2001**, *11*, 203 – 211.
- [38] Belmont, A. S. *Trends Cell Biol.* **2001**, *11*, 250 – 257.

- [39] Toomre, D.; Manstein, D. J. *Trends Cell Biol.* **2001**, *11*, 298 – 303.
- [40] Ward, W. W.; Prentice, H. J.; Roth, A. F.; Cody, C. W.; Reeves, S. C. *Photochem. Photobiol.* **1982**, *35*, 803 – 808.
- [41] Kummer, A. D.; Wiehler, J.; Rehder, H.; Kompa, C.; Steipe, B.; Michel Beyerle, M. E. *J. Phys. Chem. B* **2000**, *104*, 4791 – 4798.
- [42] Donnelly, M.; Fedeles, F.; Wirstam, M.; Siegbahn, P. E.; Zimmer, M. *J. Am. Chem. Soc.* **2001**, *123*, 4679 – 4686.
- [43] Brejc, K.; Sixma, T. K.; Kitts, P. A.; Kain, S. R.; Tsien, R. Y.; Ormö, M.; Remington, S. J. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 2306 – 2311.
- [44] van Thor, J. J.; Zanetti, G.; Ronayne, K. L.; Towrie, M. J. *Phys. Chem. B* **2005**, *109*, 16099 – 16108.
- [45] Chattoraj, M.; King, B. A.; Bublitz, G. U.; Boxer, S. G. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 8362 – 8367.
- [46] Lill, M. A.; Helms, V. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 2778 – 2781.
- [47] Ormö, M.; Cubitt, A. B.; Kallio, K.; Gross, L. A.; Tsien, R. Y.; Remington, S. J. *Science* **1996**, *273*, 1392 – 1395.
- [48] Stoner Ma, D.; Jaye, A. A.; Matousek, P.; Towrie, M.; Meech, S. R.; Tonge, P. J. *J. Am. Chem. Soc.* **2005**, *127*, 2864 – 2865.
- [49] Stoner Ma, D.; Melief, E. H.; Nappa, J.; Ronayne, K. L.; Tonge, P. J.; Meech, S. R. *J. Phys. Chem. B* **2006**, *110*, 22009 – 22018.
- [50] Vendrell, O.; Gelabert, R.; Moreno, M.; Lluch, J. M. *J. Am. Chem. Soc.* **2006**, *128*, 3564 – 3574.
- [51] Vendrell, O.; Gelabert, R.; Moreno, M.; Lluch, J. M. *J. Phys. Chem. B* **2008**, *112*, 5500 – 5511.
- [52] Di Donato, M.; van Wilderen, L. J. G. W.; van Stokkum, I. H. M.; Cohen Stuart, T.; Kennis, J. T. M.; Hellingwerf, K. J.; van Grondelle, R.; Groot, M. L. *Phys. Chem. Chem. Phys.* **2011**, *13*, 16295–16305.
- [53] Scharnagl, C.; Raupp Kossmann, R. A. *J. Phys. Chem. B* **2004**, *108*, 477 – 489.
- [54] Cui, G. L.; Lan, Z. G.; Thiel, W. *J. Am. Chem. Soc.* **2012**, *134*, 1662 – 1672.
- [55] Sinicropi, A.; Andruniow, T.; Ferre, N.; Basosi, R.; Olivucci, M. *J. Am. Chem. Soc.* **2005**, *127*, 11534 – 11535.
- [56] Weber, W.; Helms, V.; McCammon, J. A.; Langhoff, P. W. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 6177 – 6182.
- [57] Tsien, R. Y. *Annu. Rev. Biochem.* **1998**, *67*, 509 – 544.



- [58] Shaner, N. C.; Steinbach, P. A.; Tsien, R. Y. *Nat. Methods* **2005**, *2*, 905 – 909.
- [59] Tromberg, B. J.; Shah, N.; Lanning, R.; Cerussi, A.; Espinoza, J.; Pham, T.; Svaasand, L.; Butler, J. *Neoplasia* **2000**, *2*, 26 – 40.
- [60] Stamatas, G. N.; Southall, M.; Kollias, N. J. *Invest. Dermatol.* **2006**, *126*, 1753 – 1760.
- [61] Matz, M. V.; Lukyanov, K. A.; Lukyanov, S. A. *Bioessays* **2002**, *24*, 953 – 959.
- [62] Matz, M. V.; Fradkov, A. F.; Labas, Y. A.; Savitsky, A. P.; Zairaisky, A. G.; Markelov, M. L.; Lukyanov, S. A. *Nat. Biotechnol.* **1999**, *17*, 969 – 973.
- [63] Baird, G. S.; Zacharias, D. A.; Tsien, R. Y. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 11984 – 11989.
- [64] Bravaya, K. B.; Subach, O. M.; Korovina, N.; Verkhusha, V. V.; Krylov, A. I. *J. Am. Chem. Soc.* **2012**, *134*, 2807 – 2814.
- [65] Shaner, N. C.; Campbell, R. E.; Steinbach, P. A.; Giepmans, B. N. G.; Palmer, A. E.; Tsien, R. Y. *Nat. Biotechnol.* **2004**, *22*, 1567 – 1572.
- [66] Shaner, N. C.; Lin, M. Z.; McKeown, M. R.; Steinbach, P. A.; Hazelwood, K. L.; Davidson, M. W.; Tsien, R. Y. *Nat. Methods* **2008**, *5*, 545 – 551.
- [67] Bevis, B. J.; Glick, B. S. *Nat. Biotechnol.* **2002**, *20*, 83 – 87.
- [68] Henderson, J. N.; Ai, H. .; Campbell, R. E.; Remington, S. J. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 6672 – 6677.
- [69] Heim, R.; Cubitt, A. B.; Tsien, R. Y. *Nature* **1995**, *373*, 663 – 664.
- [70] Cormack, B. P.; Valdivia, R. H.; Falkow, S. *Gene* **1996**, *173*, 33 – 38.
- [71] Wachter, R. M.; King, B. A.; Heim, R.; Kallio, K.; Tsien, R. Y.; Boxer, S. G.; Remington, S. J. *Biochemistry* **1997**, *36*, 9759 – 9765.
- [72] Kneen, M.; Farinas, J.; Li, Y. X.; Verkman, A. S. *Biophys. J.* **1998**, *74*, 1591 – 1599.
- [73] Llopis, J.; McCaffery, J. M.; Miyawaki, A.; Farquhar, M. G.; Tsien, R. Y. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 6803 – 6808.
- [74] Elsliger, M. A.; Wachter, R. M.; Hanson, G. T.; Kallio, K.; Remington, S. J. *Biochemistry* **1999**, *38*, 5296 – 5301.
- [75] Shu, X.; Kallio, K.; Shi, X.; Abbyad, P.; Kanchanawong, P.; Childs, W.; Boxer, S. G.; Remington, S. J. *Biochemistry* **2007**, *46*, 12005–12013.

