



Università degli Studi “*La Sapienza*”
Facoltà di Scienze Matematiche Fisiche e Naturali
Corso di Laurea in Chimica

**STUDIO DELLE PROPRIETA’
SPETTROSCOPICHE DELLA
LUCIFERINA MEDIANTE L’USO DI
METODI QUANTO-MECCANICI E DI
DINAMICA MOLECOLARE**

Tirocinio

Relatore

Prof. Alfredo Di Nola

Co-Relatore

Prof. Osvaldo Lanzalunga

Relatore esterno

Dr. Andrea Amadei

Laureando

Simone Marocchi

matricola:698260

Anno Accademico 2008/2009

Alla mia famiglia

“Lucifero è il re del caos, ma lui odia il disordine.”

John Doe

Indice

Elenco delle figure	11
Introduzione	13
1 La Bioluminescenza nella Lucciola	17
1.1 La Lucciola	17
1.2 Il Meccanismo Reattivo	18
1.2.1 La D-luciferina	18
1.2.2 La trasformazione di D-luciferina in Ossiluciferina	19
1.3 Il mistero della Ossiluciferina	23
2 La Dinamica Molecolare	27
2.1 Generalità sulla Dinamica Molecolare	27
2.2 Nascita e sviluppo della Dinamica Molecolare	27
2.3 Equazioni del moto	29
2.4 La funzione di potenziale	29
2.5 Condizioni al contorno	32
2.6 Calcolo delle interazioni a lungo raggio	33
2.7 Vincoli in MD	36
2.8 Le condizioni iniziali	37
2.9 Integrazione delle equazioni del moto	38

2.10	Metodi per il controllo della temperatura e della pressione	40
3	Il Metodo della Matrice Perturbata (PMM)	45
3.1	Introduzione al PMM	45
3.2	Metodi Quanto-Meccanici	46
3.2.1	L'approssimazione di Born-Oppenheimer	47
3.2.2	Il Metodo di Hartree-Fock	48
3.2.3	I Metodi Post Hartree-Fock	55
3.3	Il PMM	61
3.3.1	La teoria di base	61
3.3.2	Il calcolo dello spettro di assorbimento	65
4	Risultati sullo studio del force-field	69
4.1	Parametrizzazione delle molecole	69
4.1.1	Il calcolo delle cariche	69
4.1.2	L'angolo diedro C2-C2'	70
4.2	La distribuzione degli angoli diedri C2-C2'	72
4.3	Parametri della Simulazione di Dinamica Molecolare	75
4.4	Parametri calcolo quantistico	76
5	Risultati sulle proprietà elettroniche	79
5.1	Lo spettro della D-luciferina	79
5.2	Lo spettro della Ossiluciferina	82
6	Conclusioni	93
	Bibliografia	97

Elenco delle figure

1.0.1	Lo spettro del visibile	17
1.1.1	maschio di <i>Luciola cruciata</i> , ritratto in giappone	18
1.2.1	DLH_2	19
1.2.2	a sinistra lo spettro di assorbimento della DLH_2 a $pH \sim 5$, e a destra quello a $pH \sim 11$	20
1.2.3	il meccanismo di reazione	21
1.2.4	Le due forme cheto ed enolica della $OxyLH_2$	21
1.3.1	Le possibili forme chimiche trans-C2-C2' della $OxyLH_2$	23
4.1.1	funzione esempio $y = 0.5 [1 + \cos(\frac{2\pi x}{180} - \pi)]$	71
4.2.1	distribuzione di probabilità dell'angolo diedro C2-C2' per la DLH_2 in acqua	72
4.2.2	distribuzione di probabilità dell'angolo diedro C2-C2' per la cheto- $OxyLH_2$	73
4.2.3	distribuzione di probabilità dell'angolo diedro C2-C2' per il fenolato-cheto- $OxyLH^-$	73
4.2.4	distribuzione di probabilità dell'angolo diedro C2-C2' per la enol- $OxyLH_2$ in DMSO	74
4.2.5	distribuzione di probabilità dell'angolo diedro C2-C2' per la fenolato-enol- $OxyLH^-$ in acqua	74

5.1.1	In nero lo spettro di assorbimento UV sperimentale della DLH_2 a $\text{pH}\sim 5$ ($\lambda_{Abs1} = 262 \text{ nm}$ e $\lambda_{Abs2} = 328 \text{ nm}$) e in rosso lo spettro ottenuto dal PMM ($\lambda_{PMM1} = 258 \text{ nm}$ e $\lambda_{PMM2} = 332 \text{ nm}$)	80
5.1.2	La scomposizione del massimo di assorbimento a $\lambda_{PMM1} = 258 \text{ nm}$ della DLH_2 negli assorbimenti dovuti alle singole eccitazioni	81
5.1.3	La scomposizione del massimo di assorbimento a $\lambda_{PMM2} = 332 \text{ nm}$ della DLH_2 negli assorbimenti dovuti alle singole eccitazioni	81
5.2.1	In nero lo spettro di assorbimento UV sperimentale del $DMOxyLH_2$ in DMSO ($\lambda_{Abs1} = 268 \text{ nm}$ e $\lambda_{Abs2} = 390 \text{ nm}$) e in rosso lo spettro del cheto- $OxyLH_2$ ottenuto con il PMM ($\lambda_{PMM1} = 255 \text{ nm}$ e $\lambda_{PMM2} = 345 \text{ nm}$).	84
5.2.2	La scomposizione del picco a $\lambda_{PMM1} = 255 \text{ nm}$ della cheto- $OxyLH_2$ in DMSO negli assorbimenti dovuti ai singoli stati eccitati	85
5.2.3	La scomposizione del picco a $\lambda_{PMM2} = 345 \text{ nm}$ della cheto- $OxyLH_2$ in DMSO negli assorbimenti dovuti ai singoli stati eccitati	85
5.2.4	In nero lo spettro di assorbimento UV sperimentale del $DMOxyLH^-$ in acqua a $\text{pH}\sim 7$ ($\lambda_{Abs1} = 360 \text{ nm}$ e $\lambda_{Abs2} = 487 \text{ nm}$) e in rosso lo spettro del fenolato-cheto- $OxyLH^-$ ottenuto con il PMM ($\lambda_{PMM1} = 360 \text{ nm}$ e $\lambda_{PMM2} = 487 \text{ nm}$)	86
5.2.5	La scomposizione del picco a $\lambda_{PMM1} = 360 \text{ nm}$ della fenolato-cheto- $OxyLH^-$ negli assorbimenti dovuti ai singoli stati eccitati	87
5.2.6	La scomposizione del picco a $\lambda_{PMM2} = 487 \text{ nm}$, della fenolato-cheto- $OxyLH^-$ negli assorbimenti dovuti ai singoli stati eccitati	87

5.2.7 In nero lo spettro di assorbimento UV sperimentale dell' enol- $OxyLH_2$ in DM-SO ($\lambda_{Abs} = 377 \text{ nm}$) e in rosso lo spettro dell' enol- $OxyLH_2$ ottenuto con il PMM ($\lambda_{PMM} = 362 \text{ nm}$)	88
5.2.8 In nero lo spettro del fenolato-enol- $OxyLH^-$ in acqua a pH 7,8 ($\lambda_{Abs} = 410 \text{ nm}$) e in verde lo spettro PMM del fenolato-enol- $OxyLH^-$ ($\lambda_{PMM} = 433 \text{ nm}$) . . .	89
5.2.9 La scomposizione del picco a $\lambda_{PMM} = 433 \text{ nm}$ della fenolato-enol- $OxyLH^-$ negli assorbimenti dovuti ai singoli stati eccitati	90

Introduzione

L'emissione di luce da animali e piante ha ispirato la curiosità e l'interesse di moltissime persone sin dai tempi di Aristotele (384 – 322 A.C.) e Plinio (23 – 79 D.C.). E' stato l'obiettivo della ricerca di un gran numero di naturalisti, fisici, e fisiologi. Si pensa che il termine "Bioluminescenza" (luminescenza da organismi viventi) sia stato coniato nel 1916 dal biologo e fisiologo americano E. Newton Harvey [1] che si può considerare come il capostipite degli studi di questo affascinante meccanismo. Oggi, le reazioni bioluminescenti sono utilizzate come strumento di analisi indispensabile in vari campi della scienza. Ad esempio, il sistema bioluminescente della lucciola (composto da un enzima chiamato **Luciferasi** e un coenzima chiamato **D-Luciferina**) è un sistema universalmente usato come metodo di misurazione dell'ATP all'interno di organismi viventi, mentre le foto-proteine che necessitano di Ca^{2+} , come quelle delle meduse, sono ampiamente utilizzate per monitorare la concentrazione di Ca^{2+} intracellulare che regola fondamentali processi biologici; alcuni analoghi della luciferina di *Cypridina* (un piccolo crostaceo) sono impiegati come sonde per misurare anioni superossido, sostanze importanti per i sistemi biologici ma piuttosto elusive. Tali esempi costituiscono solo alcune delle applicazioni dei sistemi bioluminescenti che sono derivati dagli studi chimici di bioluminescenza, mentre si tentava di comprendere questo fenomeno. La bioluminescenza si manifesta in organismi differenti, con diversi meccanismi di emissione di luce. Alcuni animali hanno complicati organi simili a degli occhi, nei quali l'emissione della luce è gestita dal sistema

nervoso secondo necessità o tramite stimolazioni esterne, ad esempio alcune specie di pesci e calamari. In alcuni altri, le singole cellule formano tutto l'apparato necessario all'emissione della luce, e la luce è emessa continuamente (alcuni batteri e funghi). Il sistema bioluminescente di un organismo può coinvolgere una serie di reazioni chimiche interconnesse, tuttavia la luce è emessa solo dalla reazione che produce lo stato di singoletto eccitato del substrato emettitore di luce, una reazione denominata “*reazione emetti-luce*”. Ci sono vari tipi di reazioni emetti-luce, ma tutte coinvolgono l'ossidazione di un substrato che produce l'energia per generare uno stato elettronico eccitato. Quindi ogni reazione bioluminescente è fondamentalmente una reazione chimica. In questa tesi si è effettuato uno studio teorico computazionale di alcune delle specie molecolari che sono coinvolte nel meccanismo reattivo della bioluminescenza delle lucciole (*Luciola cruciata* e *Photinus pyralis* sono le specie più comuni). Lo scopo è stato quello di creare modelli molecolari che riproducano dati sperimentali noti (in questo caso spettri di assorbimento UV) e che possano quindi essere utilizzati nella dinamica molecolare per poter permettere in futuro tramite tecniche classico quanto-meccaniche, come il **Perturbed Matrix Method** (PMM) usato in questa tesi, di far luce su i misteri ancora irrisolti nel processo chimico della bioluminescenza nelle lucciole.

Capitolo 1

La Bioluminescenza nella Lucciola

N.d.a.: Per seguire meglio il seguente capitolo è piuttosto comodo dare un'occhiata frequente allo spettro del visibile (Figura:1.0.1).

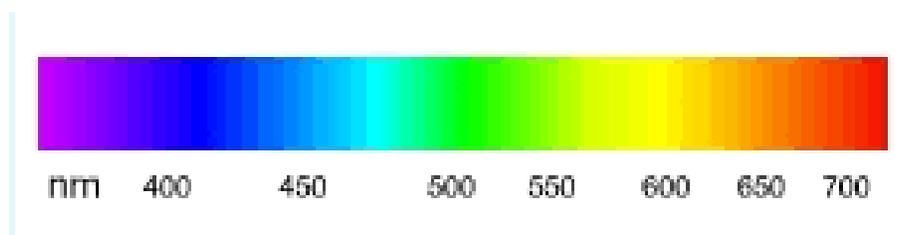


Figura 1.0.1: Lo spettro del visibile

1.1 La Lucciola

La Lucciola è un piccolo insetto della famiglia dei coleotteri, appartenenti alla famiglia dei *Lampyridae*. Comprende tre specie in Europa, di cui due sono presenti anche in Italia. Come tutti i lampiridi, le lucciole sono caratterizzate da un notevole dimorfismo sessuale. I maschi nella maggior parte delle specie sono alati (Figura:1.1.1), lunghi circa 1 cm. mentre le femmine sono attere e di aspetto larviforme, uniformemente giallo rossicce. L'organo deputato all'emissione della luce si trova nella parte terminale dell'addome. In alcune specie questo organo è molto più grande nel maschio che nella femmina e nella maggior parte dei casi entrambi i sessi

possono emettere luce. Tuttavia subentrano importanti cambiamenti morfologici tra i due sessi a seconda delle specie. In alcuni casi solo le femmine possono emettere luce, mentre in altri nessuno dei due sessi ne è capace, in altri ancora solo le larve possono emettere luce perdendo questa capacità in età adulta.



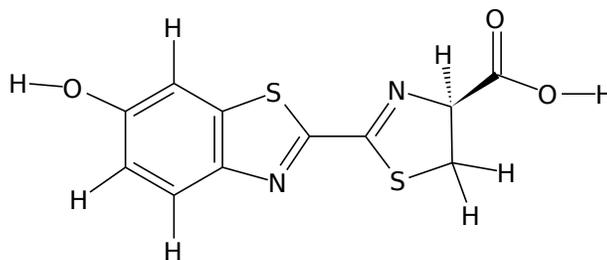
Figura 1.1.1: maschio di *Luciola cruciata*, ritratto in giappone

Le lucciole usano i segnali luminosi per rintracciarsi nell'oscurità, soprattutto nei periodi dell'accoppiamento, in cui questi segnali diventano un evidente richiamo sessuale. Nelle lucciole l'accoppiamento ha luogo allorché la femmina risponde sincronizzando la propria segnalazione luminosa (lampeggio di varia durata, a seconda della specie) a quella del proprio maschio. La luce emessa è una luce fredda, la cui lunghezza d'onda oscilla tra i 500-650 nm. L'intensità invece varia a seconda della specie.

1.2 Il Meccanismo Reattivo

1.2.1 La D-luciferina

La reazione che produce la luce, creata tra la *luciferasi* (un enzima di 548 residui) [2], e il suo substrato chiamato *D-luciferina* (DLH_2) (Figura:1.2.1), fu dimostrata da N. E. Harvey nel 1917 [1], anche se l'emissione osservata all'epoca fu debole e discontinua.

Figura 1.2.1: DLH_2

Trenta anni dopo la scoperta di Harvey, allo scopo di purificare e cristallizzare la luciferasi, Green e McElroy [2] usarono un metodo piuttosto singolare per ottenere la gran quantità di lucciole necessarie per la loro ricerca. Nella storia, ormai diventata leggendaria, loro pubblicizzarono l'acquisto di lucciole ad un centesimo per esemplare e i giovani del loro vicinato risposero in maniera talmente entusiastica a questa iniziativa che gli collezionarono un enorme numero di insetti. I cristalli della DLH_2 hanno la forma di microscopici aghi, che fondono a 205-210 °C. La DLH_2 non subisce particolari riarrangiamenti chimici al cambiare del pH tranne una normale deprotonazione, tuttavia presenta spettri di assorbimento molto differenti. In Figura:1.2.2 sono presentati in parallelo gli spettri di assorbimento UV, rispettivamente in una soluzione tampone di acetato di sodio a pH~5 e una tampone di carbonato di sodio pH~11. La DLH_2 è fluorescente e mostra un'emissione massima a 537 nm sia in condizione acide che basiche, tuttavia l'intensità della fluorescenza è più bassa in soluzione acida piuttosto che in soluzione basica. Si è scoperto che solo la DLH_2 è attiva nella reazione bioluminescente mentre la sua forma L è un forte inibitore.

1.2.2 La trasformazione di D-luciferina in Ossiluciferina

Newton McElroy [2] dimostrò la necessità della presenza di ATP come cofattore nella bioluminescenza. Infatti l'aggiunta di ATP ad una mistura di DLH_2 e luciferasi produce una

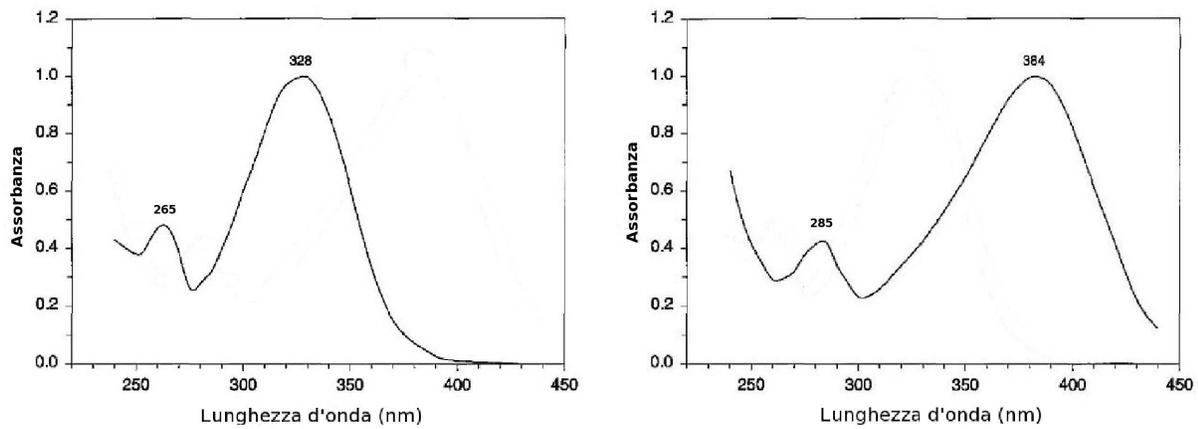


Figura 1.2.2: a sinistra lo spettro di assorbimento della DLH_2 a $\text{pH} \sim 5$, e a destra quello a $\text{pH} \sim 11$

luminescenza intensa e duratura. Nel 1949, McElroy e Strehler [3] trovarono che la reazione luminescente richiedeva Mg^{2+} in aggiunta alla DLH_2 , luciferasi e ATP . La luciferasi fu cristallizzata da Green e McElroy [2], mentre la DLH_2 fu purificata e cristallizzata per la prima volta da Bitler e McElroy [4]. Tra il 1967 e il 1977 Hopkins e McCapra [5, 6] proposero un meccanismo reattivo che coinvolgendo ossigeno molecolare portava alla formazione di un *diosietanone* intermedio. Ipotesi in seguito verificata tramite gli studi con la marcatura isotopica dell' ^{18}O effettuati da Shimomura [7] nel 1977 e quelli successivi con il ^{14}C di Wannlund [8], i quali dimostrarono che uno degli atomi di ossigeno dell'anidride carbonica prodotta nella reazione era derivato da una molecola di ossigeno. Il meccanismo ipotizzato in base ai dati ottenuti fino al 1978 (Figura:1.2.3) era il seguente. Inizialmente la DLH_2 reagisce con l' ATP per formare il *D-Luciferil-AMP* che poi reagisce a sua volta con l'ossigeno, scindendosi per formare AMP , anidride carbonica e l'*ossiluciferina* ($OxyLH_2$) che decadendo da uno stato eccitato produce la ben nota luminescenza giallo-verde il cui massimo di emissione cade a 562 nm per la lucciola della specie delle *Photinus pyralis*.

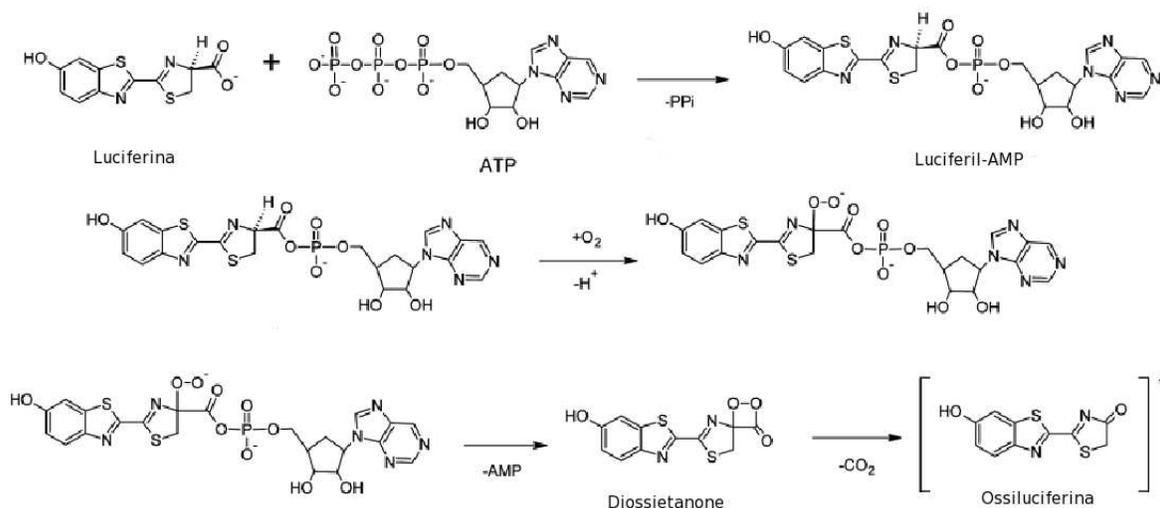
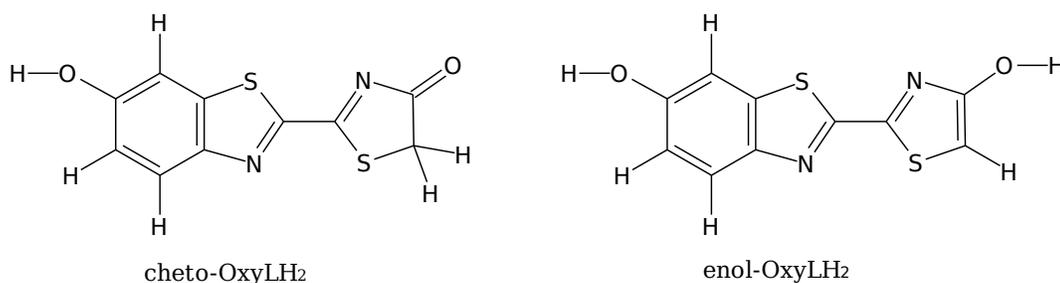


Figura 1.2.3: il meccanismo di reazione

Questa ipotesi fu confermata dal fatto che il picco di assorbimento della prima *OxyLH₂* sintetizzata in laboratorio [9] ($\lambda_{max-abs} = 382 \text{ nm}$) a pH~7.8 (quello fisiologico nella lucciola), cade molto vicino al picco di assorbimento del complesso luciferasi-*OxyLH₂* una volta che si è esaurito l'ATP [10] ($\lambda_{max-abs} = 385 \text{ nm}$). Si noti che nello spettro in fluorescenza del complesso luciferasi-*OxyLH₂* ad ATP esaurito [10] il massimo ($\lambda_{max-emis} = 523 \text{ nm}$) cade ad una distanza di ben 39 nm dalla lunghezza d'onda emessa dalla reazione bioluminescente (ricordiamo che $\lambda_{max-bio-emis} = 562 \text{ nm}$). Si pensò che questa differenza nello spettro sia dovuta alle variazioni che subisce l'intorno enzimatico della *OxyLH₂* una volta che la reazione si è conclusa.

