

Développement d'une Méthode de Chimie Quantique Mêlant Plusieurs Niveaux de Théorie : Applications à l'Etude des Etats Electroniques de Macromolécules

THÈSE

présentée et soutenue publiquement le 13 Juin 2008

pour l'obtention du titre de

Docteur de l'Université Henri Poincaré – Nancy 1

(Spécialité Chimie Informatique et Théorique)

par

Pierre-François Loos

Composition du jury

<i>Président :</i>	Mr. Claude Pouchan	Professeur, Université de Pau et des Pays del'Adour, France.
<i>Rapporteurs :</i>	Mme. Anne Milet	Professeur, Université Joseph Fourier, Grenoble, France.
	Mr. Iñaki Tuñón	Professeur, Université de Valence, Espagne.
<i>Examineurs :</i>	Mr. Jean-Louis Rivail	Professeur, Université Henri Poincaré, Nancy 1, France.
	Mr. Eric Perpète	Chercheur qualifié FNRS, Faculté Universitaire Notre-Dame de la Paix, Namur, Belgique.
	Mr. Xavier Assfeld	Professeur, Université Henri Poincaré, Nancy 1, France.

Mis en page avec la classe thloria.

Remerciements

En premier lieu, j'aimerais remercier les personnes qui ont eu le courage et la gentillesse de relire et corriger ma thèse : à savoir Elise Dumont et Adèle Laurent. Des remerciements tous particuliers à mon chef pendant ces trois années, Xavier Assfeld, avec qui j'ai eu beaucoup de plaisir à travailler mais aussi à discuter autour d'une bière ou deux ... ou trois. J'aimerais aussi remercier l'ensemble des membres de l'équipe de Chimie et Biochimie Théoriques de Nancy qui m'ont accueilli lors de ces trois longues années de souffrance (pour eux bien sur!). Un grand merci tout particulier à François Ogliaro (alias le pape des intersections coniques) avec qui j'ai eu la chance de discuter, pendant deux ans, le temps de midi. Je voudrais aussi remercier l'ensemble des gens extérieurs que j'ai rencontré pendant mon DEA et ma thèse : à savoir Arianna Fornili, Alessandro Genoni, Maurizio Sironi, David Ambrosek et Chantal Daniel. Je voudrais faire aussi un gros merci à l'équipe CTA de Namur (Belgique) pour leur gentillesse et leur amitié, et plus particulièrement un petit bisou sur le front à Julien Preat. En ce qui concerne mon travail de thèse, je remercie grandement les deux anciens étudiants en thèse de Xavier Assfeld (Nicolas Ferré et Yohann Moreau) sans lesquelles mon travail n'aurait pas été possible. Des remerciements tout particuliers à Nicolas Ferré, avec qui j'ai beaucoup discuter pendant ma thèse et qui a, à chaque fois, pris le temps de répondre à mes questions. Je tiens à remercier GRANDEMENT Jean-Louis Rivail, ainsi que le Chef alias Daniel Rinaldi, le premier pour sa très grande gentillesse et tout le reste (un certain nombre de trait noirs sur la feuille blanche qui traîne près de la cafetière, au moins), et le second par son dévouement total à la cause IBM ... et sa bonne humeur! Un petit coucou en passant à tout mes pôtes qui m'ont supportés je ne sais toujours pas comment : Alix et Pierre; Aymeric, Fanny et Liouan; Cyrille et Julie; Marine; Guilløme et Marion; Thomas, Charlotte et Anaëlle; Florent; Oliv; Greg; Joseph; Pascale; Les voi'z; Manu alias le déglingué, Valérie, Matis et Mila; Vincent Héligopthère et Delphine; Pif; Rémy; Sebastien, D'jo et Sylvain; PO, Aurélia et Lili-Rose; Eric; Thomas et Olga; Stephane; Yohann; et tout les autres! J'espère que ceux que j'ai oublié ne m'en voudront pas trop!!! Pour finir, un grand merci à toute ma famille, et logiquement à mes parents : La véro et le Tefphane ... ainsi qu'à leur conjoints respectifs Dominos et Clarisse. Special thanks à Catherine Louboutin et Francis Fernique pour l'approvisionnement en côtes de boeuf pendant 20 ans! Un petit coucou à mes grands parents (Simone, Odile et Champion), à mes oncles et tantes (les Fresses, James et Mimi, Léon et Annie, Christelle et Yves, Gerald et Dedette), à mes cousins et cousines (Etienne, Tuc, Elie, Julie, Chloé, Lison, Marin, Coline et tout particulièrement à Ben!!!)

À ma Môtman et mon Pôpa!!!

Table des matières

Introduction Générale	1
Notations et Définitions	5
0.1 Notations	5
0.2 Plage de sommations	5
0.3 Programmes	5

I Méthodologie

Chapitre 1

La Mécanique Quantique

1.1 Equation de Schrödinger	9
1.2 L'approximation de Born-Oppenheimer	10
1.3 Approximation orbitale et déterminant de Slater	11
1.4 La méthode Hartree-Fock	12
1.4.1 Les équations de Roothaan	13
1.4.2 La méthode de résolution auto-cohérente SCF	14
1.5 Les méthodes post Hartree-Fock	15
1.5.1 Les méthodes d'interaction de configuration	15
1.5.2 Les méthodes <i>Coupled-Cluster</i>	16
1.5.3 Les méthodes perturbatives de type Møller-Plesset	16

Table des matières

1.6	La théorie de la fonctionnelle de la densité	17
1.6.1	Les équations de Kohn-Sham	17
1.6.2	Les fonctionnelles d'échange-corrélation	18
1.6.3	La théorie de la fonctionnelle de la densité dépendante du temps	20
1.6.4	Réponse linéaire	21
1.6.5	Relation avec les méthodes TD-HF et CIS	22

Chapitre 2

La Mécanique Moléculaire

2.1	Les champs de forces	23
2.2	La fonction énergie potentielle	23
2.3	Champs de forces usuels	25

Chapitre 3

Les Méthodes Hybrides QM/MM

3.1	Principe général	27
3.2	Théorie	30
3.2.1	Hamiltonien QM/MM	30
3.2.2	Mechanical Embedding	30
3.2.3	Electronic Embedding	31
3.2.4	Energie QM/MM	32
3.3	Les différentes méthodes QM/MM	32
3.4	Les approches de type Link-Atom	34
3.4.1	Définition du Link-Atom	34
3.4.2	Traitement des charges proches de la frontière	34
3.5	Les approches de type ONIOM	34

II La Méthode du Champ Auto-Cohérent Local (LSCF) : Méthodologies et Développements

Chapitre 4**La Méthode *Local Self-Consistent Field* (LSCF)**

4.1	Introduction	39
4.2	Dérivation des équations LSCF	39
4.2.1	Orthogonalisations des FOs	40
4.2.2	Orthogonalisation des AOs	40
4.2.3	Modification des équations de Roothaan	42
4.2.4	Algorithme	43
4.3	Dérivées premières de l'énergie par rapport aux coordonnées nucléaires	43
4.4	Dérivées secondes de l'énergie par rapport aux coordonnées nucléaires .	45
4.5	Méthodes QM/MM et LSCF	47
4.5.1	Utilisation d'orbitales de liaison strictement localisées (SLBO) .	47
4.5.2	Les critères de localisation	48
4.5.3	Distance entre les atomes frontières	51
4.5.4	Dérivées premières des SLBOs par rapport aux coordonnées nucléaires	52

Chapitre 5**Description Avancée de la Frontière QM/MM**

5.1	La méthode <i>Optimized Local Self-Consistent Field</i> (OLSCF)	55
5.1.1	Introduction	55
5.1.2	Principe	56
5.1.3	Formalisme	57
5.1.4	Résultats	61
5.1.5	Discussion	62
5.1.6	Conclusion	65
5.2	Prise en compte des électrons de cœur et de valence	66
5.2.1	Electrons de cœur	66
5.2.2	Electrons de valence : la liaison peptidique	79

III La Méthode du Champ Auto-Cohérent Local (LSCF) : Applications à l'Etude des Etats Electroniques de Systèmes Biomimétiques

Chapitre 6

Etats Ionisés et Excités de Cœur dans les Macromolécules

6.1	Introduction	91
6.2	Théorie	92
6.3	Validation de la méthode : molécule de pentane	94
6.3.1	Influence de la base d'AO sur les énergies d'ionisation de cœur : molécule de méthane	96
6.3.2	Molécules organiques	98
6.3.3	Molécules contenant une liaison peptidique	100
6.3.4	La molécule d'alanine : influence de l'environnement biologique	103
6.4	Conclusion	105

Chapitre 7

Spectroscopie UV-vis de Macromolécules

7.1	Chromisme	109
7.2	Contexte et Problématique	114
7.3	Méthodologie	118
7.4	Résultats	119
7.4.1	Azobenzène en phase gazeuse et dans l'éthanol	119
7.4.2	Poly(acide L-glutamique) avec une chaîne latérale AB	121
7.5	Conclusions	132

Conclusion Générale et Perspectives	133
--	------------

Annexes

Annexe A**Modification des Equations CPHF pour la Méthode LSCF**

A.1	Dérivation des équations CPHF dans la base des spinorbitales	138
A.1.1	Introduction de la matrice de réponse des orbitales	138
A.1.2	Dérivation des équations séculaires et de la condition d'orthogonalité	140
A.1.3	Dérivation de la matrice de Fock	142
A.2	Dérivation des équations CPHF pour un système à couches ouvertes . .	145
A.3	Dérivation des équations CPHF pour un système à couches fermées . .	147
A.4	Contribution des atomes MM à la matrice de réponse	148

Annexe B**Expression des gradients MP2 dans le cadre de la méthode LSCF**

B.1	Introduction	151
B.2	Dérivation des gradients MP2 dans la base des spinorbitales	152
B.2.1	Forme générale des gradients MP2	152
B.2.2	Dérivée de l'énergie MP2	152
B.3	Dérivation des gradients MP2 pour un système à couche ouverte	170
B.3.1	Matrice densité à une particule	170
B.3.2	Matrice densité pondérée par l'énergie	171
B.3.3	Lagrangien MP2	173
B.3.4	Matrice densité à deux particules	174
B.4	Dérivation des gradients MP2 pour un système à couche fermée	176
B.4.1	Matrice densité à une particule	176
B.4.2	Matrice densité pondérée par l'énergie	177
B.4.3	Lagrangien MP2	178
B.4.4	Matrice densité à deux particules	179

Glossaire	181
Bibliographie	185
Table des figures	205
Liste des tableaux	209

Table des matières

Introduction Générale

Le but ultime et avoué de la chimie théorique est une description fidèle des systèmes chimiques et biochimiques. Cependant, la nature même de ces systèmes moléculaires pose un certain nombre de problèmes : de par leur taille qui peut atteindre plusieurs milliers ou millions d'atomes, mais aussi à cause de la complexité des mécanismes et processus chimiques se déroulant en leur sein. La contrainte de taille de ces édifices peut être atténuée en utilisant des méthodes peu coûteuses en mémoire et en temps de calcul comme les méthodes de mécanique moléculaire (MM). Cette classe de méthodes a l'énorme avantage de permettre d'atteindre les phénomènes dynamiques et statistiques à l'aide de simulation de dynamique moléculaire (MD pour *Molecular Dynamics*) ou Monte-Carlo (MC). Néanmoins, la plupart d'entre elles sont incapables, par nature, de décrire la formation ou la rupture de liaisons chimiques¹.

Grâce à l'augmentation de la puissance des ordinateurs et à l'efficacité toujours croissante des algorithmes, la modélisation des systèmes biochimiques à l'aide des méthodes semi-empiriques [10–30] (SE) ou DFTB [31–39] (*Density Functional based Tight Binding*) est devenu routinière, plus particulièrement avec l'avènement des algorithmes à croissance linéaire² et la paramétrisation spécifique de ces méthodes à l'aide de bases de données. Cette approche présente un bon compromis entre réactivité chimique et aspects dynamiques. En effet, la précision des méthodes SE et DFTB permet d'obtenir des informations qualitatives sur la structure et l'énergie des états de transition et des intermédiaires réactionnels, mais autorise aussi des trajectoires de MD représentatives.

Néanmoins, les nombreux processus et réactions chimiques intervenant dans ce genre de systèmes requièrent souvent la rigueur et la précision des méthodes quantiques *ab initio*. Le coût de ces calculs et le nombre d'atomes des systèmes étudiés étant tous les deux très élevés, il est impossible, encore aujourd'hui, de traiter ces systèmes à l'aide uniquement de ces approches. Néanmoins, la région où ces événements chimiques ont lieu, comme le site actif d'une enzyme, représente typiquement une faible portion de la totalité du système. Une stratégie attractive afin d'aborder l'étude de systèmes de taille importante consiste donc à exploiter la nature locale des modifications de la densité électronique lors de la formation ou de la rupture de certaines liaisons, ou bien encore lors d'une excitation ou une désexcitation électronique par absorption ou émission d'un photon. La communauté des chimistes théoriciens a donc développé un nombre important de méthodes appelées

¹Quelques exemples de champs de forces réactifs existent comme ReaxFF [1–6] (*Reactive Force Field*) par van Duin et Goddard, ainsi que celui développé par Warshell, EVB [7–9] (*Empirical Valence Bond*).

²Ces méthodes sont conçues de telle sorte que le temps de calcul augmente linéairement avec la taille du système, c'est-à-dire en $\mathcal{O}(M)$ où M est le nombre d'atomes.

mixtes ou hybrides qui consiste à diviser un système en différentes couches d'intérêts distincts vis-à-vis du phénomène physique étudié.

L'exemple le plus communément utilisé est la solvatation d'un soluté dans un solvant donné. Le nombre de molécules de solvant nécessaire à une modélisation satisfaisante de la phase condensée (solvatation explicite) est cependant gigantesque. Il a donc été développé des approches qui consistent en une solvation implicite du substrat à l'aide d'un continuum polarisable de constante diélectrique donnée (méthodes SCRF pour *Self-Consistent Reaction Field*). Ces méthodes s'appliquent tout aussi bien aux méthodes de la chimie quantique [40–46] qu'à celles de la mécanique moléculaire [41]. L'avantage mais aussi l'un des désavantages des modèles de continuum est qu'ils moyennent les interactions solvant-soluté : l'information dynamique du solvant est donc perdue.

Un moyen de recouvrer cet aspect statistique consiste à traiter le solvant à l'aide d'une méthode moins coûteuse que celle appliquée au soluté. L'utilisation d'un des nombreux modèles d'eau classiques non-polarisables (TIP3P [47], TIP4P [48], SPC [49], SPC/E [50], ST2 [51], ...) allié à des conditions limites périodiques (PBC pour *Periodic Boundary Conditions*) permet d'introduire efficacement la description explicite des molécules de solvant [52–59]. Ce couplage entre une méthode de mécanique quantique (QM) et une méthode classique MM est communément noté QM:MM³. Lorsque le sous-système traité au bas niveau de théorie est lui-même représenté à l'aide d'une méthode quantique, le schéma de partition est alors noté QM:QM' afin de différencier les deux niveaux de théorie.

Dans tous les cas considérés ci-dessus, nous n'avons pas envisagé le fait que, dans certaines situations, il est impossible de diviser rigoureusement le système en deux sous-ensembles distincts et non-joints. Il est alors indispensable de définir une frontière physique entre les deux sous-systèmes. Au milieu des années 70, il a donc été proposé par Warshell [65] de définir la jonction au niveau d'une liaison covalente, et de représenter la liaison pendante, consécutif à la séparation des deux parties, à l'aide d'une orbitale localisée. Ceci donna naissance aux méthodes dites QM/MM où le caractère « / » indique la coupure d'une liaison covalente. Une quinzaine d'années plus tard, Field, Bash et Karplus [66] proposèrent l'utilisation d'un atome d'hydrogène afin de saturer la valence des atomes possédant des liaisons pendantes. Les dix dernières années ont vu l'émergence de ces méthodes hybrides QM/MM, ainsi que d'un grand nombre de variantes qui diffèrent quasiment exclusivement par le traitement de la jonction entre les parties QM et MM. Par analogie, les extensions suivantes ont été développées : QM/QM' [67, 68], QM/QM':SCRF [69], QM/MM:SCRF [69, 70] et QM/QM'/MM [71]. Basé sur des simulations Born-Oppenheimer [72] ou Car-Parrinello [73], les approches QM/MM fournissent maintenant des trajectoires de MD de temps convenable (quelques dizaines de picosecondes).

Dans ce mémoire, nous allons nous intéresser aux méthodes hybrides mêlant plusieurs niveaux de théorie, et plus particulièrement aux méthodes mixtes QM/MM dans le cadre de la méthode *Local Self-Consistent Field* [74] (LSCF) développée à Nancy depuis plus de dix ans. La majeure partie des travaux présentés ici concernent ses développements mé-

³Il a été récemment proposé d'utiliser les modèles d'eau semi-empirique MP3-PIF [60] (*Parametrizable Interaction Function*) et MP3-MAIS [61] (*Method Adapted for Intermolecular Studies*), couplés à l'approche à croissance linéaire *Divide and Conquer* [62–64] (D&C) afin d'effectuer des simulations de MD.

thodologiques et théoriques, ainsi que l'étude de certains états électroniques de molécules de taille variable.

Ce manuscrit est composé de trois parties regroupant sept chapitres. La première partie expose les diverses méthodologies que nous avons utilisées lors des travaux de cette thèse. Le premier chapitre introduit les méthodes quantiques dont nous avons eu besoin ou que nous avons été amenés à modifier (méthode Hartree-Fock, théorie de la fonctionnelle de la densité dépendante et indépendante du temps, méthode perturbative du type Møller-Plesset), ainsi que certaines méthodes incontournables post Hartree-Fock, telle que la méthode de configuration d'interaction ou *Coupled-Cluster*. Le second chapitre présente le principe général des calculs de mécanique moléculaire, ainsi que la fonction énergie potentielle qui leur est associée. De plus, il contient une liste non exhaustive des différents champs de forces utilisés en mécanique moléculaire. Le chapitre 3 présente les méthodes de type hybrides QM/MM mêlant mécanique quantique et mécanique moléculaire. En particulier, il décrit les stratégies qui permettent le couplage des méthodes vues lors des deux premiers chapitres, en soulignant la différence entre les approches *Link-Atom* (LA) et par orbitales localisées.

La seconde partie contient une description détaillée de l'approche LSCF. Le chapitre 4 introduit l'ensemble des équations LSCF (modification du processus SCF, dérivées premières et secondes de l'énergie par rapport aux coordonnées nucléaires), ainsi que les différents points méthodologiques qui permettent son applicabilité aux calculs QM/MM. Le chapitre 5 regroupe trois développements méthodologiques qui permettent une amélioration de la description de la frontière QM/MM. Le premier consiste à introduire un degré de liberté à la liaison localisée représentant la jonction QM/MM, tandis que les deux points développés par la suite proposent de prendre en compte les électrons de cœur de l'atome frontière, ainsi que quelques électrons de valence dans le cas de la liaison peptidique.

La troisième et dernière partie regroupe deux exemples d'application de la méthodologie LSCF/MM. Dans le chapitre 6, nous discutons de la procédure d'orthogonalisation des orbitales localisées afin de déterminer de façon rigoureuse les ionisations de cœur de molécules organiques et biomimétiques. Quant à lui, le chapitre 7 contient la détermination de spectres d'absorption UV-vis de polypeptides contenant à l'extrémité de leur chaîne latérale un chromophore de type azobenzène.

Après la conclusion générale se trouve deux annexes à caractère plus théorique. La première d'entre elles détaille les modifications apportées aux équations *Coupled Perturbed Hartree-Fock* dans le cadre de la méthode LSCF. Dans la même optique, la deuxième annexe regroupe une dérivation détaillée des expressions des dérivées premières de la correction au deuxième ordre de l'énergie Møller-Plesset. Ces deux points sont en cours d'implantation.

Notations et Définitions

Dans cette section, nous définissons l'ensemble des notations que nous utiliserons dans ce manuscrit. Pour ce qui est des acronymes, un glossaire les rassemble à la fin de cette thèse. Par soucis de cohérence, l'ensemble des abréviations et des acronymes est laissé en anglais.

0.1 Notations

	Coefficients	Indices
AOs		$\mu, \nu, \lambda, \sigma$
x		Dérivation par rapport à un déplacement nucléaire x
(x)		Dérivation des intégrales sur la base des AOs seulement par rapport à un déplacement nucléaire x
α		MOs ou SLBOs de spin α , énergies de MOs α , ...
β		MOs ou SLBOs de spin β , énergies de MOs β , ...
SLBOs occupées	$(\mathbf{L})_{\mu I} = l_{\mu I}$	I, J, K, L
MOs occupées	$(\mathbf{C})_{\mu i} = c_{\mu i}$	i, j, k, l
SLBOs virtuelles	$(\mathbf{L})_{\mu A} = l_{\mu A}$	A, B, C, D
MOs virtuelles	$(\mathbf{C})_{\mu a} = c_{\mu a}$	a, b, c, d
SLBOs occ. ou virt.	$(\mathbf{L})_{\mu P} = l_{\mu P}$	P, Q, R, S
MOs occ. ou virt.	$(\mathbf{C})_{\mu p} = c_{\mu p}$	p, q, r, s
MOs ou SLBOs	$(\mathbf{D})_{\mu t} = d_{\mu t}$	t, u, v, w

0.2 Plage de sommations

- M : Nombre de noyaux du système.
- n : Nombre d'électrons du système.
- N : Nombre de fonctions de base.
- L : Nombre de FO dans le système (sauf indications contraires)
- occ : Toutes les MOs occupées ou toutes les SLBOs occupées.
- virt : Toutes les MOs virtuelles ou toutes les SLBOs virtuelles.
- all : Toutes les MOs et/ou toutes les SLBOs.

0.3 Programmes

Sauf indications contraires, l'ensemble des calculs quantiques réalisés dans cette thèse a été effectué à l'aide du logiciel Gaussian03 [75]. Une version modifiée de ce programme

Notations et Définitions

a été utilisée afin de permettre la réalisation de calculs LSCF. L'implantation, l'évolution et les modifications apportées à ce code ont été réalisées principalement par X. Assfeld, V. Rooy, N. Ferré, Y. Moreau et moi-même. Les calculs QM/MM sont réalisés par couplage entre cette version modifiée de Gaussian03 et le programme de modélisation moléculaire Tinker4.2 [76]. Ce couplage a été réalisé par N. Ferré.

Première partie

Méthodologie

1

La Mécanique Quantique

L'objectif de ce chapitre est de présenter de façon succincte les différentes méthodes que nous avons utilisé ou que nous avons été amenés à adapter à nos besoins.

1.1 Equation de Schrödinger

D'après le premier postulat de la mécanique quantique, un système physique est complètement défini à tout instant t par sa fonction d'onde Ψ (aussi appelée vecteur d'état), solution de l'équation de Schrödinger [77–83] dépendante du temps :

$$i\hbar \frac{\partial \Psi(\mathbf{r}, t)}{\partial t} = \hat{\mathcal{H}}\Psi(\mathbf{r}, t) \quad (1.1)$$

où $\hat{\mathcal{H}}$ est l'opérateur hamiltonien associé au système. Dans le cas d'un système stationnaire (c'est-à-dire ne dépendant pas explicitement du temps) et non relativiste, l'équation de Schrödinger (1.1) s'écrit suivant la forme indépendante du temps :

$$\hat{\mathcal{H}}\Psi(\mathbf{r}) = \mathcal{E}\Psi(\mathbf{r}) \quad (1.2)$$

où \mathcal{E} est l'énergie de l'état stationnaire considéré et où \mathbf{r} représente les coordonnées de l'ensemble des particules du système. Dans le cas d'un système polyatomique moléculaire composé de M noyaux et de n électrons, l'hamiltonien du système se développe (en unités atomiques) suivant l'expression :

$$\begin{aligned} \hat{\mathcal{H}} = & - \sum_i^n \frac{1}{2} \Delta_i - \sum_A^M \frac{1}{2M_A} \Delta_A - \sum_i^n \sum_A^M \frac{Z_A}{|\mathbf{R}_A - \mathbf{r}_i|} \\ & + \sum_i^n \sum_{j>i}^n \frac{1}{|\mathbf{r}_j - \mathbf{r}_i|} + \sum_A^M \sum_{B>A}^M \frac{Z_A Z_B}{|\mathbf{R}_B - \mathbf{R}_A|} \end{aligned} \quad (1.3)$$

où Δ_i est l'opérateur laplacien par rapport aux coordonnées de l'électron i . Dans cette expression, \mathbf{R}_A et \mathbf{r}_i représentent les vecteurs position du noyau A et de l'électron i . Les deux premiers termes représentent respectivement les opérateurs associés à l'énergie cinétique des électrons et des noyaux. Z_A est la charge du noyau A , tandis que M_A

est sa masse. Le terme suivant décrit l'interaction entre noyaux et électrons, tandis que les deux derniers termes correspondent à l'interaction électron-électron et noyau-noyau, respectivement.