- [76] Shi, X.; Abbyad, P.; Shu, X.; Kallio, K.; Kanchanawong, P.; Childs, W.; Remington, S. J.; Boxer, S. G. *Biochemistry* **2007**, *46*, 12014–12025.
- [77] Lin, M. Z.; Mc Keown, M. R.; Ng, H. L.; Aguilera, T. A.; Shaner, N. C.; Campbell, R. E.; Adams, S. R.; Gross, L. A.; Ma, W.; Alber, T.; Tsien, R. Y. *Chem. Biol.* **2009**, *16*, 1169 – 1179.
- [78] Shcherbakova, D. M.; Verkhusha, V. V. *Nat. Methods* **2013**, *10*, 751 – +.
- [79] Chu, J.; Haynes, R. D.; Corbel, S. Y.; Li, P.; Gonzalez Gonzalez, E.; Burg, J. S.; Ataie, N. J.; Lam, A. J.; Cranfill, P. J.; Baird, M. A.; Davidson, M. W.; Ng, H. L.; Garcia, K. C.; Contag, C. H.; Shen, K.; Blau, H. M.; Lin, M. Z. *Nat. Methods* **2014**, *11*, 572 – 578.
- [80] Piatkevich, K. D.; Malashkevich, V. N.; Morozova, K. S.; Nemkovich, N. A.; Almo, S. C.; Verkhusha, V. V. *Sci. Rep.* **2013**, *3*, 1847.
- [81] Randino, C.; Nadal Ferret, M.; Gelabert, R.; Moreno, M.; Lluch, J. M. *Theor. Chem. Acc.* **2013**, *132*, 1327.
- [82] Randino, C.; Gelabert, R.; Moreno, M.; Lluch, J. M.; Piatkevich, K. D. *Phys. Chem. Chem. Phys.* **2015**, *17*, 29363 – 29373.
- [83] Shcherbo, D.; Merzlyak, E. M.; Chepurnykh, T. V.; Fradkov, A. F.; Ermakova, G. V.; Solovieva, E. A.; Lukyanov, K. A.; Bogdanova, E. A.; Zraisky, A. G.; Lukyanov, S.; Chudakov, D. M. *Nat. Methods* **2007**, *4*, 741 – 746.
- [84] Pletnev, S.; Shcherbo, D.; Chudakov, D. M.; Pletneva, N.; Merzlyak, E. M.; Wlodawer, A.; Dauter, Z.; Pletnev, V. J. *Biol. Chem.* **2008**, *283*, 28980 – 28987.
- [85] Mildvan, A. S.; Harris, T. K.; Abeygunawardana, C. *Methods Enzymol.* **1999**, *308*, 219 – 245.
- [86] Kondo, M.; Heisler, I. A.; Stoner Ma, D.; Tonge, P. J.; Meech, S. R. *J. Am. Chem. Soc.* **2010**, *132*, 1452 – +.
- [87] Stoner Ma, D.; Jaye, A. A.; Ronayne, K. L.; Nappa, J.; Meech, S. R.; Tonge, P. J. *J. Am. Chem. Soc.* **2008**, *130*, 1227–1235.
- [88] Frenkel, D.; Smit, B. *Understanding Molecular Simulation*; Academic Press: San Diego, California, 2002.
- [89] Born, M.; Oppenheimer, R. *Ann. Phys.* **1927**, *84*, 0457 – 0484.
- [90] Sutcliffe, B. T. *Computational Techniques in Quantum Chemistry*; Reidel: Boston, 1975.
- [91] Szabo, A.; Ostlund, N. *Modern Quantum Chemistry*; Dover Publications, Inc.: Mineola, New York, 1982.
- [92] Hohenberg, P.; Kohn, W. *Phys. Rev.* **1964**, *136*, B864.

- [93] Kohn, W.; Sham, L. J. *Phys. Rev. B* **1965**, *140*, 1133.
- [94] Cohen, A. J.; Mori Sanchez, P.; Yang, W. *Chem. Rev.* **2012**, *112*, 289 – 320.
- [95] Casida, M. E. *Recent Advances in Density Functional Methods, Vol I*; World Scientific: Singapore, 1995.
- [96] Casida, M. E. *Recent Developments and Application of Modern Density Functional Theory, Theoretical and Computational Chemistry, Vol IV*; Elsevier: Amsterdam, 1996.
- [97] Stratmann, R. E.; Scuseria, G. E.; Frisch, M. J. *J. Chem. Phys.* **1998**, *109*, 8218 – 8224.
- [98] Runge, E.; Gross, E. K. U. *Phys. Rev. Lett.* **1984**, *52*, 997 – 1000.
- [99] Yanai, T.; Tew, D. P.; Handy, N. C. *Chem. Phys. Lett.* **2004**, *393*, 51 – 57.
- [100] Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648 – 5652.
- [101] Tawada, Y.; Tsuneda, T.; Yanagisawa, S.; Yanai, T.; Hirao, K. *J. Chem. Phys.* **2004**, *120*, 8425 – 8433.
- [102] Thiel, W. *WIREs Comput. Mol. Sci.* **2014**, *4*, 145 – 157.
- [103] Pople, J. A.; Santry, D. P.; Segal, G. A. *J. Chem. Phys.* **1965**, *43*, S129.
- [104] Dewar, M. J. S.; Thiel, W. J. *Am. Chem. Soc.* **1977**, *99*, 4899 – 4907.
- [105] Dewar, M. J. S.; Thiel, W. J. *Am. Chem. Soc.* **1977**, *99*, 4907 – 4917.
- [106] Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902 – 3909.
- [107] Stewart, J. J. P. *J. Comput. Chem.* **1989**, *10*, 209 – 220.
- [108] Stewart, J. J. P. *J. Comput. Chem.* **1989**, *10*, 221 – 264.
- [109] Rocha, G. B.; Freire, R. O.; Simas, A. M.; Stewart, J. J. P. *J. Comput. Chem.* **2006**, *27*, 1101 – 1111.
- [110] Stewart, J. J. P. *J. Mol. Model.* **2007**, *13*, 1173 – 1213.
- [111] Stewart, J. J. P. *J. Mol. Model.* **2013**, *19*, 1 – 32.
- [112] Kolb, M.; Thiel, W. J. *Comput. Chem.* **1993**, *14*, 775 – 789.
- [113] Weber, W.; Thiel, W. *Theor. Chem. Acc.* **2000**, *103*, 495 – 506.
- [114] Scholten, M. *Semiempirische Verfahren mit Orthogonalisierungskorrekturen: Die OM3 Methode. PhD Thesis*; Universität Düsseldorf: , 2003.
- [115] Tuttle, T.; Thiel, W. *Phys. Chem. Chem. Phys.* **2008**, *10*, 2159 – 2166.
- [116] Korth, M.; Thiel, W. J. *Chem. Theory Comput.* **2011**, *7*, 2929 – 2936.