Figura 1.2.4: Le due forme cheto ed enolica della *OxyLH₂*

Dopo molti studi ancora non è chiaro se la $OxyLH_2$ nell'istante dell'emissione sia nella sua forma cheto oppure in quella enolica e in quale stato di protonazione (Figura:1.2.4), perché sono stati riportati, negli anni successivi al 1967, dati sperimentali discordanti [11] che non permettono di definire univocamente in che stato si trovi la $OxyLH_2$, all'interno dell'enzima, prima dell'emissione. All'incirca fino al 2000, grazie agli studi di Hopkins [5], si pensava che la $OxyLH_2$ eccitata, che emette il fotone giallo-verde all'interno della proteina, fosse nello stato enolico (enol- $OxyLH_2$), Hopkins infatti trovò che il 5,5-dimetil- $OxyLH_2$ ($DMOxyLH_2$), un analogo della $OxyLH_2$ che non ha atomi di idrogeno in posizione 5 (che quindi non può avere una forma enolica), in presenza di base in DMSO ha un picco di fluorescenza nel rosso ($\sim 625\text{ nm}$), in analogia ad un intermedio della reazione bioluminescente che è il $D\text{-Luciferil-AMP}$. Perciò si era giunti alla conclusione che, se lo spettro in fluorescenza del $DMOxyLH_2$ è nel rosso, analogamente anche la forma cheto- $OxyLH_2$ dovesse emettere nel rosso, e quindi andando per esclusione l'unico candidato valido che poteva emettere nel giallo-verde doveva essere l' $OxyLH_2$ in forma enolica. Nel 2002 Branchini [12] ha confutato questa ipotesi, facendo una sorprendente scoperta. Egli vide che anche un derivato della DLH_2 noto come D-5,5-dimetil- LH_2 , riusciva comunque ad emettere, tramite la reazione bioluminescente, nel giallo-verde, se inserito in luciferasi di lucciola e nel rosso inserito in luciferasi di scarafaggio (*Pyrophorus plagiophthalmus*). Questa importante scoperta ha rimesso nell'insieme dei probabili emettitori anche la cheto- $OxyLH_2$, non chiarendo tuttavia il ruolo della enol- $OxyLH_2$. Studi computazionali effettuati da Galina Orlova [13] nel 2003 con i metodi per il calcolo quantistico **TD-DFT** [14] e **ZINDO** [15] sugli stati di eccitazione di un ragguardevole numero di forme cheto ed enoliche di $OxyLH_2$ in diversi stati di protonazione, anche non portando a conclusioni univoche, indicano che lo spettro bioluminescente dipende principalmente dallo stato

di protonazione della *OxyLH₂* e dalla polarizzazione indotta dall'ambiente enzimatico piuttosto che dall'angolo diedro tra le due parti planari della molecola (legame C2-C2') e da eventuali piccole distorsioni dalla geometria di equilibrio.

1.3 Il mistero della Ossiluciferina

La difficoltà di comprendere lo stato in cui si trova la *OxyLH₂* all'interno dell'enzima è aggravata inoltre, dal fatto che questa molecola è termolabile e tende a dare reazioni di condensazione al crescere della basicità della soluzione in cui si trova, si pensa che presenti un triplice equilibrio chimico, per un totale di 6 possibili specie (Figura:1.3.1).

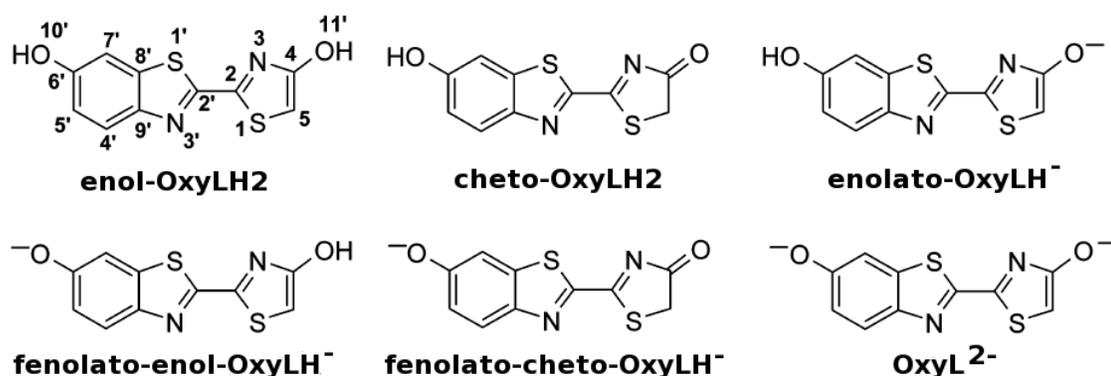


Figura 1.3.1: Le possibili forme chimiche trans-C2-C2' della *OxyLH₂*

I fattori che modificano questo intricato equilibrio sono:

- Il pH.
- La polarità del solvente.
- I legami idrogeno.
- Presenza di ioni addizionali.

- Stacking degli orbitali π

Tuttavia in questo ultimo anno sono stati pubblicati due articoli che hanno rivoluzionato lo studio della $OxyLH_2$ attraverso studi spettroscopici accoppiati ad analisi NMR sulla $OxyLH_2$ e su i suoi analoghi molto più stabili ($MOxyLH_2$ e $DMOxyLH_2$). Tali ricerche hanno permesso di identificare gli spettri di assorbimento e di emissione in fluorescenza di gran parte delle specie coinvolte in questo delicato equilibrio. Il primo di questi studi effettuato da Hirano et al. [16] ha permesso di determinare gli spettri di assorbimento e emissione della $MOxyLH_2$ (un analogo della $OxyLH_2$ metilato in posizione 5) e del $DMOxyLH_2$ in un gran numero di solventi e situazioni chimiche differenti. Il più recente invece, scritto da Naumov et al. [17] ha consentito di determinare per la prima volta la struttura a raggi-X della $OxyLH_2$ e della $MOxyLH_2$ in cristallo, e gli spettri di assorbimento ed emissione della $OxyLH_2$. Le informazioni ottenute da Naumov, Hirano [16, 17] e gli studi sulle strutture cristalline di vari tipi di luciferasi [18], sono consistenti con l'ipotesi che lo ione fenolato della forma enolica, il fenolato-enol- $OxyLH^-$ (Figura:1.3.1), sia l'emettitore principale nella luciferasi. Tuttavia va considerato che ci sono fattori interni alla proteina tra cui:

- chiusura/apertura della tasca in cui è contenuto l'emettitore;
- la presenza di molecole d'acqua in prossimità dell'emettitore;
- legami specifici con i residui dell'enzima.

Essi possono apportare variazioni molto grandi allo spettro sperimentale del possibile emettitore, ma sono stati valutati e quantificati singolarmente e non simultaneamente. Tali fattori, inoltre, non permettono ancora di identificare univocamente la specie molecolare responsabile della bioluminescenza delle lucciole. I risultati ottenuti in questo lavoro sono coerenti con gli

spettri sperimentali UV della *OxyLH₂* ottenuti a differenti valori di pH e con diversi solventi, come descritto da Naumov [17]. La perturbazione degli stati quantistici, ottenuta dalla dinamica molecolare [19] ed usata nel calcolo PMM [20, 21], ha permesso di comprendere, sia la natura delle singole eccitazioni elettroniche della molecola in soluzione nonché il modo in cui tali eccitazioni vengono modificate in energia e in intensità dall'interazione con l'intorno chimico della molecola ad ogni istante lungo una traiettoria di dinamica molecolare. Tali informazioni hanno consentito infine di riprodurre quantitativamente lo spettro d'assorbimento totale delle varie forme di *DLH₂* e *OxyLH₂* e quindi di interpretare i dati spettroscopici sperimentali.

Capitolo 2

La Dinamica Molecolare

2.1 Generalità sulla Dinamica Molecolare

La Dinamica Molecolare (MD) classica è una tecnica che permette di riprodurre e prevedere, utilizzando opportune approssimazioni, il comportamento microscopico del sistema in esame.

Gli aspetti che caratterizzano una simulazione sono riconducibili essenzialmente a :

- algoritmo di campionamento dello spazio delle fasi;
- scelta del potenziale di interazione, $V(r)$, tra le particelle del sistema.

Il modo in cui si campiona lo spazio delle fasi distingue le varie tecniche simulative. La Dinamica Molecolare classica estrae le traiettorie degli atomi integrando le equazione del moto di Newton, mentre altri metodi, utilizzano algoritmi diversi per generare l'insieme delle configurazioni. Questi ultimi possono essere divisi in due gruppi. Il primo raccoglie tutte le tecniche di ricerca sistematica, mentre nel secondo si includono i metodi di ricerca casuale (*random search methods*).

2.2 Nascita e sviluppo della Dinamica Molecolare

Le prime simulazioni di dinamica molecolare furono effettuate negli anni '50, agli albori dell'era informatica, da Alder e Wrainwright [22] , che svilupparono un metodo per sistemi di sfere

rigide. L'uso di funzioni di potenziale più complesse permise successivamente di ottenere dati simulati direttamente confrontabili con dati sperimentali, mostrando come la MD possa essere usata come un valido strumento di indagine delle proprietà di sistemi reali. Rahman e Verlet [23] furono tra i primi ad effettuare calcoli di questo tipo. Per descrivere le interazioni atomiche dell'argon fluido utilizzarono nelle simulazioni un potenziale di tipo Lennard-Jones. La simulazione di sistemi fisici sempre più complessi è stata possibile grazie soprattutto allo sviluppo di calcolatori più potenti e alla comparsa di linguaggi di programmazione ad alto livello (FORTRAN77). Tra questi sistemi, l'acqua, alcani flessibili e piccole proteine, come, per esempio, la BPTI (Bovine Pancreatic Tripsin Inhibitor). Nello stesso periodo furono inoltre sviluppati nuovi metodi per trattare in modo più accurato le interazioni elettrostatiche nella simulazione di sistemi ionici. Alla fine degli anni '70 furono effettuate le prime simulazioni di biomolecole. I successi ottenuti nella riproduzione dei dati sperimentali strutturali di proteine e macromolecole in generale, hanno portato ad una notevole espansione della MD nell'ambito degli studi di biochimica strutturale. Inoltre la MD è stata impiegata con successo per studi strutturali e dinamici di gas, liquidi, fibre polimeriche, polisaccaridi, membrane, acidi nucleici (DNA ed RNA) e proteine. Le simulazioni permettono di studiare processi chimici e biologici, di predire le proprietà di proteine modificate e le interazioni tra piccole molecole e macromolecole. Specialmente nel campo dell'ingegneria delle proteine e nello studio dei farmaci le tecniche di simulazione stanno avendo una grande importanza. Lo sviluppo esponenziale della potenza dei calcolatori (un ordine di grandezza nella velocità di calcolo ogni 6 anni), dovuto tra l'altro allo sviluppo del calcolo parallelo, ha consentito in pochi anni di aumentare il tempo di simulazione, aumentare le dimensioni del sistema simulato e migliorare i risultati ottenuti introducendo modelli sempre più complessi e accurati.

2.3 Equazioni del moto

La dinamica molecolare *classica* si limita a descrivere un sistema di particelle tramite il solo moto dei nuclei del sistema in esame. L'approssimazione di Born-Oppenheimer consente infatti la separazione del moto dei nuclei da quello degli elettroni e di trattare questi ultimi implicitamente. Il moto del sistema viene quindi ad essere descritto solo dai nuclei, dopo averne opportunamente modificato le cariche effettive per tener conto degli effetti elettronici appena menzionati. La rinuncia ad una trattazione quantistica si è dimostrata del tutto ragionevole a temperatura ambiente, per tutti gli atomi eccetto quelli più leggeri come gli atomi di idrogeno che vengono trattati tramite l'uso di vincoli olonomici [24]. Le equazioni che vengono usate per descrivere l'evoluzione del sistema sono quindi quelle di Newton, che per un sistema conservativo sono:

$$\mathbf{F}_i = m_i \mathbf{a}_i \quad (2.1)$$

e

$$\mathbf{F}_i = - \frac{\partial \langle V(\mathbf{r}) \rangle_{\mathbf{R}}}{\partial \mathbf{r}_i} \quad (2.2)$$

in cui $\langle V(\mathbf{r}) \rangle_{\mathbf{R}}$ rappresenta il campo di forze con cui interagiscono i nuclei e $\langle \dots \rangle_{\mathbf{R}}$ indica che il potenziale è mediato su tutte le coordinate elettroniche.

2.4 La funzione di potenziale

Le funzioni di potenziale usate in MD sono rappresentate da una somma di termini che dipendono dalle N coordinate nucleari r_i del sistema e da un numero (P) di parametri $s \equiv (s_1, s_2, \dots, s_p)$.

Ognuno di questi termini viene classificato in base al numero di atomi coinvolti:

$$V(\mathbf{r}, \mathbf{s}) = \underbrace{\sum_i V^{(1)}(\mathbf{r}_i; \mathbf{s})}_{\text{Inter.a1atomo}} + \underbrace{\sum_i V^{(2)}(\mathbf{r}_i, \mathbf{r}_j; \mathbf{s})}_{\text{Inter.a2atomi}} + \dots + \underbrace{\sum_{\text{tutti}} V^{(N)}(\mathbf{r}; \mathbf{s})}_{\text{Inter.aNAtomi}} \quad (2.3)$$

In molti programmi di MD viene impiegato il seguente potenziale:

$$V(r_1, r_2, \dots, r_n) = \sum_{\text{legami}} \frac{1}{2} K_d (d - d_0)^2 + \sum_{\text{angoli}} \frac{1}{2} K_\theta (\theta - \theta_0)^2 + \sum_{\text{diedri impropri}} \frac{1}{2} K_\xi (\xi - \xi_0)^2 +$$

$$+ \sum_{\text{diedri}} \frac{1}{2} K_\phi [1 + \cos(n\phi - \delta)] + \sum_{\text{coppie}(i,j)} \left[\left(\frac{C_{ij}^{12}}{r_{ij}^{12}} - \frac{C_{ij}^6}{r_{ij}^6} \right) + \frac{q_i q_j}{4\pi\epsilon_0\epsilon_r r_{ij}} \right] \quad (2.4)$$

Il primo termine (interazione a due corpi) rappresenta l'energia di vibrazione dei legami tramite un potenziale totalmente armonico. I parametri d_0 , lunghezza del legame all'equilibrio, e K_d , costante di forza del legame, sono ottenuti da dati sperimentali e/o da calcoli quantistici. Un potenziale totalmente armonico non prevede la dissociazione del legame. Per questo motivo in alcuni campi di forze si utilizza la funzione di Morse:

$$V_m(b) = \sum_{\text{legami}} D \{ e^{\{-k_d(d-d_0)\}} - 1 \}^2 \quad (2.5)$$

dove il parametro D rappresenta l'energia di dissociazione del legame. Questo è un potenziale più realistico e descrive meglio vibrazioni su valori di d più ampi e fornisce risultati più accurati quando si vogliono riprodurre gli spettri vibrazionali di sistemi molecolari. Il secondo termine (interazione a tre corpi) si riferisce all'energia dovuta alla deformazione dell'angolo di valenza. Anche in questo caso il potenziale è armonico. I parametri θ_0 e K_θ rappresentano rispettivamente l'angolo di valenza di riferimento e la costante di forza. Il terzo e quarto termine costituiscono le interazioni a 4 corpi. Il primo, di tipo armonico, viene usato per descrivere i

cosiddetti angoli diedri impropri. L'angolo diedro improprio, A-X-Y-B, è definito come l'angolo tra il piano che passa per gli atomi A, X e Y ed il piano passante per gli atomi X, Y e B e viene introdotto per mantenere sia la corretta chiralità che la planarità di determinati gruppi. Il secondo termine delle interazioni a 4 corpi si riferisce agli angoli diedri. K_ϕ rappresenta la costante di forza, n la molteplicità dei minimi e δ la fase relativa. Gli angoli diedri sono presi in accordo con la convenzione IUPAC/IUB dove ϕ costituisce l'angolo tra il piano AXY e il piano XYB ed è zero per la configurazione *cis*. L'ultimo termine nell'equazione del campo di forze descrive le interazioni di non legame tra atomi (interazione a 2 termini). E' suddiviso in due contributi distinti, van der Waals e coulombiano. La parte di van der Waals ha espressioni diverse a seconda del campo di forze utilizzato; in questo caso è rappresentata come una funzione di Lennard-Jones classica. In letteratura si trovano espressioni che differiscono dalla presente per il termine attrattivo, repulsivo o per entrambi. Buckingham, prendendo spunto dalla meccanica quantistica, propose un termine repulsivo di tipo esponenziale, be^{-ar} . Nelle interazioni elettrostatiche (contributo coulombiano), il valore della costante dielettrica relativa, ϵ_r , da utilizzare è ancora materia di discussione. Sono stati utilizzati valori che vanno da $\epsilon_r = 1$ a $\epsilon_r = 8$, mentre altri prendono ϵ_r proporzionale alla distanza r_{ij} . Altre volte il termine coulombiano è stato omissso. La parametrizzazione del campo di forza può avvenire in due modi. Un metodo elegante consiste nel tarare i parametri con i risultati di calcoli quantistici ab-initio su piccoli aggregati molecolari. I campi di forza ottenuti con questa procedura spesso non danno risultati soddisfacenti per via delle molte approssimazioni che devono essere eseguite. Alternativamente si possono determinare i parametri del campo di forza con dati sperimentali (strutture cristallografiche, energie e dinamiche di reticolo, dati spettroscopici, proprietà di liquidi tipo densità e entalpia di vaporizzazione, energia libera di solvatazione, dati di RMN, ecc.). I parametri così

ottenuti sono validi esattamente e unicamente nelle condizioni chimico fisiche in cui sono stati determinati. Questo implica che i campi di forza utilizzati nei programmi di MD siano specifici per determinati sistemi fisici e quindi il loro utilizzo per sistemi diversi potrebbe non dare risultati soddisfacenti.

2.5 Condizioni al contorno

La simulazione di un sistema molecolare immerso in un solvente richiede l'introduzione delle condizioni periodiche al contorno (o *Periodic boundary conditions*, PBC). In questo metodo la scatola centrale è circondata da repliche identiche che creano una condizione di periodicità infinita. Quando una molecola esce durante la simulazione dalla scatola originaria la sua immagine periodica rientra dalla parte opposta, mantenendo la stessa velocità e direzione. Le condizioni periodiche al contorno permettono di riprodurre sia le proprietà di sistemi periodici, come i cristalli, che quelle dei fluidi. Le interazioni all'interno delle PBC vengono trattate con la convenzione della minima immagine (MI) dove un atomo può interagire solo con i primi vicini. Per evitare l'anisotropia dovuta alla forma cubica della MI e artefatti derivanti dall'interazione di un atomo i , della scatola, con una interazione simultanea con un atomo j e con la sua immagine periodica j' , si usa un cut-off sferico di raggio R_c tale che il lato più corto della scatola sia maggiore del doppio del valore del cut-off ($L \gg 2R_c$, con L lato della scatola). La condizione che le dimensioni della scatola computazionale siano maggiori del doppio del cut-off è condizione necessaria, ma non sufficiente per evitare artefatti. Quando ad esempio si vuole simulare un soluto immerso nel solvente, è anche necessario che un atomo del soluto non interagisca con atomi di soluto appartenenti a immagini periodiche. Se questo accadesse, se cioè un atomo di soluto interagisse con atomi del soluto vicino, si starebbe simulando un sistema ad altissima

concentrazione. Affinché ciò non avvenga, occorre che la distanza di ogni atomo del soluto da ogni parete della scatola sia maggiore della metà del raggio di cut-off ($d \geq 0.5R_c$, con d distanza soluto-parete).

2.6 Calcolo delle interazioni a lungo raggio

In una simulazione di MD il 95% del tempo di calcolo è occupato dal calcolo delle interazione di non legame. E' quindi opportuno cercare un metodo che permetta di aumentare l'efficienza computazionale della simulazione agendo sulle componenti di van der Waals e coulombiana. In questo senso, sono state sviluppate varie tecniche per ridurre lo sforzo computazionale. Tra quelle più utilizzate:

- il raggio di cut-off;
- le somme di Ewald.

Il primo metodo consiste nel troncamento delle interazioni di non legame quando la distanza supera un certo valore, detto raggio di cut-off (r_c), pertanto ogni particella del sistema potrà interagire solo con le altre che sono all'interno di una sfera di raggio r_c . Ovviamente l'uso di questo metodo introduce delle approssimazioni, soprattutto nel calcolo della parte elettrostatica, che possono influenzare notevolmente le proprietà dinamiche e strutturali del sistema. Per superare questo problema, oltre ad aumentare il valore del raggio di cut-off, si può introdurre una funzione di switch che assicura la continuità delle forze alla distanza di cut-off. Anche se l'uso del raggio di cut-off soffre di queste limitazioni, rimane uno tra i metodi più diffusi per il calcolo delle forze intermolecolari. Il secondo metodo, più recente, propone l'uso del reticolo periodico, in cui si considerano tutte le interazioni tra il box centrale e le sue infinite repliche.