1.2 L'approximation de Born-Oppenheimer

La masse d'un électron étant près de deux mille fois moins importante que celle du proton, leurs mouvements relatifs peuvent être découplés. On se ramène ainsi à l'étude du comportement électronique dans le champ créé par des noyaux fixes. Ce découplage, appelé approximation de Born-Oppenheimer [84], consiste à négliger le terme d'énergie cinétique des noyaux dans l'expression (1.3) et à considérer que le terme d'interaction noyau-noyau est constant. Les termes subsistants dans l'équation (1.3) permettent de définir l'hamiltonien électronique :

$$\hat{\mathcal{H}}_{\text{elec}} = - \sum_i^n \frac{1}{2} \Delta_i - \sum_i^n \sum_A^M \frac{Z_A}{|\mathbf{R}_A - \mathbf{r}_i|} + \sum_i^n \sum_{j>i}^n \frac{1}{|\mathbf{r}_j - \mathbf{r}_i|} \quad (1.4)$$

La solution de l'équation de Schrödinger correspondante, Ψ_{elec} ,

$$\hat{\mathcal{H}}_{\text{elec}} \Psi_{\text{elec}} = \mathcal{E}_{\text{elec}} \Psi_{\text{elec}} \quad (1.5)$$

est la fonction d'onde électronique du système considéré décrivant le mouvement des électrons.

$$\Psi_{\text{elec}} = \Psi_{\text{elec}}(\{\mathbf{r}_i; \mathbf{R}_A\}) \quad (1.6)$$

Ψ_{elec} dépend explicitement des coordonnées des électrons mais de façon paramétrique des coordonnées nucléaires, tout comme l'énergie électronique :

$$\mathcal{E}_{\text{elec}} = \mathcal{E}_{\text{elec}}(\{\mathbf{R}_A\}) \quad (1.7)$$

L'énergie totale \mathcal{E}_{tot} est alors donnée par la relation (1.8) :

$$\mathcal{E}_{\text{tot}} = \mathcal{E}_{\text{elec}} + \sum_A^M \sum_{B>A}^M \frac{Z_A Z_B}{|\mathbf{R}_B - \mathbf{R}_A|} \quad (1.8)$$

Dans le cadre de l'approximation de Born-Oppenheimer, on définit l'hamiltonien nucléaire $\hat{\mathcal{H}}_{\text{nucl}}$ permettant de décrire le mouvement des noyaux dans le champ moyen des électrons.

$$\begin{aligned} \hat{\mathcal{H}}_{\text{nucl}} &= - \sum_A^M \frac{1}{2M_A} \Delta_A + \left\langle - \sum_i^n \frac{1}{2} \Delta_i - \sum_i^n \sum_A^M \frac{Z_A}{|\mathbf{R}_A - \mathbf{r}_i|} + \sum_i^n \sum_{j>i}^n \frac{1}{|\mathbf{r}_j - \mathbf{r}_i|} \right\rangle \\ &\quad + \sum_A^M \sum_{B>A}^M \frac{Z_A Z_B}{|\mathbf{R}_B - \mathbf{R}_A|} \\ &= - \sum_A^M \frac{1}{2M_A} \Delta_A + \mathcal{E}_{\text{elec}}(\{\mathbf{R}_A\}) + \sum_A^M \sum_{B>A}^M \frac{Z_A Z_B}{|\mathbf{R}_B - \mathbf{R}_A|} \\ &= - \sum_A^M \frac{1}{2M_A} \Delta_A + \mathcal{E}_{\text{tot}} \end{aligned} \quad (1.9)$$

1.3. Approximation orbitale et déterminant de Slater

où $\langle \hat{A} \rangle$ représente la moyenne de l'opérateur \hat{A} sur les coordonnées électroniques. Les fonctions d'onde Ψ_{nucl} vérifiant l'équation de Schrödinger suivante

$$\hat{\mathcal{H}}_{\text{nucl}}\Psi_{\text{nucl}} = \mathcal{E}\Psi_{\text{nucl}} \quad (1.10)$$

décrivent les mouvements de translation, rotation et vibration des noyaux.

$$\Psi_{\text{nucl}} = \Psi_{\text{nucl}}(\{\mathbf{R}_A\}) \quad (1.11)$$

Dans l'expression (1.10), \mathcal{E} représente l'énergie totale du système, donnée par la relation (1.2), dans le cadre de l'approximation de Born-Oppenheimer, incluant les composantes électronique, vibrationnelle, rotationnelle et translationnelle.

$$\Psi(\{\mathbf{r}_i; \mathbf{R}_A\}) = \Psi_{\text{elec}}(\{\mathbf{r}_i; \mathbf{R}_A\})\Psi_{\text{nucl}}(\{\mathbf{R}_A\}) \quad (1.12)$$

Dans le reste de ce chapitre, nous allons plus particulièrement nous intéresser au problème électronique [85–90]

1.3 Approximation orbitale et déterminant de Slater

L'approximation orbitale consiste à représenter la fonction d'onde multiélectronique Ψ_{elec} du système comme le produit de fonctions d'onde monoélectroniques $\{\phi_i\}_{1 \leq i \leq n}$.

$$\Psi_{\text{elec}}(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_n) \equiv \Psi_{\text{elec}}(1, 2, \dots, n) = \phi_1(1)\phi_2(2)\dots\phi_n(n) \quad (1.13)$$

L'ensemble $\{\phi_i\}_{1 \leq i \leq n}$ correspond à un jeu de spinorbitales, pouvant être décomposées comme le produit d'une fonction de spin ω (α ou β) et d'une fonction spatiale $\varphi(i)$:

$$\phi_i(i) = \omega_i(i)\varphi_i(i) \quad (1.14)$$

Etant des fermions (particules de spin demi-entier), les électrons sont régis par le principe d'exclusion de Pauli [91] stipulant que l'échange de deux particules fermioniques change le signe de la fonction d'onde du système. En combinant l'approximation orbitale et le principe d'exclusion de Pauli, la fonction d'onde Ψ_{elec} est exprimée sous la forme d'un déterminant de Slater [92], antisymétrique par nature :

$$\Psi_{\text{elec}} = \frac{1}{\sqrt{n!}} \begin{vmatrix} \phi_1(1) & \phi_2(1) & \dots & \phi_n(1) \\ \phi_1(2) & \phi_2(2) & \dots & \phi_n(2) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_1(n) & \phi_2(n) & \dots & \phi_n(n) \end{vmatrix} \equiv |\phi_1\phi_2\dots\phi_n| \quad (1.15)$$

où le facteur $\frac{1}{\sqrt{n!}}$ permet de normer le déterminant Ψ_{elec} .

1.4 La méthode Hartree-Fock

Dans le cadre de la méthode Hartree-Fock [93, 94] (HF), l'hamiltonien électronique s'écrit :

$$\begin{aligned}\hat{H}_o \equiv \hat{\mathcal{H}}_{\text{elec}} &= \sum_i^n \left[\frac{1}{2} \Delta_i - \sum_A^M \frac{Z_A}{|\mathbf{R}_A - \mathbf{r}_i|} \right] + \sum_i^n \sum_{j>i}^n \frac{1}{|\mathbf{r}_j - \mathbf{r}_i|} \\ &= \sum_i^n \hat{H}^c(i) + \sum_i^n \sum_{j>i}^n \frac{1}{r_{ij}}\end{aligned}\quad (1.16)$$

où le terme monoélectronique $\hat{H}^c(i)$, appelé hamiltonien de cœur, regroupe le terme d'énergie cinétique et d'attraction des noyaux et de l'électron i . Le dernier terme de l'équation (1.16) étant, quant à lui, un opérateur biélectronique représentant la répulsion entre électrons. Dans la suite de ce chapitre, on se placera dans le cas où le jeu de spinorbitales forme une base orthonormée, c'est-à-dire :

$$\langle \phi_i | \phi_j \rangle = \delta_{ij} \quad (1.17)$$

où δ_{ij} est le symbole de Kronecker tel que :

$$\delta_{ij} = \begin{cases} 0 & \text{si } i \neq j, \\ 1 & \text{si } i = j. \end{cases} \quad (1.18)$$

L'énergie du système est donnée par la relation (1.19).

$$E_o \equiv \mathcal{E}_{\text{elec}} = \langle \Psi_o | \hat{H}_o | \Psi_o \rangle \quad (1.19)$$

$$|\Psi_o\rangle = |\phi_1 \phi_2 \dots \phi_n\rangle \quad (1.20)$$

Le développement de cette expression en utilisant les propriétés du déterminant de Slater conduit à :

$$\begin{aligned}E_o &= \sum_i^n \langle \phi_i(1) | \hat{H}^c(1) | \phi_i(1) \rangle \\ &+ \sum_i^n \sum_{j>i}^n \left[\left\langle \phi_i(1) \phi_j(2) \left| \frac{1}{r_{12}} \right| \phi_i(1) \phi_j(2) \right\rangle - \left\langle \phi_i(1) \phi_j(2) \left| \frac{1}{r_{12}} \right| \phi_j(1) \phi_i(2) \right\rangle \right]\end{aligned}\quad (1.21)$$

On définit alors l'opérateur de Fock \hat{F} , tel que :

$$\hat{F} \phi_i = \varepsilon_i \phi_i \quad (1.22)$$

$$\hat{F}(1) = \hat{H}^c(1) + \sum_i^n \left[\hat{J}_i(1) - \hat{K}_i(1) \right] \quad (1.23)$$

où \hat{J}_i et \hat{K}_i sont respectivement les opérateurs de Coulomb et d'échange :

$$\hat{J}_i(1) \phi_j(1) = \phi_j(1) \int_{(2)} \phi_i^*(2) \frac{1}{r_{12}} \phi_i(2) d\tau_2 \quad (1.24)$$

$$\hat{K}_i(1) \phi_j(1) = \phi_i(1) \int_{(2)} \phi_i^*(2) \frac{1}{r_{12}} \phi_j(2) d\tau_2 \quad (1.25)$$

1.4.1 Les équations de Roothaan

Dans la suite de ce chapitre, nous nous placerons dans le cas d'un système à couches fermées et nous utiliserons alors le formalisme restreint ou RHF (*Restricted Hartree-Fock*), pour lequel deux électrons de spin différents dans une même orbitale moléculaire (MO pour *Molecular Orbital*) possèdent une partie d'espace identique. Dans ce cas, le nombre n représentera le nombre de MOs doublement occupées⁴. Conformément à l'approximation LCAO (*Linear Combination of Atomic Orbitals*), nous allons maintenant introduire une base de N orbitales atomiques (AOs pour *Atomic Orbitals*), préalablement définies, $\{\chi_\mu\}_{1 \leq \mu \leq N}$, tel que :

$$|\varphi_i\rangle = \sum_{\mu} c_{\mu i} |\chi_\mu\rangle \equiv \sum_{\mu} c_{\mu i} |\mu\rangle \quad (1.26)$$

Sur cette base d'AO, l'élément $\mu\nu$ de l'opérateur de Fock \hat{F} s'exprime :

$$F_{\mu\nu} = \langle \chi_\mu | \hat{F} | \chi_\nu \rangle \equiv \langle \mu | \hat{F} | \nu \rangle = H_{\mu\nu}^c + \sum_{\lambda\sigma} P_{\lambda\sigma} \left[\langle \mu\lambda | \nu\sigma \rangle - \frac{1}{2} \langle \mu\lambda | \sigma\nu \rangle \right] \quad (1.27)$$

$$= H_{\mu\nu}^c + \sum_{\lambda\sigma} P_{\lambda\sigma} \left[(\mu\nu | \lambda\sigma) - \frac{1}{2} (\mu\sigma | \lambda\nu) \right] \quad (1.28)$$

avec

$$H_{\mu\nu}^c = \langle \chi_\mu | \hat{H}^c | \chi_\nu \rangle \equiv \langle \mu | \hat{H}^c | \nu \rangle \quad (1.29)$$

$$(\chi_\mu \chi_\nu | \chi_\lambda \chi_\sigma) \equiv (\mu\nu | \lambda\sigma) = \int_{(1)} \int_{(2)} \chi_\mu(1) \chi_\nu(1) \frac{1}{r_{12}} \chi_\lambda(2) \chi_\sigma(2) d\tau_1 d\tau_2 \quad (1.30)$$

$$\langle \chi_\mu \chi_\lambda | \chi_\nu \chi_\sigma \rangle \equiv \langle \mu\lambda | \nu\sigma \rangle = \int_{(1)} \int_{(2)} \chi_\mu(1) \chi_\lambda(2) \frac{1}{r_{12}} \chi_\nu(1) \chi_\sigma(2) d\tau_1 d\tau_2 \quad (1.31)$$

On définit alors, en formalisme restreint, l'élément $\mu\nu$ de la matrice densité totale du système \mathbf{P}^T comme⁵ :

$$P_{\mu\nu}^T = 2 \sum_i^n c_{\mu i} c_{\nu i} \quad (1.32)$$

L'énergie électronique du système s'écrit donc, sur la base des AOs :

$$E = \sum_{\mu\nu} P_{\mu\nu}^T H_{\mu\nu} + \frac{1}{2} \sum_{\mu\nu\lambda\sigma} P_{\mu\nu}^T P_{\lambda\sigma}^T G_{\mu\nu\lambda\sigma} \quad (1.33)$$

Sous forme matricielle, la matrice de Fock \mathbf{F} peut se décomposer selon :

$$\mathbf{F} = \mathbf{H}^c + \mathbf{G} \quad (1.34)$$

⁴Dans le cas de systèmes à couches ouvertes, le formalisme UHF (*Unrestricted Hartree-Fock*), développé par Pople-Nesbet [95] et Berthier [96, 97] simultanément, peut être utilisé.

⁵Dans le formalisme UHF, la matrice densité est définie selon la relation $P_{\mu\nu}^T = P_{\mu\nu}^\alpha + P_{\mu\nu}^\beta = \sum_{i^\alpha}^{occ^\alpha} c_{\mu i}^\alpha c_{\nu i}^\alpha + \sum_{i^\beta}^{occ^\beta} c_{\mu i}^\beta c_{\nu i}^\beta$.

où

$$G_{\mu\nu} = \sum_{\lambda\sigma}^N P_{\lambda\sigma}^T \left[(\mu\nu|\lambda\sigma) - \frac{1}{2}(\mu\sigma|\lambda\nu) \right] \quad (1.35)$$

$$G_{\mu\nu\lambda\sigma} = (\mu\nu|\lambda\sigma) - \frac{1}{2}(\mu\sigma|\lambda\nu) \quad (1.36)$$

La condition de stationnarité de l'énergie par rapport aux coefficients $\{c_{\mu i}\}_{\substack{1 \leq \mu \leq N \\ 1 \leq i \leq N}}$:

$$\frac{\delta E}{\delta c_{\mu i}} = 0 \quad (1.37)$$

conduit aux équations de Roothaan [98] :

$$\sum_{\nu}^N F_{\mu\nu} c_{\nu i} = \sum_{\nu}^N S_{\mu\nu} c_{\nu i} \varepsilon_i \quad (1.38)$$

ou sous forme matricielle,

$$\mathbf{F} \cdot \mathbf{C} = \mathbf{S} \cdot \mathbf{C} \cdot \mathbf{E} \quad (1.39)$$

La matrice \mathbf{S} représente la matrice métrique de la base considérée, ou matrice de recouvrement, tel que $S_{\mu\nu} = \langle \mu | \nu \rangle$. De plus, la matrice \mathbf{C} regroupe les coefficients des AOs dans les MOs, tel que : $(\mathbf{C})_{\mu i} = c_{\mu i}$. On introduit alors la matrice \mathbf{X} permettant de travailler dans une base d'AOs orthogonales, tel que :

$$\mathbf{X}^\dagger \cdot \mathbf{S} \cdot \mathbf{X} = \mathbb{I} \quad (1.40)$$

où \mathbb{I} représente la matrice identité. La matrice \mathbf{X} correspond à une matrice d'orthogonalisation de la base d'AOs. Dans le cas général d'un calcul SCF, cette matrice est égale à $\mathbf{S}^{-1/2}$, appelé matrice d'orthogonalisation de Löwdin [99]. La multiplication à gauche de l'expression (1.39) par \mathbf{X}^\dagger , matrice adjointe de la matrice \mathbf{X} , permet de se ramener à un problème standard aux valeurs propres hermitiques :

$$\mathbf{F}' \cdot \mathbf{C}' = \mathbf{C}' \cdot \mathbf{E} \quad (1.41)$$

avec

$$\mathbf{F}' = \mathbf{X}^\dagger \cdot \mathbf{F} \cdot \mathbf{X} \quad (1.42)$$

$$\mathbf{C} = \mathbf{X} \cdot \mathbf{C}' \quad (1.43)$$

1.4.2 La méthode de résolution auto-cohérente SCF

La résolution de l'équation (1.41) s'effectue donc par diagonalisation de la matrice \mathbf{F}' . Cependant cette matrice dépend des coefficients \mathbf{C} recherchés. Pour résoudre ce problème, on utilise une méthode itérative appelée résolution auto-cohérente ou SCF (*Self-Consistent Field*) permettant de déterminer les coefficients $\{c_{\mu i}\}_{\substack{1 \leq \mu \leq N \\ 1 \leq i \leq N}}$. L'algorithme est résumé ci-dessous :

1. Estimation de la matrice densité \mathbf{P} .
2. Construction de la matrice de Fock $\mathbf{F} = \mathbf{H}^c + \mathbf{G}$.
3. Expression de \mathbf{F} dans la base orthogonale : $\mathbf{F}' = \mathbf{X}^\dagger \cdot \mathbf{F} \cdot \mathbf{X}$ ($\mathbf{X} = \mathbf{S}^{-1/2}$).
4. Diagonalisation de \mathbf{F}' .
5. Transformation des MOs dans la base non-orthogonale : $\mathbf{C} = \mathbf{X} \cdot \mathbf{C}'$.
6. Calcul de la nouvelle matrice densité \mathbf{P} et calcul de l'énergie $E = \frac{1}{2}Tr [\mathbf{P} (\mathbf{H}^c + \mathbf{F})]$.
7. Test de convergence sur E et/ou \mathbf{P} . Si non satisfait, retour à 2.

Cependant, les méthodes de type HF ne permettent pas d'obtenir l'énergie exacte du système, même dans le cas d'une base d'AOs complète (CBS pour *Complete Basis Set*). Ceci est dû au traitement du terme biélectronique de répulsion électron-électron, $\sum_{i<j}^n \frac{1}{r_{ij}}$, qui n'est pris en compte que par le biais d'une moyenne globale. La méthode HF fait partie des méthodes de champ moyen, c'est-à-dire qu'un électron évolue dans le champ moyen créé par les autres électrons. Nous allons voir dans les sections suivantes, comment aller au-delà de l'approximation de HF.

1.5 Les méthodes post Hartree-Fock

On définit l'énergie de corrélation E_{corr} comme la différence entre l'énergie exacte non-relativiste du système \mathcal{E}_o et l'énergie HF E_o dans la limite d'une base supposée complète d'AOs.

$$E_{\text{corr}} = \mathcal{E}_o - E_o \quad (1.44)$$

Cette énergie est toujours négative puisque le principe variationnel assure que l'énergie E_o est toujours supérieure à \mathcal{E}_o .

1.5.1 Les méthodes d'interaction de configuration

Un moyen d'obtenir l'ensemble ou un fragment de cette énergie de corrélation est de décomposer la fonction d'onde multiélectronique CI [100, 101] (*Configuration Interaction*) du système Φ_o comme une combinaison linéaire de plusieurs déterminants de Slater. Ces déterminants, construits par excitation d'un ou plusieurs électrons d'une spinorbitale occupée vers une spinorbitale vacante, sont basés sur un déterminant de référence Ψ_o (souvent de type HF).

$$\Phi = C_o \Psi_o + \sum_i^{\text{occ}} \sum_a^{\text{virt}} C_i^a \Psi_i^a + \sum_{i<j}^{\text{occ}} \sum_{a<b}^{\text{virt}} C_{ij}^{ab} \Psi_{ij}^{ab} + \sum_{i<j<k}^{\text{occ}} \sum_{a<b<c}^{\text{virt}} C_{ijk}^{abc} \Psi_{ijk}^{abc} + \dots \quad (1.45)$$

où Ψ_i^a , Ψ_{ij}^{ab} et Ψ_{ijk}^{abc} correspondent respectivement aux mono, di et tri-excitations. Ψ_i^a représente l'excitation d'un électron de la spinorbitale i vers la spinorbitale a . On montre que l'énergie obtenue en accord avec cette fonction d'onde Φ_o , $\langle \Phi_o | \hat{H}_o | \Phi_o \rangle$, est forcément supérieure à l'énergie exacte du système.

Lorsque l'ensemble des excitations est pris en compte, cette méthode est appelée *Full CI*. Cependant, cette solution, bien que très élégante, est très coûteuse à cause du nombre

important d'excitations à prendre en compte. Lorsque que l'on ne considère que les mono et les di-excitations, ce qui est généralement le cas à cause de leur part importante à l'énergie de corrélation, cette méthodologie est appelée CISD (*Configuration Interaction Single Double*). Cette dernière permet de récupérer une part importante de la corrélation dynamique du système. Cependant, le désavantage principal de celle-ci est qu'elle n'est pas *size-consistent*, ni par conséquent *size-extensive*. Une méthode est dite *size-consistent* si l'énergie de deux systèmes sans interaction est égale à la somme des énergies des deux systèmes. Le concept de *size-extensivity* illustre le fait que l'énergie d'un système doit croître linéairement avec le nombre d'atomes qui le compose.

1.5.2 Les méthodes *Coupled-Cluster*

Un moyen d'obtenir une méthode *size-consistent* est proposé par l'approche *Coupled-Cluster* (CC) [102–104] qui consiste à inclure toutes les corrections d'un type donné à l'ordre infini. Ceci s'effectue grâce à la fonction d'onde CC :

$$\Psi_{\text{CC}} = e^{\hat{T}} \Psi_o \quad (1.46)$$

$$e^{\hat{T}} = \mathbb{I} + \hat{T} + \frac{1}{2} \hat{T}^2 + \frac{1}{6} \hat{T}^3 + \mathcal{O}(\hat{T}^4) \quad (1.47)$$

où \hat{T} est l'opérateur *cluster* défini par :

$$\hat{T} = \hat{T}_1 + \hat{T}_2 + \hat{T}_3 + \dots + \hat{T}_n \quad (1.48)$$

L'opérateur \mathbf{T}_i a pour effet de générer tous les déterminants de Slater correspondant aux $i^{\text{ème}}$ excitations. Cependant, cette méthode a pour défaut de ne pas être variationnelle : une énergie inférieure à l'énergie exacte du système peut être obtenue. Ceci est aussi le cas des approches perturbatives que nous allons voir dans la section suivante.

1.5.3 Les méthodes perturbatives de type Møller-Plesset

La théorie des perturbations de Rayleigh-Schrödinger consiste à développer en série de Taylor l'énergie du système. On définit alors l'hamiltonien du système $\hat{\mathcal{H}}$ comme étant la somme d'un hamiltonien de référence \hat{H}_o et d'un opérateur de perturbation $\hat{\mathcal{V}}$.

$$\hat{\mathcal{H}} = \hat{H}_o + \hat{\mathcal{V}} \quad (1.49)$$

Dans le cas de la théorie des perturbations de Møller-Plesset à l'ordre n (MP n), l'hamiltonien de référence est donné par la relation [105, 106] :

$$\hat{H}_o = \sum_i^n \hat{F}(i) \quad (1.50)$$

En se limitant à l'ordre 2 dans le développement (MP2), on aboutit à l'équation (1.51)

$$E_{\text{tot}} = E_o + E^{(2)} \quad (1.51)$$

1.6. La théorie de la fonctionnelle de la densité

où la correction de l'énergie au deuxième ordre $E^{(2)}$ est donnée par la relation :

$$E^{(2)} = \sum_{i < j}^{\text{occ}} \sum_{a < b}^{\text{virt}} \frac{\left| \left\langle \Psi_o \left| \sum_{k < l}^n \frac{1}{r_{kl}} \right| \Psi_{ij}^{ab} \right\rangle \right|^2}{\varepsilon_i + \varepsilon_j - \varepsilon_a - \varepsilon_b} = \frac{1}{4} \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{virt}} (ia||jb) T_{ij}^{ab} \quad (1.52)$$

avec

$$T_{ij}^{ab} = (ia||jb) / D_{ij}^{ab} \quad (1.53)$$

$$D_{ij}^{ab} = \varepsilon_i + \varepsilon_j - \varepsilon_a - \varepsilon_b \quad (1.54)$$

1.6 La théorie de la fonctionnelle de la densité

L'idée centrale de la théorie de la fonctionnelle de la densité DFT (*Density Functional Theory*) est basée sur les théorèmes de Hohenberg-Kohn [107] (HK) et l'existence d'un système de référence sans interaction ayant une densité électronique $\rho_S(\mathbf{r})$ égale à celle du système réel $\rho(\mathbf{r})$ en interaction. Ce système de référence est représenté par un déterminant de Slater $\Psi(\mathbf{r})$, constitué d'orbitales monoélectroniques $\{\phi_i(\mathbf{r})\}_{1 \leq i \leq n}$.

$$\rho_S(\mathbf{r}) = \rho(\mathbf{r}) = \sum_i^n \phi_i^*(\mathbf{r}) \phi_i(\mathbf{r}) \quad (1.55)$$

Le premier théorème établit que l'état fondamental électronique d'un système physique est complètement déterminé par la connaissance de sa densité électronique $\rho(\mathbf{r})$, et ainsi la fonction d'onde de l'état fondamental est une fonctionnelle de la densité $\Psi[\rho(\mathbf{r})]$. En d'autres termes, il existe une relation bi-univoque entre la densité électronique du système et le potentiel externe $V_{\text{ext}}[\rho(\mathbf{r})] \equiv E_{\text{ne}}[\rho(\mathbf{r})]$. Le second théorème prouve l'existence d'un principe variationnel, c'est-à-dire, que l'énergie calculée avec une densité électronique $\rho'(\mathbf{r})$ différente de la densité exacte $\rho(\mathbf{r})$ conduira nécessairement à une énergie supérieure. Ces deux théorèmes permettent de décrire tout le système grâce à un outil mathématique à trois dimensions, la densité électronique $\rho(\mathbf{r})$, comparé aux théories précédentes décrivant le système avec une fonction d'onde $\Psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_n)$ à $3n$ dimensions [108–110].

1.6.1 Les équations de Kohn-Sham

D'après les travaux de Kohn et Sham [111] (KS), on peut décomposer l'énergie totale du système $E[\rho(\mathbf{r})]$ de la manière suivante :

$$E[\rho(\mathbf{r})] = V[\rho(\mathbf{r})] + T_S[\rho(\mathbf{r})] + E_{\text{xc}}[\rho(\mathbf{r})] \quad (1.56)$$

où $V[\rho(\mathbf{r})]$ représente la composante électrostatique classique décomposée en contribution noyau-électron $E_{\text{ne}}[\rho(\mathbf{r})]$ et électron-électron $J[\rho(\mathbf{r})]$.

$$V[\rho(\mathbf{r})] = E_{\text{ne}}[\rho(\mathbf{r})] + J[\rho(\mathbf{r})] = - \sum_A^M \int \frac{Z_A \rho(\mathbf{r})}{|\mathbf{r} - \mathbf{R}_A|} d\mathbf{r} + \frac{1}{2} \int \int \frac{\rho(\mathbf{r}) \rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}' \quad (1.57)$$

Chapitre 1. La Mécanique Quantique

Le deuxième terme $T_S[\rho(\mathbf{r})]$ correspond à l'énergie cinétique du système sans interaction ($E_{ne}[\rho(\mathbf{r})] = 0$).

$$T_S[\rho(\mathbf{r})] = -\frac{1}{2} \sum_i^n \int \phi_i^*(\mathbf{r}) \Delta_{\mathbf{r}} \phi_i(\mathbf{r}) d\mathbf{r} \quad (1.58)$$

Quant à $E_{xc}[\rho(\mathbf{r})]$, il regroupe la contribution de l'échange (x) et de la corrélation (c), ainsi qu'un terme $T[\rho(\mathbf{r})] - T_S[\rho(\mathbf{r})]$ correspondant à la différence entre l'énergie cinétique du système réel en interaction et celle du système référence sans interaction. Cette quantité est centrale en théorie de la fonctionnelle de la densité. En utilisant la condition de stationnarité de $E[\rho(\mathbf{r})]$ par rapport à $\rho(\mathbf{r})$,

$$\frac{\delta E[\rho(\mathbf{r})]}{\delta \rho(\mathbf{r})} = 0 \quad (1.59)$$

on déduit les équations de KS :

$$\left\{ -\frac{1}{2} \Delta_{\mathbf{r}} - \sum_A^M \frac{Z_A}{|\mathbf{r} - \mathbf{R}_A|} + \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' + \frac{\delta E_{xc}[\rho(\mathbf{r})]}{\delta \rho(\mathbf{r})} \right\} \phi_i(\mathbf{r}) = \varepsilon_i \phi_i(\mathbf{r}) \quad (1.60)$$

que l'on peut réécrire sous forme d'un problème aux valeurs propres :

$$\hat{F}_{KS} \phi_i(\mathbf{r}) = \varepsilon_i \phi_i(\mathbf{r}) \quad (1.61)$$

Ces équations se résolvent de manière auto-cohérente, tout comme les équations de HF en développant les orbitales monoélectroniques $\{\phi_i(\mathbf{r})\}_{1 \leq i \leq n}$ sur une base d'orbitales atomiques $\{\chi_{\mu}(\mathbf{r})\}_{1 \leq \mu \leq N}$,

$$\phi_i(\mathbf{r}) = \sum_{\mu}^N c_{\mu i} \chi_{\mu}(\mathbf{r}) \quad (1.62)$$

qui conduit à

$$\mathbf{F}_{KS} \cdot \mathbf{C} = \mathbf{S} \cdot \mathbf{C} \cdot \mathbf{E} \quad (1.63)$$

1.6.2 Les fonctionnelles d'échange-corrélation

On définit le potentiel d'échange-corrélation $v_{xc}(\mathbf{r})$ par la dérivée de la fonctionnelle d'échange-corrélation $E_{xc}[\rho(\mathbf{r})]$:

$$v_{xc}(\mathbf{r}) = \frac{\delta E_{xc}[\rho(\mathbf{r})]}{\delta \rho(\mathbf{r})} = \varepsilon_{xc}[\rho(\mathbf{r})] + \rho(\mathbf{r}) \frac{\partial \varepsilon_{xc}(\mathbf{r})}{\partial \rho(\mathbf{r})} \quad (1.64)$$

La connaissance de ce potentiel, intervenant dans les équations de KS, permet de déterminer de façon exacte la densité électronique, et donc les valeurs exactes de l'ensemble des observables de l'état fondamental⁶. Cependant, la nature et la forme mathématique

⁶Nous rappelons qu'à ce stade, aucune approximation ou simplification n'a été apportée dans les équations.