- [117] Dral, P. O.; Wu, X.; Spoerkel, L.; Koslowski, A.; Weber, W.; Steiger, R.; Scholten, M.; Thiel, W. J. *Chem. Theory Comput.* **2016**, *12*, 1082 – 1096.
- [118] Silva Junior, M. R.; Thiel, W. J. *Chem. Theory Comput.* **2010**, *6*, 1546 – 1564.
- [119] Heggen, B.; Lan, Z.; Thiel, W. *Phys. Chem. Chem. Phys.* **2012**, *14*, 8137 – 8146.
- [120] Spoerkel, L.; Cui, G.; Thiel, W. J. *Phys. Chem. A* **2013**, *117*, 4574 – 4583.
- [121] Spoerkel, L.; Cui, G.; Koslowski, A.; Thiel, W. J. *Phys. Chem. A* **2014**, *118*, 152 – 157.
- [122] Mac Kerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz Kuczera, J.; Yin, D.; Karplus, M. *J. Phys. Chem. B* **1998**, *102*, 3586 – 3616.
- [123] Mac Kerell, A. D.; Feig, M.; Brooks, C. L. *J. Am. Chem. Soc.* **2004**, *126*, 698 – 699.
- [124] Brooks, B. R.; Brucoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187 – 217.
- [125] Brooks, B. R.; Brooks, C. L. I. I. I.; Mackerell, A. D. J.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M. *J. Comput. Chem.* **2009**, *30*, 1545 – 1614.
- [126] Wang, J. M.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. *J. Comput. Chem.* **2005**, *26*, 114 – 114.
- [127] Wang, J. M.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. *J. Comput. Chem.* **2004**, *25*, 1157 – 1174.
- [128] Vander Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. *J. Comput. Chem.* **2005**, *26*, 1701 – 1718.
- [129] Warshel, A.; Karplus, M. *J. Am. Chem. Soc.* **1972**, *94*, 5612.
- [130] Warshel, A.; Levitt, M. *J. Mol. Biol.* **1976**, *103*, 227 – 249.
- [131] Senn, H. M.; Thiel, W. *Angew. Chem. Int. Ed.* **2009**, *48*, 1198 – 1229.

- [132] Gao, J. L.; Amara, P.; Alhambra, C.; Field, M. J. *J. Phys. Chem. A* **1998**, *102*, 4714 – 4721.
- [133] Philipp, D. M.; Friesner, R. A. *J. Comput. Chem.* **1999**, *20*, 1468 – 1494.
- [134] Murphy, R. B.; Philipp, D. M.; Friesner, R. A. *Chem. Phys. Lett.* **2000**, *321*, 113 – 120.
- [135] Garcia Viloca, M.; Gao, J. L. *Theor. Chem. Acc.* **2004**, *111*, 280 – 286.
- [136] Field, M. J.; Bash, P. A.; Karplus, M. J. *Comput. Chem.* **1990**, *11*, 700 – 733.
- [137] Singh, U. C.; Kollman, P. A. *J. Comput. Chem.* **1986**, *7*, 718 – 730.
- [138] Senn, H. M.; Thiel, W. *Curr. Opin. Chem. Biol.* **2007**, *11*, 182 – 187.
- [139] Sherwood, P.; de Vries, A. H.; Guest, M. F.; Schreckenbach, G.; Catlow, C. R. A.; French, S. A.; Sokol, A. A.; Bromley, S. T.; Thiel, W.; Turner, A. J.; Billeter, S.; Terstegen, F.; Thiel, S.; Kendrick, J.; Rogers, S. C.; Casci, J.; Watson, M.; King, F.; Karlsen, E.; Sjøvoll, M.; Fahmi, A.; Schafer, A.; Lennartz, C. *Journal of Molecular Structure-Theochem* **2003**, *632*, 1 – 28.
- [140] Sherwood, P.; de Vries, A. H.; Collins, S. J.; Greatbanks, S. P.; Burton, N. A.; Vincent, M. A.; Hillier, I. H. *Faraday Discuss.* **1997**, *106*, 79 – 92.
- [141] Bakowies, D.; Thiel, W. *J. Phys. Chem.* **1996**, *100*, 10580 – 10594.
- [142] Antes, I.; Thiel, W. *J. Phys. Chem. A* **1999**, *103*, 9290 – 9295.
- [143] *ChemShell, a Computational Chemistry Shell*; See [www.chemshell.org](http://www.chemshell.org), Last accessed on: 2 de juny de 2017.
- [144] Chandler, D. *Introduction to Modern Statistical Mechanics*; Oxford University Press: New York, 1987.
- [145] McQuarrie, D. A. *Statistical Mechanics*; University Science Books: Sausalito (California), 2000.
- [146] Verlet, L. *Phys. Rev.* **1967**, *159*, 98.
- [147] Swope, W. C.; Andersen, H. C.; Berens, P. H.; Wilson, K. R. *J. Chem. Phys.* **1982**, *76*, 637 – 649.
- [148] Hockney, R. W. *Methods in Computational Physics* **1970**, *9*, 135 – 211.
- [149] Tully, J. C. *J. Chem. Phys.* **1990**, *93*, 1061 – 1071.
- [150] Fabiano, E.; Keal, T. W.; Thiel, W. *Chem. Phys.* **2008**, *349*, 334 – 347.
- [151] Coker, D. F.; Xiao, L. *J. Chem. Phys.* **1995**, *102*, 496 – 510.
- [152] Barbatti, M.; Lan, Z.; Crespo Otero, R.; Szymczak, J. J.; Lischka, H.; Thiel, W. *J. Chem. Phys.* **2012**, *137*, 22A503.