L'energia coulombiana in un sistema periodico formato da N particelle è ottenuta sommando su tutte le coppie di atomi tra i quali uno è nel box centrale e l'altro in una delle repliche:

$$E = \frac{1}{8\pi\epsilon_0} \sum_{|n|=0}^{\star\infty} \left[\sum_{i=0}^N \sum_{j=0}^N \frac{q_i q_j}{|r_{ij} + n|} \right] \quad (2.6)$$

dove n sono le immagini periodiche, i, j le particelle e il simbolo \star indica che la sommatoria non include il termine con $i = j$ se $n = 0$. Si può dimostrare che questa sommatoria, per potenziali che vanno come $\frac{1}{r}$, può non essere totalmente convergente, cioè il suo limite può variare o divergere se l'ordine dei termini viene cambiato; pertanto questa espressione non può essere utilizzata nel calcolo dell'energia elettrostatica. Per risolvere questo problema si può considerare ogni carica circondata da una distribuzione di carica ρ^G di uguale intensità e segno opposto. Questa distribuzione, che supponiamo essere gaussiana, avrà l'effetto di schermare le interazioni tra le cariche vicine e potremmo quindi considerare l'energia di interazione di tipo *short-range*. La distribuzione di carica totale è

$$\rho_i(r) = \rho_i^q(r) + \rho_i^G(r) \quad (2.7)$$

dove $\rho_i^q(r)$ è la distribuzione della carica della particella i -esima e $\rho_i^G(r)$ è la distribuzione gaussiana corrispondente:

$$\rho_i^q(r) = q_i \delta(r - r_i) \quad (2.8)$$

Dopo aver calcolato l'energia di interazione in (2.7) nello spazio reale, per tornare alla funzione di distribuzione originaria $\rho_i(r)$, si sottrae una funzione uguale a $\rho_i^G(r)$ nello spazio reciproco, usando la trasformata di Fourier. L'espressione finale dell'energia di interazione

totale sarà quindi:

$$\begin{aligned}
E = & \frac{1}{8\pi\epsilon_0} \sum_{i=0}^N \sum_{j=0}^N \left[\sum_{|n|=0}^{\infty} \frac{q_i q_j \operatorname{erfc}(\alpha |r_{ij} + n|)}{|r_{ij} + n|} \right] + \\
& + \frac{1}{8\pi L^3} \sum_{k=0} \frac{4\pi q_i q_j}{k^2} \exp(-k^2/4\alpha^2) \cos(k \cdot r_{ij}) + \\
& - \frac{\alpha}{4\pi^{3/2}\epsilon_0} \sum_{i=1}^{\infty} q_i^2 + \frac{\left| \sum_{i=1}^N q_i r_i \right|^2}{2\epsilon_0 L^3 (2\epsilon' + 1)} \tag{2.9}
\end{aligned}$$

dove $\operatorname{erfc}(x)$ è la funzione errore complementare, che tende a zero per $x \rightarrow \infty$. Per quanto riguarda il primo termine, per α abbastanza grande, la somma si riduce ad un unico termine con $n = 0$; il secondo termine è invece una somma sui vettori reciproci $k = 2\pi n/L$. Gli ultimi due termini sono, rispettivamente, la funzione di correzione, dovuta alla presenza di un termine di auto-interazione in $\rho_i^G(r)$, e il contributo del campo depolarizzante all'energia, che è compensato dall'effetto del dielettrico esterno. Questo ultimo punto necessita di un ulteriore commento, infatti nelle somme di Ewald, i box virtuali sono visti come strati sferici concentrici ordinati dall'interno verso l'esterno; la sommatoria è ovviamente troncata dopo un certo numero di strati e questo sistema sferico è immerso in un dielettrico continuo con costante dielettrica uguale a ϵ' . L'ultimo termine nell'equazione è la somma dei contributi del campo depolarizzante e il campo esterno dovuto al dielettrico. Se la sfera è immersa in un mezzo con una costante dielettrica infinita, questo termine scompare, perché è compensato esattamente da questi due effetti. Recentemente sono stati sviluppati altri metodi della stessa famiglia, come il *Particle-Particle*, *Particle-Mesh (PPPM)* o il *Particle Mesh Ewald(PME)*[25], che basandosi sugli stessi principi sui quali sono state sviluppate le somme di Ewald, sono computazionalmente più efficienti e stabili. In particolare, in tutte le simulazioni di questo lavoro, è stato usato il PME come me-

todo di calcolo per le forze elettrostatiche. Il metodo del PME sostanzialmente è identico alle somme di Ewald, con la sola differenza che per calcolare l'energia di una certa distribuzione di carica si considerano le cariche disposte su una griglia (mesh); questo rende notevolmente più efficiente il calcolo della parte in cui è coinvolta la trasformata di Fourier, che costituisce la parte computazionalmente più pesante.

2.7 Vincoli in MD

Si possono applicare vari tipi di vincoli nelle simulazioni di MD; i più comuni sono quelli che bloccano le posizioni di determinati atomi, detti *Position Constraints*. L'uso di vincoli in MD nasce dalla necessità di eliminare le vibrazioni dei legami ad alta frequenza che richiedono passi di integrazione brevi. Infatti quando si integra l'equazione di Newton il fattore che determina il passo di integrazione da poter utilizzare è la più alta frequenza che si presenta nel sistema. Nel caso di macromolecole biologiche in soluzione le vibrazioni che coinvolgono gli atomi di idrogeno determinano le vibrazioni a più alta frequenza. Per il legame O-H la frequenza di *stretching* è dell'ordine di $10^{14} Hz$, per cui un periodo medio è dell'ordine dei 10 fs. Questo limita il passo di integrazione (*time step*) che si può usare in una MD a circa 0.5 fs, se si considera che, per riprodurre correttamente l'andamento di una funzione periodica, è necessario campionarla almeno venti volte per periodo. Introducendo un metodo che vincola questi legami o, come avviene in pratica, tutti i legami covalenti, si può arrivare ad un passo di integrazione di 2 fs, che è quello comunemente utilizzato anche in questo lavoro di tesi. Poiché le vibrazioni di legame sono praticamente disaccoppiate dalle altre vibrazioni del sistema, questo vincolo non altera il resto della dinamica, mentre ciò non è vero per le vibrazioni degli angoli di valenza. Il metodo più comunemente utilizzato è il cosiddetto SHAKE che fu introdotto [24] nel 1977. SHAKE

consiste in una procedura iterativa che dopo ogni passo di integrazione varia le posizioni iniziali $r'_i(t + \Delta t)$ degli atomi finché tutte le condizioni di vincolo non vengano soddisfatte nelle nuove posizioni $r_i(t + \Delta t)$. L'uso di SHAKE consente di aumentare il passo di integrazione di 2-4 volte, ma presenta comunque degli svantaggi:

- può dare problemi di convergenza quando viene applicato a grandi gruppi planari;
- costituisce generalmente un collo di bottiglia computazionale nella parallelizzazione dell'algoritmo;
- in alcune applicazioni il vincolo sulle distanze di legame costituisce un modello fisico non realistico.

Recentemente è stato introdotto un metodo chiamato LINCS, che supera in parte le limitazioni di SHAKE. Questo metodo è stato utilizzato nel presente lavoro di tesi.

2.8 Le condizioni iniziali

Una volta definiti il campo di forze, le condizioni al contorno e gli eventuali vincoli bisogna avere le coordinate e le posizioni iniziali del sistema che si vuole studiare con la MD. Se il tempo di osservazione fosse sufficientemente lungo i risultati ottenuti dovrebbero essere indipendenti dalle condizioni iniziali, ma poiché la maggior parte dei sistemi fisici raggiunge l'equilibrio in tempi molto più lunghi di quelli simulabili è opportuno effettuare le simulazioni partendo da una configurazione che sia il più vicino possibile a quella di equilibrio. Si possono ottenere le coordinate da dati sperimentali come raggi-X e NMR per macromolecole, mentre per un fluido, come nel nostro caso, si possono disporre le molecole in modo casuale nel box di simulazione.

Le velocità iniziali invece si possono ottenere con una semplice distribuzione maxwelliana, che dipenda dalla temperatura iniziale della simulazione:

$$\rho_{x_i} = \sqrt{\frac{m_i}{2\pi k_B T}} \exp\left(-\frac{\frac{1}{2}m_i v_{x_i}^2}{k_B T}\right) \quad (2.10)$$

dove ρ_{x_i} è la densità di probabilità della componente v_{x_i} ; chiaramente valgono equazioni analoghe per le componenti y e z .

2.9 Integrazione delle equazioni del moto

In una simulazione MD le traiettorie degli N atomi del sistema si ottengono dall'integrazione delle equazioni del moto di Newton. Vista la complessità dell'espressione del campo di forze (2.4), l'integrazione di queste equazioni non ha soluzione analitica. Per questo sono stati sviluppati nel corso degli anni diversi metodi approssimati che estraggono le traiettorie da una simulazione MD per via numerica.

Le caratteristiche che deve avere un buon algoritmo di integrazione possono essere così riassunte:

- deve permettere l'uso di un time step lungo. E' evidente che maggiore risulta il passo di integrazione, minore è il numero di integrazioni necessarie in una simulazione. D'altro canto, se il Δt è troppo grande, si rischia di non riuscire a seguire certe componenti del moto del sistema, come ad esempio le vibrazioni, portando così a risultati a volte catastrofici. E' pertanto opportuno tarare il passo di integrazione sui moti a frequenza più alta;
- deve essere reversibile temporalmente e conservare le grandezze costanti nel sistema.

Le equazioni di Newton sono reversibili temporalmente ed è pertanto auspicabile che l'algoritmo utilizzato per integrarle riproduca con fedeltà questo dato.

Gli algoritmi utilizzati si differenziano nella precisione di calcolo (ordine dell'algoritmo), nella occupazione di memoria e nel tipo di rappresentazione. La natura degli algoritmi usati deriva dallo sviluppo in serie di Taylor delle grandezze $r_i(t + \Delta t)$ e $v_i(t + \Delta t)$. In MD si adoperano generalmente algoritmi che utilizzano espansioni di Taylor fino al terzo ordine in quanto si è osservato che algoritmi di ordine inferiore sono poco accurati mentre quelli di ordine superiore sono troppo dispendiosi in termini di risorse computazionali. Gli algoritmi comunemente usati sono :

- l'algoritmo di Verlet;
- l'algoritmo leap-frog.

Il primo algoritmo è veloce, richiede poca memoria, presenta un piccolo errore per tempi lunghi ed una buona conservazione dell'energia per tempi brevi. Si ottiene da addizioni e sottrazioni di espansioni di Taylor delle coordinate $\mathbf{r}_i(t + \Delta t)$ e $\mathbf{r}_i(t - \Delta t)$ di particelle di massa m_i :

$$\mathbf{r}_i(t \pm \Delta t) = \mathbf{r}_i(t) \pm \Delta t \dot{\mathbf{r}}_i(t) \pm \frac{(\Delta t)^2}{2!} \ddot{\mathbf{r}}_i \pm \frac{(\Delta t)^3}{3!} \dddot{\mathbf{r}}_i + O(\Delta t^4) \quad (2.11)$$

da cui si ricava:

$$\mathbf{r}_i(t + \Delta t) = -\mathbf{r}_i(t - \Delta t) + 2\mathbf{r}_i(t) + \frac{(\Delta t)^2}{m_i} \mathbf{F}_i(t) \quad (2.12)$$

dove $\mathbf{F}_i(t)$ è la forza esercitata sulla particella i al tempo t e dove i termini di quarto ordine

e superiori sono stati trascurati. Questo schema di integrazione è esatto fino al quarto ordine in Δt . Le velocità in questo schema sono date da :

$$\mathbf{v}_i(t) = \dot{\mathbf{r}}_i(t) = \frac{1}{2\Delta t}[\mathbf{r}_i(t + \Delta t) - \mathbf{r}_i(t - \Delta t)] \quad (2.13)$$

in cui l'errore sulle velocità è dell'ordine di $(\Delta t)^3$.

Il leap-frog è una variante dello schema d'integrazione appena visto. Viene utilizzato in GROMACS, ed in particolare è stato impiegato in tutte le simulazioni di questo lavoro. Questo schema usa per le coordinate intervalli interi e per le velocità intervalli semi-interi.

$$\mathbf{v}_i(t + \Delta t/2) = \mathbf{v}_i(t - \frac{1}{2}\Delta t) + \frac{\mathbf{F}_i(t)\Delta t}{m_i} \quad (2.14)$$

$$\mathbf{r}_i(t + \Delta t) = \mathbf{r}_i(t) + \mathbf{v}_i(t + \Delta t/2)\Delta t \quad (2.15)$$

L'algoritmo leap-frog è molto efficiente da un punto di vista computazionale, essendo semplice e richiedendo meno tempo di CPU e memoria rispetto ad altri più avanzati.

2.10 Metodi per il controllo della temperatura e della pressione

Per poter confrontare i risultati di una simulazione MD con i dati sperimentali occorre mettersi nelle stesse condizioni in cui questi sono stati ricavati. Questo spesso costringe ad eseguire la simulazione a pressione e temperatura costanti. Nel corso degli anni sono stati sviluppati, a questo proposito, vari algoritmi. I metodi generalmente impiegati per generare insiemi statistici isotermi e isobari si possono dividere in due gruppi. Del primo fanno parte quella classe di procedure definite con il nome di *extended system methods* [26]. Questi metodi fanno uso

di meccanismi di retroazione per correggere le deviazioni di parametri controllati (temperatura e pressione) da valori medi prestabiliti, questo implica che il valore può fluttuare, ma l'ampiezza della fluttuazione può essere controllata. L'altro gruppo è rappresentato dai cosiddetti *constraint methods*, in cui le equazioni del moto vengono accoppiate ad un opportuno vincolo (per esempio, nel caso della temperatura questo vincolo è l'energia cinetica costante). Nel caso della temperatura esistono due ulteriori gruppi costituiti dai cosiddetti *stochastic methods* [27] e dai *weak coupling methods* [28]. I primi consistono nell'aggiustare la temperatura in modo da riprodurre l'effetto di un bagno termico esterno o, in modo da generare ad ogni passo una distribuzione maxwelliana di velocità in modo da riportare il valore dell'energia cinetica totale a quello stabilito; nei secondi le equazioni del moto vengono modificate in modo che ci sia un rilassamento del primo ordine in T verso un valore di riferimento T_0 secondo la classica equazione:

$$\frac{dT(t)}{dt} = \frac{(T_0 - T(t))}{\tau} \quad (2.16)$$

Nelle simulazioni eseguite nel corso di questa tesi si è fatto uso del metodo di Berendsen o dell'accoppiamento debole con un bagno termico che fa parte del gruppo dei *weak coupling methods*. Non si è utilizzato alcun metodo per il controllo della pressione poiché tutte le simulazioni sono state effettuate in un box a dimensioni fisse. Di seguito verrà illustrato brevemente il metodo di Berendsen. Uno dei pregi di questo metodo è la semplicità. L'algoritmo simula l'accoppiamento del sistema con un bagno termico esterno alla temperatura T_0 . La velocità dell'interazione tra questo bagno termostato ed il sistema è determinata da una costante di tempo, τ , che può essere liberamente scelta. L'accoppiamento viene ottenuto modificando ad ogni passo la velocità degli atomi. Questo può essere fatto moltiplicando per una opportuna

costante λ le velocità delle singole particelle. Il livellamento (*rescaling*) della temperatura T ad un valore limite T_0 seguendo una legge esponenziale può essere ottenuto come segue, tenendo conto che $T(t)$ è la temperatura ottenuta dal valore delle velocità dopo l'integrazione e $T'(t)$ è il valore di temperatura ottenuto dopo il *rescaling*:

$$\frac{dT}{dt} = \frac{1}{\tau}[T_0 - T(t)] \quad (2.17)$$

$$\Delta T = \frac{\Delta t}{\tau}[T_0 - T(t)] \quad (2.18)$$

$$T'(t) = \lambda T(t) \quad (2.19)$$

$$\Delta T = T'(t) - T(t) = (\lambda - 1)T(t) \quad (2.20)$$

da (2.18) e (2.20) si ottiene :

$$(\lambda - 1)T(t) = \frac{\Delta t}{\tau}[T_0 - T(t)] \quad (2.21)$$

e

$$\lambda = 1 + \frac{\Delta t}{\tau} \left(\frac{T_0}{T(t)} - 1 \right) \quad (2.22)$$

$$T'(t) = \left[1 + \frac{\Delta t}{\tau} \left(\frac{T_0}{T(t)} - 1 \right) \right] T(t) \quad (2.23)$$

Il metodo che è stato impiegato in questo lavoro di tesi per il controllo della temperatura è l'*isoterma gaussiano* [29]. Tale metodo appartiene ai cosiddetti *constraint methods* e utiliz-

za l'energia cinetica totale del sistema come una variabile vincolata; le equazioni del moto, utilizzando l'accoppiamento isoterma gaussiano divengono

$$\dot{x}_i = \frac{\partial H(\mathbf{x}, \mathbf{p})}{\partial p_i} \quad (2.24)$$

$$\dot{p}_i = -\frac{\partial H(\mathbf{r}, \mathbf{p})}{\partial x_i} + \gamma p_i \quad (2.25)$$

dove γ è equivalente a un coefficiente di attrito e varia implicitamente nel tempo così da garantire che il valore dell'energia cinetica sia costante, ovvero mantiene il vincolo

$$\dot{K} = \sum_{i=1} \frac{p_i \dot{p}_i}{m_i} = 0 \quad (2.26)$$

Si può dimostrare che questo metodo genera una distribuzione nello spazio delle configurazioni in accordo con la meccanica statistica.

Capitolo 3

Il Metodo della Matrice Perturbata (PMM)

3.1 Introduzione al PMM

Il metodo della matrice perturbata (PMM) [20, 21] combina la dinamica molecolare classica e le proprietà elettroniche quantistiche con lo scopo di ottenere una descrizione del comportamento dei sistemi complessi. In tale procedura, la dinamica molecolare fornisce la perturbazione che modifica le proprietà elettroniche imperturbate, calcolate mediante i metodi quanto-meccanici standard. In questo capitolo si introdurranno i metodi quanto-meccanici utilizzati per ottenere i dati “ imperturbati ” necessari al PMM. In particolare, le equazioni generali del metodo Hartree-Fock (HF), Post Hartree-Fock come il **CI** [30] (Interazione di Configurazione) o il **MCSCF** [31] (SCF Multi Configurazione) e il metodo **DFT** [14](Teoria del Funzionale della Densità). Nell’ultima sezione di questo capitolo verranno illustrate le derivazioni di base del PMM e la metodologia per caratterizzare gli stati elettronici perturbati e calcolare lo spettro di assorbimento. In questa tesi il PMM è stato utilizzato per riprodurre lo spettro UV della DLH_2 e dei suoi derivati coinvolti nel processo di bioluminescenza della lucciola.

3.2 Metodi Quanto-Meccanici

Tenendo conto degli effetti relativistici che possono essere trascurati, l'equazione di Schrödinger può descrivere il comportamento elettronico di un sistema molecolare. In generale l'equazione di Schrödinger dipendente dal tempo che prevede l'evoluzione temporale della funzione d'onda ha la forma seguente:

$$\hat{H}\Psi(\mathbf{r}, t) = i\hbar\frac{\partial}{\partial t}\Psi(\mathbf{r}, t) \quad (3.1)$$

dove \hat{H} è l'operatore hamiltoniano definito come $\hat{H} = \hat{K} + \hat{V}$ dove \hat{K} è l'operatore energia cinetica:

$$\hat{K} = -\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right) \quad (3.2)$$

con $\hbar = h/2\pi$, dove h è la costante di Planck e \hat{V} l'operatore di energia potenziale. Quindi tutte le informazioni del sistema sono contenute nella funzione d'onda $\Psi(\mathbf{r}, t)$ che soddisfa l'equazione (3.1). Se l'hamiltoniano è dipendente dal tempo, noi possiamo fattorizzare la funzione d'onda $\Psi(\mathbf{r}, t)$ come $\Psi(\mathbf{r}, t) = \psi(\mathbf{r})\chi(t)$, al fine di utilizzare il metodo della separazioni delle variabili per risolvere l'equazione (3.1), fattorizzandola nelle due equazioni:

$$i\hbar \frac{d\chi}{dt} = E\chi \quad (3.3)$$

$$\hat{H}\psi = E\psi \quad (3.4)$$

L'equazione (3.3) è facilmente risolvibile e da come soluzione $\chi(t) = e^{iEt/\hbar}$ mentre l'equazione (3.4) è conosciuta come equazione di Schrödinger indipendente dal tempo e le sue soluzioni sono chiamate *funzioni d'onda stazionarie*.

3.2.1 L'approssimazione di Born-Oppenheimer

L'approssimazione di Born-Oppenheimer (BO), che è universalmente assunta nella teoria della struttura elettronica, è un disaccoppiamento del moto elettronico da quello nucleare. L'approssimazione si basa sulla differenza di massa tra gli elettroni e i nuclei (un fattore di $10^3 - 10^5$), e assume che gli elettroni seguano i nuclei “istantaneamente” durante il movimento di questi ultimi. La derivazione completa può essere trovata in molti libri di testo [32, 33], ma qui tratteremo brevemente solo i concetti principali. Quando i gradi di libertà nucleari ed elettronici vengono separati, l'Hamiltoniano elettronico va a dipendere parametricamente dalle posizioni nucleari, con il risultato che le coordinate nucleari appaiono nell'Hamiltoniano elettronico ma non ne compaiono le derivate a nessun ordine. Il problema elettronico può pertanto essere risolto per i nuclei che sono momentaneamente bloccati in posizioni fisse nello spazio. E' importante considerare che la funzione d'onda elettronica dipende dalle coordinate nucleari, e quindi in linea di principio l'operatore di energia cinetica nucleare ha degli effetti su di essa. Tuttavia, anche se questa dipendenza è formalmente presente, è in genere considerata irrilevante a causa della grande differenza di massa tra i nuclei e gli elettroni che rende il movimento dei nuclei molto più lento di quello degli elettroni. Ciò si traduce nel fatto che la funzione d'onda elettronica

varia poco in caso di mutamenti delle posizioni nucleari, e quindi i termini che contengono le derivate prime e seconde delle coordinate nucleari, ∇_N e ∇_N^2 , quando vengono applicate all'interno della funzione d'onda elettronica possono essere trascurate. Gli errori introdotti dall'approssimazione BO sono piuttosto piccoli, quindi le soluzioni della parte elettronica sono chiamate *esatte*. Notiamo inoltre che un gran numero dei concetti basilari odierni in chimica, come ad esempio le superfici di energia potenziale e le strutture molecolari, sono risultati impliciti dell'approssimazione BO.