1.6. La théorie de la fonctionnelle de la densité

précise de ce potentiel est inconnue, ce qui donne naissance à plusieurs types d'approximation permettant d'approcher au mieux cette quantité. Généralement, l'énergie d'échange-corrélation E_{xc} est divisé en deux termes : un terme d'échange E_x et un terme de corrélation E_c , bien que cette séparation ne soit pas justifiée théoriquement. L'échange est normalement associée aux interactions entre électrons de même spin, tandis que la corrélation représente s'applique à tous les électrons.

$$E_{xc} = E_x + E_c \quad (1.65)$$

Les différentes fonctionnelles sont donc classées en tant que fonctionnelle d'échange ou de corrélation, l'association de ces deux types donnant naissance à une fonctionnelle d'échange-corrélation. On distingue trois grandes familles de fonctionnelles d'échange et de corrélation [112] :

- Les fonctionnelles de type LDA (*Local Density Approximation*) pour lesquelles la densité est supposée se comporter localement comme la densité d'un gaz uniforme d'électrons (*jellium*). De façon équivalente, la densité est considérée comme uniforme et variant de façon lente avec la position ⁷. La plus ancienne est la fonctionnelle d'énergie cinétique du gaz uniforme d'électron déterminée par Thomas-Fermi [113, 114]. Suivant ce principe, Dirac détermina la fonctionnelle d'échange du gaz uniforme d'électron [115], ce qui donna naissance à la méthode $X\alpha$ de Slater en 1951 [116]. Cette fonctionnelle peut être considérée comme une fonctionnelle LDA où l'énergie de corrélation est négligée, et l'énergie d'échange, donnée par la fonctionnelle de Dirac, est pondérée par un facteur α . Basé sur les travaux de Ceperley et Alder [117] qui utilisèrent la méthode Monte-Carlo quantique pour déterminer la fonctionnelle de corrélation du jellium, Vosko, Wilk et Nusair (VWN) trouvèrent une fonction ajustant au mieux leurs données, afin d'obtenir une fonctionnelle de corrélation [118].
- Pour décrire le caractère non-uniforme de la densité électronique, les méthodes GGA (*Generalized Gradient Approximation*) ont été introduites afin d'insérer une correction utilisant le gradient de la densité $\frac{|\nabla r\rho|}{\rho}$. Nous citerons les fonctionnelles d'échange PW86 [119] (Perdew et Wang), ainsi que celle de Becke [120] (B ou B88) ou bien encore PBE [121, 122]. Nous citerons aussi la fonctionnelle de corrélation PW91 [123] et LYP [124] (Lee, Yang et Parr), donnant naissance à la fonctionnelle BLYP [120] par combinaison avec la fonctionnelle d'échange de Becke. Il existe aussi les fonctionnelles *meta*-GGA utilisant la densité d'énergie cinétique pour correction [125].
- Enfin, les fonctionnelles de type hybride introduisent un certain pourcentage d'échange HF (c_{HF}) en accord avec la théorie de la connection adiabatique entre le système de référence sans interaction et le système réel où les électrons interagissent⁸. Nous citerons ici les fonctionnelles très connues comme B3LYP, B3PW91 [123, 128–130], BH&HLYP [131] ou encore PBE0 [132].

Les fonctionnelles hybrides sont connues comme étant les plus performantes pour la structure électronique. Cependant, leur coût en temps de calcul est bien supérieur aux fonctionnelles LDA ou GGA. Si, tout comme les fonctionnelles LDA et GGA, elles nécessitent le

⁷Nous pourrions remarquer qu'il existe aussi la version utilisant des densités différentes pour les électrons de spin α et β , nommée LSDA (Local Spin Density Approximation).

⁸Les fonctionnelles GWA [126, 127] (*GW Approximation*) introduisent ce pourcentage de façon non-uniforme à l'aide d'une fonction d'écrantage.

Chapitre 1. La Mécanique Quantique

calcul de l'énergie d'échange-corrélation sur une grille d'intégration, elles diffèrent des premières par le calcul de l'échange HF. Dans le cas des fonctionnelles hybrides, l'opérateur de Fock s'écrit :

$$\hat{F}_{\text{KS}} = \left\{ -\frac{1}{2}\Delta_{\mathbf{r}} - \sum_A^M \frac{Z_A}{|\mathbf{r} - \mathbf{R}_A|} + \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' - c_{\text{HF}} \int \frac{\rho(\mathbf{r}, \mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' + (1 - c_{\text{HF}}) \frac{\delta E_{\text{xc}}[\rho(\mathbf{r})]}{\delta \rho(\mathbf{r})} \right\} \quad (1.66)$$

où le quatrième terme représente l'échange HF pondéré par le facteur c_{HF} (taux d'échange HF).

1.6.3 La théorie de la fonctionnelle de la densité dépendante du temps

La méthode de la fonctionnelle de la densité dépendante du temps TD-DFT (*Time-Dependent Density Functional Theory*) est basée sur le théorème de Runge-Gross [133] et sa généralisation aux systèmes moléculaires par van Leeuwen [134]. Le théorème de Runge-Gross représente la version dépendante du temps du premier théorème d'HK. Il constitue la base des équations de KS dépendantes du temps. Par analogie avec la version indépendante du temps, on suppose qu'il existe un système référence dépendant du temps sans interaction et décrit par un déterminant de Slater $\Psi(\mathbf{r}, t)$ construit sur des orbitales monoélectroniques $\psi_i(\mathbf{r}, t)$, tel que :

$$\rho_S(\mathbf{r}, t) = \rho(\mathbf{r}, t) = \sum_i^n \psi_i^*(\mathbf{r}, t) \psi_i(\mathbf{r}, t) \quad (1.67)$$

On en déduit les équations de KS dépendante du temps :

$$\left\{ -\frac{1}{2}\Delta_{\mathbf{r}} - \sum_A^M \frac{Z_A}{|\mathbf{r} - \mathbf{R}_A|} + \int \frac{\rho(\mathbf{r}', t)}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' + \frac{\delta E_{\text{xc}}[\rho(\mathbf{r}, t)]}{\delta \rho(\mathbf{r}, t)} \right\} \psi_i(\mathbf{r}, t) = \hat{F}_{\text{KS}} \psi_i(\mathbf{r}, t) \quad (1.68)$$

$$= i \frac{\partial}{\partial t} \psi_i(\mathbf{r}, t)$$

qui peut s'exprimer comme un problème aux valeurs propres en développant les fonctions monoélectroniques $\{\phi_p(\mathbf{r}, t)\}_{1 \leq p \leq N}$ sur une base de fonctions, monoélectroniques elles aussi, mais indépendantes du temps, tel que :

$$\psi_p(\mathbf{r}, t) = \sum_q^N c_{pq}(t) \phi_q(\mathbf{r}) \quad (1.69)$$

En notation matricielle, ceci conduit à :

$$i \frac{\partial}{\partial t} \mathbf{C} = \mathbf{F}_{\text{KS}} \cdot \mathbf{C} \quad (1.70)$$

1.6.4 Réponse linéaire

Dans le cadre de l'approximation de la réponse linéaire, on suppose que le système est, dans son état fondamental, décrit par les équations de KS indépendantes du temps. Un champ électrique dépendant du temps est alors appliqué au système, et on cherche à déterminer la réponse linéaire (c'est-à-dire la réponse au premier ordre) du système défini par sa matrice densité \mathbf{P} et sa matrice de Fock \mathbf{F} :

$$P_{pq} = P_{pq}^{(0)} + P_{pq}^{(1)} \quad (1.71)$$

$$F_{pq} = F_{pq}^{(0)} + F_{pq}^{(1)} \quad (1.72)$$

La perturbation due au champ électrique dépendant du temps comporte deux termes. La première contribution correspond au champ électrique lui-même :

$$g_{pq} = \frac{1}{2} [f_{pq}e^{-i\omega t} + f_{qp}^*e^{i\omega t}] \quad (1.73)$$

où les éléments f_{pq} sont des opérateurs monoélectroniques décrivant le type de perturbation et tel que,

$$F_{pq}^{(1)} = \Delta F_{pq}^{(0)} + g_{pq} \quad (1.74)$$

$$\Delta F_{pq}^{(0)} = \sum_{st} \frac{\partial F_{pq}^{(0)}}{\partial P_{st}} P_{st}^{(1)} \quad (1.75)$$

La matrice densité est, quant à elle, affectée au premier ordre par :

$$P_{pq}^{(1)} = \frac{1}{2} [d_{pq} \exp^{-i\omega t} + d_{qp}^* \exp^{i\omega t}] \quad (1.76)$$

où les éléments d_{pq} représentent les densités de perturbation. En notant, $x_{ai} \equiv d_{ai}$ et $y_{ai} \equiv d_{ia}$, nous obtenons un problème aux valeurs propres non hermitien

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^* & \mathbf{A}^* \end{pmatrix} \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} = \omega \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} \quad (1.77)$$

avec

$$A_{iajb} = \delta_{ij}\delta_{ab}(\varepsilon_a - \varepsilon_b) + (ia|jb) - c_{\text{HF}}(ij|ab) + (1 - c_{\text{HF}})(ia|\hat{f}_{\text{xc}}|jb) \quad (1.78)$$

$$B_{iajb} = (ia|bj) - c_{\text{HF}}(ib|aj) + (1 - c_{\text{HF}})(ia|\hat{f}_{\text{xc}}|bj) \quad (1.79)$$

dans le cas d'une fonctionnelle d'échange-corrélation hybride. Nous noterons qu'il existe une manière simple de se ramener à un problème aux valeurs propres hermitien, en effectuant l'approximation TDA (*Tamm-Dancoff Approximation*), qui correspond à négliger la matrice \mathbf{B} dans l'expression (1.77) [135, 136]. L'équation (1.77) peut être convertie en un problème hermitien en considérant que les orbitales sont réelles :

$$(\mathbf{A} - \mathbf{B})^{1/2} \cdot (\mathbf{A} + \mathbf{B}) \cdot (\mathbf{A} - \mathbf{B})^{1/2} \cdot \mathbf{Z} = \omega^2 \mathbf{Z} \quad (1.80)$$

$$\mathbf{Z} = (\mathbf{A} - \mathbf{B})^{-1/2} \cdot (\mathbf{X} + \mathbf{Y}) \quad (1.81)$$

La matrice \mathbf{Z} , correspondant aux vecteurs propres, et les valeurs propres ω permet de déterminer les longueurs d'onde de transition entre l'état fondamental et les états excités considérés, ainsi que les moments de transition, que l'on peut faire correspondre aux forces d'oscillateurs.

1.6.5 Relation avec les méthodes TD-HF et CIS

Les éléments de matrice de la méthode *Time-Dependent Hartree-Fock* peuvent être facilement obtenus en fixant le pourcentage d'échange HF à 100 % ($c_{\text{HF}} = 1$) dans la relation (1.78). Ceci conduit à l'expression :

$$A_{iajb} = \delta_{ij}\delta_{ab}(\varepsilon_a - \varepsilon_b) + (ia||jb) \quad (1.82)$$

$$B_{iajb} = (ia||bj) \quad (1.83)$$

De même en effectuant l'approximation TDA, qui correspond à négliger la matrice \mathbf{B} dans l'expression (1.82), nous obtenons l'expression de la matrice \mathbf{A} de la méthode CIS [137] (*Configuration Interaction Single*) :

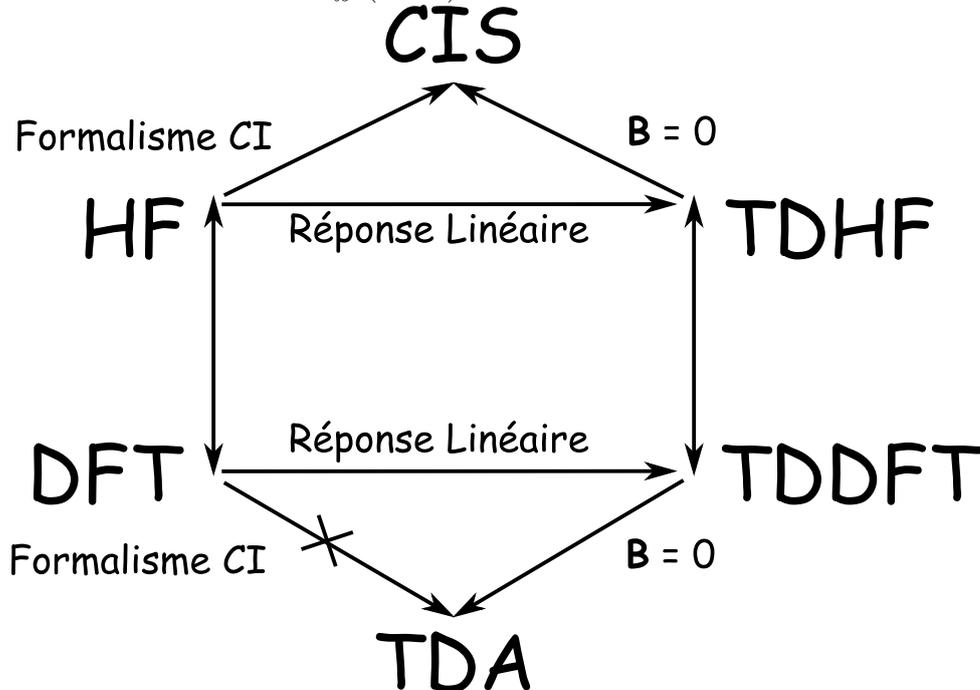
$$A_{iajb} = \delta_{ij}\delta_{ab}(\varepsilon_a - \varepsilon_b) + (ia||jb) \quad (1.84)$$

qui conduisent au problème aux valeurs propres suivant :

$$\mathbf{A} \cdot \mathbf{X} = \omega \mathbf{X} \quad (1.85)$$

Il existe donc une relation très forte entre ces trois méthodes selon le type de potentiel d'échange-corrélation, le niveau d'approximation et le formalisme utilisé (FIG. 1.1). Nous renvoyons le lecteur à l'excellent article de A. Dreuw et M. Head-Gordon [138] pour de plus amples détails sur les équations TD-DFT, TD-HF et CIS.

FIG. 1.1 – Représentation schématique de la relation entre les théories *Hartree-Fock* (HF), *Density Functional Theory* (DFT), *Time-Dependent Hartree-Fock* (TD-HF), *Time-Dependent Density Functional Theory* (TD-DFT), *Configuration Interaction Single* (CIS) et l'approximation *Tamm-Dancoff* (TDA) en TD-DFT.



2

La Mécanique Moléculaire

2.1 Les champs de forces

Les différentes méthodes de chimie quantique présentées au chapitre précédent permettent une description très précise du système, et plus particulièrement de la fonction d'onde électronique. Cependant, leur coût en temps de calcul et en ressources informatiques peut devenir très important même pour des systèmes comportant un nombre d'atomes très faible. Ce coût est principalement dû au calcul de l'énergie électronique pour une configuration nucléaire donnée, conséquence de la résolution approchée de l'équation de Schrödinger. Afin d'éviter cet obstacle, les champs de forces utilisent une fonction paramétrée, correspondant à la fonction énergie potentielle du système, dépendant explicitement des coordonnées nucléaires. Ces paramètres peuvent provenir de différentes approches, comme de valeurs expérimentales ou bien encore être ajustés sur des calculs quantiques. Dans les méthodes basées sur les champs de forces, que l'on réfère souvent par méthodes de mécanique moléculaire (MM), le système est décrit par un ensemble d'atomes représentés par leurs différents paramètres. Ces atomes interagissent alors par un ensemble de potentiels prédéfinis, définissant la fonction énergie potentielle du champ de forces considéré [139, 140].

2.2 La fonction énergie potentielle

De manière générale, la fonction d'énergie potentielle d'un système moléculaire comportant M atomes de position $\{\mathbf{r}_A\}_{1 \leq A \leq M}$ est définie par la relation :

$$E^{\text{MM}}(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_M) = \underbrace{V_{\text{liaisons}} + V_{\text{angles}} + V_{\text{dièdres}} + V_{\text{impropres}}}_{\text{liés}} + \underbrace{V_{\text{vdw}} + V_{\text{elec}}}_{\text{non-liés}} \quad (2.1)$$

où V_{liaisons} , V_{angles} , $V_{\text{dièdres}}$ et $V_{\text{impropres}}$ représentent respectivement l'énergie potentielle des liaisons, des angles, des dièdres et des dièdres impropres. Ces quatre termes sont regroupés dans les potentiels concernant les termes liés. On définit aussi les termes correspondant aux interactions non-liées, V_{vdw} et V_{elec} , qui sont les potentiels d'interaction de van der Waals (vdw) et électrostatiques. Dans les champs de forces les plus communs, les termes

Chapitre 2. La Mécanique Moléculaire

liés suivent, en général, les relations suivantes :

$$V_{\text{liaisons}} = \sum_{\text{liaisons}} \frac{1}{2} k_r (r - r_o)^2 \quad (2.2)$$

$$V_{\text{angles}} = \sum_{\text{angles}} \frac{1}{2} k_\theta (\theta - \theta_o)^2 \quad (2.3)$$

$$V_{\text{dièdres}} = \sum_{\text{dièdres}} \sum_n \frac{V_n}{2} [1 + \cos(n\phi - \gamma)] \quad (2.4)$$

$$V_{\text{impropres}} = \sum_{\text{impropres}} \frac{1}{2} k_i d^2 \quad (2.5)$$

Le terme de liaisons (resp. d'angles) est un potentiel harmonique de constante de force k_r (resp. k_θ) et de valeur à l'équilibre égale à r_o (resp. θ_o). Le potentiel correspondant aux dièdres est développé en série de Fourier ce qui permet de reproduire le caractère périodique de ce potentiel. Les valeurs de V_n permettent de reproduire le profil de torsion, tandis que ϕ représente l'angle de torsion et γ le déphasage à l'origine. Dans la plupart des champs de forces, on considère aussi les angles dièdres impropres, de potentiel $V_{\text{impropres}}$ et de constante de raideur k_i . Ce potentiel est utilisé pour empêcher la pyramidalisation d'un atome hybridé sp^2 . d représente alors la distance entre le plan défini par les trois premiers atomes et le dernier. Considérons maintenant les termes non-liés. Leur expression est donnée par :

$$V_{\text{vdw}} = \sum_{A>B} 4\varepsilon_{AB} \left[\left(\frac{\sigma_{AB}}{R_{AB}} \right)^{12} - \left(\frac{\sigma_{AB}}{R_{AB}} \right)^6 \right] \quad (2.6)$$

$$V_{\text{elec}} = \sum_{A>B} \frac{q_A q_B}{R_{AB}} \quad (2.7)$$

où V_{vdw} représente le potentiel d'interaction de Lennard-Jones (LJ) 6-12 et V_{elec} l'interaction coulombienne entre deux charges q_A et q_B . Le terme en r^{-12} du potentiel de LJ correspond à la répulsion à très courte portée (non-interpénétrabilité des nuages électroniques ou violation du principe de Pauli et répulsion noyau-noyau), tandis que le terme en r^{-6} reproduit l'attraction à longue portée principalement due aux forces de dispersion⁹. Dans la majeure partie des cas, les paramètres $\{\varepsilon_{AB}\}_{\substack{1 \leq A \leq M \\ 1 \leq B \leq M}}$ et $\{\sigma_{AB}\}_{\substack{1 \leq A \leq M \\ 1 \leq B \leq M}}$ entre deux atomes A et B sont calculés d'après les règles de combinaisons de Lorentz-Berthelot¹⁰ :

$$\varepsilon_{AB} = \sqrt{\varepsilon_A \varepsilon_B} \quad (2.8)$$

$$\sigma_{AB} = \frac{\sigma_A + \sigma_B}{2} \quad (2.9)$$

⁹Dans les versions précédentes des champs de forces, un terme prenant en compte l'interaction spécifique par liaison hydrogène (LH) était présent, de forme $\sum_{\text{LH}} \left[\frac{C_{AB}}{R_{AB}^{12}} - \frac{D_{AB}}{R_{AB}^{10}} \right]$.

¹⁰Le champ de forces OPLS (Section 2.3) utilise une moyenne géométrique pour calculer à la fois ε_{AB} ($\varepsilon_{AB} = \sqrt{\varepsilon_A \varepsilon_B}$), mais aussi σ_{AB} ($\sigma_{AB} = \sqrt{\sigma_A \sigma_B}$).

Les interactions non-liées suivent, dans la plupart des cas, la condition 1-4. Celle-ci stipule que les interactions entre deux atomes sont prises en compte s'ils sont séparés par au moins trois liaisons. Dans le cas d'une séparation de trois liaisons, les interactions électrostatiques et de vdw sont pondérées par un facteur préalablement défini, afin d'illustrer que les deux atomes interagissent déjà par un angle dièdre. Pour les séparations supérieures à trois liaisons, l'interaction totale est prise en compte.

2.3 Champs de forces usuels

En mécanique moléculaire, il existe un grand nombre de champs de forces, plus ou moins spécialisés dans un domaine d'applications précis. Nous ne citerons ici que les plus connus :

- Le champ de forces AMBER [141–147] (*Assisted Model Building with Energy Refinement*) est le champ de forces le plus connu, et l'un des plus utilisés. Il a été développé par le groupe de Peter Kollmann, puis de David Case. Il existe plusieurs versions de ce champ de forces (*ff96*, *ff99* ou bien encore *ff03*). Ce champ de forces est très spécifique à l'étude des systèmes biologiques, et plus particulièrement aux protéines et aux acides nucléiques¹¹. La version GAFF (*Generalized AMBER Force Field*) de ce champ de forces a été développée pour les molécules organiques.
- Le champ de forces CHARMM [148–153] (*Chemistry at HARvard using Molecular Mechanics*), développé par Karplus et collaborateurs, est lui aussi principalement utilisé pour les systèmes biologiques. Sa fonction énergie potentielle est quasi-identique à celle du champ de forces AMBER.
- Le champ de forces OPLS [154–160] (*Optimized Potential for Liquid Simulation*), basé sur certains paramètres de AMBER, est plutôt utilisé pour les composés organiques. Il est développé par le groupe de Jørgensen.
- Les champs de forces MM1 [161], MM2 [162], MM3 [163–167, 167–170] et MM4 [171–174] de Allinger et collaborateurs. qui sont des champs de forces beaucoup plus généraux. Ceux-ci utilisent une fonction énergie potentielle très différente des trois champs de forces ci-dessus. Par exemple, le champ de forces MM3 contient des termes d'interaction additionnels [163] (liaison-angle, liaison-torsion, angle-angle) comparé aux champs de forces usuels.
- Dreiding [175] et UFF [176] (*Universal Force Field*), implanté dans le programme Gaussian 03 et développé par Goddard et collaborateurs. Ces deux champs de forces, dits universels, déterminent les différents paramètres à l'aide des propriétés atomiques des éléments (rayons atomiques, potentiels d'ionisation, électronégativités, polarisabilités, etc). Ils permettent donc de couvrir l'ensemble des éléments de la classification périodique. Bien qu'ils permettent d'obtenir des géométries qualitativement bonnes, les résultats concernant les énergies conformationnelles sont souvent moins concluants.
- AMOEBA (*Atomic Multipole Optimized Energetics for Biomolecular Applications*), champ de forces polarisable basé sur des multipôles atomiques. [177–181]

¹¹Dans le cas des acides nucléiques, un potentiel de LJ 10-12 est utilisé

Chapitre 2. La Mécanique Moléculaire

Nous allons maintenant nous intéresser aux méthodes qui combinent ces champs de forces MM à des méthodes de nature quantique.

3

Les Méthodes Hybrides QM/MM

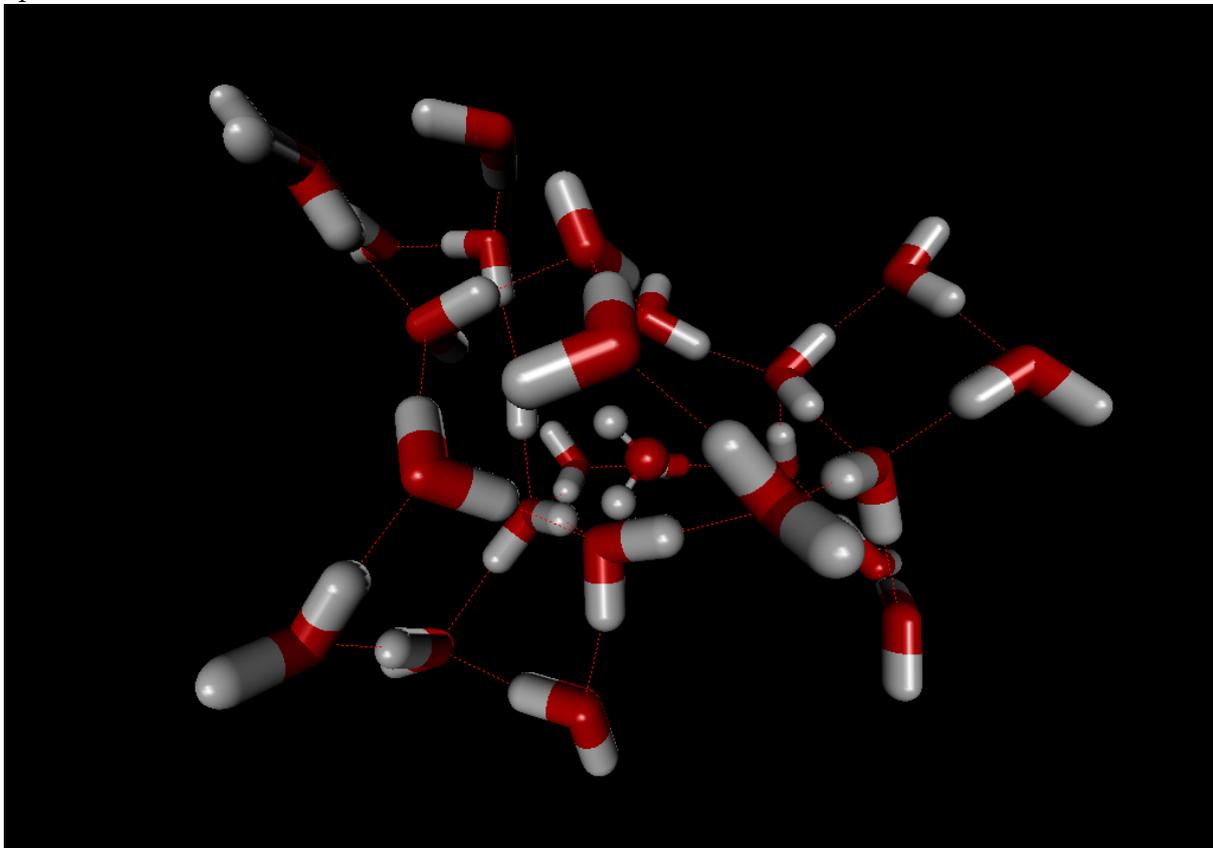
3.1 Principe général

Le principe fondamental des méthodes hybrides QM/MM [65, 66, 182, 183] consiste à diviser spatialement le système étudié en deux parties. Ces méthodes utilisent la nature locale des modifications de la densité électronique, conséquence d'une réaction chimique ou d'une excitation électronique¹². Ceci est généralement motivé par deux raisons principales. Premièrement, le système peut être de taille trop importante pour être traité à l'aide de la mécanique quantique, même avec une méthode peu coûteuse. Deuxièmement, il se peut aussi que l'utilisation d'un champ de forces ne soit pas adaptée au phénomène chimique que l'on veut étudier, comme par exemple dans le cas de la formation ou de la coupure d'une liaison. La méthodologie QM/MM consiste donc à traiter une région localisée, de faible taille, à l'aide de la mécanique quantique qui assure une description correcte des modifications chimiques de cette partie. La partie restante du système, jouant un rôle mineur et/ou secondaire sur le phénomène étudié, peut alors être modélisée par une méthode de moindre coût, comme la mécanique moléculaire qui autorise la prise en compte d'un nombre d'atomes important.