- [153] Lan, Z.; Lu, Y.; Fabiano, E.; Thiel, W. *ChemPhysChem* **2011**, *12*, 1989 – 1998.
- [154] Santoro, F.; Petrongolo, C.; Granucci, G.; Persico, M. *Chem. Phys.* **2000**, *259*, 193 – 200.
- [155] Hack, M. D.; Jasper, A. W.; Volobuev, Y. L.; Schwenke, D. W.; Truhlar, D. G. *J. Phys. Chem. A* **1999**, *103*, 6309 – 6326.
- [156] Fazzi, D.; Barbatti, M.; Thiel, W. *J. Am. Chem. Soc.* **2016**, *138*, 4502 – 4511.
- [157] Colbert, D. T.; Miller, W. H. *J. Chem. Phys.* **1992**, *96*, 1982 – 1991.
- [158] Sakurai, S. S.; Napolitano, J. J. *Modern Quantum Mechanics*; Pearson: San Fransico, CA, 2013.
- [159] Kohlhoff, K. J.; Shukla, D.; Lawrenz, M.; Bowman, G. R.; Konerding, D. E.; Belov, D.; Altman, R. B.; Pande, V. S. *Nat. Chem.* **2014**, *6*, 15 – 21.
- [160] Li, H.; Robertson, A. D.; Jensen, J. H. *Proteins: Struct., Funct., Bioinf.* **2005**, *61*, 704 – 721.
- [161] Olsson, M. H. M.; Sondergaard, C. R.; Rostkowski, M.; Jensen, J. H. *J. Chem. Theory Comput.* **2011**, *7*, 525 – 537.
- [162] Bas, D. C.; Rogers, D. M.; Jensen, J. H. *Proteins: Struct., Funct., Bioinf.* **2008**, *73*, 765 – 783.
- [163] Sondergaard, C. R.; Olsson, M. H. M.; Rostkowski, M.; Jensen, J. H. *J. Chem. Theory Comput.* **2011**, *7*, 2284 – 2295.
- [164] Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926 – 935.
- [165] Reuter, N.; Lin, H.; Thiel, W. *J. Phys. Chem. B* **2002**, *106*, 6310 – 6321.
- [166] Simkovitch, R.; Huppert, A.; Huppert, D.; Remington, S. J.; Miller, Y. *J. Phys. Chem. B* **2013**, *117*, 11921 – 11931.
- [167] Cheng, C. L.; Zhang, M. Z.; Zhao, G. J. *RSC Adv.* **2014**, *4*, 6513 – 6516.
- [168] Wineman Fisher, V.; Simkovitch, R.; Shomer, S.; Gepshtein, R.; Huppert, D.; Saif, M.; Kallio, K.; Remington, S. J.; Miller, Y. *Phys. Chem. Chem. Phys.* **2014**, *16*, 11196 – 11208.
- [169] Thiel, W. *MNDO99 Program, Version 7.0*; Max-Planck-Institut für Kohlenforschung, Mülheim, Germany, 2005.
- [170] 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.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.;

- Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09 Revision C.01*; Gaussian Inc. Wallingford CT, 2009.
- [171] Sanchez Garcia, E.; Doerr, M.; Thiel, W. J. *Comput. Chem.* **2010**, *31*, 1603 – 1612.
- [172] Nadal Ferret, M.; Gelabert, R.; Moreno, M.; Lluch, J. M. *J. Chem. Theory Comput.* **2013**, *9*, 1731 – 1742.
- [173] Schaefer, L. V.; Groenhof, G.; Klingen, A. R.; Ullmann, G. M.; Boggio Pasqua, M.; Robb, M. A.; Grubmueller, H. *Angew. Chem. Int. Ed.* **2007**, *46*, 530 – 536.
- [174] Parac, M.; Doerr, M.; Marian, C. M.; Thiel, W. J. *Comput. Chem.* **2010**, *31*, 90 – 106.
- [175] Amat, P.; Nifosi, R. J. *Chem. Theory Comput.* **2013**, *9*, 497 – 508.
- [176] Beerepoot, M. T. P.; Steindal, A. H.; Kongsted, J.; Brandsdal, B. O.; Frediani, L.; Ruud, K.; Olsen, J. M. H. *Phys. Chem. Chem. Phys.* **2013**, *15*, 4735 – 4743.
- [177] Murugan, N. A.; Kongsted, J.; Rinkevicius, Z.; Agren, H. *Phys. Chem. Chem. Phys.* **2012**, *14*, 1107 – 1112.
- [178] Petrone, A.; Caruso, P.; Tenuta, S.; Rega, N. *Phys. Chem. Chem. Phys.* **2013**, *15*, 20536 – 20544.
- [179] Randino, C.; Moreno, M.; Gelabert, R.; Lluch, J. M. *J. Phys. Chem. B* **2012**, *116*, 14302 – 14310.
- [180] Rhee, Y. M.; Head Gordon, M. J. *Phys. Chem. A* **2007**, *111*, 5314 – 5326.
- [181] Shao, Y. *et al. Mol. Phys.* **2015**, *113*, 184 – 215.
- [182] Dreuw, A.; Head Gordon, M. *Chem. Rev.* **2005**, *105*, 4009 – 4037.
- [183] Martin, R. L. *J. Chem. Phys.* **2003**, *118*, 4775 – 4777.
- [184] Foster, J. P.; Weinhold, F. J. *Am. Chem. Soc.* **1980**, *102*, 7211 – 7218.
- [185] Armengol, P.; Gelabert, R.; Moreno, M.; Lluch, J. M. *Org. Biomol. Chem.* **2014**, *12*, 9845 – 9852.