3.2.2 Il Metodo di Hartree-Fock

Come puntualizzato nel precedente paragrafo rimane il problema di trovare un metodo semplice che possa provvedere a fornire, in maniera qualitativamente corretta, soluzioni approssimate dell'equazione di Schrödinger multi-elettronica. Il metodo di Hartree-Fock (HF) fornisce tale soluzione e allo stesso tempo soddisfa anche la necessità di un modello gestibile della struttura elettronica di sistemi a molti elettroni. Spesso nei problemi di fisica matematica una buona funzione di prova per la risoluzione di un sistema multi-elettronico è il cosiddetto *prodotto di Hartree* (3.5)

$$\Theta(1, 2, 3, \dots, n) = \phi_1(1)\phi_2(2)\phi_3(3) \dots \phi_n(n) \quad (3.5)$$

dove le funzioni ϕ_i , gli spin-orbitale, sono funzioni d'onda che descrivono un singolo elettrone. Questo modello era quello usato agli albori della meccanica quantistica per ottenere le strutture elettroniche degli atomi, ma è stato surclassato da metodi successivi. Infatti il prodotto di Hartree non soddisfa il principio di antisimmetria (o principio di Pauli), e quindi sono necessarie ulteriori modifiche a questa funzioni d'onda per renderla antisimmetrica. Una funzione

d'onda antisimmetrica scritta in forma di determinante, è il *Determinante di Slater* (3.6)

$$\Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) = (N!)^{-1/2} \begin{vmatrix} \phi_i(\mathbf{x}_1) & \phi_j(\mathbf{x}_1) & \cdots & \phi_k(\mathbf{x}_1) \\ \phi_i(\mathbf{x}_2) & \phi_j(\mathbf{x}_2) & \cdots & \phi_k(\mathbf{x}_2) \\ \vdots & \vdots & & \vdots \\ \phi_i(\mathbf{x}_N) & \phi_j(\mathbf{x}_N) & \cdots & \phi_k(\mathbf{x}_N) \end{vmatrix} \quad (3.6)$$

dove il fattore $(N!)^{-1/2}$ è un fattore di normalizzazione e gli orbitali di spin $(\phi_i, \phi_j, \dots, \phi_k)$ sono gli N spin-orbitale occupati da N elettroni, in cui tuttavia non viene specificato quale elettrone è in quale orbitale. I determinanti di Slater obbediscono al principio di antisimmetria perché scambiare due righe qualsiasi del determinante, che corrisponde di fatto allo scambio di due elettroni, modifica il segno del determinante e quindi della funzione d'onda. Ovviamente diventa poco pratico scrivere il determinante di Slater come nell'equazione (3.6) ogni volta che deve essere rappresentato, e quindi si preferisce utilizzare una abbreviazione, mostrando solamente gli elementi diagonali del determinante:

$$\Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) = (N!)^{-1/2} \det|\phi_i(\mathbf{x}_1)\phi_j(\mathbf{x}_2) \cdots \phi_k(\mathbf{x}_N)\rangle = (N!)^{-1/2} \det|\phi_i\phi_j \cdots \phi_k\rangle \quad (3.7)$$

L'Hamiltoniano elettronico per un sistema a molti elettroni (un sistema atomico o molecolare) può essere scritto in termini a zero, uno e due elettroni come segue:

$$\hat{H} = \hat{h}_0 + \sum_i \hat{h}_i + \sum_{i < j} \hat{g}_{ij} \quad (3.8)$$

dove

$$\hat{h}_0 = \sum_{\mu < \nu} \frac{Z_\mu Z_\nu}{|\mathbf{r}_\mu - \mathbf{r}_\nu|} = \sum_{\mu < \nu} \frac{Z_\mu Z_\nu}{r_{\mu\nu}} \quad (3.9)$$

è una costante additiva derivante dalle coordinate nucleari fisse. L'equazione

$$\hat{h}_i = -\frac{1}{2}\nabla_i^2 - \sum_{\mu} \frac{Z_\mu}{r_{i\mu}} \quad (3.10)$$

dipende dalle coordinate di un singolo elettrone; mentre l'operatore a due elettroni

$$\hat{g}_{ij} = \frac{1}{r_{ij}} \quad (3.11)$$

descrive l'interazione tra due elettroni i e j . Il valore atteso di un hamiltoniano che opera su un determinante di Slater prende la forma (3.12)

$$E[\Psi] = \langle \Psi | \hat{H} | \Psi \rangle = \sum_i \langle \phi_i | \hat{h} | \phi_i \rangle + \frac{1}{2} \sum_i \langle \phi_i | \hat{J} - \hat{K} | \phi_i \rangle \quad (3.12)$$

dove \hat{J} e \hat{K} sono rispettivamente gli operatori di *Coulomb* e di *Scambio*. I determinanti di Slater sono una ragionevole approssimazione di una funzione d'onda a molti elettroni. Tuttavia, gli orbitali che definiscono questa funzione d'onda devono ancora essere determinati esattamente. Pertanto è possibile applicare il principio variazionale all'espressione dell'energia ottenuta con una funzione d'onda iniziale, con il presupposto implicito che gli orbitali che portano all'energia più bassa siano i migliori possibili. Di frequente nella ricerca di valori limite, ci si avvicina alla meta desiderata utilizzando il calcolo differenziale. Gli orbitali sono quindi modificati di una piccola (infinitesima) quantità, $\phi_i \Rightarrow \phi_i + \delta\phi_i$. Questo comporta un cambiamento

della funzione d'onda totale $\Psi \Rightarrow \Psi + \delta\Psi$, nonché nell'espressione dell'energia. Rimpiazzando ϕ_i con $\phi_i + \delta\phi_i$ nell'equazione (3.12), è possibile ottenere la variazione al primo ordine dell'energia $\delta^{(1)}E[\Psi]$

$$\begin{aligned}
 \delta^{(1)}E[\Psi] &= \sum_i \langle \delta\phi_i | \hat{h} | \phi_i \rangle + \sum_{i,j} [\langle \delta\phi_i \phi_j | \hat{g} | \phi_i \phi_j \rangle - \langle \delta\phi_i \phi_j | \hat{g} | \phi_j \phi_i \rangle] + c.c. \\
 &= \sum_i \langle \delta\phi_i | \hat{h} | \phi_i \rangle + \sum_i \langle \delta\phi_i | \hat{J} - \hat{K} | \phi_i \rangle + c.c. \\
 &= \sum_i \langle \phi_i | \hat{F} | \phi_i \rangle + c.c.
 \end{aligned} \tag{3.13}$$

dove *c.c.* sta per complesso coniugato e \hat{F} è l' *Operatore di Fock*, definito come

$$\hat{F} = \hat{h} + \hat{J} - \hat{K} \tag{3.14}$$

Scrivendo $\delta\phi_i$ in termini di orbitali molecolari (MO) occupati ϕ_j e non occupati (chiamati anche *orbitali virtuali*) ϕ_a , si può ottenere l'equazione (3.15).

$$\langle \phi_a | \hat{F} | \phi_i \rangle = 0 \tag{3.15}$$

che è chiamata *equazione di Hartree-Fock*. Tale equazione può anche essere riscritta come

$$\hat{F}\phi_i = \sum_k \lambda_{ki} \phi_k \tag{3.16}$$

dove k varia su gli indici dei MO occupati e λ_{ji} definisce una matrice hermitiana visto che \hat{F} è hermitiana. Introducendo una matrice \mathbf{V} che diagonalizza la precedente matrice, $\boldsymbol{\lambda}$, otteniamo

i nuovi MO che lasciano invariata Ψ .

$$\tilde{\phi}_j = \sum_i \phi_i V_{ij} \quad (3.17)$$

Quindi moltiplicando l'equazione (3.16) per la matrice V e sommando su i è possibile ottenere la forma più conosciuta delle equazioni di Hartree-Fock (3.18),

$$\hat{F}\tilde{\phi}_j = \epsilon_j \tilde{\phi}_j \quad (3.18)$$

Queste equazioni sono chiamate *equazioni canoniche di Hartree-Fock*, i corrispondenti orbitali sono chiamati orbitali canonici e gli autovalori ϵ_i rappresentano le energie degli orbitali che sono date da

$$\epsilon_i = \langle \tilde{\phi}_i | \hat{F} | \tilde{\phi}_i \rangle \quad (3.19)$$

e l'energia totale è espressa come

$$E = \sum_i \epsilon_i - \frac{1}{2} \sum_i \langle \tilde{\phi}_i | \hat{J} - \hat{K} | \tilde{\phi}_i \rangle \quad (3.20)$$

Si noti infine che l'energia totale HF non è la somma delle singole energie degli orbitali canonici e che comunque gli elettroni sono caratterizzati anche dal loro spin s . L'elettrone ha un numero quantico di spin $s = \frac{1}{2}$, che identifica la lunghezza del vettore di spin, la componente di questo vettore è quantizzata e può prendere uno dei possibili valori $m_s = \pm \frac{1}{2}$. Per gli spin elettronici, al fine di dare significato alla teoria MO, dobbiamo lasciare che gli orbitali ottenuti dipendano dallo spin. Visto che ci sono solo due possibili casi, è sufficiente classificare i spin-orbitali ϕ come ϕ^α o ϕ^β , dipendenti dallo spin:

$$\begin{aligned}\phi_k^\alpha &= \psi_k^\alpha(\mathbf{r})\alpha \\ \phi_k^\beta &= \psi_k^\beta(\mathbf{r})\beta\end{aligned}\tag{3.21}$$

dove ψ sono orbitali spaziali, ossia funzioni che dipendono solo dalle coordinate spaziali, dove tuttavia la parte spaziale di uno spin-orbitale può dipendere dal fatto che la parte di spin è α o β . Finché l'hamiltoniana non dipende esplicitamente dallo spin, essa può commutare con gli operatori di spin, e quindi ognuna delle sue autofunzioni dovrebbe essere anche una autofunzione degli operatori di spin. Allo stesso modo, una funzione d'onda multi-elettronica che è una autofunzione di un Hamiltoniano e non dipende esplicitamente dallo spin è anche un autofunzione degli operatori di spin totale di \hat{S}^2 e \hat{S}_z , cioè deve soddisfare le relazioni:

$$\hat{S}^2\Psi = S(S+1)\Psi\tag{3.22}$$

$$\hat{S}_z\Psi = M_s\Psi\tag{3.23}$$

L'implementazione di queste equazioni è conosciuta come *Hartree-Fock non ristretto* (UHF) [30]. Il determinante di Slater in generale non soddisfa l'equazione (3.22) ma il corretto comportamento degli spin della funzione d'onda può essere ottenuto in alcuni casi tramite il metodo computazionalmente complesso noto con il nome di *Hartree-Fock non ristretto a shell aperta* (ROHF) [34]. Un importante situazione si ha quando lo stato elettronico è uno stato di singoletto, cioè $S = 0$, in cui la molteplicità di spin vale 1 e si ha un numero pari di elettroni. Imponendo che gli orbitali accoppiati abbiano la medesima parte spaziale e si differenzino solo nella parte di spin, sono possibili importanti semplificazioni:

$$\begin{aligned}\phi_k^\alpha &= \psi_k(\mathbf{r})\alpha \\ \phi_k^\beta &= \psi_k(\mathbf{r})\beta\end{aligned}\tag{3.24}$$

Il determinante è quindi un'autofunzione dell'operatore \hat{S}_z con autovalore 0. Segue inoltre che esso debba essere un'autofunzione anche di \hat{S}^2 . L'approccio utilizzato nelle equazioni (3.24) assicura quindi le corrette proprietà degli spin nella funzione d'onda di singoletto e la metodologia derivante viene chiamata *Hartree-Fock ristretto* (RHF) [30]. L'RHF è di gran lunga la variante più utilizzata del metodo di Hartree-Fock. Ciò è dovuto al fatto che la maggior parte dei principali gruppi di molecole stabili possiede un numero pari di elettroni e uno stato fondamentale di singoletto. Se lo schema UHF fosse applicato su molecole a shell chiuso le soluzioni ottenute soddisferebbero ancora le equazioni (3.22) e (3.24), ma il calcolo sarebbe, computazionalmente parlando, molto più gravoso del metodo RHF. Per risolvere le equazioni di HF è ampiamente utilizzata una tecnica basata sull'espansione della parte spaziale dei MO $\psi_i(\mathbf{r})$ come combinazione lineare di orbitali atomici $\chi_p(\mathbf{r})$:

$$\psi_i(\mathbf{r}) = \sum_{p=1}^N C_{pi} \chi_p(\mathbf{r})\tag{3.25}$$

identificata dall'acronimo LCAO (*Linear Combination Atomic Orbital*). L'espansione di funzioni sconosciute in una serie di funzioni di base note è un potente approccio comunemente utilizzato in molti settori della matematica, del resto la LCAO è anche molto interessante dal punto di vista puramente chimico. Nel metodo LCAO l'energia deve essere minimizzata variando i coefficienti C_{pi} , che specificano il coefficiente che moltiplica il p -esimo orbitale atomico

dello i -esimo orbitale molecolare. Seguendo questo approccio il problema può essere successivamente risolto tramite le *equazioni di Roothaan-Hall* [35]. Per risolvere queste equazioni il metodo più comune è la procedura del *campo auto-consistente* (SCF) [36]. Si noti che le equazioni HF potrebbero essere risolte esattamente se fosse possibile costruire l'operatore di Fock, e quindi la matrice di Fock, che una volta diagonalizzata fornirebbe gli orbitali molecolari. In ogni caso l'operatore di Fock è definito proprio da quella stessa serie di orbitali che dovrebbe infine fornire, cioè le soluzioni delle equazioni sono proprio quelle necessarie per definire le equazioni del problema. Per risolvere questo inconveniente viene adottata una procedura iterativa:

1. Si adotta un set di prova di orbitali, senza alcuna particolare giustificazione, per l'inizio del processo.
2. Usando questi orbitali è possibile costruire un operatore di Fock approssimato che può quindi essere diagonalizzato per ottenere un nuovo set di orbitali.
3. I vecchi orbitali utilizzati per la costruzione dell'operatore di Fock vengono rimpiazzati da i nuovi orbitali ottenuti e si diagonalizza di nuovo per ottenere un nuovo set di orbitali.

Dopo un certo numero di iterazioni si può constatare che il set di orbitali varia pochissimo da un passo iterativo a quello successivo. Arrivati a questo punto, si può osservare che gli orbitali così ottenuti soddisfano le equazioni canoniche di Hartree-Fock, ottenendo la soluzione al problema.

3.2.3 I Metodi Post Hartree-Fock

Anche se l'approssimazione di HF ci dà dei risultati notevoli non tiene conto dell'*energia di correlazione elettronica*. Sono presenti diverse procedure in letteratura per migliorare calcoli

basati su un primo approccio HF, tali metodi sono chiamati *Post HF*.

CI e MCSCF

In questo paragrafo vengono presentati i concetti principali del metodo dell' *Interazione di Configurazione* (CI) [30] e del più sofisticato approccio *SCF Multi Configurazione* (MCSCF).

Il metodo CI si basa sul concetto che una qualunque funzione d'onda per N elettroni possa essere espressa esattamente come una combinazione lineare di tutti i possibili determinanti di Slater ad N elettroni (base polielettronica) formati partendo da un set completo di spin-orbitale monoelettronici (base monoelettronica) e usando in seguito il metodo variazionale. Se la base fosse completa si potrebbe ottenere in linea teorica l'esatta energia dello stato fondamentale ma anche di tutti gli stati elettronici eccitati del sistema. Se il sistema molecolare di interesse ha un numero pari di elettroni e può essere rappresentato da un singolo determinante RHF, $|\Psi_0\rangle$, è possibile ottenere gli altri determinanti modificando la struttura di questo primo determinante. Quindi un insieme possibile di determinanti dovrebbe comprendere $|\Psi_0\rangle$, i determinanti singolarmente eccitati $|\Psi_i^a\rangle$, che differiscono da $|\Psi_0\rangle$ perché lo spin-orbitale occupato ϕ_i viene sostituito con il precedente orbitale virtuale ϕ_a , i determinanti doppiamente eccitati $|\Psi_{ij}^{ab}\rangle$, etc., fino ad includere tutti i possibili determinanti eccitati N -esimi. Queste funzioni d'onda multi-elettroniche vengono utilizzate come base per espandere la funzione d'onda multi-elettronica esatta $|\Phi_0\rangle$ come:

$$|\Phi_0\rangle = c_0|\Psi_0\rangle + c_S|S\rangle + c_D|D\rangle + c_T|T\rangle + c_Q|Q\rangle + \dots \quad (3.26)$$

dove i $|S\rangle$ rappresentano i termini in cui sono coinvolte singole eccitazioni, $|D\rangle$ i termini in cui ci sono doppie eccitazioni, e così via. Data la funzione di prova definita nell'equazione

(3.26), le energie corrispondenti dei vari stati elettronici sono ottenute utilizzando il metodo variazionale lineare, che consiste nel formare la rappresentazione matriciale dell'hamiltoniano nella base delle funzioni d'onda utilizzate nell'espansione (3.26), e in seguito calcolare gli autovalori di tale matrice hamiltoniana. Se nell'espansione della funzione d'onda esatta vengono considerate tutte le infinite eccitazioni possibili di un sistema molecolare, la matrice hamiltoniana che ne deriva prende il nome di *matrice CI completa*, e il metodo è chiamato *Full-CI*. Di solito, per motivi ovvi, nell'espansione della funzione d'onda vengono considerate solo le singole eccitazioni, le doppie eccitazioni, o le singole e doppie eccitazioni e quindi le metodologie corrispondenti sono chiamate rispettivamente SCI, DCI e SDCI. Si noti che nel metodo CI solo i coefficienti dell'espansione (3.26) vengono ottimizzati. In generale la serie che determina la funzione d'onda CI converge molto lentamente quando si usano gli orbitali canonici di HF, ma qualunque troncamento a quella serie produce degli errori apprezzabili soprattutto quando si hanno molti elettroni. Esiste una serie di metodi computazionali detti *Multi Configurazionali* (MCSCF) in cui vengono inclusi nella serie CI solamente pochi determinanti. Non sarebbe possibile scegliere i determinanti da includere se si usassero gli orbitali canonici forniti da un calcolo HF in quanto, come abbiamo detto, la risultante serie CI converge lentamente con essi. Il metodo MCSCF quindi non si limita a determinare i coefficienti di mescolamento dei determinanti (come il Full-CI o il SDCI), ma ottimizza simultaneamente anche gli orbitali molecolari. Quindi la funzione d'onda MCSCF è un'espansione CI troncata

$$|\Psi_{MCSCF}\rangle = \sum_I c_I |\Psi_I\rangle \quad (3.27)$$

in cui vengono ottimizzati sia i coefficienti dell'espansione c_I che gli orbitali contenuti in $|\Psi_I\rangle$. Uno dei maggiori problemi dei calcoli MCSCF è legato alla costruzione della struttura

generale della funzione d'onda. Questo non può essere fatto senza una conoscenza a priori della struttura elettronica. Un metodo ampiamente utilizzato per costruire funzioni d'onda MCSCF è chiamato CASSCF [31]. Tale metodo parte dalla separazione di tutti gli orbitali molecolari occupati in sottogruppi chiamati orbitali molecolari attivi e inattivi. Gli orbitali molecolari attivi possono essere quelli di valenza, quelli maggiormente implicati nelle reazioni chimiche di interesse o quelli virtuali se stiamo cercando di caratterizzare anche gli stati eccitati. Gli orbitali molecolari inattivi vengono lasciati doppiamente occupati in tutte le configurazioni che sono utilizzate per costruire le funzioni d'onda CASSCF. I restanti elettroni liberi occupano in tutti i modi possibili il sottospazio dei soli orbitali attivi.