Lorsque cette partition correspond à un système solvant-soluté le couplage entre la partie QM et MM, que nous noterons QM:MM, est relativement simple car les deux sous-systèmes sont bien définis et interagissent exclusivement par des interactions non-liées (électrostatique et vdw). Cependant, lorsque l'on veut étudier des systèmes macromoléculaires, une attention toute spéciale est exigée car la jonction entre les deux parties s'effectue par l'intermédiaire d'une liaison covalente. Nous noterons ce couplage QM/MM. Plusieurs traitements de la frontière QM/MM ont été envisagés afin de décrire proprement cette situation.

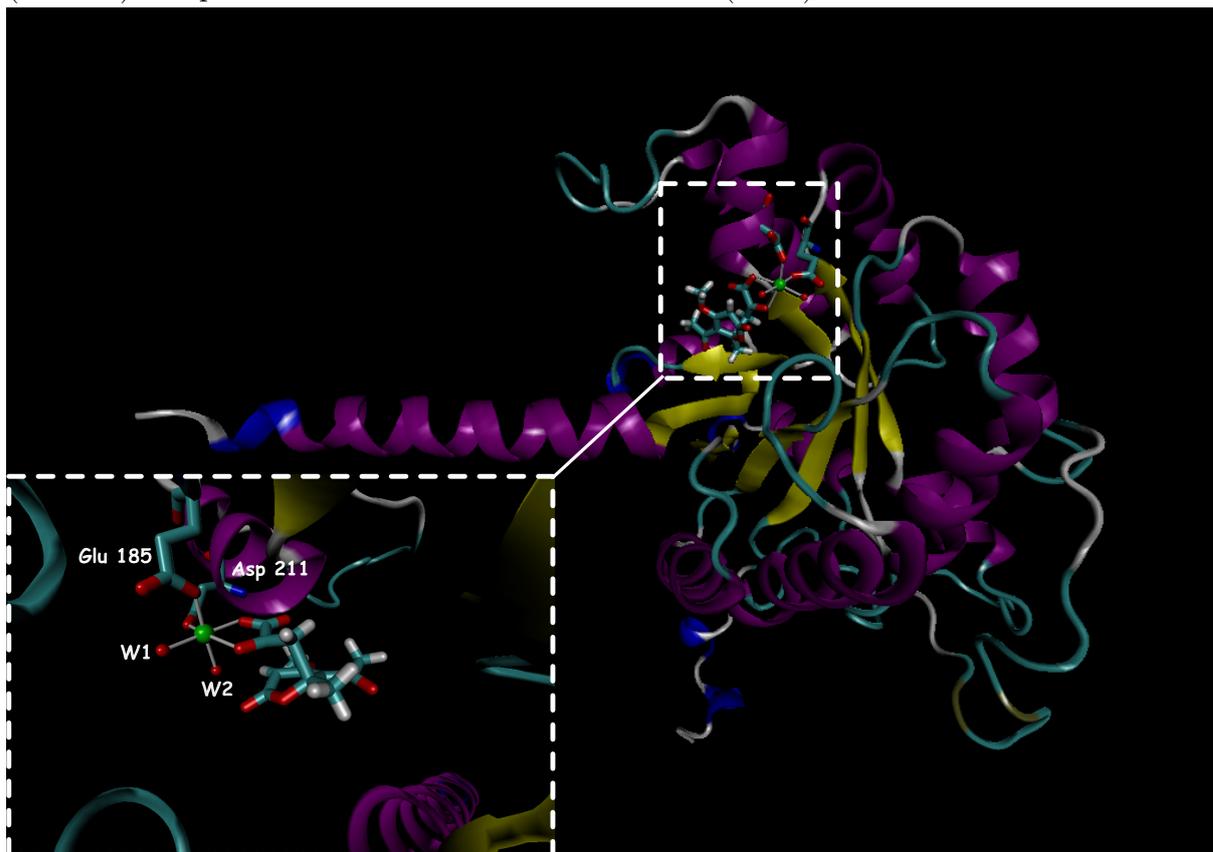
¹²Notons qu'il existe certaines méthodes, considérées comme hybrides, utilisant un découpage énergétique de l'espace des MOs. Nous citerons la méthode MMVB (*Molecular Mechanics Valence Bond*) développée par Robb *et al.* [184].

FIG. 3.1 – Cluster d'eau — Partition QM:MM : la molécule d'eau traitée au niveau QM est représentée en boules et bâtons, tandis que la partie MM est représentée en bâtons épais.



3.1. Principe général

FIG. 3.2 – Macrophomate synthase (MPS) [185–187] — Partition QM/MM : le site actif de l'enzyme est traité au niveau QM (bâtons fins). Il correspond à deux molécules d'eau (W1 et W2), deux acides aminés (Glu 185 et Asp 211), ainsi que deux oxygènes de la molécule de pyruvate. Ils forment un complexe octaédrique autour du magnésium (représenté en vert). Nous représentons l'état de transition hypothétique de la cycloaddition [4+2] (Diels-Alderase) entre le pyruvate (diénophile) et la pyrone (diène). L'existence de cet état de transition est sujet à controverse [188]. Le reste de l'enzyme est traité au niveau MM (*cartoon*). D'après l'entrée de la *Protein Data Bank* (PDB) 1IZC.



3.2 Théorie

3.2.1 Hamiltonien QM/MM

Dans le cas de la méthodologie QM/MM, l'hamiltonien total du système $\hat{\mathcal{H}}$ s'écrit, sous la forme conventionnelle, de la manière suivante :

$$\hat{\mathcal{H}} = \hat{\mathcal{H}}^{\text{QM}} + \hat{\mathcal{H}}^{\text{QM/MM}} + \hat{\mathcal{H}}^{\text{MM}} \quad (3.1)$$

où $\hat{\mathcal{H}}^{\text{QM}}$ et $\hat{\mathcal{H}}^{\text{MM}}$ sont les hamiltoniens de la partie QM et MM, respectivement. L'hamiltonien $\hat{\mathcal{H}}^{\text{QM/MM}}$ correspond à l'interaction QM/MM entre les parties QM et MM. Comme précédemment, on peut séparer l'hamiltonien $\hat{\mathcal{H}}^{\text{QM}}$ en deux parties correspondant à l'hamiltonien de la partie électronique et nucléaire ($\hat{\mathcal{H}}_{\text{elec}}^{\text{QM}}$ et $\hat{\mathcal{H}}_{\text{nucl}}^{\text{QM}}$) :

$$\hat{\mathcal{H}}^{\text{QM}} = \hat{\mathcal{H}}_{\text{elec}}^{\text{QM}} + \hat{\mathcal{H}}_{\text{nucl}}^{\text{QM}} \quad (3.2)$$

où

$$\hat{\mathcal{H}}_{\text{elec}}^{\text{QM}} = -\frac{1}{2} \sum_i^n \Delta_i - \sum_i^n \sum_A^M \frac{Z_A}{|\mathbf{R}_A - \mathbf{r}_i|} + \sum_i^n \sum_{j>i}^n \frac{1}{|\mathbf{r}_j - \mathbf{r}_i|} \quad (3.3)$$

$$\hat{\mathcal{H}}_{\text{nucl}}^{\text{QM}} = \sum_A^M \sum_{B>A}^M \frac{Z_A Z_B}{|\mathbf{R}_B - \mathbf{R}_A|} \quad (3.4)$$

L'hamiltonien $\hat{\mathcal{H}}^{\text{QM/MM}}$ s'explique suivant :

$$\hat{\mathcal{H}}^{\text{QM/MM}} = V_{\text{elec}}^{\text{QM/MM}} + V_{\text{nucl}}^{\text{QM/MM}} + V_{\text{vdw}}^{\text{QM/MM}} \quad (3.5)$$

où $V_{\text{elec}}^{\text{QM/MM}}$, $V_{\text{nucl}}^{\text{QM/MM}}$, $V_{\text{vdw}}^{\text{QM/MM}}$ sont les potentiels d'interaction électron-charge, noyau-charge et de van der Waals (vdw), respectivement. $V_{\text{vdw}}^{\text{QM/MM}}$ représente l'interaction de vdw donnée par la relation (2.6), entre l'ensemble des atomes (QM et MM). Ceci nécessite donc la connaissance des paramètres $\{\varepsilon_A, \sigma_A\}_{A \in \text{QM}}$ pour les atomes de la partie QM. Le choix le plus commun est d'utiliser les paramètres du champ de forces MM considéré pour les atomes de la partie QM. Cependant ceci n'est pas toujours possible, et une reparamétrisation des atomes est parfois nécessaire [189]. Quant à lui, $V_{\text{nucl}}^{\text{QM/MM}}$ est donné par la relation :

$$V_{\text{nucl}}^{\text{QM/MM}} = \sum_A^M \sum_{B \in \text{MM}} \frac{Z_A q_B}{|\mathbf{R}_B - \mathbf{R}_A|} \quad (3.6)$$

Pour ce qui est du potentiel d'interaction électrostatique entre les atomes QM et MM, il existe différents niveaux d'approximation permettant de prendre en compte ce couplage que nous détaillons ci-dessous.

3.2.2 Mechanical Embedding

La méthode *Mechanical Embedding* (ME) calcule l'interaction électrostatique au bas niveau de théorie. Dans le cas d'un calcul QM/MM ou QM :MM, le potentiel $V_{\text{elec}}^{\text{QM/MM}}$

est donc calculé grâce à la relation (2.7). Il faut donc disposer d'un jeu de charges classiques pour les atomes quantiques. On peut résoudre ce problème en utilisant les charges natives du champ de forces pour les atomes de la partie QM. Néanmoins, ceci n'est pas toujours possible car un jeu de paramètres MM n'est pas toujours disponible. Un moyen simple et peu coûteux d'obtenir un jeu de charges pour les atomes QM consiste à dériver les charges du potentiel électrostatique de la partie quantique. Cette procédure permet, par exemple, d'obtenir les charges Merz-Kollman [190,191] (MK), CHELP [192] (*Charges from Electrostatic Potential*) ou CHELPG [193] (*Charges from Electrostatic Potential, Grid Method*). La limitation la plus sérieuse et la plus reconnue de la méthode ME est que les charges de la partie MM ne polarisent pas la fonction d'onde quantique. En d'autres termes, pour une configuration nucléaire donnée, la présence de l'environnement MM n'influence pas la fonction d'onde de la partie quantique. Ceci est particulièrement important lorsque le phénomène chimique étudié engendre un transfert de charges. De plus, cette polarisation de la fonction d'onde peut, dans certains cas, modifier l'ordre énergétique d'un ensemble d'états, comme par exemple dans le cas d'un calcul TD-DFT. Cependant, cette technique a l'avantage d'être très simple à utiliser et à implanter. La version originale de la méthode IMOMM (*Integrated Molecular Orbital Molecular Mechanics*) basée sur le principe ONIOM (*Our own N-layer Integrated molecular Orbital and molecular Mechanics*) par Morokuma et collaborateurs [194,195], utilise la méthodologie ME pour calculer l'interaction électrostatique entre les deux sous-parties.

3.2.3 Electronic Embedding

Contrairement au schéma ME, le schéma *Electronic Embedding* (EE) ne requiert pas la connaissance des charges ponctuelles de la partie QM car l'interaction électrostatique est calculée au haut niveau de théorie par la formule :

$$\hat{V}_{\text{elec}}^{\text{QM/MM}} = - \sum_i^n \sum_{B \in \text{MM}} \frac{q_B}{|\mathbf{R}_B - \mathbf{r}_i|} \quad (3.7)$$

Lorsque cette interaction est ajoutée à l'hamiltonien de cœur de la partie quantique, et donc pris en compte lors du processus de convergence SCF, elle permet une polarisation de la fonction d'onde quantique par les charges de la partie traitée par la mécanique moléculaire. Ceci est un phénomène majeur qui n'est pas présent dans le schéma ME. Il est reproché au schéma EE de demander un effort au niveau du temps de calcul (et aussi de l'implantation) plus important que dans le cas du ME. Cependant, ce surcoût est souvent négligeable par rapport au calcul QM (le calcul des intégrales biélectroniques étant beaucoup plus coûteux par exemple). Ce schéma est maintenant le plus utilisé dans le cas de calculs QM/MM et QM:MM, comme dans la nouvelle version d'IMOMM [196,197], et même, très récemment dans le cas de la méthode IMOMO [198] (*Integrated Molecular Orbital Molecular Orbital*) basé sur une partition QM:QM'. Pour une discussion très complète sur ce sujet, nous renvoyons le lecteur à l'excellent article de Hai Lin et Don Truhlar [199].

3.2.4 Energie QM/MM

En accord avec la décomposition de l'hamiltonien précédente, on définit alors l'énergie totale E d'un système QM/MM (ou QM:MM) comme la somme des énergies de la partie QM, E^{QM} , de la partie MM, E^{MM} , et de l'énergie d'interaction entre les deux parties $E^{\text{QM/MM}}$:

$$\begin{aligned} E &= E^{\text{QM}} + E^{\text{MM}} + E^{\text{QM/MM}} \\ &= E_{\text{elec}}^{\text{QM}} + E_{\text{nuc}}^{\text{QM}} + E^{\text{MM}} + E_{\text{elec}}^{\text{QM/MM}} + E_{\text{nuc}}^{\text{QM/MM}} + E_{\text{vdw}}^{\text{QM/MM}} \end{aligned} \quad (3.8)$$

$E^{\text{QM/MM}}$ est définie par la relation :

$$E^{\text{QM/MM}} = \langle \Psi | - \sum_i^n \sum_{B \in \text{MM}} \frac{q_B}{|\mathbf{R}_B - \mathbf{r}_i|} | \Psi \rangle + V_{\text{nuc}}^{\text{QM/MM}} + V_{\text{vdw}}^{\text{QM/MM}} \quad (3.9)$$

où le premier terme revient à calculer l'interaction électrostatique au niveau EE. L'énergie électronique de la partie quantique $E_{\text{elec}}^{\text{QM}}$ est donnée par la relation (1.33). Cette expression diffère selon les méthodes utilisées pour traiter les problèmes de la coupure d'une liaison covalente dans le cas d'un couplage QM/MM.

3.3 Les différentes méthodes QM/MM

La méthode la plus simple et la plus utilisée est la méthode basée sur un *Link-Atom* (LA) utilisant un atome monovalent afin de saturer les liaisons coupées au niveau de la jonction QM/MM (FIG. 3.3a). Le LA est, dans la majeure partie des cas, un atome d'hydrogène (HLA) [66, 69, 183, 194, 195, 200–204] ou un atome paramétré comme dans le cas de la méthode *Pseudobond* [205] (PB), ou *Adjusted Connection Atom* [206] (ACA), ou bien encore dans la philosophie de la méthode *Quantum Capping Potential* [207] (QCP).

La méthode PB de Zhang *et al.* utilise, dans sa version originale [205], une pseudo-liaison entre la partie active (QM) et inactive (MM). Un atome monovalent associé à un pseudo-potentiel effectif de cœur est construit afin d'effectuer la jonction entre les deux sous-systèmes. Cet atome contient sept électrons de valence, une charge nucléaire égale à sept et la base d'AOs de l'atome de fluor. Les électrons de cœur sont pris en compte dans le pseudo-potentiel effectif qui est ajusté afin de reproduire les paramètres géométriques d'une liaison C(sp³)-C(sp³). Cette stratégie a été par la suite étendue à un nombre plus important de cas [208, 209] : C(sp³)-C(sp²), C(sp³)-N(sp³) et C(sp³)-C(sp², carbonyle)¹³. Dans la méthode ACA, les paramètres semi-empiriques (MNDO, AM1 et PM3) de l'atome frontière sont modifiés afin de reproduire les paramètres géométriques et électroniques d'un groupement méthyle, tandis que la méthode QCP utilise un atome monovalent possédant un potentiel effectif de groupe pour représenter à la fois ses électrons de cœur mais aussi ses électrons de valence.

Une deuxième classe de méthodes utilise des orbitales gelées (FOs pour *Frozen Orbitals*) pour décrire la frontière. Par exemple la méthode *Local Self-Consistent Field*

¹³Une version utilisant un *design atom* à cinq électrons de valence a été récemment proposée [210].

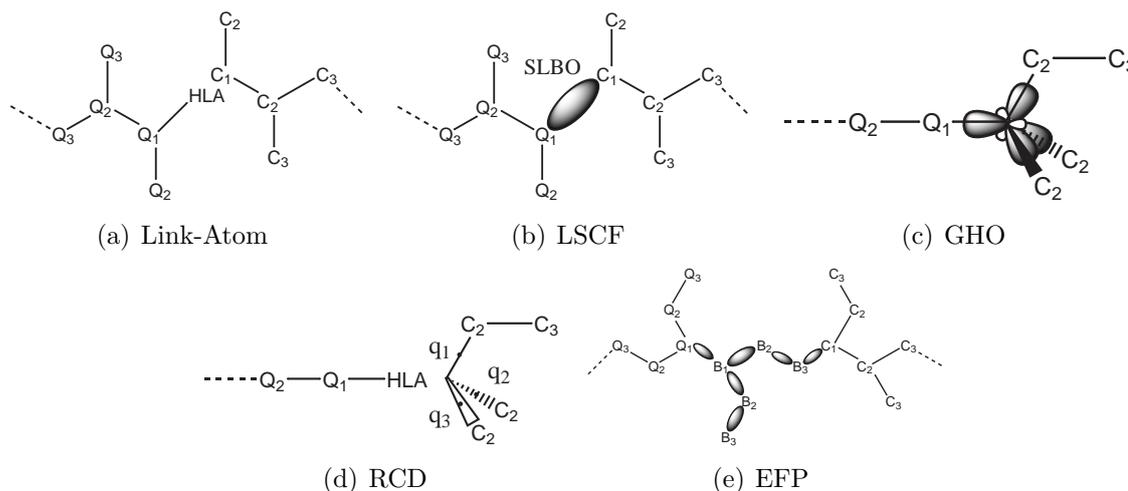
3.3. Les différentes méthodes QM/MM

(LSCF) [74, 211–215], emploie des orbitales localisées pour représenter la liaison entre les atomes frontières (FIG. 3.3b). En utilisant la même idée, Friesner et collaborateurs ont développé une "frozen orbital QM/MM interface methodology" [216–218]. Une extension de la méthode LSCF, appelée *Generalized Hybrid Orbital* [219–223] (GHO), utilise un jeu de quatre orbitales hybrides sp^3 pour chaque atome frontière (FIG. 3.3c). L'ensemble de ces quatre orbitales forme un jeu de quatre fonctions orthonormées. L'orbitale orientée vers la partie quantique est appelée orbitale active et participe au processus d'optimisation de la fonction d'onde. Les trois orbitales hybrides pointant vers les atomes MM directement connectés à l'atome frontière sont appelées orbitales auxiliaires. Leur hybridation est déterminée par la géométrie locale des trois atomes MM connectés à l'atome frontière, et leur occupation est fixée par rapport à la charge classique q_Y de l'atome frontière $(1 - \frac{q_Y}{3})$. Afin de d'assurer l'orthogonalité entre les orbitales hybrides et les MOs optimisées variationnellement, plusieurs méthodes d'orthogonalisation ont été proposées [222, 224]. L'atome frontière est traité à l'aide d'une base minimale de valence, et possède une charge nucléaire égale à quatre. Afin d'atténuer les effets induits par ce mauvais équilibre des bases à l'interface QM/MM, une re-paramétrisation des intégrales monoélectroniques de l'opérateur énergie cinétique impliquant les atomes proches de la frontière a été réalisée.

La méthode *Effective Fragment Potential* (EFP) définit, quant à elle, une région tampon constituée d'orbitales localisées gelées entre les parties QM et MM [225] (FIG. 3.3e).

En combinant les approches LA et GHO, Lin *et al.* ont développé la méthode RCD [226] (*Redistributed Charge and Dipole*) qui consiste à utiliser un HLA et à redistribuer la charge de l'atome C_1 sur les liaisons C_1-C_2 (représentant les orbitales de GHO) en conservant la charge et le dipôle de cette liaison¹⁴ (FIG. 3.3d).

FIG. 3.3 – Différentes méthodologies QM/MM : Link-Atom, LSCF, GHO, RCD et EFP. Les atomes de la partie quantique sont notés Q et ceux de la partie classique C. Dans le cas de la méthode EFP, les atomes de la partie tampon sont notés B.



¹⁴Il a été développé récemment une variante de cette méthode [227] permettant de déterminer les charges à la frontière de manière auto-cohérente suivant la méthode QEq [228]

3.4 Les approches de type Link-Atom

3.4.1 Définition du Link-Atom

Dans les méthodes LA, un atome d'hydrogène (HLA) est ajouté à la partie quantique afin de saturer les liaisons coupées. Cependant, cet atome apporte des degrés de liberté additionnels au système qui ne devraient pas être présents. En effet, ce HLA n'appartient ni à la partie QM, ni à la partie MM, entraînant quelques ambiguïtés dans la définition de l'énergie du système. Afin d'éliminer ces degrés de liberté, la position du HLA est contrainte sur la ligne définie par les atomes Q_1 et C_1 . Conformément aux développements de Morokuma et collaborateurs [67, 201], la distance $R(Q_1 - \text{HLA})$ est modifiée suivant la relation :

$$R(Q_1 - \text{HLA}) = c_{\text{HLA}} R(Q_1 - C_1) \quad (3.10)$$

Le facteur c_{HLA} peut être défini de plusieurs manières, mais est souvent déterminé par le rapport des rayons covalents R_o des atomes de la liaison Q_1-C_1 :

$$c_{\text{HLA}} = \frac{R_o(Q_1 - \text{HLA})}{R_o(Q_1 - C_1)} \quad (3.11)$$

Sa valeur est proche de 0,71 dans la plupart des cas.

3.4.2 Traitement des charges proches de la frontière

Dans le cas des méthodes LA, la proximité de l'atome C_1 et du HLA ($\simeq 0,5 \text{ \AA}$) provoque une sur-polarisation de la fonction d'onde électronique. Afin de pallier cet inconvénient, la charge de l'atome C_1 est souvent mise à zéro, ainsi que dans certains cas, les charges des atomes C_2 et C_3 . Ceci correspondant à l'option par défaut dans le logiciel Gaussian03 [75] pour les calculs ONIOM. Ce logiciel utilise aussi des facteurs d'amortissement différents pour les charges des trois jeux d'atomes C_1 , C_2 et C_3 , ce qui entraîne la non-neutralité de la partie MM, pouvant introduire des artefacts dans la description du système. D'autres groupes conseillent de les redistribuer [226, 227] ou bien encore de les délocaliser partiellement à l'aide de fonctions gaussiennes [204, 229].

3.5 Les approches de type ONIOM

L'acronyme ONIOM désigne une approche générale permettant d'effectuer des calculs utilisant plusieurs « couches » (trois au maximum dans le cas des implantations standards), correspondant aux diverses parties du système, à différents niveaux de théories, que ce soit *ab initio*, semi-empirique ou bien encore en mécanique moléculaire. Le principe du LA est utilisé dans le cas d'une coupure selon une liaison covalente.

Le principe de base consiste à définir un système modèle nécessitant d'être traité à un haut niveau de théorie (HN) d'énergie $E_{\text{modèle}}^{\text{HN}}$ à l'intérieur du système global (dit système réel). Ce système réel est traité au bas niveau de théorie (BN), d'énergie $E_{\text{réel}}^{\text{BN}}$, tandis que le système modèle est traité à la fois au HN ($E_{\text{modèle}}^{\text{HN}}$) et au BN ($E_{\text{modèle}}^{\text{BN}}$) de théorie. La correction due au reste du système sur le système modèle est alors estimée par la différence

3.5. Les approches de type ONIOM

d'énergie entre le système réel au BN et le système modèle au BN ($E_{\text{réel}}^{\text{BN}} - E_{\text{modèle}}^{\text{BN}}$). L'énergie totale E^{ONIOM} du système réel est alors définie par :

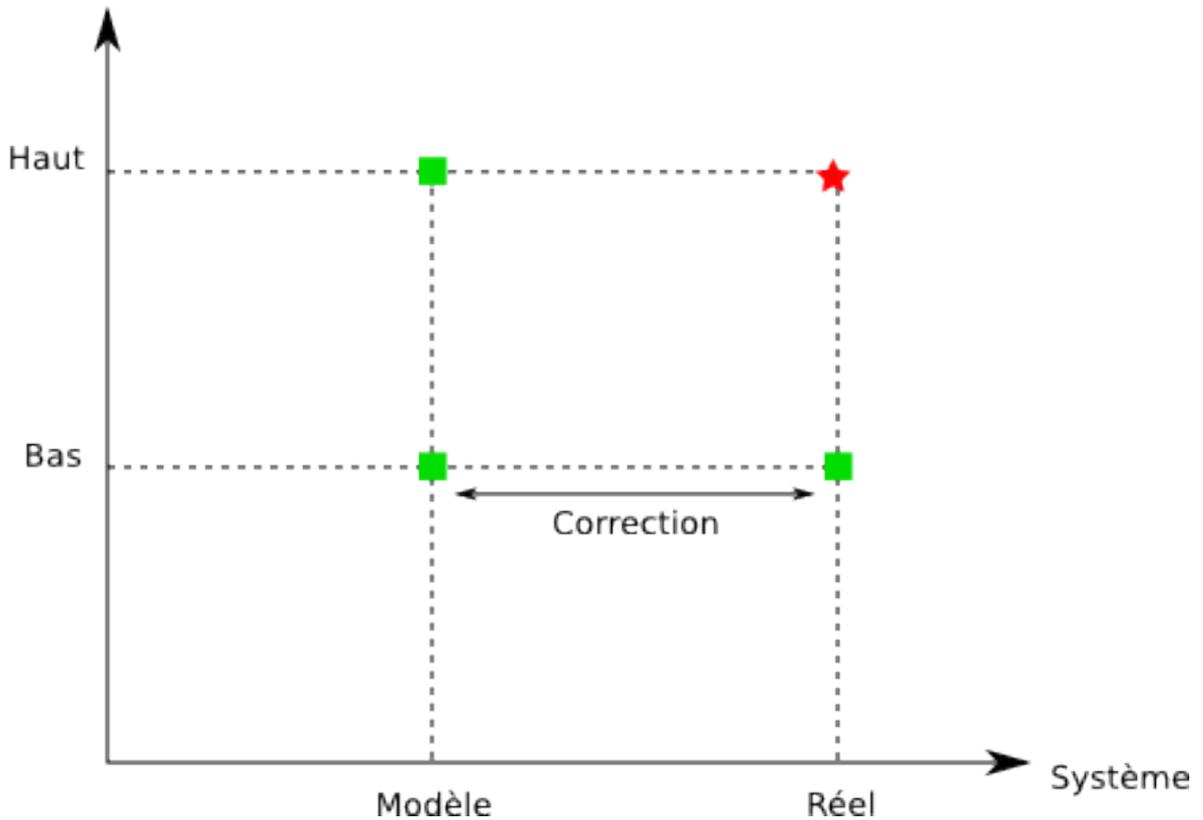
$$E^{\text{ONIOM}} = E_{\text{modèle}}^{\text{HN}} + E_{\text{réel}}^{\text{BN}} - E_{\text{modèle}}^{\text{BN}} \quad (3.12)$$

Cette approche est très générale et permet d'utiliser n'importe quel type de niveau de théorie (FIG. 3.4). Cependant, la correction est calculée au bas niveau de théorie. Le principe des calculs ONIOM a donné naissance aux méthodes IMOMO (QM/QM ou QM:QM) et IMOMM (QM/MM ou QM:MM) ou bien encore celles à trois couches combinant les deux approches (QM/QM/MM ou QM/QM :MM par exemple). Enfin, nous ferons noter qu'il existe une version EE des méthodes IMOMM vérifiant la définition suivante :

$$E^{\text{IMOMM,EE}} = E_{\text{modèle}}^{\text{HN,EE}} + E_{\text{réel}}^{\text{BN}} - E_{\text{modèle}}^{\text{BN,EE}} \quad (3.13)$$

FIG. 3.4 – Principe d'un calcul ONIOM. Les calculs marqués par le symbole ■ sont effectués, tandis que l'énergie du calcul notée par ★ est extrapolée.