- [186] Armengol, P.; Gelabert, R.; Moreno, M.; Lluch, J. M. *Phys. Chem. Chem. Phys.* **2016**, *18*, 16964 – 16976.
- [187] Granucci, G.; Persico, M.; Zocante, A. *J. Chem. Phys.* **2010**, *133*, 134111.
- [188] Spoerkel, L.; Thiel, W. *J. Chem. Phys.* **2016**, *144*, 194108.
- [189] Spitzner, N.; Lohr, F.; Pfeiffer, S.; Koumanov, A.; Karshikoff, A.; Ruterjans, H. *Eu. Biophys. J.* **2001**, *30*, 186 – 197.
- [190] Bell, A. F.; He, X.; Wachter, R. M.; Tonge, P. J. *Biochemistry* **2000**, *39*, 4423–4431.
- [191] Nielsen, S. B.; Lapierre, A.; Andersen, J. U.; Pedersen, U. V.; Tomita, S.; Andersen, L. H. *Phys. Rev. Lett.* **2001**, *87*, 228102.
- [192] Fayngold, M.; Fayngold, V. *Quantum Mechanics and Quantum Information*; Wiley-WCH: , 2013.
- [193] Popper, K. *Logik der Forschung*; Mohr Siebeck: Heidelberg, 1935.
- [194] Zhang, Q.; Chen, X.; Cui, G.; Fang, W. H.; Thiel, W. *Angew. Chem. Int. Ed.* **2014**, *53*, 8649 – 8653.
- [195] Ortiz Sanchez, J. M.; Gelabert, R.; Moreno, M.; Lluch, J. M. *ChemPhysChem* **2010**, *11*, 3696 – 3703.
- [196] Moreno, M.; Manuel-Ortiz Sanchez, J.; Gelabert, R.; Lluch, J. M. *Phys. Chem. Chem. Phys.* **2013**, *15*, 20236 – 20246.
- [197] Levine, B. G.; Ko, C.; Quenneville, J.; Martinez, T. J. *Mol. Phys.* **2006**, *104*, 1039 – 1051.
- [198] Sadeghian, K.; Schuetz, M. *J. Am. Chem. Soc.* **2007**, *129*, 4068 – 4074.
- [199] Werner, U.; Mitric, R.; Suzuki, T.; Bonacic Koutecky, V. *Chem. Phys.* **2008**, *349*, 319 – 324.
- [200] Tapavicza, E.; Tavernelli, I.; Rothlisberger, U. *Phys. Rev. Lett.* **2007**, *98*, 023001.
- [201] Armengol, P.; Gelabert, R.; Moreno, M.; Lluch, J. M. *J. Phys. Chem. B* **2015**, *119*, 2274 – 2291.
- [202] Armengol, P.; Gelabert, R.; Moreno, M.; Lluch, J. M. *ChemPhysChem* **2015**, *16*, 2134 – 2139.





V

APÈNDIX



# List of Abbreviations

<b>ALDA</b>	Adiabatic Local Density Aproximation
<b>BO</b>	Born Oppenheimer
<b>CI</b>	Conical Intersection
<b>CIS</b>	Configuration Interaction (with) Singles (excitations)
<b>CISD</b>	Configuration Interaction (with) Singles (and) Doubles (excitations)
<b>CISDT</b>	Configuration Interaction (with) Singles, Doubles (and) Triples (excitations)
<b>CISDTQ</b>	Configuration Interaction (with) Singles, Doubles, Triples (and) Quadruples (excitations)
<b>CNDO</b>	Complete Neglect (of) Differential Overlap
<b>CPU</b>	Central Processing Unit
<b>DFT</b>	Density Functional Theory
<b>EE</b>	Electronic Embedding
<b>ESPT</b>	Excited State Proton Transfer
<b>FP</b>	Fluorescent Protein
<b>FSA</b>	Fewest Switches Approach
<b>GFP</b>	Green Fluorescent Protein
<b>HBDI</b>	4-( <i>p</i> -hydroxybenzildie)imidazolid-5-ona)
<b>HOMO</b>	Highest Occupied Molecular Orbital
<b>INDO</b>	Intermediate Neglect (of) Differential Overlap
<b>IR</b>	Infrared

<b>KIE</b>	<b>Kinetic Isotope Effect</b>
<b>LBHB</b>	<b>Low Barrier Hydrogen Bond</b>
<b>LUMO</b>	<b>Lowest Unoccupied Molecular Orbital</b>
<b>MD</b>	<b>Molecular Dynamics</b>
<b>ME</b>	<b>Mechanical Embedding</b>
<b>MM</b>	<b>Molecular Mechanics</b>
<b>MR</b>	<b>Multi Referential</b>
<b>NBO</b>	<b>Natural Bond Orbital</b>
<b>NDDO</b>	<b>Neglect (of) Diatomic Differential Overlap</b>
<b>NMR</b>	<b>Nuclear Magnetic Resonance</b>
<b>NTO</b>	<b>Natural Transition Orbital</b>
<b>PBC</b>	<b>Periodic Boundary Conditions</b>
<b>PCM</b>	<b>Polarizable Continuum Model</b>
<b>PDB</b>	<b>Protein Data Bank</b>
<b>PES</b>	<b>Potential Energy Surface</b>
<b>QD</b>	<b>Quantum Dynamics</b>
<b>QM</b>	<b>Quantum Mechanics</b>
<b>QY</b>	<b>Quantum Yield</b>
<b>RFP</b>	<b>Red Fluorescent Protein</b>
<b>RMS</b>	<b>Root Mean Square</b>
<b>SIE</b>	<b>Self Interaction Error</b>
<b>SOS</b>	<b>Scaled Opposite Spin</b>
<b>TD-</b>	<b>Time Dependent</b>

*wt-*      wild type



# Índex de figures

1.1	En l'exemple de Hawking i Mlodinov, les possibles lleis científiques que desenvolupés el peix des de dins la peixera, que serien diferents a les nostres, serien igualment vàlides per al seu model de realitat. Extret de Hawking <i>et al.</i> <sup>4</sup> . . . .	5
1.2	D'esquerra a dreta: la medusa <i>Aequorea victoria</i> , l'estructura de barril $\beta$ de la GFP i l'estructura molecular del cromòfor .	8
1.3	Fons: Imatge microscòpica de la GFP expressada en teixit vegetal. Primer pla: Ratolins en els quals ha estat expressada la GFP al costat de ratolins no modificats, tots il·luminats amb llum ultraviolada. . . . .	9
1.4	Representació esquemàtica del mecanisme de fluorescència de la GFP amb els possibles estats A, I o B implicats. L'excitació de l'espècie A porta a A*, que evoluciona fins a I* per mitja de l'ESPT. L'espècie I* és l'emissora de la fluorescència ja que la transformació de I* a B* és un procés lent. Extret de Lill <i>et al.</i> <sup>46</sup> . . . . .	10
1.5	Representació de l'estructura del cromòfor juntament amb els altres actors del cable de protons que fa possible la ESPT: la molècula d'aigua, etiquetada com a W22 per raons de nomenclatura, la Treonina 203, la Serina 205 i el Glutamat 222, que serà l'acceptor final del protó. Extret de Di Donato <i>et al.</i> <sup>52</sup>	11
1.6	Les diferents estructures dels cromòfors comporten diferents longituds d'ona d'absorció i d'emissió. Es mostren els cromòfors de les principals famílies de proteïnes amb la zona de l'espectre on absorbeixen (barra superior) i on emeten (barra inferior). Extret de Acharya <i>et al.</i> <sup>12</sup> . . . . .	12
1.7	Representació de l'estructura del cromòfor de les RFPs. S'hi pot apreciar l'extensió de la conjugació pel grup C=N acilimínic. També s'hi representa el residu Phe62, que també influeix a la posició del màxim de la banda d'absorció. . . .	14
1.8	Representació esquemàtica del procés proposat per Kondo <i>et al.</i> : després d'absorbir la radiació de 411 nm es dona una transferència ultrarràpida ( 100 fs) i sense barrera, seguida d'una relaxació que ocorre en l'ordre dels picosegons. Extret de Kondo <i>et al.</i> <sup>86</sup> . . . . .	17