II DFT

Come già detto in precedenza i calcoli HF non tengono conto dell' *energia di correlazione elettronica*, ossia quel contributo energetico strettamente attribuibile all'effettiva interazione istantanea fra coppie di elettroni (intrinsecamente trascurata in qualsiasi modello a particelle indipendenti). In prima approssimazione questa limitazione può essere aggirata utilizzando metodi basati sulla teoria delle perturbazioni a molti corpi, ad esempio la Møller-Plesset (MP-n) [37] o i metodi *Coupled Cluster* [38, 39]. Una strada alternativa, computazionalmente molto più conveniente rispetto ai metodi perturbativi appena nominati, è la teoria basata sul *funzionale della densità* (DFT). Questa teoria porta alla sostituzione della complicata funzione d'onda multi-elettronica $\Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N)$ e dell'equazione di Schrödinger associata con la più semplice densità elettronica $\rho(\mathbf{r})$ e il suo schema di calcolo associato. All'inizio utilizzando la densità elettronica la correlazione elettronica è presa in considerazione implicitamente. Se N è il numero di elettroni del sistema, la *densità di particella* (o elettronica) è definita nell'equazione

(3.28),

$$\rho(\mathbf{r}) = N \int \dots \int |\Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N)|^2 d\mathbf{r} d\mathbf{x}_1 d\mathbf{x}_2 \dots d\mathbf{x}_N \quad (3.28)$$

dove $\int \rho(\mathbf{r}) d\mathbf{r} = N$. La teoria è derivata dall'idea originale di Thomas e Fermi del 1920 [40, 41], in cui si assumeva che gli elettroni erano distribuiti uniformemente nello spazio delle fasi a sei dimensioni, alla frequenza di due elettroni per ogni elemento di volume h^3 , e che vi fosse un campo di potenziale effettivo generato dalle cariche nucleari e dalla stessa distribuzione elettronica. I due teoremi di Hohenberg e Kohn [42] modificano la precedente intuizione fino a portarla a fornire predizioni quantitative per atomi, molecole e nella fisica dello stato solido. Essi dimostrarono, enunciando il loro primo teorema, che la densità elettronica dello stato fondamentale non degenerare di un sistema di elettroni permette di determinare univocamente tutte le proprietà dello stato fondamentale, e che quindi è possibile rappresentare l'energia come un funzionale della densità, come segue:

$$\begin{aligned} E[\rho] &= V_{ne}[\rho] + T[\rho] + V_{ee}[\rho] \\ &= \int \rho(\mathbf{r}) v(\mathbf{r}) d\mathbf{r} + T[\rho] + V_{ee}[\rho] \end{aligned} \quad (3.29)$$

dove $v(\mathbf{r})$ è il potenziale esterno in cui interagiscono gli elettroni, $T[\rho]$ l'energia cinetica e $V_{ee}[\rho]$ è l'energia di interazione elettrone-elettrone che contiene l'interazione di Coulomb $J[\rho]$. Il secondo teorema di Hohenberg-Kohn introduce il principio variazionale, inoltre restringe la teoria al solo stato fondamentale. Tale teorema enuncia che per una densità di prova $\tilde{\rho}(\mathbf{r})$ tale

che $\tilde{\rho}(\mathbf{r}) \geq 0$ e $\int \tilde{\rho}(\mathbf{r}) d\mathbf{r} = N$ valga che,

$$E_0 \leq E[\tilde{\rho}] \quad (3.30)$$

dove $E[\tilde{\rho}]$ è il funzionale dell'energia dell'equazione (3.29). L'importanza di questo teorema è nella sua praticità, vale a dire che, se si assume di possedere una buona rappresentazione del funzionale, allora è possibile cercare di minimizzarlo per ottenere i migliori orbitali e la migliore struttura. Si noti che l'uso del principio variazionale in questo processo limita l'uso del DFT per il calcolo del solo stato fondamentale. L'utilizzo attuale del DFT in chimica computazionale si basa sull'idea di Kohn e Sham [43] di considerare la funzione d'onda determinata per N elettroni non interagenti come N orbitali ϕ_i e quindi utilizzare su questi orbitali le *equazioni di Kohn-Sham* e risolvere il problema DFT per una data struttura molecolare. Queste equazioni danno la densità esatta, una volta che la parte di scambio e di correlazione del funzionale sono stati determinati. Inoltre tali equazioni sembrano simili alle equazioni SCF, quindi gli orbitali di Kohn-Sham solitamente sono espansi in termini dello stesso set atomico di base. Esistono altri modi di rappresentare il funzionale, come ad esempio l'*approssimazione locale della densità* (LDA) [44, 45], la *correzione dello scambio di Becke* [46] o il *potenziale di correlazione di Lee-Yang-Parr* (LYP) [47]. Va osservato che il DFT è fondamentalmente un metodo semi-empirico, perché probabilmente non sarà mai possibile trovare un funzionale che valuti in maniera esatta gli integrali di correlazione e di scambio. Tuttavia il DFT funziona molto meglio di altri metodi semi-empirici perché i parametri, che definiscono il funzionale, non sono specifici per ogni molecola. Inoltre negli ultimi anni il DFT è stato applicato con ottimi risultati su molti sistemi molecolari di dimensioni differenti. Recentemente il metodo DFT è diventato adatto anche per valutare le energie di eccitazione elettroniche. Tutto questo è possibile applicando una

perturbazione tempo dipendente (TD) [14] sullo stato fondamentale ben descritto dal normale DFT [48].

3.3 II PMM

La trattazione principale del PMM è ampiamente descritta nei seguenti articoli [20, 21, 49–52].

In questa sezione verrà descritta la teoria e le implicazioni fisiche del PMM. Infine, si mostrerà come utilizzare le analisi statistiche su i risultati del PMM, con lo scopo di caratterizzare gli stati quantistici perturbati del *centro quantistico* (QC).

3.3.1 La teoria di base

Definendo con \mathbf{r}_n le coordinate del QC e con \mathbf{x} le coordinate degli atomi che generano il campo perturbante (classico), si può scrivere il seguente operatore Hamiltoniano elettronico complessivo¹ $\hat{\mathcal{H}}$ come

$$\hat{\mathcal{H}}(\mathbf{r}_n, \mathbf{x}) = \hat{H}^0(\mathbf{r}_n) + \hat{V}_{QC,e}(\mathbf{r}_n, \mathbf{x}) + \hat{\mathcal{H}}_e^0(\mathbf{x}) \quad (3.31)$$

dove \hat{H}^0 è l'operatore Hamiltoniano del QC isolato, $\hat{V}_{QC,e}$ rappresenta l'operatore di interazione del QC con l'ambiente e $\hat{\mathcal{H}}_e^0$ è l'operatore Hamiltoniano dell'ambiente (escludendo le interazioni QC-ambiente). Per ogni configurazione atomica definita da \mathbf{r}_n e \mathbf{x} si assume sempre che lo stato elettronico complessivo Ψ_T possa essere espresso dal prodotto di una funzione d'onda Φ che descriva il QC e una funzione d'onda che descriva l'ambiente Ψ_e , ossia $\Psi_T \cong \Phi\Psi_e$. Tale ipotesi può essere considerata come una approssimazione eccellente fintanto che il QC e l'ambiente possano essere ben distinti e trattati come appartenenti a diversi sistemi fisici, condizione che viene rispettata nel caso in cui non si considerino reazioni chimiche o in-

¹con Hamiltoniano elettronico complessivo si intende l'operatore dell'energia elettronica dove le interazioni nucleo-nucleo sono incluse

terazioni molto strette tra il QC e l'ambiente. Se si assume inoltre che per ogni stato eccitato del QC lo stato fondamentale dell'ambiente rimanga praticamente invariato, per ogni stato quantico complessivo in cui solamente il QC è in uno stato eccitato si ha che

$$\hat{\mathcal{H}}\Psi_{T,n} = \mathcal{H}_n\Psi_{T,n} \quad (3.32)$$

$$\Psi_{T,n} = \Phi_n\Psi_{e,0} \quad (3.33)$$

il quale implica

$$\hat{\mathcal{H}}_{\Psi_{e,0}}\Phi_n = \mathcal{H}_n\Phi_n \quad (3.34)$$

$$\hat{\mathcal{H}}_{\Psi_{e,0}} = \langle \Psi_{e,0} | \hat{\mathcal{H}} | \Psi_{e,0} \rangle \quad (3.35)$$

in cui \mathcal{H}_n è l'autovalore dell'Hamiltoniano elettronico complessivo. Inoltre, dalla definizione dell'operatore Hamiltoniano si ha

$$\hat{\mathcal{H}}_{\Psi_{e,0}} = \hat{H}^0 + \hat{V}_{\Psi_{e,0}} + \mathcal{H}_{e,0} \quad (3.36)$$

$$\hat{V}_{\Psi_{e,0}} = \langle \Psi_{e,0} | \hat{V}_{QC,e} | \Psi_{e,0} \rangle \quad (3.37)$$

$$\mathcal{H}_{e,0} = \langle \Psi_{e,0} | \hat{\mathcal{H}}_e^0 | \Psi_{e,0} \rangle \quad (3.38)$$

Si noti che $\hat{V}_{\Psi_{e,0}}$ corrisponde alla perturbazione del QC ad opera della distribuzione di carica dell'ambiente descritta dalla funzione d'onda $\Psi_{e,0}$ (cioè lo stato fondamentale perturbato dell'ambiente). Quindi, separando il termine dovuto alle cariche atomiche dell'ambiente $\hat{V}_{\Psi_{e,0}}^{ac}$

da quelli di ordine superiore, espressi da $\hat{V}_{\Psi_{e,0}}^{ac}$, si ottiene

$$\hat{\mathcal{H}}_{\Psi_{e,0}} = \hat{H}^0 + \hat{V}_{\Psi_{e,0}}^{ac} + \Delta\hat{V}_{\Psi_{e,0}} + \mathcal{H}_{e,0} \quad (3.39)$$

Fornendolo invece in notazione matriciale otteniamo

$$\tilde{\mathcal{H}}_{\Psi_{e,0}} = \tilde{H}^0 + \tilde{V}_{\Psi_{e,0}}^{ac} + \Delta\tilde{V}_{\Psi_{e,0}} + \mathcal{H}_{e,0}\tilde{I} \quad (3.40)$$

in cui si usa il `basis-set` imperturbato del QC per esprimere tutte le matrici (cioè la matrice Hamiltoniana \tilde{H}^0 del QC è diagonale). L'espressione dell'operatore ottenuta si può sviluppare ulteriormente considerando che per un QC interagente con un set di cariche puntiformi, le cariche atomiche, le interazioni di dispersione a coppia e le repulsioni a corto raggio non sono presenti e quindi, assumendo $\Delta\tilde{V}_{\Psi_{e,0}} \cong \Delta V_{\Psi_{e,0}}\tilde{I}$ (ossia che gli autovettori di $\tilde{\mathcal{H}}_{\Psi_{e,0}}$ coincidano con gli autovettori di $\tilde{H}^0 + \tilde{V}_{\Psi_{e,0}}^{ac}$), si può scrivere

$$\begin{aligned} \tilde{\mathcal{H}}_{\Psi_{e,0}} &\cong \tilde{H}^0 + \tilde{V}_{\Psi_{e,0}}^{ac} + (\Delta V_{\Psi_{e,0}} + \mathcal{H}_{e,0})\tilde{I} \\ &\cong \tilde{H}^0 + \tilde{V}_{\Psi_{e,0}}^{ac} + (\mathcal{U}'_{env} + \mathcal{U}'_{D,R})\tilde{I} \end{aligned} \quad (3.41)$$

dove, all'interno dell'approssimazione utilizzata $\Delta V_{\Psi_{e,0}} + \mathcal{H}_{e,0} \cong \mathcal{U}'_{env} + \mathcal{U}'_{D,R}$ il termine \mathcal{U}'_{env} è l'energia potenziale dell'ambiente (escludendo tutte le interazioni semi-classiche con il QC) e il termine $\mathcal{U}'_{D,R}$ rappresenta le interazioni a coppia di repulsione e dispersione tra il QC e l'ambiente generate dal campo di forze utilizzato nella simulazione di MD (si noti che il potenziale di interazione $\tilde{\Delta V}_{\Psi_{e,0}}$ non è in generale equivalente al termine di interazione di dispersione e repulsione come per $\mathcal{U}'_{D,R}$ perché in questo caso possono essere coinvolte energie di rilassamento intramolecolari). Infine espandendo $\tilde{V}_{\Psi_{e,0}}^{ac}$ in multipoli del QC e approssimando

tutti i termini da i quadrupoli in poi tramite un potenziale a corto raggio (cioé $\Delta V_{\Psi_{e,0}}^{ac} \tilde{I}$ con \tilde{I} la matrice identità), si ottiene [49–52]

$$\begin{aligned} \tilde{H} &= \tilde{\mathcal{H}}_{\Psi_{e,0}}(\mathbf{r}_n, \mathbf{x}) - \mathcal{U}'_{env} \tilde{I} = \tilde{H}^0(\mathbf{r}_n) + \tilde{V}(\mathbf{r}_0, \mathbf{x}) \\ &\cong \tilde{H}^0(\mathbf{r}_n) + q_T \mathcal{V}(\mathbf{r}_0, \mathbf{x}) \tilde{I} + \tilde{Z}_1(\mathbf{E}(\mathbf{r}_0, \mathbf{x}), \mathbf{r}_n) + \Delta V(\mathbf{r}_n, \mathbf{x}) \tilde{I} \end{aligned} \quad (3.42)$$

in cui q_T è la carica totale del QC, $\mathcal{V}(\mathbf{r}_0, \mathbf{x})$ e $\mathbf{E}(\mathbf{r}_0, \mathbf{x})$ sono rispettivamente il potenziale elettrico perturbante e il campo elettrico prodotto dalle cariche atomiche dell'ambiente, per una data geometria del QC \mathbf{r}_0 (tipicamente il centro di massa o geometrico del QC), $\tilde{Z}_1(\mathbf{E}, \mathbf{r}_n)$ è la matrice di perturbazione dell'energia data esplicitamente da $[\tilde{Z}_1]_{l,l'} = -\mathbf{E} \cdot \langle \Phi_l^0 | \hat{\boldsymbol{\mu}} | \Phi_{l'}^0 \rangle$ in cui $\hat{\boldsymbol{\mu}}$ è l'operatore di dipolo, e $\Delta V(\mathbf{r}_n, \mathbf{x}) = \Delta V_{\Psi_{e,0}}^{ac} + \mathcal{U}'_{DR}$ approssima tutti i termini di ordine superiore come un potenziale a corto raggio (nel caso in cui il QC fosse solo una parte di una molecola più grande, ΔV potrebbe includere una costante additiva corrispondente alla possibile variazione dell'energia di riferimento). L'equazione (3.42) mostra la matrice Hamiltoniana perturbata del QC \tilde{H} e fornisce l'espressione del PMM per un QC interagente in maniera semi-classica con l'ambiente. Esprimendo tale equazione in notazione matriciale

$$\begin{aligned} \hat{H} &= \hat{H}^0 + \hat{V} \\ &\cong \hat{H}^0 + q_T \mathcal{V} - \mathbf{E} \cdot \hat{\boldsymbol{\mu}} + \Delta V \end{aligned} \quad (3.43)$$

con l'operatore di perturbazione \hat{V} fisicamente coincidente alla perturbazione dovuta al campo generato dallo stato fondamentale dell'ambiente agente sul QC. Gli autovettori di \tilde{H} , cioè gli autostati di \hat{H} , possono essere quindi utilizzati per ottenere le proprietà elettroniche

perturbate del QC. Deve essere sottolineato tuttavia che le equazioni (3.42) e (3.43), basandosi sull'approssimazione che l'operatore di perturbazione possa essere troncato al termine dipolare includendo esplicitamente i termini di ordine superiore come un potenziale a corto raggio indipendente dallo stato elettronico, possono non essere adeguate per descrivere il tipo di interazioni QC-ambiente in cui sia necessario esplicitare termini di ordine superiore o operatori multipolari distribuiti. D'altronde, quando è possibile ignorare, stati elettronici eccitati ad alta energia l'approssimazione utilizzata nelle equazioni (3.42) e (3.43) è tipicamente piuttosto accurata. L'introduzione di termini multipolari, benchè possibile in teoria [50, 52], richiederebbe l'utilizzo di multipoli di transizione che allo stato attuale dei software sono meno accurati dei corrispondenti dipoli, il che si tradurrebbe in un PMM di peggior qualità.

3.3.2 Il calcolo dello spettro di assorbimento

In questo paragrafo verrà descritto come attraverso il calcolo degli stati elettronici perturbati del QC durante la simulazione, limitando l'attenzione ad un sistema con un solo cromoforo, cioè considerando un singolo QC immerso in un ambiente atomistico semi-classico, sia possibile ricostruire lo spettro di assorbimento UV complessivo. Tale approccio è stato utilizzato in questo lavoro di tesi, in cui il QC è definito dalla DLH_2 e dai suoi derivati simulati a diluizione infinita in acqua e DMSO. Se si effettua una dinamica molecolare, o più in generale un campionamento dello spazio delle fasi, viene prodotta, durante la simulazione, una 'sequenza' di potenziali elettrici \mathcal{V} e campi elettrici \mathbf{E} perturbanti da cui si può ottenere la matrice Hamiltoniana complessiva perturbata $\tilde{H}(\mathbf{r}_n, \mathbf{x})$. Quindi, la diagonalizzazione di $\tilde{H}(\mathbf{r}_n, \mathbf{x})$ effettuata lungo questo campionamento di configurazioni produrrà una 'traiettoria' di autovalori (energie) e autovettori (stati elettronici c_i) perturbati del QC. Attraverso gli autovettori perturbati è possibile calcolare i dipoli di transizione perturbati ($\mu_{i,j}$), necessari per ottenere lo spettro di assorbimento,

mediante la seguente equazione

$$\begin{aligned}\boldsymbol{\mu}_{i,j} &= \langle \Phi_i | \hat{\boldsymbol{\mu}} | \Phi_j \rangle \\ &= \mathbf{c}_i^T \tilde{\Gamma}_x^0 \mathbf{c}_j \hat{i} + \mathbf{c}_i^T \tilde{\Gamma}_y^0 \mathbf{c}_j \hat{j} + \mathbf{c}_i^T \tilde{\Gamma}_z^0 \mathbf{c}_j \hat{k}\end{aligned}\quad (3.44)$$

in cui le matrici $\tilde{\Gamma}^0$ dei dipoli sono espresse nella base imperturbata, ossia

$$\left[\tilde{\Gamma}_x^0 \right]_{l,m} = \langle \Phi_l^0 | \hat{\mu}_x | \Phi_m^0 \rangle \quad (3.45)$$

$$\left[\tilde{\Gamma}_y^0 \right]_{l,m} = \langle \Phi_l^0 | \hat{\mu}_y | \Phi_m^0 \rangle \quad (3.46)$$

$$\left[\tilde{\Gamma}_z^0 \right]_{l,m} = \langle \Phi_l^0 | \hat{\mu}_z | \Phi_m^0 \rangle \quad (3.47)$$

Una volta ottenuti i dipoli di transizione è possibile ottenere lo spettro elettronico (considerando una densità unitaria di energia per unità di frequenza e trascurando il contributo vibrazionale) attraverso il coefficiente di assorbimento molecolare,

$$\epsilon_{i,j}(\nu) = \frac{B_{i,j}(\nu) \rho(\nu) h \nu}{c} \quad (3.48)$$

$$B_{i,j} = \frac{|\boldsymbol{\mu}_{i,j}|^2}{6\epsilon_0 \hbar^2} \quad (3.49)$$

in cui $B_{i,j}(\nu)$ è il coefficiente di Einstein dipendente dalla frequenza ν . Nelle equazioni (3.49,3.48), ϵ_0 è la costante dielettrica nel vuoto, $\rho(\nu)$ è la densità di probabilità di assorbimento in funzione della frequenza ν , cioè la densità di probabilità di trovare un cromoforo che assorba energia in un intervallo infinitesimale di frequenza centrato su ν (tale funzione viene ricostruita attraverso l'analisi della traiettoria di MD su cui viene applicato il PMM), h è la

costante di Plank, con $\hbar = h/2\pi$, e c è la velocità della luce. La conoscenza delle funzioni $\epsilon_{0,j}(\nu)$ calcolate con 0 che rappresenta lo stato fondamentale per tutti i possibili stati eccitati j permette la ricostruzione dello spettro di assorbimento complessivo del QC perturbato. Occorre ricordare che gli autostati della matrice Hamiltoniana complessiva perturbata $\tilde{H}(\mathbf{r}_n, \mathbf{x})$, sono una ricombinazione degli autostati imperturbati del QC, perciò benchè in alcuni casi sia possibile identificare delle corrispondenze dirette tra autostati perturbati e imperturbati, quando i vari autostati non rimescolino eccessivamente, in generale tale corrispondenza non è possibile.

Capitolo 4

Risultati sullo studio del force-field

4.1 Parametrizzazione delle molecole

4.1.1 Il calcolo delle cariche

Le molecole utilizzate in questo lavoro di tesi non appartengono al database di GROMACS [19], quindi sono state costruite e parametrizzate dal nostro gruppo secondo il seguente metodo. Sono state per prima cosa costruite le strutture delle molecole con il software di modellizzazione molecolare 3D Hyperchem [53]. Le geometrie ottenute sono state poi ottimizzate con il software GAUSSIAN [54], un programma per la modellizzazione di strutture elettroniche, utilizzando il *basis-set* cc-pVTZ, e il metodo DFT denominato B3LYP [55]. Le strutture così ottimizzate sono state poi utilizzate per costruire una griglia di punti in cui è stato calcolato il potenziale elettrico con un passo di 0.3 Å e una distanza massima di ogni punto calcolato all'atomo più vicino non superiore ai 3.0 Å con il programma di modellizzazione di strutture elettroniche chiamato GAMESS [56] con il *basis-set* cc-pVTZ, e il metodo B3LYP [55]. I valori del potenziale elettrico in questi punti sono stati poi utilizzati per calcolare le cariche puntiformi su ogni atomo, con una procedura di fitting chiamata RESP [57], in modo tale che riproducano al meglio la superficie di potenziale elettrico.

4.1.2 L'angolo diedro C2-C2'

Una volta ottenute le cariche elettriche su i vari atomi e la geometria di equilibrio della molecola, il passo successivo è stato quello di trovare le costanti di forza adeguate a descrivere il comportamento intramolecolare. Gran parte di queste costanti sono state prese per analogia da molecole simili già parametrizzate nel `force-field` di GROMACS [19], tuttavia la costante di forza K_ϕ dell'angolo diedro passante per il legame C2-C2' (paragrafo:2.4), necessaria per utilizzare la funzione di potenziale tipica della MD (l'equazione (2.4)), non era mai stata parametrizzata per le molecole di LH_2 e $OxyLH_2$. Per ottenere K_ϕ è stato utilizzato il seguente metodo:

1. Sono state fatte varie ottimizzazioni di geometria della stessa molecola tenendo fisso unicamente l'angolo diedro passante per il legame C2-C2' partendo da un angolo diedro di 0 gradi (molecola in posizione *trans*) fino a 180 gradi, con un passo di 15 gradi.
2. Per ogni geometria ottimizzata sono state calcolate le energie totali a cui sono state sottratte le energie di interazione elettrostatiche e di Lennard-Jones di tutte le possibili coppie di atomi la cui distanza varia al variare del solo angolo diedro C2-C2' e quindi costruite scegliendo atomi appartenenti alle 2 parti planari della molecola (anello tiazolico e anello benzotiazolico) separate dal legame C2 e C2'.
3. Una volta posti questi valori di energia in kJ/mol all'interno di un grafico, il fitting di tale grafico con la funzione periodica $y = K_\phi \left[1 + \cos \left(\frac{2\pi x}{180} - \pi \right) \right]$ (Figura:4.1.1), utilizzando come unico parametro K_ϕ , permette, in linea di principio di identificare il valore di questa costante.