Niveau de théorie



Nous allons maintenant nous intéresser aux méthodes utilisant des orbitales gelées, et plus particulièrement à la méthode LSCF.

Deuxième partie

La Méthode du Champ Auto-Cohérent Local (LSCF) : Méthodologies et Développements

4

La Méthode *Local Self-Consistent Field* (LSCF)

4.1 Introduction

Le principe fondamental de la méthode LSCF (*Local Self-Consistent Field*), développée initialement pour les calculs semi-empiriques [211–213, 230] puis *ab initio* [70, 74, 214, 215, 231–241], est l’optimisation de la fonction d’onde multiélectronique sous contrainte. En effet, dans cette dernière, il est possible de geler une partie des spinorbitales durant le processus SCF. Ces spinorbitales sont fixées par l’utilisateur et représentent une densité électronique invariante, ressentie par le reste du système. Dans l’approximation LCAO, les seules données nécessaires à la résolution des équations de Hartree-Fock [93, 94] (ou de KS [107, 111] dans le cas de la DFT) sont les coefficients des AOs dans les orbitales gelées (FOs) et leurs nombres d’occupation. Il n’y a pas de restriction sur la nature des spinorbitales. En effet, celles-ci peuvent être occupées ou vacantes, et utilisées dans le formalisme restreint (RHF) ou non restreint [95–97] (UHF). Dans la section suivante, nous présentons comment la procédure itérative standard SCF doit être modifiée pour prendre en compte ces contraintes.

4.2 Dérivation des équations LSCF

Considérons un système à couches fermées contenant $2n$ électrons et plaçons-nous, pour simplifier l’étude, dans le formalisme restreint. Les MOs variationnelles du système moléculaire seront développées sur une base contenant N AOs $\{|\mu\rangle\}_{1\leq\mu\leq N}$ selon :

$$|p\rangle = \sum_{\mu}^N c_{\mu p} |\mu\rangle \quad (4.1)$$

De même, considérons L orbitales gelées (FOs) $\{|l_P\rangle\}_{1\leq P\leq L}$ développées sur la même base d’AOs :

$$|l_P\rangle = \sum_{\mu}^N l_{\mu P} |\mu\rangle \quad (4.2)$$

4.2.1 Orthogonalisations des FOs

Afin d'éviter de travailler dans une base de MOs non-orthogonales (ce qui augmente le nombre d'intégrales à calculer), les MOs optimisées variationnellement doivent être mutuellement orthogonales, mais aussi orthogonales aux L FOs. Un moyen simple d'assurer ce point est de construire ces MOs sur une base d'AOs à la fois mutuellement orthogonales et orthogonales aux FOs. La première étape de la méthode consiste donc à orthogonaliser mutuellement les L FOs pour obtenir un jeu de L fonctions orthonormées (s'il n'existe pas de dépendance linéaire initialement). Il existe plusieurs types de méthodes permettant d'orthogonaliser ces fonctions à partir d'une matrice d'orthogonalisation \mathbf{X} :

- La méthode la plus courante est la méthode de Löwdin [242]. Elle consiste à utiliser comme matrice d'orthogonalisation \mathbf{X} l'inverse de la racine carrée de la matrice de recouvrement. Notons \mathbf{D} la matrice de recouvrement entre les FOs, $D_{PQ} = \langle l_P | l_Q \rangle$. La matrice de Löwdin est alors égale à $\mathbf{X} = \mathbf{D}^{-1/2}$ [85]. Cette procédure est employée dans le processus SCF afin de travailler dans une base d'AOs mutuellement orthogonales.
- La procédure d'orthogonalisation de Gram-Schmidt (GS) consiste à prendre les vecteurs un à un, et à retrancher successivement les projections des vecteurs précédents pour former un vecteur orthogonal aux précédents. Nous verrons dans la suite que ce procédé est très utile car il permet de laisser inchangé le premier vecteur considéré. Cependant, il est préférable d'utiliser cette procédure dans le cas d'un nombre relativement faible de vecteurs à orthogonaliser. En effet, lorsque l'ensemble devient trop important, des instabilités numériques peuvent intervenir. C'est pour cela que la procédure de Löwdin est préférée dans le cas des calculs SCF. Néanmoins, le nombre de FOs étant en général faible, cette méthode est tout à fait applicable dans notre cas.

Comme il n'y a pas de restriction sur le nombre d'occupation des FOs, la méthode de Löwdin entraîne des rotations sur l'ensemble de l'espace défini par les FOs. Dans le cas où des FOs occupées et vacantes sont définies par l'utilisateur, ceci entraîne donc une rotation entre l'espace occupé et l'espace vacant. Ceci a pour effet de peupler de façon partielle la FO définie comme vide, conduisant à une description erronée du système.

Dans ce cas précis, il faudra donc utiliser une stratégie d'orthogonalisation différente, comme la méthode de GS ¹⁵. Nous verrons dans le cas de l'étude des ionisations de cœur [238], qu'une procédure spécifique a été développée afin de préserver la cohérence dans la représentation du système étudié (Chapitre 6).

4.2.2 Orthogonalisation des AOs

Dans la suite, nous considérerons que les coefficients des FOs ($\{l_{\mu P}\}_{1 \leq P \leq L}$) correspondent au jeu de fonctions orthonormées. L'étape suivante consiste à construire une nouvelle base orthonormée et orthogonale aux L FOs. Pour cela, les N fonctions de base sont orthogonalisées aux FOs en les projetant sur le sous-espace vectoriel complémentaire

¹⁵Une autre méthode consiste à effectuer une orthogonalisation de Löwdin sur les FOs occupées, puis d'utiliser la procédure de GS pour orthogonaliser les FOs vacantes aux FOs occupées précédemment orthogonalisées.

des FOs [243]. Soit $|\tilde{\mu}\rangle$ une telle fonction :

$$|\tilde{\mu}\rangle = \left(1 - \sum_P^L S_{\mu P}^2\right)^{-1/2} \left(|\mu\rangle - \sum_P^L |l_P\rangle \langle l_P|\mu\rangle\right) \quad (4.3)$$

où $S_{\mu P}$ représente l'intégrale de recouvrement $\langle \mu|l_P\rangle$. Ceci permet de définir une matrice \mathbf{M} transformant les AOs de départ $\{|\mu\rangle\}_{1 \leq \mu \leq N}$ en un jeu de fonctions orthogonales aux FOs $\{|\tilde{\mu}\rangle\}_{1 \leq \mu \leq N}$, de terme général :

$$M_{\mu\nu} = \left(1 - \sum_P^L S_{\mu P}^2\right)^{-1/2} \left(\delta_{\mu\nu} - \sum_P^L \sum_{\lambda}^N l_{\nu P} l_{\lambda P} S_{\lambda\mu}\right) \quad (4.4)$$

L'ensemble $\{|\tilde{\mu}\rangle\}_{1 \leq \mu \leq N}$ ne forme pas une base de fonctions mutuellement orthogonales. Ceci peut être résolu en procédant à une orthogonalisation canonique [86] transformant le jeu précédent en un jeu $\{|\mu'\rangle\}_{1 \leq \mu' \leq (N-L)}$ à l'aide de la matrice \mathbf{X} :

$$\mathbf{X} = \mathbf{U} \cdot \tilde{\mathbf{s}}^{-1/2} \quad (4.5)$$

où \mathbf{U} représente la matrice des vecteurs propres de la matrice $\tilde{\mathbf{S}}$ (matrice de recouvrement entre $\{|\tilde{\mu}\rangle\}_{1 \leq \mu \leq N}$) et $\tilde{\mathbf{s}}$, cette même matrice dans la base des vecteurs propres (c'est-à-dire la matrice diagonale). Cependant, cette matrice $\tilde{\mathbf{s}}$ possède L valeurs propres nulles correspondant à L dépendances linéaires introduites par les L FOs. Il nous faut alors éliminer les L colonnes correspondant à ces valeurs propres.

La combinaison de ces deux transformations permet de définir la matrice \mathbf{B} de dimension $(N \times (N - L))$, matrice de passage vers une base de $(N - L)$ AOs mutuellement orthogonales, et orthogonales aux L FOs, $\{|\mu'\rangle\}_{1 \leq \mu' \leq N}$:

$$\mathbf{B} = \mathbf{M} \cdot \mathbf{X} \quad (4.6)$$

Cette matrice \mathbf{B} , ainsi définie, permet de remplacer la matrice de passage $\mathbf{S}^{-1/2}$ du processus SCF standard, et d'effectuer l'optimisation de la fonction d'onde dans le sous-espace complémentaire aux L FOs précédemment définies. Dans ce contexte, l'opérateur de Fock \mathbf{F} est défini suivant la relation classique :

$$F_{\mu\nu} = H_{\mu\nu}^c + \sum_{\lambda\sigma} P_{\lambda\sigma}^T \left[(\mu\nu|\lambda\sigma) - \frac{1}{2}(\mu\sigma|\lambda\nu) \right] \quad (4.7)$$

où nous avons employé les notations des chimistes $(\mu\nu|\lambda\sigma)$ pour les intégrales biélectroniques. \mathbf{P}^T représente la matrice densité totale du système associant deux contributions : l'une de la partie gelée \mathbf{P}^G et l'autre de la partie variationnelle \mathbf{P}^Q , tel que :

$$P_{\mu\nu}^T = P_{\mu\nu}^Q + P_{\mu\nu}^G = 2 \sum_i^{\text{occ}} c_{\mu i} c_{\nu i} + 2 \sum_J^{\text{occ}} l_{\mu J} l_{\nu J} \quad (4.8)$$

La matrice de Fock \mathbf{F} , de dimension $(N \times N)$ est alors transformée dans la base des $\{|\mu'\rangle\}_{1 \leq \mu' \leq (N-L)}$ grâce à la matrice \mathbf{B} pour donner une matrice \mathbf{F}' de dimension $((N -$

Chapitre 4. La Méthode *Local Self-Consistent Field* (LSCF)

$L) \times (N - L)$) par la transformation $\mathbf{F}' = \mathbf{B}^\dagger \cdot \mathbf{F} \cdot \mathbf{B}$. Il ne reste alors qu'à diagonaliser cette dernière pour obtenir la matrice des $(N - L)$ vecteurs propres \mathbf{C}' et des $(N - L)$ valeurs propres formant la matrice diagonale ϵ , puis à repasser dans la base initiale par $\mathbf{C} = \mathbf{B} \cdot \mathbf{C}'$ pour calculer la matrice densité correspondant à la contribution des MOs, \mathbf{P}^Q . En ajoutant à cette dernière la contribution invariante due aux FOs, \mathbf{P}^G , on obtient la matrice densité totale qui permet alors de vérifier la convergence du calcul LSCF en calculant l'énergie par l'expression :

$$E = \sum_{\mu\nu} P_{\mu\nu}^T H_{\mu\nu}^c + \frac{1}{2} \sum_{\mu\nu\lambda\sigma} P_{\mu\nu}^T P_{\lambda\sigma}^T \left[(\mu\nu|\lambda\sigma) - \frac{1}{2}(\mu\sigma|\lambda\nu) \right] \quad (4.9)$$

Après convergence, on obtient donc $(N - L)$ MOs qui forment avec les L FOs un ensemble de N MOs orthonormées, dont n sont doublement occupées.

4.2.3 Modification des équations de Roothaan

La condition de stationnarité de l'énergie par rapport aux coefficients des MOs variationnelles $\{c_{\mu p}\}_{\substack{1 \leq \mu \leq N \\ 1 \leq p \leq N-L}}$,

$$\frac{\delta E}{\delta c_{\mu p}} = 0 \quad (4.10)$$

et les contraintes d'orthogonalité entre l'ensemble des orbitales du système (MOs et FOs),

$$\sum_{\mu\nu} c_{\mu p} S_{\mu\nu} c_{\nu q} = \delta_{pq} \quad (4.11)$$

$$\sum_{\mu\nu} l_{\mu P} S_{\mu\nu} l_{\nu Q} = \delta_{PQ} \quad (4.12)$$

$$\sum_{\mu\nu} c_{\mu p} S_{\mu\nu} l_{\nu Q} = 0 \quad (4.13)$$

conduit à la modification des équations de Roothaan [214, 244] :

$$\mathbf{F} \cdot \mathbf{C} = \mathbf{S} \cdot \mathbf{C} \cdot \mathbf{E} + \mathbf{S} \cdot \mathbf{L} \cdot \mathbf{\Lambda} \quad (4.14)$$

où \mathbf{L} représente la matrice des FOs et $\{\Lambda_{pQ}\}_{\substack{1 \leq p \leq (N-L) \\ 1 \leq Q \leq L}}$ les multiplicateurs de Lagrange du bloc MOs-FOs, tel que :

$$\Lambda_{pQ} \equiv \varepsilon_{pQ} = \sum_{\mu\nu} c_{\mu p} F_{\mu\nu} l_{\nu Q} \quad (4.15)$$

En utilisant le fait que $\mathbf{C} = \mathbf{B} \cdot \mathbf{C}'$ et que $\mathbf{X}^\dagger \cdot \mathbf{S} \cdot \mathbf{X} = \mathbb{I}$, puis en multipliant à gauche par \mathbf{B}^\dagger , on obtient :

$$\mathbf{F}' \cdot \mathbf{C}' = \mathbf{C}' \cdot \mathbf{E} \quad (4.16)$$

avec $\mathbf{F}' = \mathbf{B}^\dagger \cdot \mathbf{F} \cdot \mathbf{B}$ et $\mathbf{C}'^\dagger \cdot \mathbf{S} \cdot \mathbf{L} = \mathbb{O}$ (orthogonalité entre MOs et FOs). Nous retombons sur un problème aux valeurs propres hermitiques, qui se résout par diagonalisation de la matrice \mathbf{F}' . Nous renvoyons le lecteur aux Ref. [244] et [214] pour de plus amples informations sur la modification des équations de Roothaan, et l'introduction des multiplicateurs de Lagrange.

4.3. Dérivées premières de l'énergie par rapport aux coordonnées nucléaires

4.2.4 Algorithme

Nous résumons ici l'algorithme de la méthode LSCF implanté dans Gaussian03 [75] :

1. Orthogonaliser mutuellement les FOs.
2. Construire la matrice \mathbf{M} , orthogonalisant la base des AOs aux FOs.
3. Construire la matrice \mathbf{X} . Transformer les fonctions obtenues en 2. en une base de fonctions orthogonales et linéairement indépendantes.
4. Calculer $\mathbf{B} = \mathbf{M} \cdot \mathbf{X}$.
5. Définir la matrice de Fock \mathbf{F} dans la base originale.
6. Calculer $\mathbf{F}' = \mathbf{B}^\dagger \cdot \mathbf{F} \cdot \mathbf{B}$.
7. Diagonaliser \mathbf{F}' : $\epsilon = \mathbf{C}'^\dagger \cdot \mathbf{F}' \cdot \mathbf{C}'$ où ϵ est la matrice diagonale des valeurs propres.
8. Transformer les vecteurs propres dans la base originale : $\mathbf{C} = \mathbf{B} \cdot \mathbf{C}'$.
9. Calculer \mathbf{P}^Q .
10. Test de sortie. Si non satisfait, retour à 5.

4.3 Dérivées premières de l'énergie par rapport aux coordonnées nucléaires

Afin de déterminer les forces qui s'exercent sur les atomes constituant le système, et d'effectuer ainsi des optimisations de géométries, la connaissance des dérivées premières de l'énergie par rapport à un déplacement nucléaire x est nécessaire [86, 245–247]. Dans la suite de cette thèse, nous utiliserons les notations du livre de Yamaguchi *et al.* [248]. Les travaux concernant les dérivées premières de l'énergie sont décrits dans la thèse de Nicolas Ferré [244]. L'exposant x correspond à la dérivation d'une quantité suivant un déplacement nucléaire x , tandis que l'exposant (x) correspond seulement à la dérivation des intégrales sur la base des AOs (pas des coefficients) intervenant dans cette même quantité. Par exemple :

$$\begin{aligned}
 F_{\mu\nu}^x &= H_{\mu\nu}^{(x)} + \sum_{\lambda\sigma} P_{\lambda\sigma}^T G_{\mu\nu\lambda\sigma}^{(x)} + \sum_{\lambda\sigma} (P_{\lambda\sigma}^T)^x G_{\mu\nu\lambda\sigma} \\
 &= F_{\mu\nu}^{(x)} + \sum_{\lambda\sigma} (P_{\lambda\sigma}^T)^x G_{\mu\nu\lambda\sigma} \\
 H_{\mu\nu}^x &\equiv H_{\mu\nu}^{(x)}
 \end{aligned}
 \tag{4.17}$$

La détermination des dérivées premières de l'énergie dans le cas de la méthode LSCF impose quelques modifications par rapport à l'approche générale [249], dû à la nature non-variationnelle des FOs. En effet, les coefficients des MOs $\{c_{\mu p}\}_{\substack{1 \leq \mu \leq N \\ 1 \leq p \leq (N-L)}}$ sont bien obtenus en accord avec la minimisation de l'énergie LSCF. Cependant, ce n'est pas le cas pour les coefficients $\{l_{\mu P}\}_{\substack{1 \leq \mu \leq N \\ 1 \leq P \leq L}}$ des FOs qui restent, par définition, inchangés. L'expression de

Chapitre 4. La Méthode *Local Self-Consistent Field* (LSCF)

la dérivée de l'énergie par rapport à x , E^x , peut donc s'exprimer à l'aide des dérivées chaînées :

$$\begin{aligned} E^x &= E^{(x)} + \sum_{\mu} \sum_t^{\text{all}} \frac{\partial E}{\partial d_{\mu t}} d_{\mu t}^x \\ &= E^{(x)} + \sum_{\mu} \sum_p^{\text{all}} \frac{\partial E}{\partial c_{\mu p}} c_{\mu p}^x + \sum_{\mu} \sum_P^{\text{all}} \frac{\partial E}{\partial l_{\mu P}} l_{\mu P}^x \end{aligned} \quad (4.18)$$

Or, comme $\partial E / \partial c_{\mu p} = 0$ et $\partial E / \partial l_{\mu P} \neq 0$, il en découle :

$$E^x = E^{(x)} + \sum_{\mu} \sum_P^{\text{all}} \frac{\partial E}{\partial l_{\mu P}} l_{\mu P}^x \quad (4.19)$$

Nous pouvons donc affirmer que l'expression donnant E^x ne fera intervenir que les dérivées par rapport à x des FOs. La dérivation par rapport à x de la relation (4.9) conduit à [86] :

$$E^x = \sum_{\mu\nu} P_{\mu\nu}^T H_{\mu\nu}^{(x)} + \frac{1}{2} \sum_{\mu\nu\lambda\sigma} P_{\mu\nu}^T P_{\lambda\sigma}^T G_{\mu\nu\lambda\sigma}^{(x)} + \sum_{\mu\nu} (P_{\mu\nu}^T)^x F_{\mu\nu} \quad (4.20)$$

qui peut s'écrire en utilisant l'expression (4.14) des équations de Roothaan dans le cas de la méthode LSCF,

$$\begin{aligned} E^x &= \sum_{\mu\nu} P_{\mu\nu}^T H_{\mu\nu}^{(x)} + \frac{1}{2} \sum_{\mu\nu\lambda\sigma} P_{\mu\nu}^T P_{\lambda\sigma}^T G_{\mu\nu\lambda\sigma}^{(x)} - \sum_{\mu\nu} W_{\mu\nu} [I] S_{\mu\nu}^{(x)} \\ &+ \sum_{\mu\nu} \sum_I^{\text{occ}} l_{\mu I}^x F_{\mu\nu} l_{\nu I} + \sum_{\mu\nu} \sum_I^{\text{occ}} l_{\mu I} F_{\mu\nu} l_{\nu I}^x \\ &- \sum_{\mu\nu} \sum_P^{\text{all}} \left[l_{\mu P}^x (\mathbf{S} \cdot \mathbf{P}^Q \cdot \mathbf{F})_{\mu\nu} l_{\nu P} + l_{\mu P} (\mathbf{F} \cdot \mathbf{P}^Q \cdot \mathbf{S})_{\mu\nu} l_{\nu P}^x \right] \end{aligned} \quad (4.21)$$

où

$$W_{\mu\nu} [I] = \sum_i^{\text{occ}} \varepsilon_i c_{\mu i} c_{\nu i} + 2 \sum_i^{\text{occ}} \sum_P^{\text{all}} \varepsilon_i P c_{\mu i} l_{\nu P} \quad (4.22)$$

est la matrice densité pondérée par l'énergie correspondant aux blocs MOs-MOs et MOs-FOs. Ceci conduit à :

$$\begin{aligned} E^x &= \sum_{\mu\nu} P_{\mu\nu}^T H_{\mu\nu}^{(x)} + \frac{1}{2} \sum_{\mu\nu\lambda\sigma} P_{\mu\nu}^T P_{\lambda\sigma}^T G_{\mu\nu\lambda\sigma}^{(x)} - \sum_{\mu\nu} W_{\mu\nu} [I] S_{\mu\nu}^{(x)} \\ &+ \sum_{\mu\nu} (P_{\mu\nu}^G)^x F_{\mu\nu} - \sum_{\mu\nu} \sum_i^{\text{occ}} \sum_P^{\text{all}} \left[\varepsilon_i P l_{\mu P}^x S_{\mu\nu} c_{\nu i} + \varepsilon_i P c_{\mu i} S_{\mu\nu} l_{\nu P}^x \right] \end{aligned} \quad (4.23)$$

4.4. Dérivées secondes de l'énergie par rapport aux coordonnées nucléaires

L'expression de E^x peut s'expliciter de la manière suivante :

$$\begin{aligned}
 E^x = & \sum_{\mu\nu} P_{\mu\nu}^T H_{\mu\nu}^{(x)} + \frac{1}{2} \sum_{\mu\nu\lambda\sigma} P_{\mu\nu}^T P_{\lambda\sigma}^T G_{\mu\nu\lambda\sigma}^{(x)} - \sum_{\mu\nu} W_{\mu\nu} [I] S_{\mu\nu}^{(x)} \\
 & + \sum_{\mu\nu} (P_{\mu\nu}^G)^x F_{\mu\nu} - \sum_{\mu\nu} (W_{\mu\nu} [II])^x S_{\mu\nu}
 \end{aligned} \tag{4.24}$$

avec

$$(P_{\mu\nu}^G)^x = \sum_I^{\text{occ}} l_{\mu I}^x l_{\nu I} + \sum_I l_{\mu I} l_{\nu I}^x \tag{4.25}$$

$$(W_{\mu\nu} [II])^x = \sum_i^{\text{occ}} \sum_P^{\text{all}} [\varepsilon_{iP} l_{\mu P}^x c_{\nu i} + \varepsilon_{iP} c_{\mu i} l_{\nu P}^x] \tag{4.26}$$

où $(W_{\mu\nu} [II])^x$ correspond à la dérivée du bloc MOs-FOs de la matrice densité pondérée par l'énergie. Nous remarquons que l'expression (4.24) nécessite bien la connaissance des dérivées des FOs par rapport aux coordonnées nucléaires. Nous verrons dans la section 4.5 que l'utilisation d'orbitales strictement localisées (SLOs pour *Strictly Localized Orbitals*) permet de déterminer cette quantité, ainsi que les dérivées d'ordre supérieur.

4.4 Dérivées secondes de l'énergie par rapport aux coordonnées nucléaires

Afin de permettre la détermination de la nature des points caractéristiques de la surface d'énergie potentielle du système, mais aussi d'effectuer des optimisations fiables des états de transition et des structures correspondant à un minimum de la courbe d'énergie potentielle, il est préférable de pouvoir effectuer des calculs de fréquences. Plusieurs stratégies peuvent être adoptées : la détermination des modes normaux et de leurs valeurs propres associées peut s'effectuer par calculs numériques. Cependant, cette méthode reste très coûteuse même dans le cas de systèmes de très petite taille. Une deuxième solution consiste à effectuer une partie du calcul de la matrice hessienne au niveau numérique (préférentiellement de petite taille), et l'autre au niveau analytique [250,251]. Cependant, cette méthode n'est valable que si le nombre d'atomes dans la partie calculée numériquement reste très faible, ce qui est rarement le cas. Il est alors préférable d'utiliser une méthodologie permettant d'obtenir des fréquences purement analytiques, comme nous allons le voir dans cette section.