2.1	En el tractament <i>charge-shift</i> , la càrrega del primer àtom de la part MM ( $M^1$ ) que fa frontera amb la part QM ( $Q^1$ ), es reparteix amb els àtoms MM contigus ( $M^2$ ). Parelles de càrregues addicionals es situen a prop dels àtoms $M^2$ per a restaurar els dipols $M^1$ - $M^2$ . . . . .	30
2.2	Representació model de dues PES en funció dels vectors $\vec{g}$ i $\vec{h}$ . La CI es localitza en el punt on es troben i es produeix la degeneració. . . . .	36
4.1	Representació esquemàtica de (a) un protocol estàndard per a la realització d'una simulació MD i de (b) el protocol <i>task-farming</i> . . . . .	49
5.1	Esquerra: Representació gràfica del model utilitzat, que inclou tots els àtoms de la proteïna i la gota d'aigües de 60 Å de diàmetre. Dreta: Etiquetats per A, B i C, el cromòfor i els residus propers al mateix per a les proteïnes <i>wt</i> -GFP, S65T/H148D i E222Q/H148D, respectivament. El camí del cable de protons està ressaltat en verd. . . . .	52
5.2	Distribució de la distància entre els oxígens del cromòfor i l'Asp148 al llarg del nanosegon de simulació MD per les mutants E222Q/H148D (blau) i S65T/H148D (verd). . . . .	55
5.3	Distribució de la l'angle dièdre entre els dos anells del cromòfor al llarg del nanosegon de simulació MD per les mutants E222Q/H148D (blau) i S65T/H148D (verd). . . . .	56
5.4	Espectres d'absorció calculats per a la <i>wt</i> -GFP (verd), E222Q/H148D (blau) i S65T/H148D (lila). . . . .	57
5.5	Representació de la longitud d'ona d'absorció de l'estat fotoactiu en funció de la distància entre cromòfor i Asp148. Es pot observar una correlació negativa. . . . .	58
5.6	Orbitals HOMO (superior) i LUMO (inferior) involucrats en la excitació $\pi\pi^*$ de la mutant S65T/H148D, descrita a nivell TD-DFT. Els orbitals per a la mutant E222Q/H148D són essencialment iguals i no es mostren per evitar redundància. . . . .	59
5.7	Variació de la càrrega dels diferents fragments del cromòfor durant el procés de fotoexcitació calculada a partir de les càrregues NBO de l'estat fonamental i excitat fotoactiu. En blau es mostren els fragments que perden càrrega electrònica i en vermell els fragments que en guanyen. . . . .	59
5.8	Etiquetatge dels àtoms en els cromòfors de la mNeptune i mCardinal. Els àtoms sense etiqueta (amb els seus paràmetres associats) es poden consultar a la bibliografia. <sup>165</sup> . . . . .	61
5.9	Espectre d'absorció calculat a nivell QM/MM a partir dels 1000 fotogrames extrets dels 100 ns de la simulació MM-MD per a les proteïnes mNeptune (vermell) i mCardinal (blau). Les etiquetes I i II identifiquen els diferents pics. . . . .	64

5.10	Representació de la part QM presa en els càlculs QM/MM i NTOs associats a les 3 transicions amb forces d'oscil·lador considerables, amb la seva correspondència a les bandes de l'espectre d'absorció de la mNeptune. . . . .	67
5.11	Representació de la part QM presa en els càlculs QM/MM i dels principals orbitals moleculars canònics involucrats en les excitacions de la Banda I i II de la mNeptune. . . . .	68
5.12	Espectres d'absorció de les 5 proteïnes estudiades en aquest capítol: la <i>wt</i> -GFP (verd), es dues dobles mutants S65T/H148D i E222Q/H148D (blau clar i lila respectivament) i les RFPs mNeptune i mCardinal (vermell i blau, respectivament). . .	69
5.13	Representació gràfica del cromòfor de les RFPs on s'ha etiquetat els àtoms a partir dels quals es defineix l'angle <i>tilt</i> $\tau$ (angle díedre format pels àtoms $\alpha - \beta - \gamma - \delta$ ) i <i>twist</i> $\phi$ (angle díedre format pels àtoms $\beta - \gamma - \delta - \epsilon$ ) . . . . .	70
5.14	Anàlisi de la planaritat del cromòfor en les simulacions de MD en termes de la distribució dels angles $\tau$ i $\phi$ per a la mNeptune (panell esquerra) i mCardinal (panell dret). . . .	71
5.15	Panell esquerra: dibuix d'un fotograma representatiu de la simulació MD de la mNeptune. S'hi mostra el cromòfor, l'oxigen de la Phe62, la Gly41 i les aigües que ocupen el forat que deixa aquesta última. Panell dret: dibuix d'un fotograma representatiu de la simulació MD de la mCardinal. S'hi mostra el cromòfor, la Gln41, l'oxigen de la Phe62 i l'aigua més propera al mateix. . . . .	72
5.16	Distribució de distàncies al llarg dels 100 ns de simulació MD. En blau es mostra la distància entre l'oxigen de la Phe62 i el nitrogen de la Gly41 per la mCardinal i en vermell la distància entre el mateix oxigen i el $C_\alpha$ per la mNeptune. . . . .	73
5.17	Estructura de solvatació de l'oxigen de la Phe62 per a la mNeptune (panell esquerra) i la mCardinal (panell dret). Es mostren les funcions de distribució radial per a les aigües més properes a l'oxigen de la Phe62 en cada cas. . . . .	73
5.18	Esquematzació de les possibles estratègies a seguir per aconseguir un desplaçament cap al roig de la longitud d'ona d'absorció en les FFPs. Es mostra el cromòfor de les RFPs, però qualitativament el cromòfor de les GFPs és equivalent. . . . .	75
5.19	Espectre d'absorció calculat per a les mutants proposades. S'ha utilitzat una convolució amb funcions gaussianes per reduir el soroll estadístic amb una $\omega=10$ nm. . . . .	78
5.20	Espectre d'absorció calculat per a les mutants proposades amb els mètodes OM2 (CIS, CISD i CISDT), on també s'hi inclouen els espectres calculats a nivell TD-DFT. S'ha utilitzat una convolució amb funcions gaussianes per reduir el soroll estadístic amb una $\omega=10$ nm. . . . .	80