Purtroppo a causa dell'estrema vicinanza dei 2 atomi di zolfo nei tipi di molecola studiate, le costanti C^6 e C^{12} per il potenziale di interazione Lennard-Jones fornite dal `force-field` di GROMACS mal si adattavano a descrivere l'interazione zolfo-zolfo e quindi il profilo di energia ottenuto alla fine del punto 2 dello schema precedente non era simmetrico come sarebbe dovuto essere.

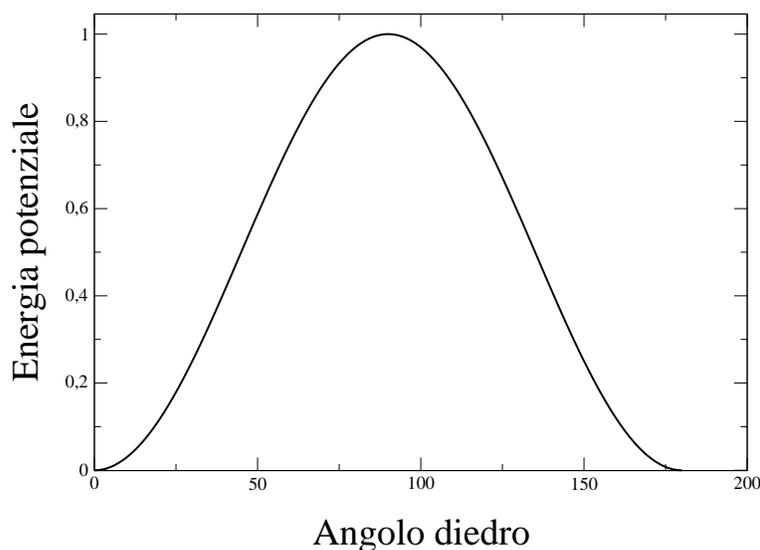


Figura 4.1.1: funzione esempio $y = 0.5 \left[1 + \cos \left(\frac{2\pi x}{180} - \pi \right) \right]$

Per superare questo ulteriore problema tale profilo di energia è stato ricalcolato tralasciando l'interazione Lennard-Jones zolfo-zolfo. Il nuovo profilo di energia così ottenuto è stato usato per il fitting con la funzione (4.1),

$$y = K_{\phi} \left[1 + \cos \left(\frac{2\pi x}{180} - \pi \right) \right] - \left(\frac{C^{12}}{r^{12}} - \frac{C^6}{r^6} \right) \quad (4.1)$$

in cui la distanza zolfo-zolfo r è determinata dall'angolo diedro x , ottenendo contemporaneamente i tre parametri C^6 , C^{12} (dell'interazione zolfo-zolfo) e K_{ϕ} dell'angolo diedro C2-C2'.

Le costanti di forza K_{ϕ} trovate sono riportate nella Tabella:4.1.1.

	K_ϕ (kJ/mol)
DLH_2	13.934
cheto- $OxyLH_2$	17.028
fenolato-cheto- $OxyLH^-$	33.205
enol- $OxyLH_2$	15.071
fenolato-enol- $OxyLH^-$	20.151

Tabella 4.1.1: costanti di forza K_ϕ dell'angolo diedro C2-C2' per le molecole studiate

4.2 La distribuzione degli angoli diedri C2-C2'

Per completezza in questa sezione verranno mostrate le distribuzioni (normalizzate) degli angoli diedri C2-C2' così come ottenute dalle simulazioni. Si è preso come angolo 0 la configurazione *trans*, e come mostrano tutti gli istogrammi (da Figura:4.2.1 - 4.2.5), durante le simulazioni le molecole non riescono mai ad accedere alla configurazione *cis*. Le costanti di forza calcolate come descritto nel Paragrafo:4.1.2 e le interazioni con il solvente producono delle oscillazioni intorno alla posizione di equilibrio (*trans*).

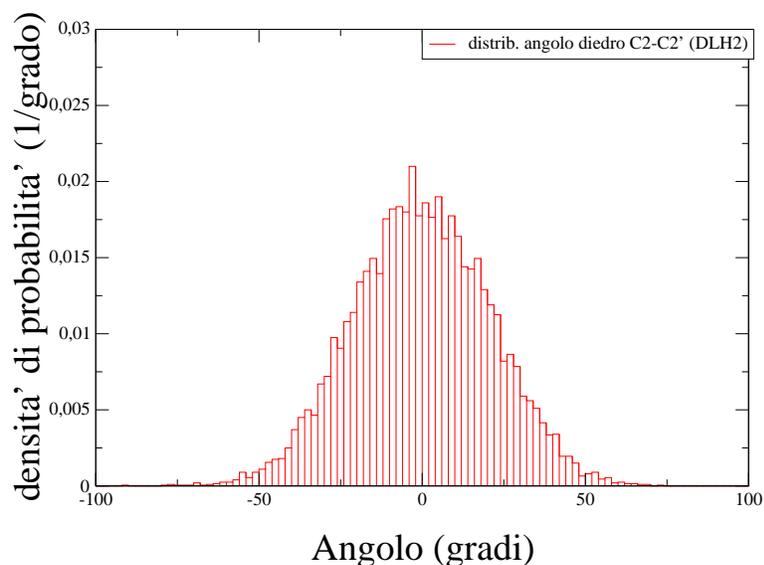
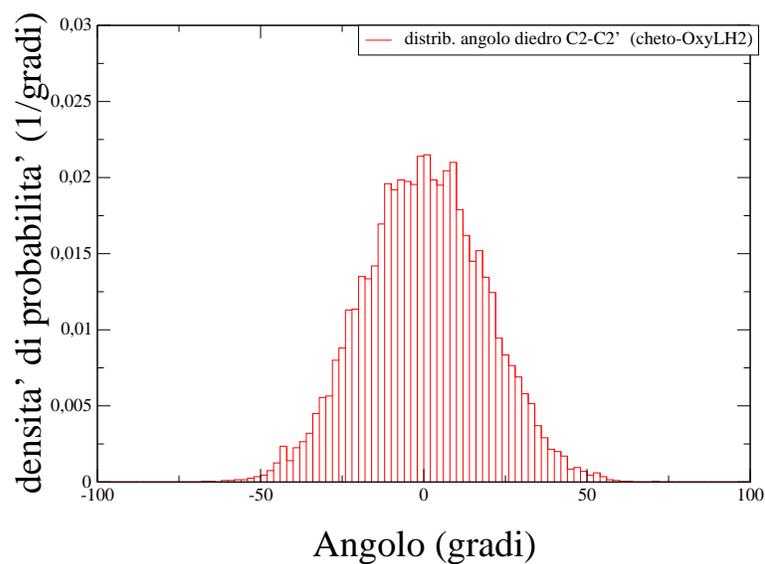
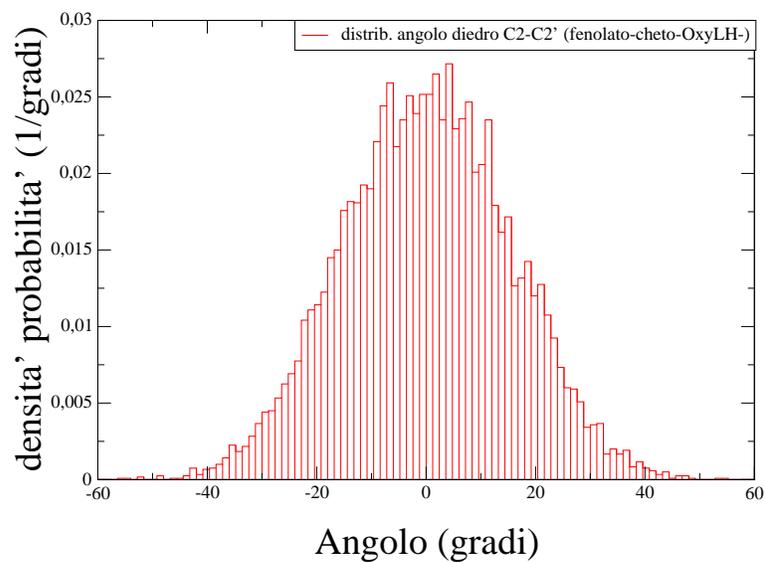


Figura 4.2.1: distribuzione di probabilità dell'angolo diedro C2-C2' per la DLH_2 in acqua

Figura 4.2.2: distribuzione di probabilita' dell'angolo diedro C2-C2' per la cheto-OxyLH₂Figura 4.2.3: distribuzione di probabilita' dell'angolo diedro C2-C2' per il fenolato-cheto-OxyLH⁻

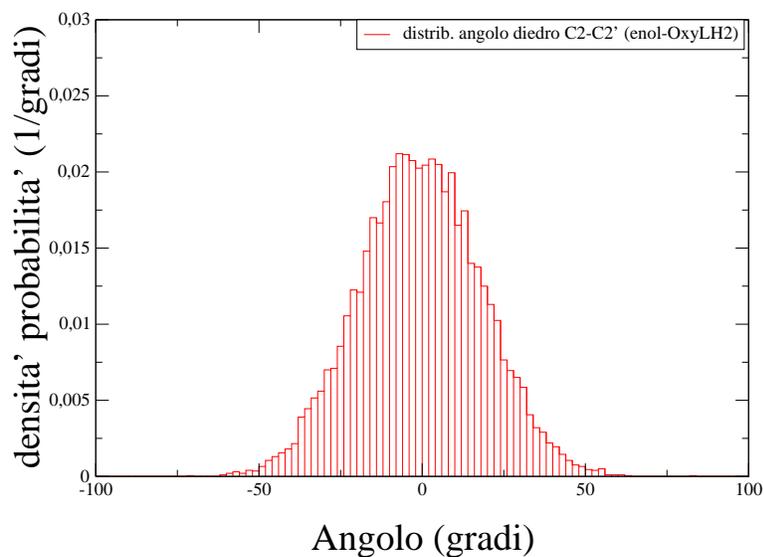


Figura 4.2.4: distribuzione di probabilità dell'angolo diedro C2-C2' per la enol-*OxyLH*₂ in DMSO

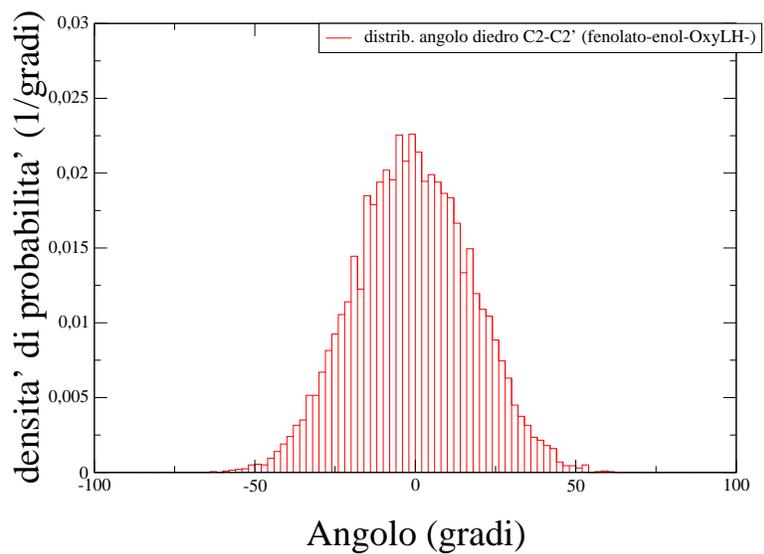


Figura 4.2.5: distribuzione di probabilità dell'angolo diedro C2-C2' per la fenolato-enol-*OxyLH*⁻ in acqua

4.3 Parametri della Simulazione di Dinamica Molecolare

Le molecole che sono state simulate sono le seguenti:

- la DLH_2 (Figura:1.2.1) in acqua con tutti i protoni;
- la cheto- $OxyLH_2$ (Figura:1.3.1) in acqua e in DMSO;
- il fenolato-cheto- $OxyLH^-$ (Figura:1.3.1) in acqua, utilizzando come controione un atomo di Na^+ ;
- l'enol- $OxyLH_2$ (Figura:1.3.1) in acqua e in DMSO;
- il fenolato-enol- $OxyLH^-$ (Figura:1.3.1) in acqua, utilizzando come controione un atomo di Na^+ .

Tutte le simulazioni, in acqua e in DMSO, sono state effettuate con il pacchetto simulativo GROMACS [19], le velocità iniziali delle particelle sono state generate con una semplice distribuzione maxwelliana, dipendente dalla temperatura della simulazione. Tutte le traiettorie simulate sono di 10 ns (tranne quella della DLH_2 che è di 20 ns) e sono ottenute usando 10000000 di passi ed un tempuscolo di integrazione di 1 fs . La temperatura di tutte le simulazioni è di 300 K. Come metodo per il controllo della temperatura è stato utilizzato l'*accoppiamento isoterma gaussiano* (descritto nel paragrafo:2.10), per tutte le simulazioni è stato usato l'algoritmo LINCS per vincolare le lunghezze di legame ed inoltre è stato bloccato il centro di massa del soluto. Le interazioni a corto raggio sono state calcolate entro un raggio di cut-off di 0.9 nm . L'energia elettrostatica a lungo raggio è stata calcolata utilizzando il metodo Particle Mesh Ewald [25]. Per le simulazioni in acqua è stato costruito un box cubico di 674 molecole, 673 molecole di acqua più quella di interesse al centro del box, ad una densità di 55.5 mol/L (tipica

densità dell'acqua liquida a temperatura ambiente). Per le simulazioni in DMSO è stato costruito un box cubico di 163 molecole, 162 di DMSO più quella di interesse al centro del box, ad una densità di 14.08 mol/L (tipica densità del DMSO a temperatura ambiente).

4.4 Parametri calcolo quantistico

Come spiegato nel capitolo 3, per effettuare il calcolo PMM sono necessarie:

1. la matrice Hamiltoniana imperturbata definita dalle energie degli stati elettronici imperturbati (stati elettronici delle molecola isolata).
2. i dipoli dello stato fondamentale e degli stati eccitati imperturbati.
3. Tutti i possibili dipoli di transizione tra gli stati elettronici imperturbati.

Per ottenere tutte queste informazioni è stato utilizzato il software di chimica quantistica DALTON [58], utilizzando il *basis-set* 6-31+G* e il metodo TD-DFT [14] con il funzionale B3LYP [55]. Infine la base imperturbata utilizzata nel calcolo PMM ha incluso lo stato fondamentale e i primi 11 stati eccitati garantendo una ottima convergenza nella regione di spettro di interesse.

Capitolo 5

Risultati sulle proprietà elettroniche

In questo lavoro di tesi le metodologie teorico computazionali descritte nei capitoli precedenti sono state utilizzate per studiare il comportamento spettroscopico della D-luciferina (DLH_2) e di alcuni suoi derivati in acqua e in DMSO. Tale studio si propone da un lato di verificare l'accuratezza quantitativa ed eventualmente i possibili limiti dei metodi teorico-computazionali utilizzati e dall'altro, se ottenuta una sufficiente accuratezza, di interpretare gli spettri sperimentali in termini delle eccitazioni principali coinvolte nelle regioni spettrali di interesse.

5.1 Lo spettro della D-luciferina

Per verificare l'accuratezza del metodo usato è stato riprodotto lo spettro della DLH_2 in acqua (pH~5), condizione in cui questa molecola non è soggetta a nessun equilibrio chimico di rilievo ed è stabilmente nella sua forma protonata [59]. Nel grafico (Figura:5.1.1) confrontiamo lo spettro sperimentale della DLH_2 in acqua con lo spettro corrispondente ottenuto dal metodo PMM/MD. Da tale grafico (Figura:5.1.1) è evidente l'ottimo livello di riproduzione dello spettro sperimentale ottenuto dalla nostra metodologia. Si noti che il calcolo teorico non riproduce quasi esattamente solo la posizione del picco principale ma anche la posizione del picco secondario a più alta energia, così come la forma e l'ampiezza dello spettro sperimentale.

L'intensità del picco secondario è maggiore nello spettro sperimentale che in quello PMM, probabilmente a causa della parziale sovrapposizione di eccitazioni di ordine superiore trascurate nel nostro calcolo (nel presente lavoro abbiamo utilizzato una base imperturbata definita dai primi 12 stati elettronici).

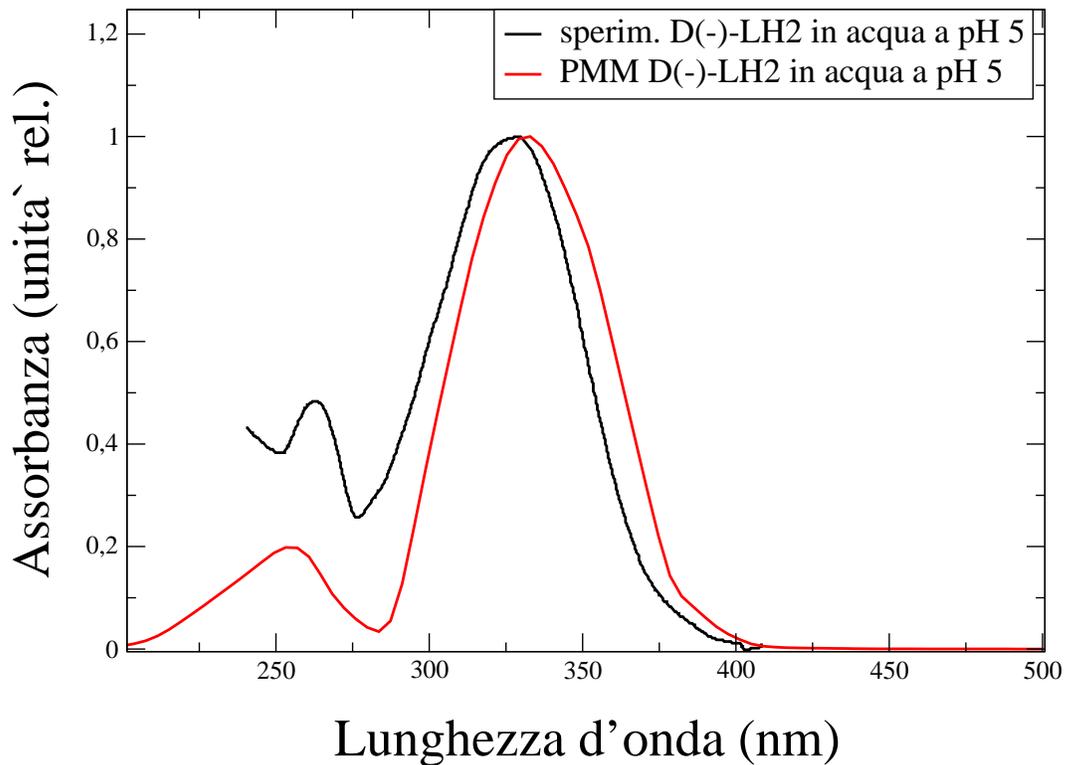


Figura 5.1.1: In nero lo spettro di assorbimento UV sperimentale della DLH_2 a $\text{pH} \sim 5$ ($\lambda_{Abs1} = 262 \text{ nm}$ e $\lambda_{Abs2} = 328 \text{ nm}$) e in rosso lo spettro ottenuto dal PMM ($\lambda_{PMM1} = 258 \text{ nm}$ e $\lambda_{PMM2} = 332 \text{ nm}$)

Verificata la sostanziale accuratezza nel riprodurre gli aspetti quantitativi fondamentali della DLH_2 in acqua, è stato possibile caratterizzare le singole eccitazioni coinvolte nello spettro complessivo studiato e quindi identificare per ognuno dei picchi rilevati i corrispondenti stati eccitati principali (Figura:5.1.2, Figura:5.1.3).

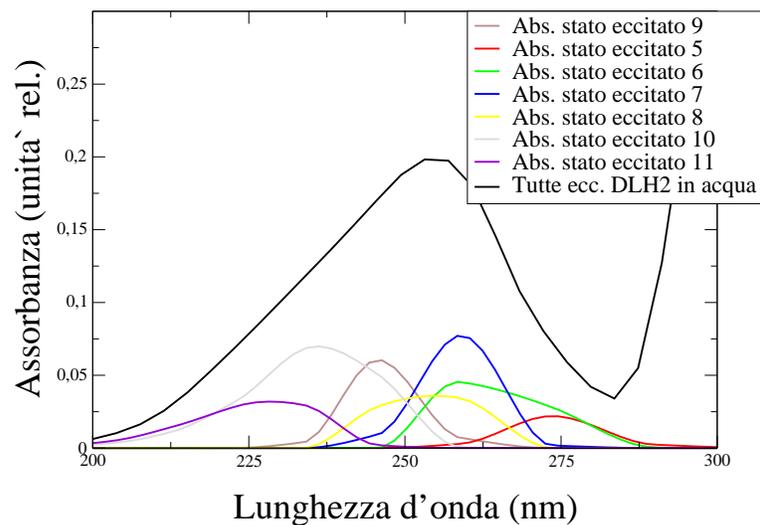


Figura 5.1.2: La scomposizione del massimo di assorbimento a $\lambda_{PMM1} = 258 \text{ nm}$ della DLH_2 negli assorbimenti dovuti alle singole eccitazioni

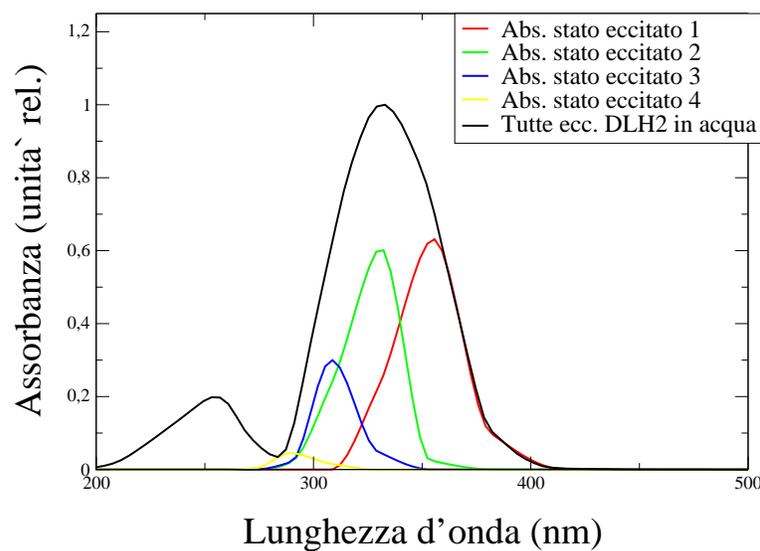


Figura 5.1.3: La scomposizione del massimo di assorbimento a $\lambda_{PMM2} = 332 \text{ nm}$ della DLH_2 negli assorbimenti dovuti alle singole eccitazioni

E' interessante notare che i risultati ottenuti dal metodo PMM/MD mostrano che il picco principale della DLH_2 ($\lambda_{PMM} = 332 \text{ nm}$) è il risultato della combinazione delle prime 4 eccitazioni e quindi non è riconducibile alla sola eccitazione 0->1 come invece ipotizzato sulla base di calcoli *ab initio* fatti in alcuni lavori computazionali precedenti [13]. Inoltre si può notare come il picco secondario (Figura:5.1.2) sia la combinazione di tutte le eccitazioni superiori alla quarta, tali eccitazioni diventano sempre più dense ma diminuiscono via via in intensità. Tale comportamento è stato riscontrato per tutte le molecole studiate in questo lavoro di tesi.