La dérivation de la relation (4.24) en accord avec un déplacement nucléaire y conduit

Chapitre 4. La Méthode *Local Self-Consistent Field* (LSCF)

aux dérivées secondes de l'énergie par rapport aux coordonnées nucléaires [86, 252] :

$$\begin{aligned}
E^{xy} &= \sum_{\mu\nu} P_{\mu\nu}^T H_{\mu\nu}^{(xy)} + \frac{1}{2} \sum_{\mu\nu\lambda\sigma} P_{\mu\nu}^T P_{\lambda\sigma}^T G_{\mu\nu\lambda\sigma}^{(xy)} - \sum_{\mu\nu} W_{\mu\nu} [I] S_{\mu\nu}^{(xy)} \\
&+ \sum_{\mu\nu} (P_{\mu\nu}^T)^y H_{\mu\nu}^{(x)} + \sum_{\mu\nu\lambda\sigma} (P_{\mu\nu}^T)^y P_{\lambda\sigma}^T G_{\mu\nu\lambda\sigma}^{(x)} - \sum_{\mu\nu} (W_{\mu\nu} [I])^y S_{\mu\nu}^{(x)} \\
&+ \sum_{\mu\nu} (P_{\mu\nu}^G)^{xy} F_{\mu\nu} + \sum_{\mu\nu} (P_{\mu\nu}^G)^x F_{\mu\nu}^y \\
&- \sum_{\mu\nu} (W_{\mu\nu} [II])^{xy} S_{\mu\nu} - \sum_{\mu\nu} (W_{\mu\nu} [II])^x S_{\mu\nu}^{(y)}
\end{aligned} \tag{4.27}$$

Ceci conduit à la relation (4.28) donnant la dérivée seconde de l'énergie LSCF par rapport aux coordonnées nucléaires :

$$\begin{aligned}
E^{xy} &= \sum_{\mu\nu} P_{\mu\nu}^T H_{\mu\nu}^{(xy)} + \frac{1}{2} \sum_{\mu\nu\lambda\sigma} P_{\mu\nu}^T P_{\lambda\sigma}^T G_{\mu\nu\lambda\sigma}^{(xy)} - \sum_{\mu\nu} W_{\mu\nu} [I] S_{\mu\nu}^{(xy)} \\
&+ \sum_{\mu\nu} (P_{\mu\nu}^T)^y F_{\mu\nu}^{(x)} - \sum_{\mu\nu} (W_{\mu\nu} [I])^y S_{\mu\nu}^{(x)} \\
&+ \sum_{\mu\nu} (P_{\mu\nu}^G)^{xy} F_{\mu\nu} + \sum_{\mu\nu} (P_{\mu\nu}^G)^x F_{\mu\nu}^{(y)} + \sum_{\mu\nu\lambda\sigma} (P_{\mu\nu}^G)^x (P_{\lambda\sigma}^T)^y G_{\mu\nu\lambda\sigma} \\
&- \sum_{\mu\nu} (W_{\mu\nu} [II])^{xy} S_{\mu\nu} - \sum_{\mu\nu} (W_{\mu\nu} [II])^x S_{\mu\nu}^{(y)}
\end{aligned} \tag{4.28}$$

avec

$$\begin{aligned}
(W_{\mu\nu} [I])^x &= \sum_i^{\text{occ}} [\varepsilon_i^x c_{\mu i} c_{\nu i} + \varepsilon_i c_{\mu i}^x c_{\nu i} + \varepsilon_i c_{\mu i} c_{\nu i}^x] \\
&+ 2 \sum_i^{\text{occ}} \sum_P^{\text{all}} [\varepsilon_{iP}^x c_{\mu i} l_{\nu P} + \varepsilon_{iP} c_{\mu i}^x l_{\nu P} + \varepsilon_{iP} c_{\mu i} l_{\nu P}^x] \\
&= \sum_i^{\text{occ}} \left[\varepsilon_i^x c_{\mu i} c_{\nu i} + \sum_t^{\text{all}} \varepsilon_i U_{ti}^x d_{\mu t} c_{\nu i} + \sum_t^{\text{all}} \varepsilon_i U_{ti}^x c_{\mu i} d_{\nu t} \right] \\
&+ 2 \sum_i^{\text{occ}} \sum_P^{\text{all}} \left[\varepsilon_{iP}^x c_{\mu i} l_{\nu P} + \sum_t^{\text{all}} \varepsilon_{iP} U_{ti}^x d_{\mu t} l_{\nu P} + \sum_t^{\text{all}} \varepsilon_{iP} U_{tP}^x c_{\mu i} d_{\nu t} \right]
\end{aligned} \tag{4.29}$$

et

$$\begin{aligned}
 (W_{\mu\nu} [II])^{xy} &= \sum_i^{\text{occ}} \sum_P^{\text{all}} [\varepsilon_{iP}^y l_{\mu P}^x c_{\nu i} + \varepsilon_{iP} l_{\mu P}^x c_{\nu i}^y + \varepsilon_{iP} l_{\mu P}^{xy} c_{\nu i}] \\
 &+ \sum_i^{\text{occ}} \sum_P^{\text{all}} [\varepsilon_{iP}^y c_{\mu i} l_{\nu P}^x + \varepsilon_{iP} c_{\mu i}^y l_{\nu P}^x + \varepsilon_{iP} c_{\mu i} l_{\nu P}^{xy}] \\
 &= \sum_i^{\text{occ}} \sum_P^{\text{all}} \left[\varepsilon_{iP}^y l_{\mu P}^x c_{\nu i} + \sum_{tu}^{\text{all}} \varepsilon_{iP} U_{tP}^x U_{ui}^y d_{\mu t} d_{\nu u} + \varepsilon_{iP} l_{\mu P}^{xy} c_{\nu i} \right] \\
 &+ \sum_i^{\text{occ}} \sum_P^{\text{all}} \left[\sum_t^{\text{all}} \varepsilon_{iP}^y U_{tP}^x c_{\mu i} d_{\nu t} + \sum_{tu}^{\text{all}} \varepsilon_{iP} U_{ti}^y U_{uP}^x d_{\mu i} d_{\nu u} + \varepsilon_{iP} c_{\mu i} l_{\nu P}^{xy} \right]
 \end{aligned} \tag{4.30}$$

Nous avons introduit la matrice \mathbf{U}^x , matrice de réponse des orbitales à une perturbation x , définie par :

$$\begin{aligned}
 |t\rangle^x &= \left(\sum_{\mu} d_{\mu t} |\mu\rangle \right)^x = \sum_{\mu} d_{\mu t} |\mu\rangle^x + \sum_{\mu} d_{\mu t}^x |\mu\rangle \\
 &= |t\rangle^{(x)} + \sum_{\mu} \sum_u^{\text{all}} U_{ut}^x d_{\mu u} |\mu\rangle = |t\rangle^{(x)} + \sum_u^{\text{all}} U_{ut}^x |u\rangle
 \end{aligned} \tag{4.31}$$

Cette matrice est déterminée par la méthode CPHF [252] (*Coupled-Perturbed Hartree-Fock*). La modification de cette méthode dans le cas de la méthode LSCF est donnée en Annexe A. Cette matrice est aussi nécessaire dans le cas de la détermination des dérivées premières pour les méthodes de type Møller-Plesset [105]. Nous détaillerons la dérivation des gradients MP2 dans le cadre de la méthode LSCF en Annexe B.

4.5 Méthodes QM/MM et LSCF

4.5.1 Utilisation d'orbitales de liaison strictement localisées (SLBO)

L'utilisation de la méthode LSCF ne limite pas la nature des MOs gelées. En effet, celles-ci peuvent être délocalisées ou bien localisées. D'autre part, il n'est pas obligatoire que les FOs soient occupées. L'utilisation d'orbitales vacantes a permis d'utiliser la méthode LSCF pour calculer des énergies d'ionisation de cœur [231, 238] comme nous le verrons dans le chapitre 6. Dans la suite de cette partie, nous nous focaliserons sur l'utilisation d'orbitales de liaison strictement localisées (SLBOs pour *Strictly Localized Bond Orbitals*). Ces orbitales de liaison $\{|l_P\rangle\}_{1 \leq P \leq L}$ ont la particularité d'être définies entre deux atomes X_P et Y_P , tel que :

$$|l_P\rangle = \sum_{\mu \in X_P} l_{\mu P} |\mu\rangle + \sum_{\nu \in Y_P} l_{\nu P} |\nu\rangle \tag{4.32}$$

Dans le cadre de la méthode LSCF, nous utilisons une SLBO doublement occupée $|l_I\rangle$ afin de décrire la jonction entre la partie QM et la partie MM. La SLBO est localisée

entre deux atomes : l'un appartenant à la partie quantique (X_I), et l'autre à la partie classique (Y_I). L'atome de la partie classique Y_I possède un caractère hybride. En effet, ce dernier possède certains paramètres MM du champ de forces utilisé (TAB. 4.1), mais aussi un jeu de fonctions de base, identique à celui qu'il aurait s'il était dans la partie QM. De plus, pour prendre en compte le fait que chaque atome frontière participe à la SLBO en donnant un électron¹⁶, une charge +1 est ajoutée à l'atome classique en plus de sa charge classique MM.

La SLBO est déterminée sur un système modèle contenant la liaison chimique d'intérêt, et appliquée sur le système réel en accord avec le principe de transférabilité, après rotation et renormation. La SLBO peut être obtenue grâce aux nombreux critères de localisation disponibles tels que Pipek-Mezey [253] (PM), Boys-Foster [254–256] (BF), Edmiston-Ruedenberg [257–259] (ER), Weinstein-Paucz (WP) [260, 261] (PM), Magnasco-Perico [262, 263] (MP) ou ELMO [235, 264–269] (*Extremely Localized Molecular Orbitals*). Dans la section suivante, nous allons expliquer comment obtenir ce type de fonctions.

4.5.2 Les critères de localisation

Les spinorbitales obtenues à l'aide de la méthode de résolution SCF sont appelées spinorbitales canoniques. Cet ensemble est unique et défini comme le jeu de fonctions diagonalisant la matrice de Fock \mathbf{F} . De manière plus « chimique », il correspond à un jeu de spinorbitales délocalisées sur l'ensemble de la molécule. Cependant, nous allons montrer qu'il est possible de définir d'autres ensembles de spinorbitales qui conservent les observables, et en particulier l'énergie du système.

Invariance de l'énergie par transformation unitaire

Considérons le jeu de spinorbitale $\{\phi'_i\}_{1 \leq i \leq n}$ obtenu à partir des spinorbitales canoniques $\{\phi_i\}_{1 \leq i \leq n}$, tel que :

$$\phi'_i = \sum_j^n \phi_j U_{ji} \quad (4.33)$$

La matrice \mathbf{U} représente une transformation unitaire ($\mathbf{U}^{-1} = \mathbf{U}^\dagger$) qui conserve l'orthonormalité des spinorbitales, c'est-à-dire que si le jeu $\{\phi_i\}_{1 \leq i \leq n}$ correspond à un ensemble de n fonctions orthonormées, il en sera de même pour le nouveau jeu $\{\phi'_i\}_{1 \leq i \leq n}$. Définissons la matrice \mathbf{A} :

$$\mathbf{A} = \begin{pmatrix} \phi_1(1) & \phi_2(1) & \dots & \phi_n(1) \\ \phi_1(2) & \phi_2(2) & \dots & \phi_n(2) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_1(n) & \phi_2(n) & \dots & \phi_n(n) \end{pmatrix} \quad (4.34)$$

Suivant cette définition, le déterminant de Slater du système $|\Psi_o\rangle$ s'écrit :

$$|\Psi_o\rangle = \frac{1}{\sqrt{n!}} \det(\mathbf{A}) \quad (4.35)$$

¹⁶Dans le cas des liaisons datives, l'un des atomes peut donner deux électrons et l'autre zéro.

On a alors :

$$\mathbf{A}' = \mathbf{A} \cdot \mathbf{U} \quad (4.36)$$

Sachant que $\det(\mathbf{A} \cdot \mathbf{B}) = \det(\mathbf{A}) \det(\mathbf{B})$, on a :

$$\det \mathbf{A}' = \det \mathbf{A} \det \mathbf{U} \quad (4.37)$$

De façon équivalente,

$$|\Psi'_o\rangle = \det(\mathbf{U}) |\Psi_o\rangle \quad (4.38)$$

où $|\Psi'_o\rangle$ représente le déterminant de Slater construit sur la base des $\{\phi'_i\}_{1 \leq i \leq n}$. Comme la matrice \mathbf{U} correspond à une transformation unitaire,

$$\mathbf{U}^{-1} \cdot \mathbf{U} = \mathbf{U}^\dagger \cdot \mathbf{U} = \mathbb{I} \quad (4.39)$$

$$\det(\mathbf{U}^\dagger \cdot \mathbf{U}) = \det(\mathbf{U}^\dagger) \det(\mathbf{U}) = \det(\mathbf{U})^* \det(\mathbf{U}) = |\det(\mathbf{U})|^2 = \det(\mathbb{I}) = 1 \quad (4.40)$$

$$\exists \theta \in [0, 2\pi[, \det(\mathbf{U}) = e^{i\theta} \quad (4.41)$$

Nous en déduisons que les deux déterminants ne diffèrent que par un facteur de phase $e^{i\theta}$. De plus, la matrice \mathbf{U} étant dans la majeure partie des cas réelle, $e^{i\theta} = \pm 1$. Nous pouvons donc conclure que la fonction d'onde construite à l'aide du jeu $\{\phi'_i\}_{1 \leq i \leq n}$ conserve la même signification physique (et en particulier la même énergie) que la fonction d'onde construite sur la base des spinorbitales canoniques. Nous renvoyons le lecteur au livre de Attila Szabo et Neil S. Ostlund [86] pour une dérivation détaillée de ces équations (p. 120). En choisissant astucieusement la transformation unitaire \mathbf{U} , il est donc possible de déterminer des MOs possédant un caractère localisé, contrairement aux MOs canoniques. L'ensemble des critères de localisation consiste à minimiser ou maximiser une quantité physique $\langle \Omega \rangle$ à l'aide de rotations 2×2 successives dans l'espace des MOs occupées. Cette méthodologie est aussi utilisée dans l'espace des fonctions vacantes afin de localiser les orbitales virtuelles [270], comme requis dans les méthodes post-HF à croissance linéaire, comme LMP2 [271–279] (*Local MP2*) ou LCC [280–285] (*Local Coupled-Cluster*).

Les critères externes

Les deux premiers critères que nous allons présenter font partie des critères externes, c'est-à-dire que l'utilisateur doit spécifier les atomes A et B entre lesquels il veut localiser une MO à partir des MOs canoniques.

Weinstein-Pauncz Le critère de Weinstein-Pauncz [260, 261] (WP) consiste à maximiser la quantité $\langle \Omega \rangle_{\text{WP}}$ qui représente la population de recouvrement, au sens de Mulliken [286], entre les deux atomes A et B :

$$\langle \Omega \rangle_{\text{WP}} = \sum_{\mu \in A} \sum_{\nu \in B} c_{\mu i} S_{\mu\nu} c_{\nu i} \quad (4.42)$$

Chapitre 4. La Méthode *Local Self-Consistent Field* (LSCF)

Magnasco-Perico Quant au critère de Magnasco-Perico [262, 263] (MP), il consiste à maximiser la quantité $\langle \Omega \rangle_{\text{MP}}$ donnée par la relation (4.43).

$$\langle \Omega \rangle_{\text{MP}} = \sum_{\mu \in \{A, B\}} \sum_{\nu \in \{A, B\}} c_{\mu i} S_{\mu\nu} c_{\nu i} \quad (4.43)$$

Ceci correspond à maximiser la population de recouvrement entre les atomes A et B , de manière identique au critère de WP, mais incorpore aussi les populations électroniques des deux atomes. De manière générale, ces deux critères produisent des orbitales localisées (LO pour *Localized Orbital*) très similaires. Cependant, lorsque la différence d'électronégativité entre les atomes A et B est importante, les LOs ont un caractère plus ionique avec le second critère, conséquence de la prise en compte des populations électroniques des atomes [287]. Du point de vue pratique, ce dernier est toujours utilisé comme affinement du critère WP, c'est-à-dire que les LOs de départ de la méthode de MP correspondent à celles obtenues suivant la méthode de WP. Nous allons maintenant nous intéresser aux critères internes qui ont la particularité de localiser l'ensemble des MOs.

Les critères internes

Edmiston-Ruedenberg Le critère d'Edmiston-Ruedenberg [257–259] (ER) utilise l'inverse de la distance entre deux électrons comme opérateur. Il maximise donc l'énergie de *self-repulsion* $\langle \Omega \rangle_{\text{ER}}$:

$$\langle \Omega \rangle_{\text{ER}} = \sum_i^n \left\langle \phi'_i \phi'_i \left| \frac{1}{|\mathbf{r}_2 - \mathbf{r}_1|} \right| \phi'_i \phi'_i \right\rangle \quad (4.44)$$

von Niessen La méthode de von Niessen [288] (VN) est une variante de la localisation d'ER, qui utilise la fonction de Dirac δ de la distance entre les électrons suivant la relation (4.45) :

$$\langle \Omega \rangle_{\text{VN}} = \sum_i^n \langle \phi'_i \phi'_i | \delta(\mathbf{r}_2 - \mathbf{r}_1) | \phi'_i \phi'_i \rangle \quad (4.45)$$

Elle correspond à déterminer un jeu de LOs qui maximise la *self-charge*.

Boys-Foster Le critère de Boys-Foster [254–256] (BF) consiste à minimiser la distance au carré entre deux électrons :

$$\langle \Omega \rangle_{\text{BF}} = \sum_i^n \langle \phi'_i \phi'_i | (\mathbf{r}_2 - \mathbf{r}_1)^2 | \phi'_i \phi'_i \rangle \quad (4.46)$$

Physiquement, ceci correspond à déterminer un jeu d'orbitales le moins spatialement étendues, *i.e.* le plus compactes possible. En d'autres termes, la méthode de BF consiste à maximiser la distance entre les centroïdes des orbitales. Contrairement au critère d'ER, cette approche a la particularité de ne pas conserver la séparation entre les MOs $|\sigma\rangle$ et

$|\pi\rangle$. En effet, on dit communément que cette méthode produit des liaisons « banane » (τ et τ'), combinaison symétrique et antisymétrique des liaisons σ et π :

$$|\tau\rangle = \frac{1}{\sqrt{2}} (|\sigma\rangle + |\pi\rangle) \quad (4.47)$$

$$|\tau'\rangle = \frac{1}{\sqrt{2}} (|\sigma\rangle - |\pi\rangle) \quad (4.48)$$

La Ref. [289] apporte une comparaison sur les critères d'ER et de BF.

Pipek-Mezey Ce critère a l'originalité de ne pas localiser les MOs par rapport à des propriétés coulombiennes, comme c'est le cas pour les autres critères externes. Le critère de localisation de Pipek-Mezey [253, 290] consiste à maximiser la somme des charges atomiques de Mulliken [286] : le nombre d'atomes sur lesquels chaque MO est localisée est donc minimisé. La charge atomique de la $i^{\text{ème}}$ MO de l'atome A est donnée par :

$$q_i(A) = \sum_{\mu \in A} \sum_{\nu} c_{\mu i} c_{\nu i} S_{\mu\nu} \quad (4.49)$$

La fonction à maximiser est donc :

$$\langle \Omega \rangle_{\text{PM}} = \sum_A^M [q_i(A)]^2 \quad (4.50)$$

Contrairement à BF, la méthode PM permet de respecter la séparation σ - π .

Utilisation de SLO dans le cadre de la méthode LSCF

A partir des MOs canoniques obtenues par un calcul SCF, le critère de localisation choisi permet d'obtenir des MOs localisées (LMOs pour *Localized Molecular Orbitals*). Ces LMOs sont alors transformées en SLOs (*Strictly Localized Orbitals*) en annulant ce que l'on nomme les queues d'orthogonalisation-délocalisation. Ceci correspond aux coefficients des AOs n'appartenant pas aux deux atomes entre lesquels est localisée la SLO. Ils possèdent généralement une valeur très faible mais non nulle afin de vérifier l'orthogonalité entres les LMOs. Les SLOs sont ensuite appelées SLBO si elle possède un caractère liant.

La méthode ELMO permet d'obtenir directement ce type d'orbitale grâce à une contrainte de la matrice des coefficients \mathbf{C} . En effet, elle permet l'optimisation de la fonction en fixant certains coefficients à zéro. Ceci permet d'obtenir « les meilleures orbitales localisées » et d'éviter l'annulation arbitraire des queues d'orthogonalisation-délocalisation. Cependant, cette méthode requiert un algorithme du deuxième ordre afin d'obtenir une convergence acceptable du processus SCF (type Newton-Raphson).

4.5.3 Distance entre les atomes frontières

La seule présence de la SLBO l_I entre les deux atomes de la frontière X_I et Y_I ne permet pas d'obtenir une distance de liaison acceptable, et ceci quelque soit le couple d'atomes

Chapitre 4. La Méthode *Local Self-Consistent Field* (LSCF)

TAB. 4.1 – Détails des interactions QM/MM. Les atomes de la partie quantique sont notés Q, et ceux de la partie MM sont notés C. Z_A et q_A représentent respectivement la charge nucléaire et la charge classique de l'atome A . $P_{\mu\nu}^T$ est l'élément $\mu\nu$ de la matrice densité totale du système.

		Termes liés
Liaisons		Y-C ₁
Angles		X-Y-C ₁ , Y-C ₁ -C ₂
Torsions		Q ₁ -X-Y-C ₁ , X-Y-C ₁ -C ₂ , Y-C ₁ -C ₂ -C ₃
		Termes non-liés
vdw ^a		Q _i -Y _k , Q _i -C _B , Y _k -C _B , Y _k -Y _l ^b
Electrostatiques	Z-Z	Z _{Q_i} -Z _{Y_k} , Z _{Y_k} -Z _{Y₁} ^b
	Z-q	Z _{Q_i} -q _{Y_k} , Z _{Q_i} -q _{C_B} , Z _{Y_k} -q _{C_B} , Z _{Y_k} -q _{Y₁} ^b
	q-q ^a	q _{Y_k} -q _{C_B} , q _{Y_k} -q _{Y₁} ^b
	e-Z	$\sum_k \sum_{\mu\nu} P_{\mu\nu}^T \left\langle \mu \left \frac{Z_{Y_k}}{r_k - r} \right \nu \right\rangle$
e-q	$\sum_k \sum_{\mu\nu} P_{\mu\nu}^T \left\langle \mu \left \frac{q_{Y_k}}{r_k - r} \right \nu \right\rangle, \sum_{A \in \text{MM}} \sum_{\mu\nu} P_{\mu\nu}^T \left\langle \mu \left \frac{q_{C_A}}{R_A - r} \right \nu \right\rangle$	

^aCondition 1-4 standard des champs de forces.

^b $k \neq l$

frontières considéré. Afin de résoudre ce problème, Ferré *et al.* ont développé un potentiel de liaison frontière (FBP pour *Frontier Bond Potential*) contenant cinq paramètres, de la forme [214, 244] :

$$E_{X_I Y_I} = (A + Br + Cr^2) e^{Dr} + \frac{E}{r} \quad (4.51)$$

où r représente la distance entre les atomes X_I et Y_I . Le premier terme correspond à la prise en compte de la variation du recouvrement entre les AOs des deux atomes en fonction de leur distance r , tandis que le deuxième terme permet de corriger l'interaction électrostatique entre ces deux atomes. En effet, l'atome Y_I possède une charge nucléaire proche de +1, alors que sa charge nucléaire devrait être égale à celle de l'atome quantique. Les cinq paramètres $\{A, B, C, D, E\}$ ont été déterminés par ajustement sur des calculs purement quantiques. Ils sont très peu dépendants de la base des AOs considérée, mais davantage de la nature de l'atome (C, O ou N), et de son état d'hybridation. Les valeurs sont regroupées dans la Ref. [214].

4.5.4 Dérivées premières des SLBOs par rapport aux coordonnées nucléaires

Comme nous l'avons vu dans les sections 4.3 et 4.4, la détermination des dérivées premières et secondes de l'énergie par rapport aux coordonnées nucléaires nécessite la connaissance de dérivées des FOs par rapport à un déplacement élémentaire x , $\{l_{\mu P}^x\}_{1 \leq P \leq L}$. Dans le cas de l'utilisation de SLBOs, cette quantité peut être déterminée par différenciation

numérique :

$$l_{\mu P}^x(x) = \frac{l_{\mu P}(x + \delta x) - l_{\mu P}(x - \delta x)}{2\delta x} \quad (4.52)$$

où $l_{\mu P}(x + \delta x)$ représente le coefficient des AOs μ dans la SLBO $|l_P\rangle$ lors d'un déplacement δx suivant la direction x . En effet, les coefficients de la SLBO $|l_P\rangle$ dépendent uniquement des coordonnées des deux atomes X_P et Y_P . De plus, comme leur nombre est relativement restreint, la détermination de leurs dérivées numériques représente une part négligeable du temps de calcul par rapport aux autres opérations. En effectuant cette différenciation numérique par rapport à deux directions différentes (x et y par exemple), l'expression (4.52) permet aussi de déterminer les dérivées d'ordre supérieur dans le cas des dérivées secondes de l'énergies (Section 4.4).

5

Description Avancée de la Frontière QM/MM

Publications

- **Self-Consistent Strictly Localized Orbitals**
P.-F. Loos, X. Assfeld
J. Chem. Theor. Comput., **2007**, *3*, 1047–1053
- **Frozen core orbitals as an alternative to specific frontier bond potential in hybrid Quantum Mechanics/Molecular Mechanics methods**
A. Fornili, P.-F. Loos, M. Sironi, X. Assfeld
Chem. Phys. Lett., **2006**, *427*, 236–240
- **Removing the extra frontier parameters in QM/MM methods : a tentative with the Local Self-Consistent Field approach**
P.-F. Loos, A. Fornili, M. Sironi, X. Assfeld
Comp. Lett. **2007**, *4*, 473–486
- **On the frontier bond location in the QM/MM description of the peptides and proteins**
P.-F. Loos, X. Assfeld
AIP Conference Proceedings, **2007**, *963*, 308–315

5.1 La méthode *Optimized Local Self-Consistent Field* (OLSCF)

5.1.1 Introduction

Depuis quelques années, de nouvelles méthodes hybrides QM/QM' ont été développées avec pour base la méthodologie ONIOM [67, 68, 198, 291] ou sur d'autres prin-

cipes [292–296]. Celles-ci consistent à découper le système de façon similaire aux méthodes QM/MM mais à traiter les deux parties grâce à des méthodes quantiques différentes. La partie précédemment traitée à l’aide de la mécanique moléculaire est alors décrite par une méthode quantique moins coûteuse (en semi-empirique par exemple) que celle de la partie QM d’intérêt. Dans ce cas, il existe, comme en QM/MM, une interaction électrostatique entre les deux parties ayant pour effet de polariser la fonction d’onde. Toutefois, à cette polarisation, il s’ajoute un terme de transfert électronique dû à la nature quantique de la partie précédemment traitée par mécanique moléculaire. En effet, cette dernière possède maintenant une densité électronique. Il va donc y avoir transfert de charge entre les deux parties par l’intermédiaire des fluctuations de densité électronique de chacune des parties.

Bien que très robuste pour les calculs QM/MM, la méthode LSCF ne permet pas le transfert électronique nécessaire à une bonne description de la jonction entre les deux fragments décrits quantiquement. Ceci est une conséquence directe de l’absence de degré de liberté de la SLBO décrivant l’interface. Dans cet exposé, nous présentons une modification de la méthode LSCF, permettant la réadaptation d’une ou plusieurs SLBOs, en accord avec les variations de densité électronique. Cette méthode, que nous nommerons OLSCF (*Optimized Local Self-Consistent Field*), pourrait permettre, dans l’avenir, de modéliser le transfert de charge entre les deux parties quantiques lors d’un calcul QM/QM’. Afin de valider la méthode, nous travaillerons sur des systèmes modèles, traités de façon totalement quantique, possédant une ou plusieurs SLBOs.

5.1.2 Principe

Notre but est de permettre aux SLBOs, définies sur les fonctions de base des deux atomes appartenant à la liaison, de rester strictement localisées mais d’avoir quelques degrés de liberté afin de s’ajuster en accord avec l’environnement qui l’entoure. Nous avons décidé de mélanger la SLBO doublement occupée avec son orbitale anti-liante correspondante que nous appellerons SLABO (*Strictly Localized Anti-Bonding Orbital*). Chaque frontière sera donc décrite par une SCSLBO (*Self-Consistent Strictly Localized Bonding Orbital*) doublement occupée, combinaison linéaire de la SLBO et de la SLABO¹⁷. Afin d’éviter la délocalisation de la SCSLBO, nous n’orthogonalisons pas mutuellement les SLBOs ainsi que les SLABOs. Les orbitales moléculaires variationnelles sont mutuellement orthogonales et orthogonales à chaque couple {SLBO, SLABO}.

Le mélange d’orbitales localisées occupées et vacantes a déjà été utilisé dans d’autres contextes. Plusieurs études ont été réalisées dans les années 1980, principalement par P. Surjân [297–302], et même avant dans les années 1960 et 1970 par J.-P Malrieu avec la méthode PCILO (*Perturbation Configuration Interaction using Localized Orbitals*) [303–305] et autres [306–309]. Dans cette étude, nous serons moins ambitieux car seules quelques liaisons seront traitées suivant ce principe.

¹⁷Nous donnons un degré de liberté électronique à la liaison.

5.1.3 Formalisme

Considérons un système à couches fermées contenant $2n$ électrons. Par souci de simplicité, nous nous placerons dans le formalisme RHF. La généralisation au formalisme UHF est triviale. Dans cette sous-section, nous allons définir et montrer comment obtenir les quatre types d'orbitales caractérisant la méthode OLSCF. Premièrement, nous rappellerons succinctement comment obtenir les SLBOs. Nous étendrons ce formalisme aux SLABOs. Puis, nous développerons les étapes successives permettant d'obtenir les MOs variationnelles du système. Enfin, le dernier paragraphe concernera la détermination auto-cohérente des SCSLBOs.