- 6.1 Perfils d'energia potencial per a la transferència protònica per a la mutant S65T/H148D-GFP per a diferents estats electrònics a nivell TD-DFT/CHARMM27. Esquerra: Perfil realitzat a partir de l'estructura cristal·lina del PDB però en el model de solvatació QM/MM explicat. S'hi mostra l'estat fonamental ( $S_0$ ) i els primers 3 estats excitats ( $S_1$ ,  $S_2$  i  $S_3$ , essent el segon l'estat fotoactiu (etiquetat com a  $\pi\pi^*$ ). Dreta: Perfil realitzat a partir de l'estructura optimitzada a l'estat fonamental a nivell QM/MM. S'hi mostra l'estat fonamental,  $S_0$ , i l'estat fotoactiu,  $S_1$  en aquest cas (etiquetat com a  $\pi\pi^*$ ). Per tal de facilitar la comparació, en línies puntejades, també s'ha representat els estats corresponents de l'estructura cristal·lina del PDB. . . . . 88
- 6.2 Distribució de l'angle que forma el grup carboxílic de l'Asp148 amb el pla del cromòfor durant el 1,0 ns de simulació QM/MM. 92
- 6.3 Anàlisi de l'energètica de la transferència protònica a l'estat fonamental i excitat fotoactiu, realitzada a partir dels 500 perfils per a la transferència entre el cromòfor i l'Asp148, obtinguts a nivell TD-DFT/CHARMM27. L'eix d'abscisses representa la exo- o endoergicitat del procés a l'estat fonamental ( $\Delta E$ ) per a cada perfil i el d'ordenades fa el mateix per a l'estat excitat ( $\Delta E^*$ ). Els punts vermells representen aquells perfils etiquetats com a "no reactius" ( $\Delta E^* > 0$ ), els punts blaus aquells etiquetats com a "reactius" ( $\Delta E^* < 0$ ) i els punts verds aquells en els quals, ja a l'estat fonamental, el protó es troba a l'Asp148 ( $\Delta E < 0$ ). Els histogrames que acompanyen l'eix d'abscisses i ordenades representen la distribució de cada valor ( $\Delta E$  i  $\Delta E^*$ ). . . . . 93
- 6.4 Histograma de la distribució de distàncies entre els oxígens fenòlics del cromòfor i de la Tyr145 durant el 1,0 ns de simulació dinàmica QM/MM. . . . . 95
- 6.5 Funcions de distribució radial de les molècules d'aigua respecte els diferents oxígens indicats a la llegenda, obtingudes en la simulació QM/MM de la mutant S65T/H148D . . . . . 96
- 6.6 Perfils d'energia potencial per l'estat fonamental (vermell) i l'estat fotoactiu (verd) obtinguts a nivell DFT/CHARMM27 i TDDFT/CHARMM27, respectivament, a partir de 3 fotogrames seleccionats de la simulació dinàmica QM/MM. Al panell esquerra, es mostra un perfil típic "no reactiu" ( $\Delta E^* > 0$ ). Al panell central es mostra un cas típic "reactiu" ( $\Delta E^* < 0$ ). Al panell dret es mostra un perfil "reactiu" amb una  $\Delta E^*$  considerable. En vermell es mostren els perfils a l'estat fonamental i en verd el perfil a l'estat fotoactiu. Les línies horitzontals negres representen l'energia de punt 0 i les línies discontinues representen la densitat de probabilitat ( $|\psi|^2$ ) per a les funcions d'ona del protó. . . . . 97

- 6.7 Representació de les funcions de probabilitat de supervivència en funció del temps, calculades sobre els mateixos perfils de la Figura 6.6. Les línies contínues representen la dinàmica del protó i les discontinues les del deuteró. . . . . 99
- 6.8 Aquí es mostra la superposició d'un fotograma representatiu de la simulació QM/MM per les dues mutants, S65T/H148D i S65T/H148D/Y145F, de color blau fosc i cian, respectivament. Al panell esquerre es pot observar com la Phe145 tendeix a estar més allunyada en la triple mutant que la Tyr145 en la doble. Això causa la distorsió de la fulla  $\beta$  que implica la coplanaritat de l'Asp148 amb el cromòfor. Com es pot observar, en el cas de la doble mutant, l'Asp148 roman essencialment perpendicular al mateix. Al panell dret, on es mostra el mateix fotograma però des d'un altre punt de vista, es pot apreciar l'anomenada distorsió de la fulla  $\beta$  causada per la pèrdua de la interacció de la Tyr145 amb el cromòfor. 101
- 6.9 Distribució de l'angle díedre entre el pla del cromòfor i el grup carboxílic de l'Asp148 al llarg del 1,0 ns de simulació QM/MM per a les mutants S65T/H148D (línia contínua) i S65T/H148D/Y145F (línia discontinua). . . . . 102
- 6.10 Funcions de distribució radial de les molècules d'aigua respecte els diferents oxígens indicats a la llegenda, obtingudes en la simulació QM/MM de la triple mutant. . . . . 103
- 6.11 Anàlisi de l'energètica de la transferència protònica a l'estat fonamental i excitat fotoactiu, realitzada a partir dels 500 perfils per a la transferència entre el cromòfor i l'Asp148, obtinguts a nivell TD-DFT/CHARMM27, per a les mutants S65T/H148D (vermell) i S65T/H148D/Y145F (verd). L'eix d'abcises representa la exo- o endoergicitat del procés a l'estat fonamental ( $\Delta E$ ) per a cada perfil i el d'ordenades fa el mateix per a l'estat excitat ( $\Delta E^*$ ). . . . . 104
- 6.12 Perfils d'energia potencial obtinguts a nivell TDDFT/CHARMM27 a partir de 3 fotogrames seleccionats de la simulació dinàmica QM/MM per a la triple mutant. Al panell esquerre es mostra un perfil especialment endoèrgic, al panell central es mostra un cas típic, i al panell dret se'n mostra un d'especialment exoèrgic. Les línies horitzontals negres representen l'energia de punt 0 i les línies discontinues representen la densitat de probabilitat ( $|\Psi|^2$ ) per a les funcions d'ona del protó. . . . . 104
- 6.13 Etiquetatge dels àtoms que s'utilitzarà a la discussió que segueix. . . . . 106