5.2 Lo spettro della Ossiluciferina

Come già descritto nel paragrafo (1.3), al contrario della DLH_2 , la $OxyLH_2$, molecola responsabile dell'emissione luminosa nell'enzima luciferasi, è una specie chimica piuttosto mutevole in soluzione che presenta un triplice equilibrio chimico e una bassa stabilità termica. Per indagare in dettaglio il comportamento spettroscopico complessivo della $OxyLH_2$ sono state studiate tutte le sue principali forme che coesistono in soluzione e che sono inoltre di rilievo nel processo enzimatico della bioluminescenza. A causa della sua instabilità il confronto con i dati sperimentali è estremamente difficile. Fortunatamente di recente Hirano et al. [16] e Naumov et al. [17] hanno ottenuto gli spettri UV in soluzione di alcune forme della $OxyLH_2$ [17] (enol- $OxyLH_2$ e il fenolato-enol- $OxyLH^-$) e di alcuni suoi analoghi (la $MOxyLH_2$ e la $DMOxyLH_2$) [16], essenzialmente corrispondenti alle forme di $OxyLH_2$ non rilevabili in soluzione (cheto- $OxyLH_2$ e fenolato-cheto- $OxyLH^-$).

Nei grafici seguenti sono riportati, analogamente al caso della DLH_2 , i confronti degli spettri ottenuti attraverso il metodo PMM/MD con i corrispettivi spettri sperimentali per le seguenti forme della $OxyLH_2$:

1. la cheto- $OxyLH_2$ (Figura:1.3.1) in DMSO, spettroscopicamente corrispondente ai dati sperimentali sul $DMOxyLH_2$ in DMSO [16].
2. il fenolato-cheto- $OxyLH^-$ (Figura:1.3.1) in acqua, spettroscopicamente corrispondente ai dati sperimentali sul fenolato- $DMOxyLH^-$ in acqua (pH~7) [16].
3. l'enol- $OxyLH_2$ (Figura:1.3.1) in DMSO il cui spettro sperimentale è noto [17].
4. il fenolato-enol- $OxyLH^-$ (Figura:1.3.1) in acqua il cui spettro sperimentale è noto [17].

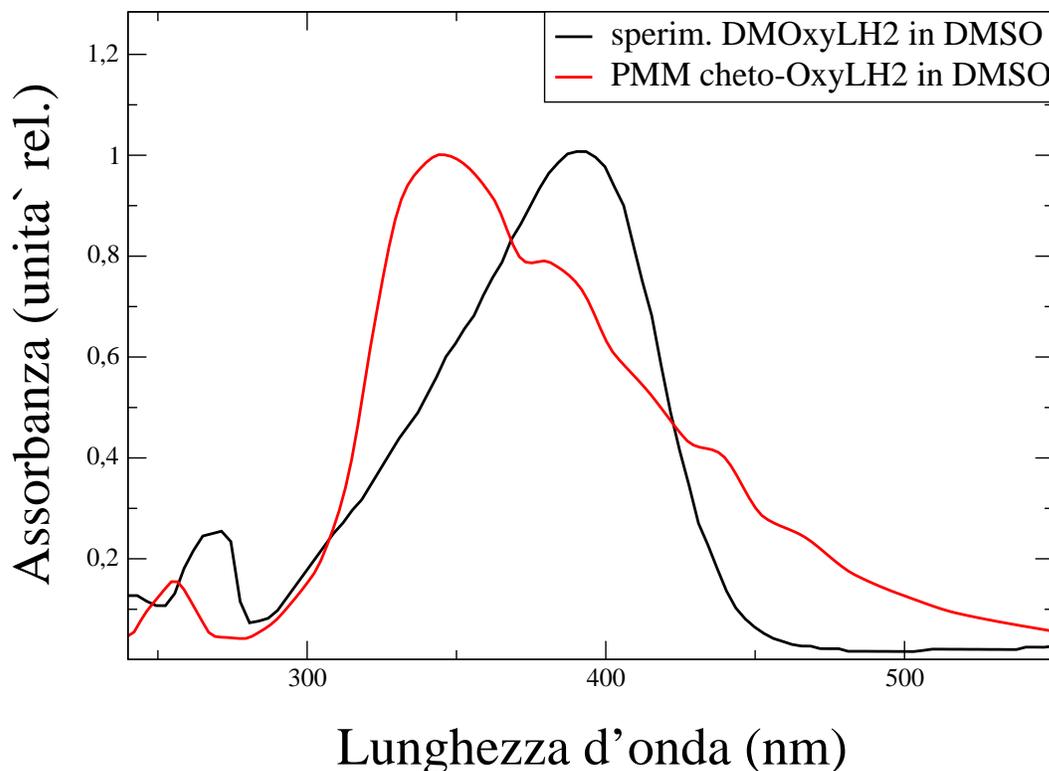
1. cheto-*OxyLH*₂ in DMSO.

Figura 5.2.1: In nero lo spettro di assorbimento UV sperimentale del *DMOxyLH*₂ in DMSO ($\lambda_{Abs1} = 268 \text{ nm}$ e $\lambda_{Abs2} = 390 \text{ nm}$) e in rosso lo spettro del cheto-*OxyLH*₂ ottenuto con il PMM ($\lambda_{PMM1} = 255 \text{ nm}$ e $\lambda_{PMM2} = 345 \text{ nm}$).

Nella Figura:5.2.1 si può notare come, benchè si stiano confrontando spettri di molecole differenti, lo spettro di assorbimento PMM della cheto-*OxyLH*₂ in DMSO (spettro non rivelabile sperimentalmente) sia molto simile per forma, ampiezza e posizione dei picchi (lo spettro PMM mostra uno shift a energie più alte) a quello della *DMOxyLH*₂ in DMSO. I risultati ottenuti in Figura:5.2.2 e Figura:5.2.3 mostrano che il picco principale della cheto-*OxyLH*₂ ($\lambda_{PMM2} = 345 \text{ nm}$) è il risultato della combinazione delle prime 5 eccitazioni e il picco secondario ($\lambda_{PMM1} = 255 \text{ nm}$) è la combinazione di tutte le eccitazioni superiori alla quinta, anche se buona parte del picco è essenzialmente costituito solo alla sesta eccitazione, mentre le altre

diventano più dense diminuendo via via di intensità.

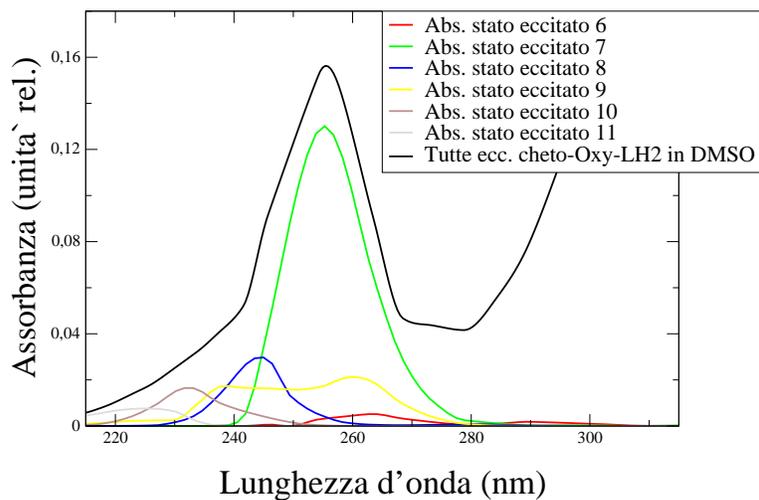


Figura 5.2.2: La scomposizione del picco a $\lambda_{PM1} = 255 \text{ nm}$ della cheto-*OxyLH*₂ in DMSO negli assorbimenti dovuti ai singoli stati eccitati

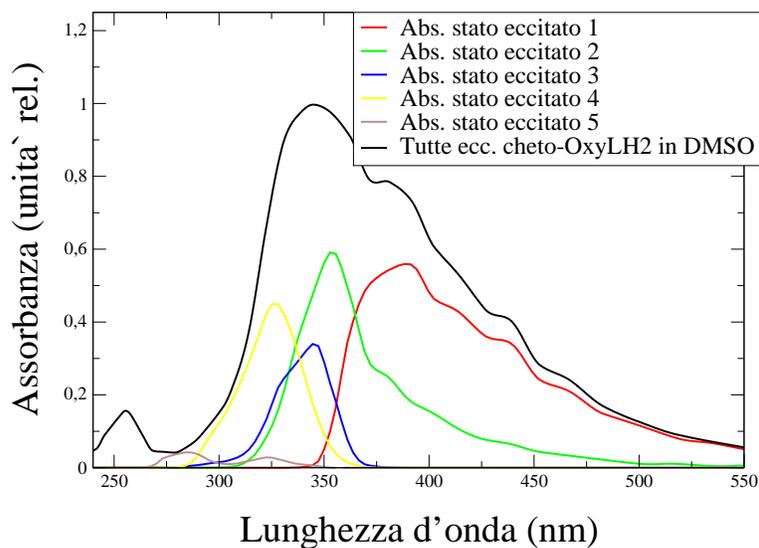


Figura 5.2.3: La scomposizione del picco a $\lambda_{PM2} = 345 \text{ nm}$ della cheto-*OxyLH*₂ in DMSO negli assorbimenti dovuti ai singoli stati eccitati

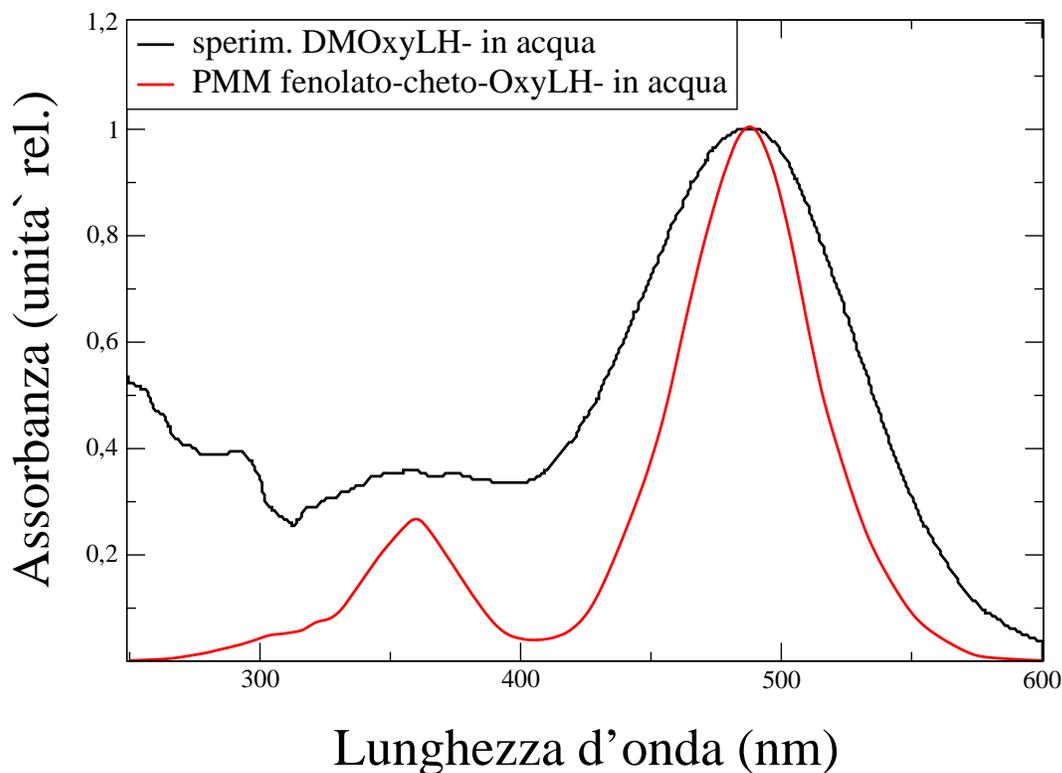
2. fenolato-cheto-*OxyLH*₂ in acqua.

Figura 5.2.4: In nero lo spettro di assorbimento UV sperimentale del $DMOxyLH^-$ in acqua a $pH \sim 7$ ($\lambda_{Abs1} = 360 \text{ nm}$ e $\lambda_{Abs2} = 487 \text{ nm}$) e in rosso lo spettro del fenolato-cheto- $OxyLH^-$ ottenuto con il PMM ($\lambda_{PMM1} = 360 \text{ nm}$ e a $\lambda_{PMM2} = 487 \text{ nm}$)

La Figura:5.2.4 mostra che lo spettro PMM del fenolato-cheto- $OxyLH^-$ riproduce fedelmente lo spettro sperimentale del $DMOxyLH^-$ (Figura:5.2.4), mostrando come, nel caso dell'anione l'aggiunta dei due metili sostituiti in posizione 5 del $DMOxyLH^-$ non produce le grandi variazioni osservate nel confronto cheto- $OxyLH_2 \leftrightarrow DMOxyLH_2$. Dalle scomposizioni dei picchi (Figure:5.2.5 - 5.2.6) si può notare che in questo caso le eccitazioni che vanno a costituire il picco principale (prima e seconda eccitazione) e il secondario (essenzialmente terza e quarta eccitazione) sono ridistribuite in maniera molto differente rispetto al caso della molecola non carica (Figure:5.2.2 - 5.2.3).

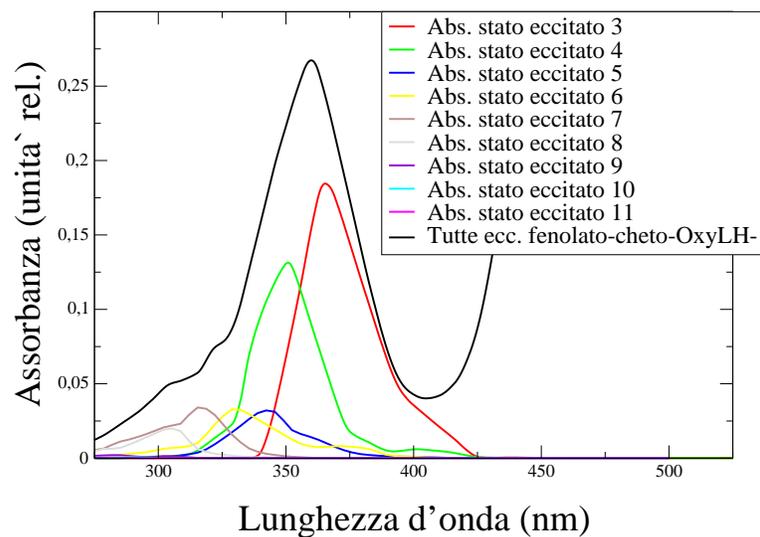


Figura 5.2.5: La scomposizione del picco a $\lambda_{PMM1} = 360 \text{ nm}$ della fenolato-cheto-*OxyLH*⁻ negli assorbimenti dovuti ai singoli stati eccitati

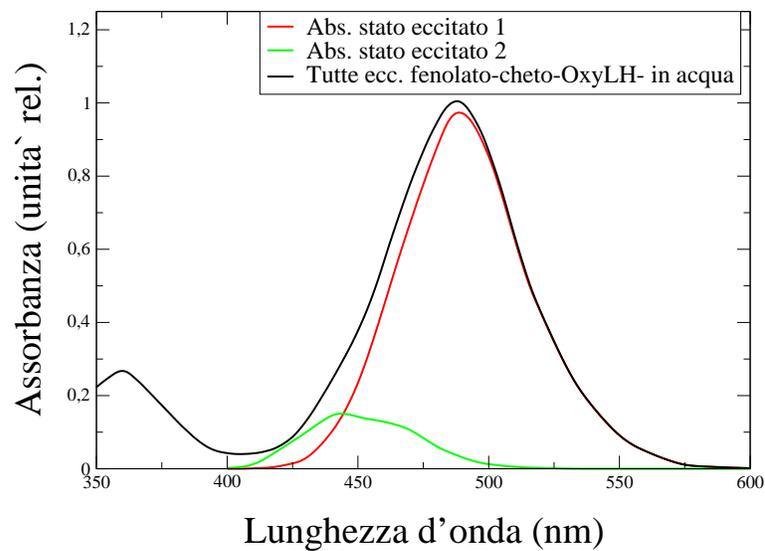


Figura 5.2.6: La scomposizione del picco a $\lambda_{PMM2} = 487 \text{ nm}$, della fenolato-cheto-*OxyLH*⁻ negli assorbimenti dovuti ai singoli stati eccitati

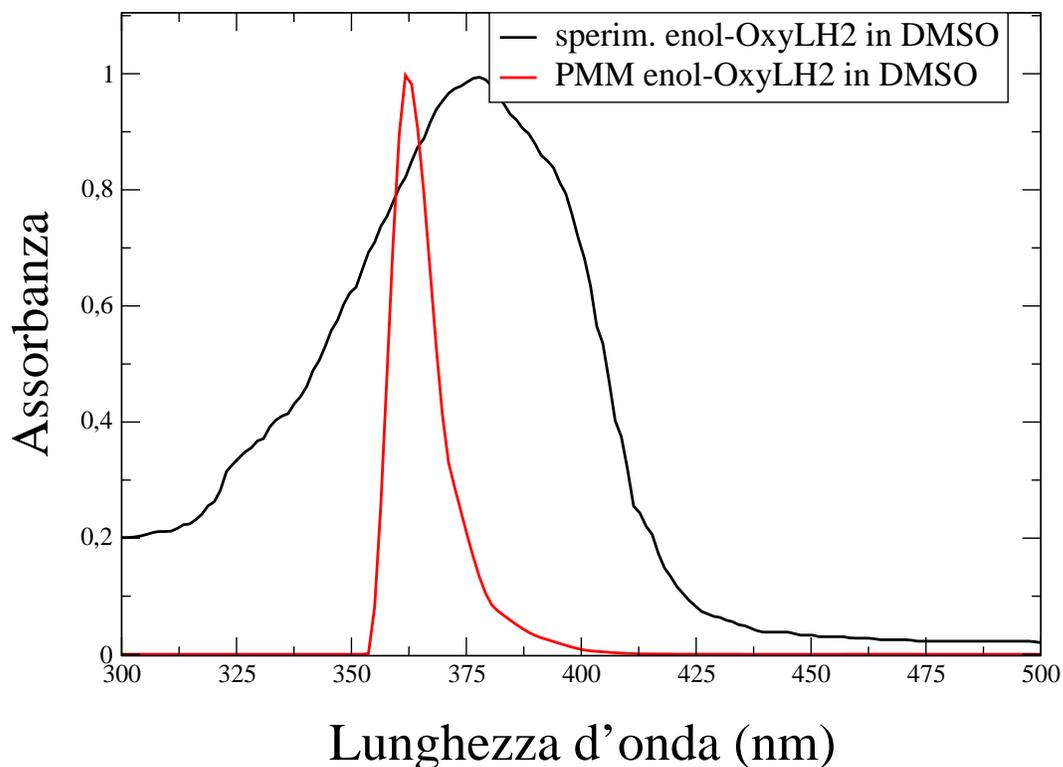
3. enol-*OxyLH*₂ in DMSO.

Figura 5.2.7: In nero lo spettro di assorbimento UV sperimentale dell'enol-*OxyLH*₂ in DMSO ($\lambda_{Abs} = 377 \text{ nm}$) e in rosso lo spettro dell'enol-*OxyLH*₂ ottenuto con il PMM ($\lambda_{PMM} = 362 \text{ nm}$)

Benchè riesca a riprodurre il massimo di assorbimento con un errore di 10 *nm*, tra tutti gli spettri studiati in questo lavoro di tesi, quello ottenuto per l'enol-*OxyLH*₂ in DMSO (Figura:5.2.7) è senz'altro il più peculiare visto che il picco è determinato essenzialmente dalla sola eccitazione 0->1 (la scomposizione del picco non viene mostrata perchè identica al grafico della Figura:5.2.7). In questo caso specifico è necessario aggiungere che l'attendibilità del risultato ottenuto non è molto elevata data la scarsa qualità nel riprodurre la forma e l'ampiezza dello spettro sperimentale.