SLBO Soit une base d'AOs $\{\chi_\mu\}_{1 \leq \mu \leq N}$ et un nombre L de SLBO doublement occupées $\{|l_I\rangle\}_{1 \leq I \leq L}$ développées sur deux atomes X_I et Y_I définissant une liaison I .

$$|l_I\rangle = \sum_{\mu \in X_I} l_{\mu I} |\mu\rangle + \sum_{\nu \in Y_I} l_{\nu I} |\nu\rangle \quad (5.1)$$

Ces SLBOs sont obtenues suivant les procédures usuelles de localisation [253,255,260–263]. La stricte localisation de ces fonctions est obtenue en annulant les coefficients des AOs n'appartenant pas aux deux atomes X_I et Y_I . La relation (5.1) peut être écrite de façon plus simple en introduisant des orbitales hybrides (HOs pour *Hybrid Orbitals*) :

$$|l_I\rangle = l_{X_I I} |h_I^{X_I}\rangle + l_{Y_I I} |h_I^{Y_I}\rangle \quad (5.2)$$

où $h_I^{X_I}$ est la HO portée par l'atome X_I et utilisée dans la $I^{\text{ème}}$ SLBO $|l_I\rangle$.

SLABO Les SLABOs correspondantes, notées $\{|l_I^*\rangle\}_{1 \leq I \leq L}$, peuvent être obtenues suivant différentes manières :

1. Suivant la même manière que les SLBOs, en utilisant les critères de localisation sur l'espace des orbitales virtuelles [253, 255, 260–263, 310]. Cependant, cette méthode requiert une base d'AOs de petite taille, même pour les molécules comportant un nombre d'atomes peu important.
2. Lorsque la taille de la base d'AOs devient trop importante, le nombre de MOs virtuelles augmente beaucoup plus vite que l'espace occupé. La convergence de la localisation devient alors très difficile à atteindre. Dans ce cas, nous pouvons projeter la SLABO obtenue pour une base d'AOs de plus faible taille (PB pour petite base) sur la base d'AOs de taille plus importante (GB pour grande base).

$$|l_I^*\rangle = \sum_{\mu\nu}^{\text{GB}} |\mu\rangle (\mathbf{S}^{-1})_{\mu\nu} \langle \nu | l_I^* \rangle_{\text{PB}} \quad (5.3)$$

où $\mathbf{S}_{\mu\nu}$ représente le recouvrement des AOs χ_μ et χ_ν de la GB.

3. Les SLABOs peuvent aussi être obtenues grâce à la méthode plus élaborée proposée par Head-Gordon et collaborateurs [311]. Cette méthode localise uniquement les orbitales virtuelles de valence par les procédures standards (BF ou PM par exemple) et projette le reste des orbitales virtuelles sur la base des AOs.

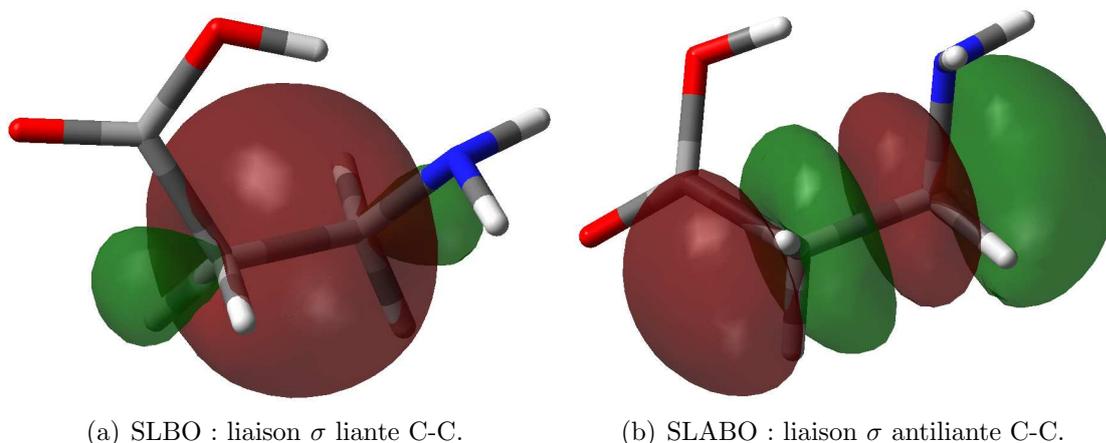
4. Finalement, il est aussi possible d'obtenir l'orbitale antiliante directement à partir de l'orbitale liante associée [298, 312]. Les signes des coefficients des fonctions de base de l'un des atomes sont alors inversés, et la fonction obtenue est normée.

$$\begin{aligned} |l_I^*\rangle &= N_I^* (l_{X_I I} |h_I^{X_I}\rangle - l_{Y_I I} |h_I^{Y_I}\rangle) \\ &= l_{X_I I}^* |h_I^{X_I}\rangle + l_{Y_I I}^* |h_I^{Y_I}\rangle \end{aligned} \quad (5.4)$$

La figure 5.1 présente la SLBO et la SLABO de la liaison C-C centrale de la molécule de β -alanine. Remarquons que dans les trois premiers cas, la SLBO et la SLABO correspondante sont développées sur des HO différentes, alors que la quatrième solution développe la SLBO et la SLABO à partir des mêmes HO. Dans cette section, nous ne considérerons que les deux premières solutions car :

- Elles sont faciles à obtenir.
- D'après les calculs préliminaires, elles donnent les résultats les plus concluants.

FIG. 5.1 – SLBO et SLABO de la liaison C-C centrale de la molécule de β -alanine obtenues par la méthode de PM au niveau B3LYP/6-311++G**. La localisation de la SLBO et de la SLABO a été effectuée sur l'espace des fonctions occupées et sur l'espace des fonctions virtuelles, respectivement.



MO Pour déterminer les MOs variationnelles, nous utiliserons la même procédure que dans la méthode LSCF. La seule exception est que les couples {SLBO, SLABO} ne sont pas mutuellement orthogonaux, afin d'éviter leur délocalisation¹⁸. Le projecteur est alors modifié en conséquence :

$$|\tilde{\mu}\rangle = \left(1 - \sum_{PQ} S_{\mu P} (\mathbf{D}^{-1})_{PQ} S_{\mu Q}\right)^{-1/2} \left(|\mu\rangle - \sum_{PQ} |l_P\rangle (\mathbf{D}^{-1})_{PQ} \langle l_Q|\mu\rangle\right) \quad (5.5)$$

¹⁸L'orthogonalisation entraînerait, par rotation, de développer les SLBOs sur des fonctions de base d'autres atomes.

5.1. La méthode *Optimized Local Self-Consistent Field* (OLSCF)

où \mathbf{D} est la matrice de recouvrement entre deux SLOs l_P et l_Q , tel que $D_{QP} = \langle l_P | l_Q \rangle$. $S_{\mu Q}$ représente l'intégrale de recouvrement entre une AO χ_μ et une SLO l_Q , $\langle \mu | l_Q \rangle$. Ceci définit la transformation matricielle \mathbf{M} selon :

$$M_{\mu\nu} = \left(1 - \sum_{PQ}^{2L} S_{\mu P} (\mathbf{D}^{-1})_{PQ} S_{\mu Q} \right)^{-1/2} \left(\delta_{\mu\nu} - \sum_{PQ}^{2L} \sum_{\lambda} l_{\nu P} (\mathbf{D}^{-1})_{PQ} l_{\lambda Q} S_{\lambda\mu} \right) \quad (5.6)$$

Cette projection est alors suivie d'une orthogonalisation canonique, permettant d'éliminer les $2L$ dépendances linéaires introduites par les L couples {SLBO, SLABO}. Notons \mathbf{X} cette matrice. La conjoncture de ces deux opérations conduit à la transformation \mathbf{B} définie par la relation $\mathbf{B} = \mathbf{M} \cdot \mathbf{X}$, transformant la base initiale de N AOs non orthogonales en une base de $N - 2L$ AOs mutuellement orthogonales et orthogonales aux $2L$ SLOs. Cette matrice \mathbf{B} joue un rôle strictement identique à la matrice d'orthogonalisation de Löwdin $\mathbf{X} = \mathbf{S}^{-1/2}$ dans le processus SCF.

FIG. 5.2 – Structure schématique de la matrice de recouvrement lors d'un calcul OLSCF.

$$\begin{array}{c} \dots \\ \vdots \\ \langle \mu' | \\ \langle \nu' | \\ \vdots \\ \vdots \\ \langle l_P | \\ \langle l_Q | \\ \vdots \end{array} \begin{pmatrix} \dots & |\mu'\rangle & |\nu'\rangle & \dots & \dots & |l_P\rangle & |l_Q\rangle & \dots \\ \dots & 0 & 0 & \dots & \dots & 0 & 0 & \dots \\ 0 & 1 & 0 & \dots & \dots & 0 & 0 & \dots \\ 0 & 0 & 1 & \dots & \dots & 0 & 0 & \dots \\ \vdots & \vdots & \vdots & \ddots & & \vdots & \vdots & \\ \vdots & \vdots & \vdots & & \ddots & \vdots & \vdots & \\ 0 & 0 & 0 & \dots & \dots & 1 & D_{PQ} & \dots \\ 0 & 0 & 0 & \dots & \dots & D_{PQ} & 1 & \dots \\ \vdots & \vdots & \vdots & & & \vdots & \vdots & \ddots \end{pmatrix}$$

L'autre point important dans la méthode OLSCF est la définition de la matrice densité. Tout comme dans la méthode LSCF, grâce à l'orthogonalité du bloc des MOs et des SLOs (FIG. 5.2), il est possible de séparer les contributions correspondantes aux MOs \mathbf{P}^Q et aux SLOs \mathbf{P}^L dans l'expression de la matrice densité totale \mathbf{P}^T

$$P_{\mu\nu}^T = P_{\mu\nu}^Q + P_{\mu\nu}^L = 2 \sum_i^{\text{occ}} c_{\mu i} c_{\nu i} + 2 \sum_J^L l_{\mu J} \tilde{l}_{\nu J} \quad (5.7)$$

Nous avons introduit ici la notion d'orbitales duales [264] (DOs pour *Dual Orbitals*) de coefficients $\tilde{l}_{\mu P}$ définie par :

$$\tilde{l}_{\mu P} = \sum_Q^{2L} l_{\mu Q} (\mathbf{D}^{-1})_{QP} \quad (5.8)$$

Ces orbitales localisées $\{|\tilde{l}_Q\rangle\}_{1 \leq Q \leq 2L}$ ont la propriété d'être orthogonales à l'ensemble des fonctions $\{|l_P\rangle\}_{1 \leq P \leq 2L}$ ($\langle l_P | \tilde{l}_Q \rangle = \delta_{PQ}$) bien que les deux familles ne soit pas mutuellement orthogonales ($\langle l_P | l_Q \rangle = D_{PQ}$ et $\langle l_P | \tilde{l}_Q \rangle = \tilde{D}_{PQ} \neq D_{PQ}$). L'ensemble des SLBOs et

Chapitre 5. Description Avancée de la Frontière QM/MM

de leurs orbitales duales forment alors une base bi-orthogonale (base duale) [313–315]. Il existe une formulation des éléments de matrice entre deux déterminants construits sur des spinorbitales non orthogonales [242] permettant l'évaluation de l'énergie et des équations de Roothaan [316]. Cependant, cette approche nécessite de nombreux changements dans la formulation du problème, tant au point de vue théorique que technique, et augmente le temps de calcul. Nous avons donc choisi d'utiliser le passage, totalement équivalent, en base bi-orthogonale¹⁹.

On montre que les expressions de la matrice de Fock et de l'énergie sont alors identiques à celles des équations (4.7) et (4.9) en remplaçant l'expression de la matrice densité par celle de l'expression (5.7). Ayant effectué cette modification, il est donc possible de diagonaliser le Fockien obtenu et de déterminer, de façon identique à la méthode LSCF, à l'aide la matrice de passage \mathbf{B} , les MOs correspondantes et leur valeurs propres. Cependant le nombre de fonctions obtenues ne sera plus $(N - L)$ mais bien $(N - 2L)$, la dimension de l'espace des orbitales localisées ayant été doublée.

SCSLBO Pour chaque couple {SLBO, SLABO} composé des orbitales $|l_I\rangle$ et $|l_I^*\rangle$ développées sur les fonctions de base des deux atomes X_I et Y_I , nous construisons la matrice (2×2) suivante :

$$\mathbf{F}^I = \begin{pmatrix} F_{11}^I & F_{12}^I \\ F_{12}^I & F_{22}^I \end{pmatrix} = \begin{pmatrix} \langle l_I | \hat{F} | l_I \rangle & \langle l_I^* | \hat{F} | l_I \rangle \\ \langle l_I | \hat{F} | l_I^* \rangle & \langle l_I^* | \hat{F} | l_I^* \rangle \end{pmatrix} \quad (5.9)$$

où \hat{F} représente l'opérateur de Fock défini par (4.7), et prend en compte l'influence de l'ensemble du système. Il faut alors diagonaliser ce Fockien en résolvant le système :

$$\begin{pmatrix} F_{11}^I & F_{12}^I \\ F_{12}^I & F_{22}^I \end{pmatrix} \begin{pmatrix} C_1^I & -C_2^I \\ C_2^I & C_1^I \end{pmatrix} = \begin{pmatrix} \kappa_1^I & 0 \\ 0 & \kappa_2^I \end{pmatrix} \begin{pmatrix} C_1^I & -C_2^I \\ C_2^I & C_1^I \end{pmatrix} \quad (5.10)$$

où $C_1^{I2} + C_2^{I2} = 1$ grâce à la relation d'orthogonalité $\langle l_I | l_I^* \rangle = 0$, et où κ_1^I et κ_2^I représentent les valeurs propres de la matrice \mathbf{F}^I . Le résultat de la diagonalisation de (5.9) permet de déterminer les deux coefficients C_1^I et C_2^I , conduisant au nouveau jeu d'orbitales localisées :

$$|l_I\rangle_{\text{new}} = C_1^I \cdot |l_I\rangle_{\text{old}} + C_2^I \cdot |l_I^*\rangle_{\text{old}} \quad (5.11)$$

$$|l_I^*\rangle_{\text{new}} = -C_2^I \cdot |l_I\rangle_{\text{old}} + C_1^I \cdot |l_I^*\rangle_{\text{old}} \quad (5.12)$$

Ces nouveaux couples étant déterminés, une nouvelle évaluation de la matrice \mathbf{P}^L est alors possible. En ajoutant cette dernière contribution à la matrice densité \mathbf{P}^Q , on obtient alors une nouvelle estimation de la matrice densité totale \mathbf{P}^T , ce qui permet de calculer la nouvelle énergie du système par (4.9) et de vérifier la convergence de la procédure.

¹⁹Si la matrice de recouvrement est très différente de la matrice unité, les fonctions duales seront alors, elles aussi, très différentes des fonctions de l'espace direct. Mais si la matrice de recouvrement est proche de la matrice unité, les fonctions duales seront proches des fonctions de départ.

5.1. La méthode *Optimized Local Self-Consistent Field* (OLSCF)

Algorithme Nous résumons ici les modifications apportées à l'algorithme LSCF afin d'effectuer un calcul OLSCF. Cet algorithme a été implanté dans Gaussian03 [75] :

1. Construire la matrice \mathbf{M} , orthogonalisant la base des AOs aux FOs en utilisant la relation (5.6).
2. Construire la matrice \mathbf{X} . Transformer les fonctions obtenues en 1. en une base de fonctions orthogonales et linéairement indépendantes.
3. Calculer $\mathbf{B} = \mathbf{M} \cdot \mathbf{X}$.
4. Définir la matrice de Fock \mathbf{F} dans la base originale en prenant en compte la matrice densité en base non orthogonale donnée par la relation (5.7).
5. Calculer $\mathbf{F}' = \mathbf{B}^\dagger \cdot \mathbf{F} \cdot \mathbf{B}$.
6. Diagonaliser \mathbf{F}' : $\epsilon = \mathbf{C}'^\dagger \cdot \mathbf{F}' \cdot \mathbf{C}'$ où ϵ est la matrice diagonale des valeurs propres.
7. Transformer les vecteurs propres dans la base originale : $\mathbf{C} = \mathbf{B} \cdot \mathbf{C}'$.
8. Calculer \mathbf{P}^Q .
9. Pour chaque couple {SLBO, SLABO}, diagonaliser le Fockien (2×2) donné par (5.9) pour déterminer les nouveaux couples.
10. Calculer \mathbf{P}^L puis calculer \mathbf{P}^T .
11. Test de sortie. Si non satisfait, retour à 4.

5.1.4 Résultats

Dans l'optique de démontrer que ce nouveau formalisme permet bien à une SLBO de se réajuster en accord avec des contraintes internes ou des effets de variation de l'environnement, nous avons procédé à trois séries de calculs dans des situations plus ou moins modèles. Ces différents cas vont nous permettre de montrer que cette méthode est applicable à la fois dans le cas de systèmes réels, mais aussi dans des situations plus artificielles, pouvant être présentes dans le cas des méthodes hybrides. Premièrement, nous avons considéré les molécules $\text{H}_3\text{C-CX}_3$ ($\text{X} = \text{H}, \text{Li}, \text{BH}_2, \text{NH}_3, \text{OH}$ et F) pour lesquelles la SLBO initiale et la SLABO ont été déterminées sur la molécule d'éthane. Deux niveaux de théorie ont été considérés (HF/6-31G* et HF/6-311G**) et deux types de schéma de localisation (BF et PM) ont été utilisés. La polarité (P_I) de la SCSLBO $|l_I\rangle$ optimisée est calculée comme la moitié de la population de Mulliken (q_{C_X}) due à cette SCSLBO sur l'atome de carbone lié aux substituants X ($\text{CX}_3 \equiv C_X$).

$$P_I = \frac{q_{C_X}}{2} = \sum_{\mu \in C_X} \sum_{\nu} l_{\mu I} l_{\nu I} S_{\mu\nu} \quad (5.13)$$

Pour comparaison, la polarité de la SLBO obtenue avec les critères BF ou PM sur la fonction d'onde SCF optimisée est aussi reportée. Les géométries ont été optimisées au niveau HF. Les polarités sont regroupées dans la table 5.1.

La deuxième série de calculs concerne l'influence du solvant. Celui-ci sera représenté sous forme d'un continuum diélectrique à l'aide d'une méthode SCRF (*Self-Consistent Reaction Field*) basée sur la méthodologie PCM (*Polarizable Continuum Method*). La

SLBO initiale est toujours déterminée sur la molécule d'éthane au niveau B3LYP/6-311++G** à l'aide du critère de localisation PM, puis transférée sur la liaison C-C centrale de la molécule de β -alanine dans sa forme neutre et zwitterionique ($\text{H}_2\text{N}-\text{CH}_2-\text{CH}_2-\text{COOH}$ et $\text{H}_3\text{N}^+-\text{CH}_2-\text{CH}_2-\text{COO}^-$, respectivement). Les calculs sont effectués en phase gazeuse et en phase aqueuse au niveau B3LYP/6-311++G**²⁰. Les géométries sont optimisées au même niveau de théorie dans le solvant, avec et sans SLO. Les polarités des SCSLBO sont données dans la table 5.2. Pour comparaison, nous avons aussi effectué des calculs purement quantique au niveau de théorie B3LYP/6-311++G**, suivi d'une localisation PM afin d'obtenir une SLBO. La polarité de cette SLBO centrale est aussi reportée dans la table 5.2. Ces polarités correspondent à la moitié de la population de Mulliken venant de la SCSLBO de l'atome de carbone portant le groupement NH_2 .

Finalement, l'effet d'une charge ponctuelle classique ($q = |e|$) a été étudié dans la troisième série de calculs. La molécule choisie est celle d'éthane. La SLBO de départ est celle précédemment utilisée et est déterminée au niveau RHF/6-311G** avec le critère PM. La charge est placée sur la ligne définie par les deux atomes de carbone de cette molécule. La distance entre cette charge et l'atome de carbone le plus proche varie entre 1 et 4 Å. La polarité représente la moitié de la population électronique due à la SCSLBO de l'atome de carbone le plus proche de la charge classique. Les résultats sont représentés sur la figure 5.4.

5.1.5 Discussion

D'après les résultats de la table 5.1, on remarque de la polarité (*i.e.* la moitié de la population électronique due à la SCSLBO seulement de l'atome de carbone lié au substituant X) augmente avec l'électronégativité des substituants X . Ces résultats sont en parfait accord avec le « sens chimique » usuel (FIG. 5.3). Bien qu'elles ne soient pas exactement identiques, la tendance est la même quelque soit le niveau de théorie (double- ζ ou triple- ζ), le critère de localisation (BF ou PM), mais aussi quelque soit la SLBO (projetée de la base double- ζ ou non). Cependant, on remarque que la SLO construite à partir du critère PM donne lieu à une polarisation plus importante que celle obtenue par le critère BF. Afin d'obtenir un point de comparaison, nous avons calculé, sur toutes les molécules, les LMOs obtenues à partir d'un calcul SCF standard sur la liaison C-C à chaque niveau de théorie, et pour chaque critère de localisation (BF et PM). Ces LMOs sont alors transformées en SLBO, et leurs polarités sont données dans la deuxième partie du tableau 5.1. Si l'on compare les polarités des SCSLBOs et des SLBOs, on remarque que leurs valeurs sont similaires, bien que non identiques. Pour les SLBOs, l'augmentation de la population électronique avec l'augmentation de l'électronégativité des ligands n'est pas aussi importante que celle des SCSLBOs. De plus, dans le cas des ligands CH_3 et F, la polarité des SLBOs décroît alors que les variations inverses sont observées dans le cas des SCSLBOs. Ceci est dû au fait que les SLBOs sont obtenues par transformation unitaire de l'ensemble des MOs occupées de la molécule, et donc reflètent l'effet global des ligands sur la fonction d'onde. Cet effet ne peut être reproduit de façon exacte par une unique fonction (la

²⁰Un jeu de fonctions diffuses a été ajouté car leur utilisation est courante dans le cadre des calculs en phase solvatée.

5.1. La méthode *Optimized Local Self-Consistent Field* (OLSCF)

TAB. 5.1 – Polarité de l'orbitale de liaison C-C (en électron) pour l'atome de carbone lié au substituant X dans la molécule CH₃CX₃.

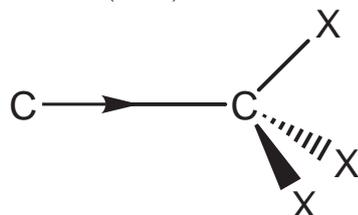
X	Pipek-Mezey			Boys-Foster	
	HF/ 6-31G*	HF/ 6-311G**a	HF/ 6-311G**b	HF/ 6-31G*	HF/ 6-311G**b
SCSLBO					
H	0,500	0,500	0,500	0,500	0,500
Li	0,479	0,481	0,472	0,490	0,484
BH ₂	0,521	0,519	0,529	0,504	0,507
CH ₃	0,529	0,524	0,539	0,509	0,516
NH ₂	0,559	0,553	0,576	0,524	0,539
OH	0,586	0,581	0,612	0,538	0,560
F	0,592	0,594	0,621	0,545	0,571
SLBO					
H	0,500	0,500	0,500	0,500	0,500
Li	0,415	0,460	0,460	0,430	0,475
BH ₂	0,518	0,544	0,544	0,519	0,548
CH ₃	0,498	0,529	0,529	0,498	0,529
NH ₂	0,527	0,580	0,580	0,524	0,575
OH	0,546	0,579	0,579	0,544	0,574
F	0,523	0,559	0,559	0,515	0,553

^aSLABO obtenue au niveau HF/6-311G**.

^bSLABO obtenue au niveau HF/6-31G* et projetée sur la base 6-311G**.

SCSLBO). Néanmoins, nous pouvons conclure que les SCSLBOs reproduisent la tendance attendue, et qu'elles s'ajustent de façon convaincante à la modification chimique d'une molécule.

FIG. 5.3 – Polarisation de la SLBO de la liaison C-C par les substituants électro-négatifs X. Elle sera quantifiée par la relation (5.13).



Les valeurs de la table 5.2 montrent que la SCSLBO et la SLBO sont davantage modifiées par les effets de solvant dans la forme zwitterionique que dans la forme neutre de la molécule de β -alanine. La différence de population électronique entre la phase gaz et en solution est faible dans les deux cas. Pour la forme zwitterionique, la variation de population électronique induite par le solvant est, au moins, dix fois supérieure à celle de la forme neutre. Nous pouvons mettre cet effet en parallèle avec l'énergie libre électrostatique de solvatation qui est aussi faible pour la forme neutre (14,5 kcal/mol) comparé à la forme zwitterionique (59,7 kcal/mol). La polarisation de la SCSLBO est moindre que celle de la SLBO, bien qu'elle présente les mêmes tendances. Ceci peut être attribué, comme précédemment, au caractère « global » de la SLBO, comparé à la nature locale de la SCSLBO. Tout comme dans le paragraphe précédent, nous remarquons toutefois que la SCSLBO réagit correctement aux perturbations du champ électrique extérieur, produit par le continuum diélectrique.

TAB. 5.2 – Polarité de l'orbitale de liaison C-C (en électron) pour l'atome de carbone porté par le groupement NH_2 de la molécule de β -alanine, due à la SCSLBO et à la SLBO(PM)^a.

Forme	Phase	SCSLBO	SLBO(PM)
Neutre	Gaz	0,495	0,501
	Solvant (H_2O)	0,495	0,493
	Δ_{solv}	+0,000	-0,008
Zwitterionique	Gaz	0,553	0,637
	Solvant (H_2O)	0,528	0,552
	Δ_{solv}	-0,025	-0,085

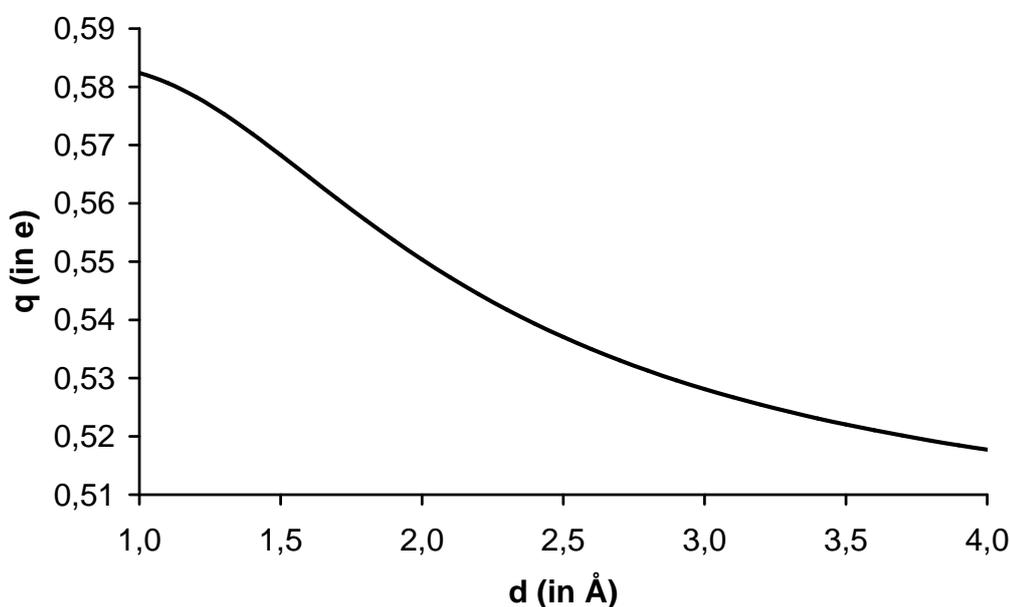
^a Δ_{solv} est la différence de polarité entre la phase aqueuse et la phase gaz.