- 6.14 En vermell, representació del nombre total de trajectòries que es transfereixen de manera no adiabàtica des de l'estat excitat fotoactiu al fonamental via la HT-CI ( $N_{HT}$ ) o la P-CI ( $N_P$ ). Superposada, en marró, el nombre de trajectòries que es transferixen via la P-CI. La diferència entre les capses vermelles i marrons és la quantitat de sistemes que es transfereixen per mitjà de la HT-CI. En blau, es representa la funció fracció de població de l'estat excitat fotoactiu en funció del temps. . . . . 107
- 6.15 Representació gràfica de l'estructura de les dues CIs que porten a l'estat fonamental de manera no radiativa: (HT), que comporta la rotació de l'enllaç  $C_4C_6$  amb la torsió del  $H_7$  i (P), que comporta la piramidalització del  $C_2$ . . . . . 108
- 6.16 Valors dels angles  $\theta$  i  $\rho$  en el moment de la transició a l'estat fonamental. Els punts vermells corresponen a estructures tipus HT-CI i els punts blaus a estructures tipus P-CI. Els punts grisos corresponen a estructures que tenen un caràcter mixt. . . . . 109
- 6.17 Coordenada de reacció a l'estat excitat fotoactiu calculada a nivell QM/MM per a la torsió del díedre  $H_7C_6C_4N_5$  ( $\theta$ ). Les línies contínues corresponen a l'estat excitat fotoactiu (on s'ha optimitzat les estructures) i les discontinües a l'estat fonamental. Les línies blaves corresponen al càlcul MRCISD-OM2/CHARMM27 i les vermelles al càlcul TD-DFT(CAM-B3LYP)/CHARMM27. . . . . 110
- 6.18 Coordenada de reacció a l'estat excitat fotoactiu calculada a nivell MRCISD-OM2/CHARMM27 per a la torsió del díedre  $C_4C_2O_3N_1$  ( $\rho$ ). La línia vermella correspon a l'energia de l'estat fonamental i la línia blava a l'energia de l'estat excitat fotoactiu . . . . . 111
- 6.19 Panell esquerre: Espectres de fluorescència calculats a diferents temps indicats ( $\Delta t$ ) després de la fotoexcitació. Panell dret: Evolució temporal de la fluorescència on els punts més foscos representen més intensitat. La línia discontinüa de color verd segueix aproximadament el màxim de fluorescència (la moda, no la mitjana, de la distribució). . . . . 112
- 6.20 Evolució temporal de l'energia d'emissió mitjana. La línia contínua és una funció biexponencial ajustada. . . . . 113
- 6.21 Intensitat de la fluorescència resolta en el temps a diferents longituds d'ona seleccionades. Extret de Kondo *et al.*<sup>86</sup> . . . . 114
- 6.22 Simulació de la intensitat de fluorescència resolta en el temps publicada per Kondo *et al.*<sup>86</sup> El bloc gris representa la fracció de temps fora del límit de detecció experimental. . . . . 114
- 6.23 Evolució temporal de la distància del protó al centre de la distància donador-acceptor (Equació 6.5, vermell) i de la distància entre el cromòfor i l'Asp148 (blau) . . . . . 115

- 
- 6.24 Evolució temporal dels angles  $\tau$  (díedre  $C_8C_6C_4C_2$ , vermell),  $\phi$  (díedre  $C_{10}C_8C_6C_4$ , blau) i  $\varphi$  (díedre  $O_{16}O_{15}C_{11}C_{12}$ , verd). 117
- 6.25 Panell esquerre: histogramització de l'evolució temporal de la distància entre l'oxigen fenòlic del cromòfor i de la Tyr145 al llarg de la dinàmica. Panell dret: evolució temporal de la mitjana d'aquest valor. . . . . 117
- 6.26 Posició del protó (usant l'Equació 6.5) en el moment del trànsit per alguna de les CIs, ja siguin del tipus HT-CI (vermell) o P-CI (blau). . . . . 118
- 6.27 Relació entre la posició del protó (Equació 6.5) i l'energia d'emissió, en les etapes primerenques del procés. Els panells superiors descriuen el valor de l'energia d'emissió de cada estructura de la simulació MD respecte el paràmetre  $r_H$  pels temps indicats. En el panell inferior es representen tots els temps alhora en els seus colors corresponents, així com el camí seguit per l'energia d'excitació mitjana durant l'interval seleccionat, en color vermell i direcció descendent. 119



## Índex de taules

5.1	Paràmetres refinats per al cromòfor de les RFPs mNeptune i mCardinal. La resta de paràmetres, prèviament publicats, es poden trobar a la bibliografia. <sup>165</sup> . . . . .	62
5.2	Relació dels estats TD-DFT amb els SOS-CIS(D). Per a fer tal relació s'ha analitzat els pesos i participacions de cadascuna de les excitacions pels dos mètodes, que no es mostren aquí. Les excitacions marcades amb un asterisc són aquelles que més contribueixen a les bandes de l'espectre calculat. . . . .	66
5.3	Càrregues NBO en el cromòfor de les RFPs per l'estat electrònic fonamental i excitat corresponent a la Banda I. . . . .	70
5.4	Relació de residus més propers als fragments fenòlic i acilimínic del cromòfor de la mCardinal amb les seves distàncies, preses de l'estructura cristal·lina resolta per rajos X. L'etiquetatge dels àtoms és el propi del camp de forces. La indicació "fora" significa que el residu en qüestió apunta en direcció oposada al cromòfor. . . . .	76
5.5	Màxims de les longituds d'ona d'absorció per a cadascuna de les mutants estudiades i per a cada mètode de càlcul emprat. . . . .	78
6.1	Longituds d'ona d'absorció (en nm) de la forma neutre i aniònica de l'HBDI en aigua i del cromòfor en la doble mutant S65T/H148D . . . . .	90