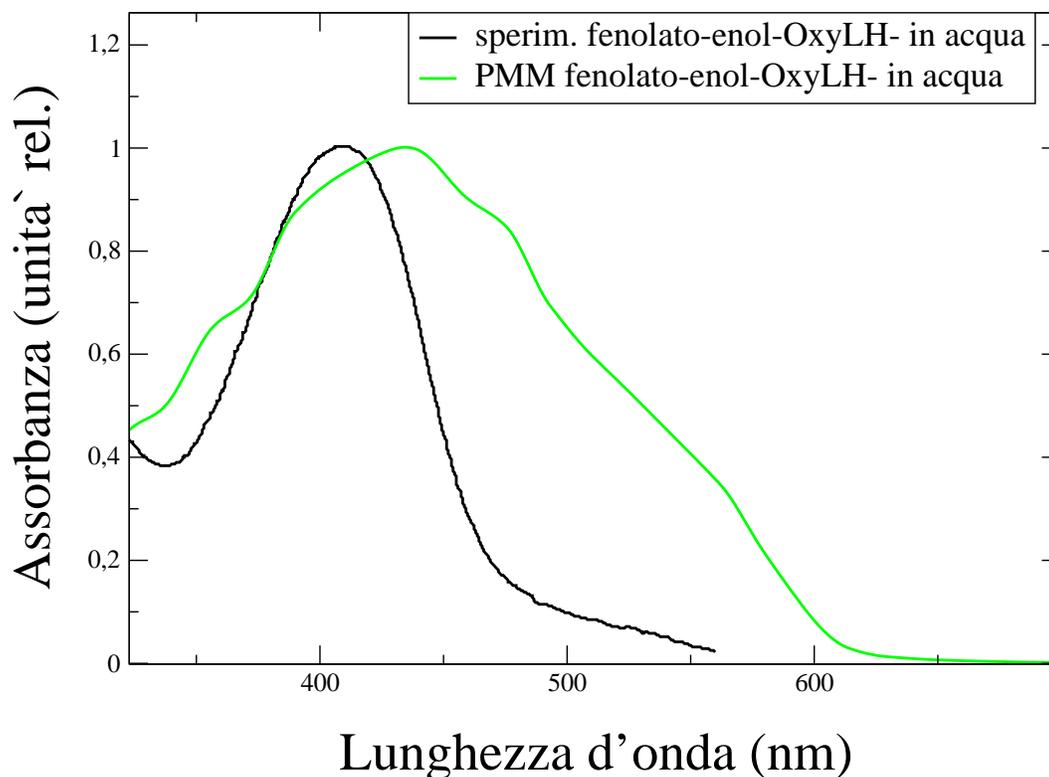
4. fenolato-enol-*OxyLH*⁻ in acqua.

Figura 5.2.8: In nero lo spettro del fenolato-enol-*OxyLH*⁻ in acqua a pH 7,8 ($\lambda_{Abs} = 410 \text{ nm}$) e in verde lo spettro PMM del fenolato-enol-*OxyLH*⁻ ($\lambda_{PMM} = 433 \text{ nm}$)

Contrariamente allo spettro PMM del suo analogo neutro (Figura:5.2.7), l'ampiezza del picco della fenolato-enol-*OxyLH*⁻ in acqua (Figura:5.2.8) supera quella del picco sperimentale. Dalla scomposizione dello spettro PMM (Figura:5.2.9) si può vedere come per questo anione, il picco sia formato dalla somma delle prime 5 eccitazioni, e non solamente dalla prima eccitazione.

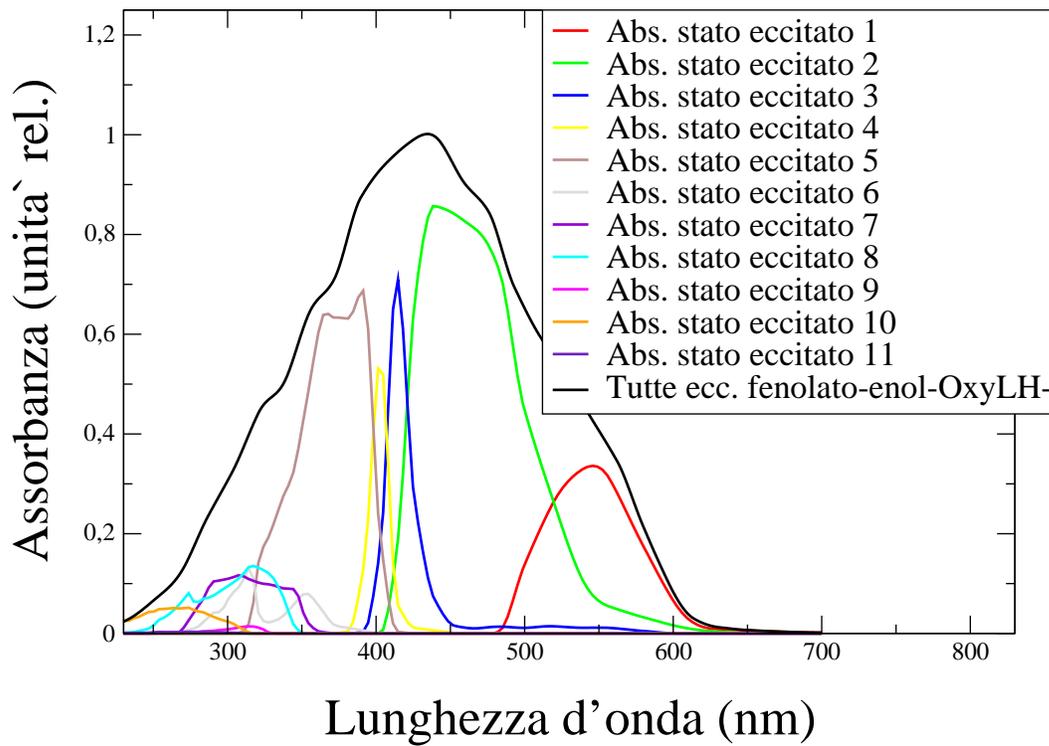


Figura 5.2.9: La scomposizione del picco a $\lambda_{PMM} = 433 \text{ nm}$ della fenolato-enol-*OxyLH*⁻ negli assorbimenti dovuti ai singoli stati eccitati

Tutti i grafici mostrano che anche per la *OxyLH*₂ il calcolo PMM/MD riproduce in generale piuttosto bene gli spettri sperimentali, ancorché con una accuratezza mediamente inferiore al caso della *DLH*₂ (Figura:5.1.1). Ciò è probabilmente dovuto a una minore qualità del calcolo *ab initio* per gli stati elettronici imperturbati, in particolare per le due forme enoliche (Figure 5.2.7 - 5.2.8), per le quali la metodologia TD-DFT (B3LYP) implementata su DALTON non dà risultati particolarmente accurati. Nonostante ciò il calcolo teorico riproduce sostanzialmente, entro un ragionevole errore, la posizione del picco principale e, almeno per le due forme chetoniche, ne riproduce quantitativamente anche la forma e l'ampiezza. E' da notare che per queste ultime, gli spettri calcolati riproducono piuttosto bene i picchi secondari (ad energia più

alta) e la differenza tra questi e il loro corrispondente picco principale. La buona riproduzione quantitativa degli spettri complessivi delle varie forme di $OxyLH_2$ permette, come nel caso della DLH_2 , di utilizzare la procedura teorico-computazionale per scomporre gli spettri nelle singole eccitazioni coinvolte. Complessivamente i dati mostrano che, come per la DLH_2 , anche per le varie forme di $OxyLH_2$, eccetto la enol- $OxyLH_2$, il picco principale è costituito dalla sovrapposizione di diverse eccitazioni (Figure 5.2.3 - 5.2.6 - 5.2.9).

Capitolo 6

Conclusioni

In questo lavoro di tesi è stato caratterizzato in dettaglio, tramite la metodologia PMM/MD, lo spettro di assorbimento UV della D-luciferina e di varie forme della Ossiluciferina in soluzione che sono implicate nei processi biochimici alla base della bioluminescenza delle lucciole. I risultati ottenuti mostrano che la metodologia PMM/MD riproduce in modo piuttosto accurato gli spettri sperimentali delle molecole studiate, fornendo informazioni quantitative sulla posizione, la forma e l'ampiezza dei picchi di assorbimento. L'analisi del picco principale di assorbimento dei vari spettri ha mostrato che in generale, tale picco è dato dalla combinazione degli assorbimenti di varie eccitazioni. Il metodo PMM/MD si basa sulla combinazione della dinamica molecolare classica e delle proprietà elettroniche allo scopo di ottenere una descrizione del comportamento quantistico dei sistemi complessi (cromoforo in solvente o in proteina). In tale procedura, la dinamica molecolare fornisce la perturbazione che modifica gli stati elettronici del centro quantistico da cui si ottengono le proprietà elettroniche (perturbate) lungo la traiettoria della simulazione. Per ottenere le traiettorie di MD, per le molecole utilizzate in questa tesi, è stato necessario parametrizzare il `force-field` della D-luciferina (DLH_2) e delle varie forme di Ossiluciferina ($OxyLH_2$) a partire da calcoli di struttura elettronica sulla molecola isolata. Le tecniche e i risultati ottenuti in questo lavoro di tesi potranno fornire in

futuro una base computazionale per affrontare un eventuale studio degli eventi chimico-fisici di assorbimento ed emissione della Ossiluciferina all'interno dell'enzima luciferasi, riuscendo forse a dimostrare quale delle possibili forme è la responsabile principale dell'emissione di luce in vivo. Le numerose informazioni ottenute dal metodo PMM/MD potrebbero inoltre essere utilizzate per studiare in dettaglio le interazioni dinamiche tra cofattore e enzima, compreso il ruolo del solvente durante l'emissione.

Bibliografia

- [1] E. N. Harvey. What substance is the source of the light in the fireflies?, *Science*, **1917**, *46*, 241–243.
- [2] A. A. Green, e W. D. McElroy. Crystalline firefly luciferase, *Biochim. Biophys. Acta*, **1956**, *20*, 170–176.
- [3] W. D. McElroy e B. L. Strehler. Factors influencing the response of the bioluminescent reaction to adenosine triphosphate, *Arch. Biochem.*, **1949**, *22*, 420–433.
- [4] B. Bitler, e W. D. McElroy. The preparation and properties of crystalline firefly luciferin, *Arch. Biochem. Biophys.*, **1957**, *72*, 358–368.
- [5] T. A. Hopkins, H. H. Seliger, E. H. White e M. W. Cass. The chemiluminescence of firefly luciferin. A model for the bioluminescence reaction and identification of the product excited state, *J. Am. Chem. Soc.*, **1967**, *89*, 7148–7150.
- [6] F. McCapra. Alternative mechanism for dioxetane decomposition, *Chem. Commun.*, **1977**, pp 946–948.
- [7] O. Shimomura, T. Goto e F. H. Johnson. Source of oxygen in the CO_2 produced in the bioluminescent oxidation of firefly luciferin, *Proc. Natl. Acad. Sci. USA*, **1977**, *74*, 2799–2802.

- [8] J. Wannlund, M. DeLuca, K. Stempel e P. D. Boyer. Use of ^{14}C carboxyl-luciferin in determining the mechanism of the firefly luciferase catalyzed reactions, *Biochem. Biophys. Res. Commun.*, **1978**, *81*, 987–992.
- [9] N. Suzuki, M. Sato, K. Nishikawa e T. Goto. Synthesis and spectral properties of 2-(6'-hydroxybenzothiazol-2'-yl)-4-hydroxythiazol, a possible emitter species in the firefly bioluminescence, *Tetrahedron Lett.*, **1969**, *55*, 4683–4684.
- [10] B. J. Gates e M. DeLuca. The production of oxyluciferin during the firefly luciferase light reaction, *Arch. Biochem. Biophys.*, **1975**, *169*, 616–621.
- [11] O. Shimomura. *Bioluminescence*. World Scientific Publishing, 2006.
- [12] B. R. Branchini, M. H. Murtiashaw, A. M. Rachele, N. C. Portier, M. C. Ruggiero e J. G. Stroh. Yellow-green and red firefly bioluminescence from 5,5-dimethoxyoxyluciferin, *J.Am.Chem.Soc.*, **2002**, *124*, 2112–2113.
- [13] G. Orlova, J. D. Goddard e L. Y. Brovko. Theoretical Study of the Amazing Firefly Bioluminescence: The Formation and Structures of the Light Emitters, *J.Am.Chem.Soc.*, **2003**, *125*, 6962–6971.
- [14] M. E. Casida, C. Jamorski, K. C. Casida e D. R. Salahub. Molecular excitation energies to high-lying bound states from time-dependent density-functional response theory: characterization and correction of the time-dependent local density approximation ionization threshold, *J. Chem. Phys.*, **1998**, *108*, 4439–4449.
- [15] M. Zerner. *Reviews in Computational Chemistry*, volume 2. VCH, New York, 1991. 313-365.

- [16] T. Hirano, Y. Hasumi, K. Ohtsuka, S. Maki, H. Niwa, M. Yamaji, e D. Hashizume. Spectroscopic Studies of the Light-Color Modulation Mechanism of Firefly (Beetle) Bioluminescence, *J.Am.Chem.Soc.*, **2009**, *131*, 2385–2396.
- [17] P. Naumov, Y. Ozawa, K. Ohkubo e S. Fukuzumi. Structure and Spectroscopy of Oxyluciferin, the Light Emitter of the Firefly Bioluminescence, *J.Am.Chem.Soc.*, **2009**, *131*, 11590–11605.
- [18] T. Nakatsu, S. Ichiyama, J. Hiratake, A. Saldanha, N. Kobashi, K. Sakata e H. Kato. Structural basis for the spectral difference in luciferase bioluminescence, *Nature*, **2006**, *440*, 372–376.
- [19] D. van der Spoel, R. an Drunen, H. J. C. Berendsen. *GRoningen MACHine for Chemical Simulations*. Department of Biophysical Chemistry, BIOSON Research Institute, Nijenborgh 4 NL-9717 AG Groningen, e-mail to gromacs@chem.rug.nl., 1994.
- [20] M. Aschi, R. Spezia, A. Di Nola e A. Amadei. A first-principle method to model perturbed electronic wavefunctions: the effect of an external homogeneous field, *Chem. Phys. Lett.*, **2001**, *344*(3-4), 374–380.
- [21] A. Amadei, M. D'Alessandro, M. D'Abramo e M. Aschi. Theoretical characterization of electronic states in interacting chemical systems, *J. Chem. Phys.*, **2009**, *130*(084109).
- [22] B. J. Alder e T. E. Wainwright. Studies in Molecular Dynamics. I. General Method, *J. Chem. Phys.*, **1959**, *31*, 459–466.
- [23] L. Verlet. Computer Experiments on Classical Fluids, *Phys. Rev.*, **1967**, *159*(1), 98–103.

- [24] J. P. Ryckaert, G. Ciccotti, H. J. C. Berendsen. Numerical integration of the Cartesian equations of motion of a system with constraints; Molecular dynamics of a n-alkanes, *J. Comp. Phys.*, **1977**, 23(3), 327–341.
- [25] T. A. Darden, D. York, L. Pedersen. Particle mesh Ewald: an $n \log(n)$ method for Ewald sums in large systems, *J. Chem. Phys.*, **1993**, 98(12), 10089–10092.
- [26] S. Nosé . A molecular dynamics method for simulations in the canonical ensemble, *Mol. Phys.*, **1984**, 100(1), 191–198.
- [27] H. C. Andersen. Molecular dynamics simulation at constant pressure and temperature, *J. Chem. Phys.*, **1980**, 72(4), 2384–2394.
- [28] H. J. C. Berendsen, J. P. M. Postma ,W. F. van Gunsteren e A. Di Nola. Molecular dynamics with coupling to an external bath, *J. Chem. Phys.*, **1984**, 81, 3684–3691.
- [29] A. Amadei. *Theoretical models for fluid thermodynamics based on the quasi-Gaussian entropy theory*. Groningen, 1998.
- [30] A. Szabo e N. S. Ostlund. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*. Dover, 1996.
- [31] B. O. Roos. The complete active space self-consistent-field method and its applications in electronic structure calculations, *Adv. Chem. Phys.*, **1987**, 69, 399–427.
- [32] G. C. Schatz e M. A. Ratner. *Quantum Mechanics in Chemistry*. Prentice Hall College Div, 1993.
- [33] D. V. George. *Principles of Quantum Chemistry*. Pergamon Press Inc. New York, 1972.

- [34] F. Jensen. *Introduction To Computational Chemistry*. John Wiley e Sons Inc., 2007.
- [35] C. C. J. Roothan. New developments in molecular orbital theory, *Rev. Mod. Phys.*, **1951**, 23, 69–89.
- [36] R. Carbo e M. Klobukowski, Ed. *Self-Consistent Field: Theory and Applications (Studies in Physical and Theoretical Chemistry)*, volume 70. Elsevier, 1990.
- [37] C. Møller and M. S. Plesset. Note on an Approximation Treatment for Many-Electron Systems, *Phys. Rev.*, **1934**, 46, 618–622.
- [38] P. R. Taylor. *Lecture Notes in Quantum Chemistry: European Summer School*. Springer-Verlag, Berlin, 1994.
- [39] W. Kutzelnigg. *Methods of Electronic Structure Theory*. Plenum Press, New York, 1977.
- [40] L. H. Thomas. The calculation of atomic fields, *Proc. Camb. Phil. Soc.*, **1927**, 23, 542–548.
- [41] E. Fermi. Un metodo statistico per la determinazione di alcune proprietà dell'atomo, *Rend. Accad. Lincei*, **1927**, 6, 602–607.
- [42] P. Hohenberg e W. Kohn. Inhomogeneous electron gas, *Phys. Rev.*, **1964**, 136, B864–B871.
- [43] W. Kohn and L. J. Sham. Self-consistent equations including exchange and correlation effects, *Phys. Rev.*, **1965**, 140, A1133–A1138.
- [44] R. G. Parr e W. Yang. *Density-Functional Theory of atoms and molecules*. Oxford University Press, Oxford, 1989.

- [45] S. J. Vosko and L. Wilk and M. Nusair. Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis, *Can. J. Phys.*, **1980**, 58, 1200–1211.
- [46] A. D. Becke. Correlation energy of an inhomogeneous electron gas: a coordinate-space model, *J. Chem. Phys.*, **1988**, 88, 1053–1062.
- [47] C. Lee and W. Yang and R. G. Parr. Development of the Colle-Salvetti correlation energy formula into a functional of the electron density, *Phys. Rev. B*, **1988**, 37, 785–789.
- [48] R. Bauernschmitt and R. Ahlrichs. Treatment of electronic excitations within the adiabatic approximation of time dependent density functional theory, *Chem. Phys. Lett.*, **1996**, 256, 454–464.
- [49] A. Amadei, M. D’Abramo, C. Zazza e M. Aschi. Electronic properties of formaldehyde: a theoretical study, *Chem. Phys. Lett.*, **2003**, 381, 187–193.
- [50] R. Spezia, M. Aschi, A. Di Nola e A. Amadei. Extension of the perturbed matrix method: application to a water molecule, *Chem. Phys. Lett.*, **2002**, 365, 450–456.
- [51] A. Amadei, M. D’Alessandro e M. Aschi. Statistical mechanical modeling of chemical reactions in complex systems: the reaction free energy surface, *J. Phys. Chem. B*, **2004**, 108, 16250–16254.
- [52] A. Amadei, F. Marinelli, M. D’Abramo, M. D’Alessandro, M. Anselmi, A. Di Nola e M. Aschi. Theoretical modeling of vibro-electronic quantum states in complex molecular systems: solvated carbon monoxide, a test case, *J. Chem. Phys.*, **2005**, 122(12), 124506.1–124506.10.

- [53] Hypercube, Inc. *HyperChem Professional 8.0*. 1115 NW 4th St. Gainesville, FL 32608 (USA), email: info@hyper.com, 1985-2007.
- [54] M. J. Frisch. Gaussian 03, *Gaussian, Inc., Pittsburgh PA, 1998*.
- [55] A. D. Becke. A new mixing of Hartree-Fock and local density-functional theories, *J. Chem. Phys.*, **1993**, *98*, 1372–1377.
- [56] M. W. Schmidt, K. K. Baldridge, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. J. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. Su, T. L. Windus, M. Dupuis e J. A. Montgomery. The **General Atomic and Molecular Electronic Structure System**, *J. Comput. Chem.*, **1993**, *14*(11), 1347–1363. sito web '<http://www.msg.ameslab.gov/GAMESS/>'.
- [57] C. I. Bayly, P. Cieplak, W. D. Cornell, e P. A. Kollman. A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges: The RESP Model, *J. Phys. Chem.*, **1993**, *97*, 10269–10280.
- [58] Collaborazione interuniversitaria. DALTON, a molecular electronic structure program, Release 2.0, sito web <http://www.kjemi.uio.no/software/dalton/dalton.html>, **2005**.
- [59] W. D. McElroy, H. H. Seliger. *Light and Life*. The Johns Hopkins Press, Baltimore, 1961.

Ringraziamenti

In ambito universitario

Isaac Newton una volta ha detto :

“ Ho fatto le mie scoperte perché sono potuto salire sulle spalle di giganti ”.

Ovviamente in questa tesi io non ho scoperto nulla di importante nè inventato niente di nuovo, tuttavia ho provato anche io l'ebbrezza di viaggiare sulle spalle di giganti. Giganti che dedicano la loro vita e i loro sogni alla comprensione del mondo che abbiamo intorno. Un mondo meraviglioso nella sua complessità. Per avermi aiutato e sopportato durante quest'anno di lavoro, facendomi sentire parte attiva di un gruppo di ricerca, vorrei ringraziare, in ordine alfabetico, i seguenti “Giganti”:

- Andrea Amadei;
- Massimiliano Anselmi;
- Massimiliano Aschi;
- Alfredo Di Nola.

Vorrei fare inoltre un ringraziamento speciale a Massimiliano Anselmi, che è stato l'ideatore e sviluppatore dell'argomento di questa tesi, guidandomi e dividendo ogni giorno con me le

gioie e i dolori affrontati nello svolgimento di questo lavoro. Dopo aver condiviso tanto spero con orgoglio di poterlo considerare mio amico.

In ambito affettivo

Quest'ultimo anno mi sono sentito come un contenitore sballottato freneticamente da un fiume in piena che non riuscivo nè a controllare nè a contenere, un fiume fatto di sogni, amore e amicizia. Tra i molti affluenti di questo fiume i principali sono i *Babbion's*, i *Freschi*, la mia amica Aurora, il Dott. Alfredo Carboni, la mia burocrate preferita Emma, e ovviamente la mia fidanzata Assia e la mia famiglia. Grazie per continuare a sballottarmi e riempirmi ogni giorno con il vostro affetto dando un senso alla mia vita.

P. s. per la mia famiglia = Vorrei farvi notare che questo risultato non appartiene solo a me, ma è la fine di una sfida cominciata insieme anni fa, perciò sentitevi laureati quanto e più di me. Non vedo l'ora di affrontare **insieme** le nostre prossime avventure.

Grazie, Grazie, Grazie ...