Finalement, nous avons décidé d'étudier l'effet d'une charge ponctuelle classique sur la polarité de la SCSLBO. Nous remarquons (FIG. 5.4) que plus l'interaction est importante (*i.e.* la distance entre la charges ponctuelle et la molécule est faible), plus la polarisation

5.1. La méthode *Optimized Local Self-Consistent Field* (OLSCF)

de la SCSLBO est importante. Ceci est exactement le comportement attendu dans ce genre de situation, et est prometteur pour l'implantation de cette méthode au niveau QM/MM et/ou QM/QM'.

FIG. 5.4 – Moitié de la population électronique (q , en électron) due à la SCSLBO de l'atome de carbone le plus proche de la charge classique (*i.e.* polarité de la liaison), fonction de la distance (d en Å) entre la charge classique positive et l'atome de carbone de la molécule d'éthane le plus proche de celle-ci. La charge ponctuelle est placée sur la ligne définie par les deux atomes de carbone.



5.1.6 Conclusion

Dans un premier temps, nous avons proposé une modification de la méthode LSCF, basée uniquement sur des calculs quantiques. Ceci a permis de démontrer l'ajustement de la SCSLBO, combinaison linéaire de la SLBO et de la SLABO, en accord avec les variations de densité électronique de l'environnement proche ou de l'état quantique. Nous avons nommé cette méthode OLSCF, pour *Optimized Local Self-Consistent Field*. Pour chaque SLBO, une SLABO est définie, et la combinaison de ces deux fonctions permet d'attribuer un degré de liberté, responsable de la relaxation de l'orbitale. La nouvelle méthodologie a été testée sur des modifications de la molécule elle-même, ainsi que par rapport à l'influence d'un solvant et à des charges ponctuelles classiques. Dans ces trois situations, il a été

montré que la SCSLBO s’ajuste de façon convaincante à ces perturbations.

5.2 Prise en compte des électrons de cœur et de valence

5.2.1 Electrons de cœur

Introduction

Pour assurer des résultats fiables et une description précise de la frontière QM/MM, toutes les méthodes QM/MM introduisent un traitement spécial des paramètres proches de la liaison frontière. Dans les méthodes de type LA, des traitements spéciaux des charges à la frontière sont utilisés. En effet, la proximité de ces charges MM avec le HLA induit une sur-polarisation de la fonction d’onde [183, 204, 206, 226, 229, 317–319]. Ces charges ponctuelles doivent alors être mise à zéro, modifiées, redistribuées [226] ou bien encore partiellement délocalisées [204]. Dans le cas des approches des atomes paramétrés (QCP, ACA et PB) [205–207], les paramètres semi-empiriques ou du pseudo-potential sont ajustés pour reproduire les paramètres géométriques du système modèle. Dans le cas de la méthode LSCF, un potentiel à cinq paramètres a été spécialement développé pour assurer une distance de liaison correcte à la frontière QM/MM, ainsi que la forme de la surface d’énergie potentielle [214]. Friesner et collaborateurs ont paramétré une librairie de densités gelées pour les chaînes latérales de tous les acides aminés et positionnent, entre les deux atomes frontières, une charge classique préalablement déterminée afin de reproduire des valeurs géométriques et énergétiques de référence [216–218]. Dans la méthode GHO, une base de valence minimale et une re-paramétrisation des intégrales monoélectroniques a été utilisée pour l’atome frontière [222]. Tous ces exemples montrent qu’une attention toute particulière doit être apportée au niveau de la jonction des deux sous-systèmes.

Cette paramétrisation spécifique, nécessaire pour obtenir une description correcte de la frontière, peut être considérée comme un effet secondaire de l’idée initiale appliquée pour connecter la partie QM et MM. Plusieurs effets collatéraux en résultent : les paramètres peuvent être dépendants du champ de forces, de la base utilisée, ou bien encore de la méthode quantique utilisée. Ceci peut alors entraîner une transférabilité moindre, impliquant une description erronée pour certains systèmes. Afin d’obtenir une plus grande généralité des approches hybrides, il est donc désirable d’aller au-delà de cette paramétrisation, d’autant plus que le contenu physique de ces paramètres additionnels est parfois mal défini, ou bien même absent.

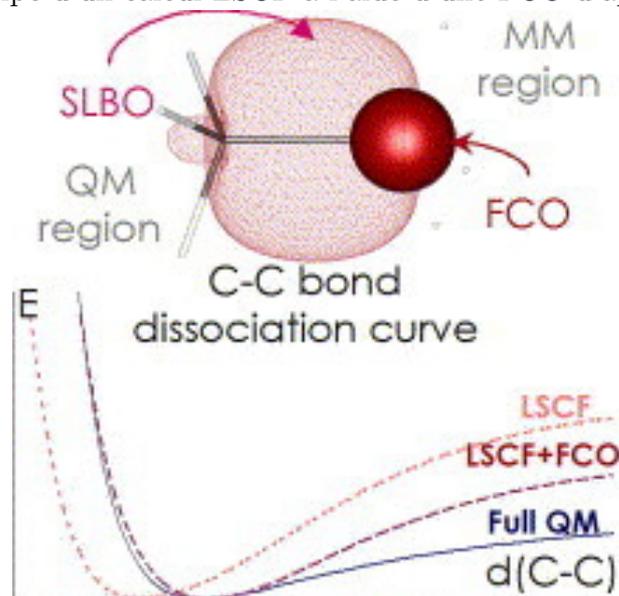
Dans le cas de la méthode LSCF, le potentiel de la liaison frontière (FBP pour *Frontier Bond Potential*), donné par la relation (4.51) corrige principalement deux erreurs, comme nous l’avons expliqué dans la section 4.5.3 : le premier terme correspond à la prise en compte de la variation du recouvrement entre les AOs des deux atomes, tandis que le deuxième terme permet de corriger l’interaction électrostatique entre ces deux atomes. Afin de substituer ce potentiel frontière empirique par une quantité de nature *ab initio*, il nous faut déterminer comment améliorer la description de l’atome frontière Y. Pour ce qui est de l’erreur sur l’interaction électrostatique entre les atomes X et Y, nous pouvons

5.2. Prise en compte des électrons de cœur et de valence

envisager d'augmenter la charge de l'atome Y afin qu'elle se rapproche le plus possible de celle qu'il posséderait s'il était de nature purement quantique. La solution la plus simple consiste donc à introduire de façon explicite les électrons de cœur des atomes frontières LSCF. Deux solutions sont alors envisageables : la première consiste à utiliser une orbitale de cœur gelée (FCO pour *Frozen Core Orbital*). Cette approche a l'avantage majeur de remplacer le potentiel frontière empirique, par un « objet » de nature quantique. De manière analogue à la SLBO qui permet de reproduire une densité correcte au niveau de la frontière QM/MM, la FCO est utilisée pour représenter la densité des électrons de cœur des atomes frontières. Néanmoins, l'introduction de la FCO, de par son caractère gelé, ne permettra pas de corriger l'erreur dû à l'invariance du recouvrement entre les AOs des deux atomes. Une deuxième solution consiste à décrire les électrons de cœur de la même façon que n'importe quels électrons de la partie QM. Ces électrons sont décrits par une orbitale de cœur obtenue de façon auto-cohérente, que nous nommerons SCCO (*Self-Consistent Core Orbital*). Cette approche a l'avantage, par rapport à celle utilisant des FCOs, d'apporter une plus grande flexibilité à la fonction d'onde quantique (pas de contrainte supplémentaire apportée par des orbitales gelées) et d'introduire la variation du recouvrement entre les AOs des atomes X et Y grâce à sa nature variationnelle (tout au moins pour les électrons de cœur. De plus, la détermination de la FCO, de manière équivalente à la SLBO, n'est pas requise. Cependant, l'approche SCCO aura pour défaut de produire une CO plus diffuse que celle utilisée dans le cas de la méthodologie FCO. En effet, la FCO étant déterminée sur un système modèle (atomique ou moléculaire), l'atome sur lequel elle est localisée possède une charge nucléaire identique à celle que l'atome Y posséderait s'il était de nature totalement quantique. En revanche, la SCCO est obtenue de manière auto-cohérente et donc localisée sur l'atome quanto-classique Y qui possède une charge nucléaire moins importante (égale au nombre d'électrons de cœur + 1).

L'approche FCO a déjà été étudiée dans la Ref. [215]. Il a été montré que les FCO sont particulièrement transférables, et que, tout comme les SLBOs, le choix de la méthode de détermination de celles-ci n'a qu'une faible influence sur la qualité des résultats géométriques et énergétiques [215, 235]. Cependant, la détermination de cette orbitale de cœur était nécessaire (FIG. 5.5). Dans cette section, notre but est de montrer que la méthode LSCF/MM, combinée avec le schéma SCCO, est une méthode générale, sans paramètres additionnels, qui peut être transférable et applicable à la plupart des systèmes chimiques et biologiques.

FIG. 5.5 – Principe d'un calcul LSCF à l'aide d'une FCO d'après la Ref. [215].



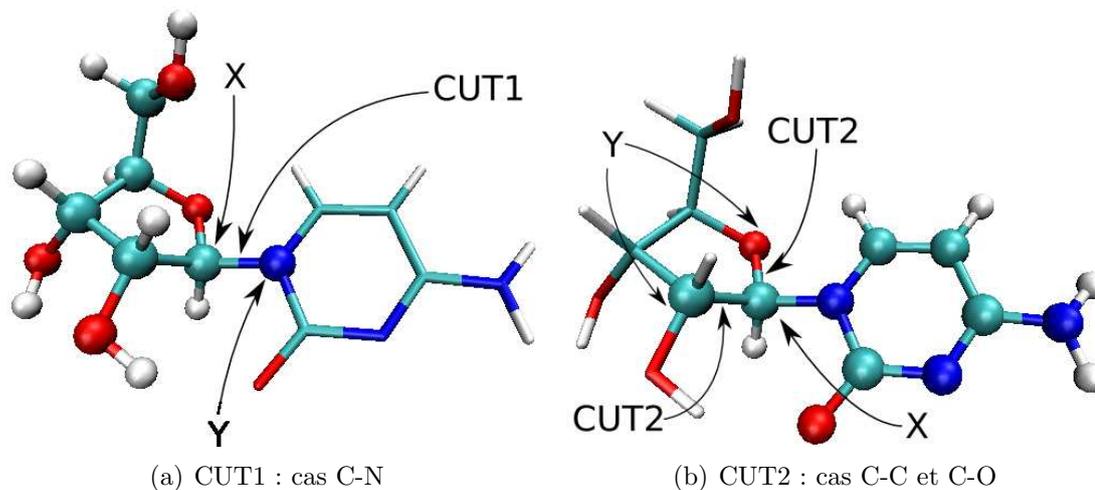
Description et détails des calculs

Dans le but de comparer les schémas SCCO et FCO, nous avons effectué plusieurs séries de calculs. Premièrement, des calculs LSCF servant de référence ont été réalisés sur des petites molécules (éthane, méthylamine et méthanol) possédant les trois liaisons les plus courantes en chimie organique (C-C, C-N et C-O). Les SLBOs ont été déterminées au niveau RHF/6-311G**, conformément au critère WP [260,261], suivi de la suppression des queues d'orthogonalisation-délocalisation. Pour ce qui est de la FCO, elle a été déterminée grâce à la théorie ELMO (*Extremely Localized Molecular Orbital*) [235, 265–269]. Le calcul a été effectué sur les molécules cibles suivant le critère de localisation de la structure de Lewis (voir Ref. [215] pour plus d'informations). Ce point est motivé par la plus grande transférabilité des orbitales localisées venant de la théorie ELMO [235]. Nous ne nous intéresserons pas ici à l'effet de la FCO car il a été montré que la méthode de détermination de la FCO (venant d'un calcul ROHF (*Restricted Open-Shell Hartree-Fock*) atomique ou d'un calcul ELMO dans un environnement moléculaire) conduit à des résultats indiscernables [215].

Afin de vérifier l'applicabilité de la nouvelle méthode sur des systèmes biologiques plus grands, nous avons effectué des calculs QM/MM sur le nucléoside cytidine (FIG. 5.6). Deux localisations de la frontière QM/MM ont été considérées, correspondant aux trois cas étudiés ci-dessus. Le champ de forces AMBER *ff99* [144, 147] a été utilisé, et pour assurer l'électroneutralité de la partie MM, les charges classiques ont été obtenues grâce au protocole RESP [320, 321] (*Restreint Electrostatic Potential*).

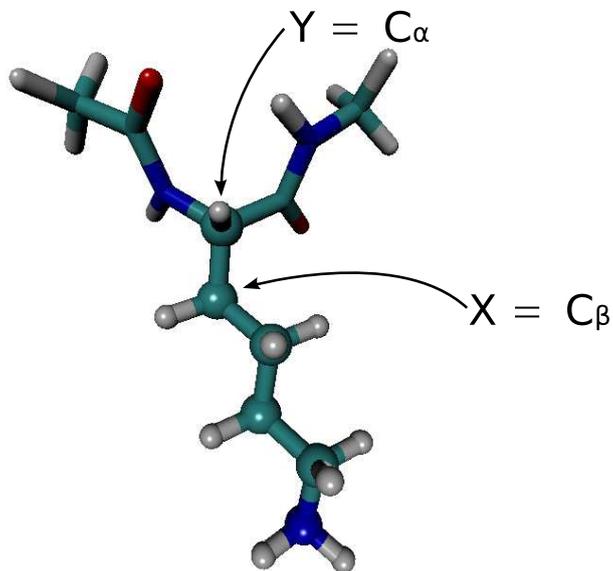
5.2. Prise en compte des électrons de cœur et de valence

FIG. 5.6 – Partition du nucléoside cytidine en partie QM (boules et bâtons) et MM (bâtons fins). Deux localisations de la frontière QM/MM sont considérées.

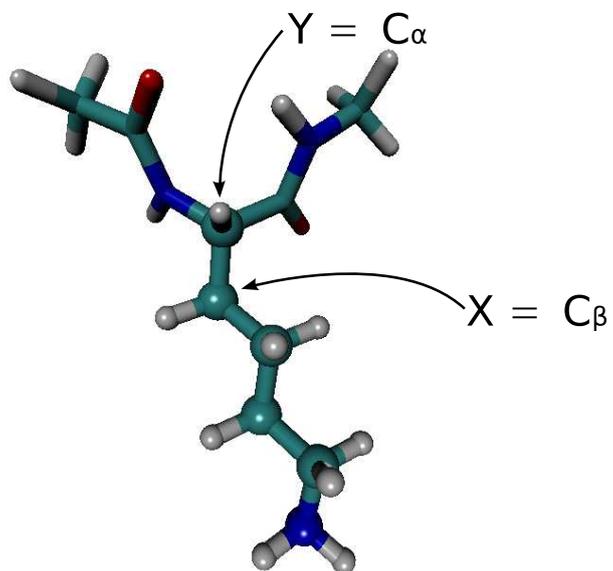


Puis, la précision de la méthode SCCO a été vérifiée sur des valeurs énergétiques. Nous avons effectué les calculs sur les deux états de protonation de la lysine protégée par des groupements acétyle et N-méthylacétamide sur le N-terminal et le C-terminal, respectivement (FIG. 5.7). La frontière QM/MM est localisée entre les carbones C_α et C_β de la lysine. Le champ de forces AMBER *ff99* et le protocole RESP ont aussi été utilisés ici, de la même manière que ci-dessus.

FIG. 5.7 – Partition du tripeptide Ace-Lys-NMe en partie QM (boules et bâtons) et MM (bâtons fins). La frontière QM/MM est localisée entre les atomes C_α et C_β de la lysine.



(a) Forme protonée



(b) Forme neutre

Résultats et Discussion

Molécules organiques Dans ce paragraphe, nous allons étudier l'effet du champ de forces et du critère de localisation de la SLBO sur les paramètres géométriques de molécules organiques de petite taille. Plus particulièrement, nous nous intéresserons à l'influence des charges classiques obtenues de différentes manières, ainsi qu'aux types d'interactions électrostatiques entre la partie QM et MM (EE ou ME).

Afin d'étudier l'influence intrinsèque du champ de forces hybride sur les paramètres géométriques de la frontière QM/MM, nous avons effectué des calculs LSCF sur les molécules d'éthane, de méthylamine et de méthanol. Les frontières ont été placées, respectivement, sur les liaisons C-C, C-N et C-O et, dans un premier temps, les atomes d'hydrogène MM ont été omis afin de fournir un modèle indépendant du champ de forces. Comme chaque atome Y participe à la SLBO par un électron, une charge nucléaire +1 a été assignée à cet atome (dans la suite, nous nous référerons à ces calculs par la notation LSCF/+1). Cependant, lorsque que les schémas FCO et SCCO sont utilisés, les électrons de cœur doivent être pris en compte, et donc, en conséquence, une charge nucléaire +3 est portée par l'atome frontière pour assurer l'électroneutralité du système (calculs LSCF/+3).

Dans la table 5.3, nous reportons les distances d'équilibre et les angles pour les trois molécules modèles au niveau LSCF(WP)/+1, LSCF(WP)/+3 FCO et LSCF(WP)/+3 SCCO (FIG. 5.8). On peut remarquer que le schéma LSCF/+1 donne des déviations raisonnables par rapport aux valeurs RHF pour les angles (RMSD_\angle), mais sous-estime largement la distance C-Y, justifiant le développement du FBP dans la méthodologie LSCF/MM précédente [214] (voir section 4.5). La description explicite des électrons de cœur de l'atome Y réduit la déviation entre les calculs LSCF/+3 et les calculs RHF, sans affecter les autres degrés de liberté. La méthode SCCO, en plus d'éviter la détermination de la FCO, donne des distances de liaison en meilleur accord avec les résultats RHF, comparé au niveau de théorie FCO et ceci quelque soit le type de liaison considéré. Plus précisément, la méthode SCCO réduit la valeur du RMSD_d de 0,016, 0,024 et 0,001 Å pour les molécules d'éthane, de méthylamine et de méthanol, respectivement, par rapport à la méthode FCO. Ce point est principalement dû à l'amélioration des longueurs de liaisons C-Y. Le changement du critère de localisation n'affecte pas les conclusions précédentes. Par exemple, la distance C-C dans la molécule d'éthane est 1,557, 1,554, 1,559 et 1,562 Å pour les critères de Magnasco-Perico [262, 263], de Boys-Foster [255], de Pipek-Mezey [253] et de Weinstein-Pauncz [260, 261], respectivement.

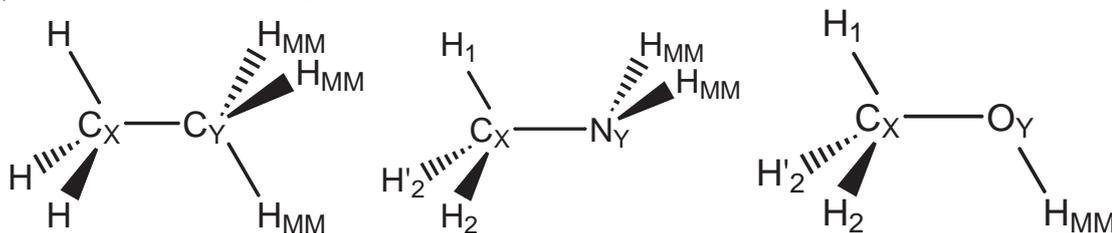
TAB. 5.3 – Sélection de distances (en Å) et d'angles (en degrés) calculés pour CH₃-R (R = CH₃, NH₂, OH) utilisant les niveaux de théorie RHF, LSCF(WP)/+1, LSCF(WP)/+3 FCO and LSCF(WP)/+3 SCCO avec la base 6-311G**. Nous reportons aussi la fréquence de vibration des liaisons X-Y (en cm⁻¹). Pour les calculs LSCF, les déviations par rapport aux valeurs RHF sont reportées entre parenthèses, avec les RMSD correspondant aux distances RMSD_d et aux angles RMSD_∠ (en Å et en degrés) pour la structure du fragment CH₃-Y.

	RHF	LSCF(WP)/+1	LSCF(WP)/+3	
			FCO	SCCO
Ethane (R = CH ₃)				
d(C-C)	1,527	1,118(-0,409)	1,585(0,057)	1,562(0,035)
ν(C-C)	1052	3598(2546)	1414(362)	1229(177)
d(C-H)	1,086	1,104(0,018)	1,094(0,008)	1,094(0,007)
∠(H-C-C)	111,2	114,0(2,8)	114,0(2,8)	114,6(3,4)
∠(H-C-H)	107,7	104,6(-3,1)	104,6(-3,1)	103,9(-3,8)
RMSD _d	—	0,290	0,041	0,025
RMSD _∠	—	3,0	3,0	3,6
Méthylamine (R = NH ₂)				
d(C-N)	1,454	1,030(-0,424)	1,439(-0,015)	1,454(0,000)
ν(C-N)	1137	4698(3561)	1592(456)	1574(437)
d(C-H) ^a	1,088	1,110(0,022)	1,097(0,009)	1,096(0,008)
∠(H ₁ -C-N)	114,6	115,2(0,6)	116,2(1,6)	116,5(1,9)
∠(H ₂ -C-N)	109,4	114,6(5,2)	113,3(3,9)	113,4(4,0)
∠(H ₁ -C-H ₂)	108,0	103,8(-4,2)	104,4(-3,6)	104,3(-3,7)
∠(H ₂ -C-H' ₂)	107,3	103,6(-3,7)	104,1(-3,1)	103,6(-3,7)
RMSD _d	—	0,300	0,030	0,006
RMSD _∠	—	3,8	3,2	3,4
Méthanol (R = OH)				
d(C-O)	1,399	0,947(-0,453)	1,323(-0,077)	1,356(-0,043)
ν(C-O)	1179	6089(4910)	1632(453)	1625(446)
d(C-H) ^a	1,085	1,116(0,031)	1,100(0,015)	1,098(0,013)
∠(H ₁ -C-O)	107,4	113,7(6,3)	111,4(4,1)	110,7(3,3)
∠(H ₂ -C-O)	112,0	116,1(4,1)	115,8(3,8)	115,6(3,6)
∠(H ₁ -C-H ₂)	108,3	103,0(-5,3)	104,1(-4,2)	104,3(-4,0)
∠(H ₂ -C-H' ₂)	108,7	103,2(-5,5)	104,3(-4,4)	105,1(-3,6)
RMSD _d	—	0,321	0,033	0,032
RMSD _∠	—	5,3	4,1	3,6

^aMoyenne sur le groupement méthyle

5.2. Prise en compte des électrons de cœur et de valence

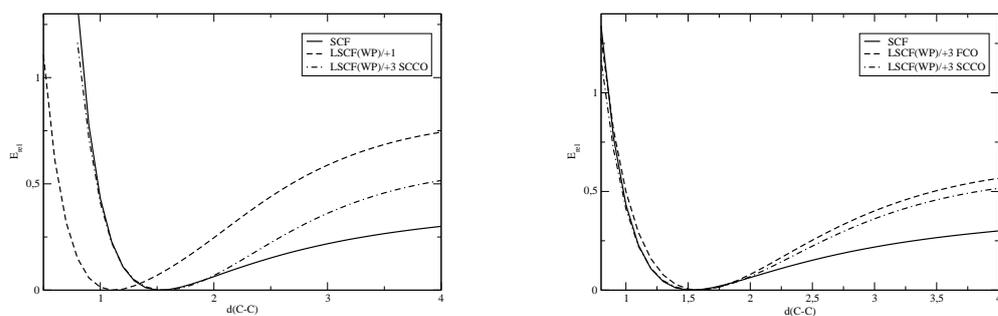
FIG. 5.8 – Numérotation des atomes et partition des fragments $\text{CH}_3\text{-R}$ ($\text{R} = \text{CH}_3, \text{NH}_2, \text{OH}$). Les atomes d'hydrogène notés H_{MM} correspondent à ceux de la partie MM.



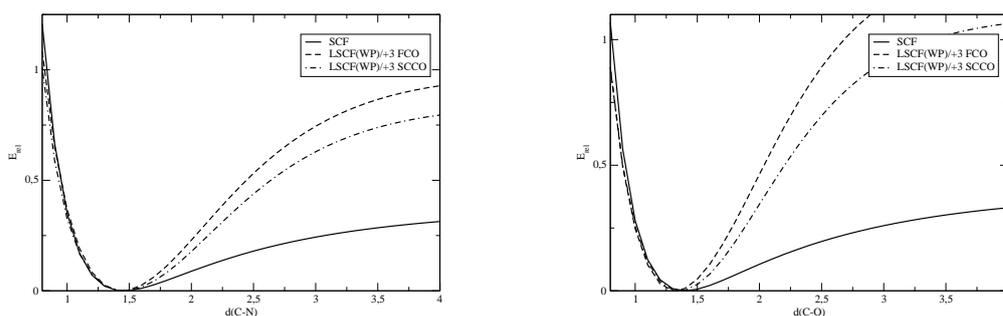
Sur la figure 5.9, nous avons reporté les profils énergétiques de la molécule d'éthane, de méthylamine et de méthanol obtenu par un étirement rigide le long de la liaison C-Y ($\text{Y} = \text{C}, \text{N}, \text{O}$) au niveau de théorie RHF, FCO et SCCO. Les autres degrés de liberté sont maintenus à leurs valeurs d'équilibre. Comme nous pouvons le voir, la méthodologie LSCF/+3 améliore de façon très significative la position des minima de la courbe d'énergie potentielle, et présente un profil énergétique similaire pour les faibles valeurs de la distance interatomique. Aux plus grandes distances interatomiques, le niveau de théorie LSCF/+3 présente un comportement différent, comparé aux calculs RHF, dû au fait que l'on néglige le réajustement de la SLBO. D'un point de vue géométrique, les premiers résultats prenant en compte le réajustement de la SLBO sont encourageants [237]. Une autre source d'erreurs peut être attribuée à la charge nucléaire de l'atome frontière Y. En effet, sa charge nucléaire reste égale à +3 quelque soit la nature de l'atome (C, N ou O). Ceci entraîne une déviation systématique entre les profils de dissociation purement quantique et QM/MM pour les distances de liaison plus grandes (supérieure à 2 Å). Ces déviations augmentent d'ailleurs en allant de C à O, conséquence de l'augmentation de la charge nucléaire de l'atome Y dans les calculs quantiques de référence. Cependant, le comportement aux fortes distances ne pose pas de problèmes au niveau pratique : lors d'un calcul QM/MM, la liaison frontière n'est pas censée se dissocier ou bien même évoluer dans cette gamme de longueur. En revanche, ceci peut être une source d'erreurs dans le cas d'une dynamique moléculaire (MD pour *Molecular Dynamics*), où les atomes vibrent autour de leurs positions d'équilibre. Afin d'évaluer l'erreur commise par l'introduction de l'atome frontière Y, nous avons calculé les fréquences de vibration des liaisons C-C, C-N et C-O à leurs positions d'équilibre et dans l'approximation harmonique (TAB. 5.3). Nous remarquons que les résultats fournis par la méthodologie LSCF/+1 sont très éloignés des valeurs RHF à cause du fort caractère C-H de la liaison frontière. Néanmoins, les deux approches LSCF/+3 permettent de diminuer significativement ces écarts, avec une déviation maximale de 362, 456 et 453 cm^{-1} pour les liaisons C-C, C-N et C-O, respectivement. De plus, la méthodologie SCCO assure une diminution systématique de ces erreurs par rapport à la méthode FCO. Dans le cas de l'éthane, l'approche SCCO réduit d'un facteur deux la déviation entre les calculs QM/MM et *ab initio* avec une différence de 177 cm^{-1} . Enfin, on peut noter que ces déviations augmentent en allant du carbone à l'oxygène, et ceci quelque soit la méthodologie QM/MM utilisée. Ce point s'explique par l'augmentation de la charge nucléaire de ces atomes qui n'est pas reproduit dans les calculs QM/MM, l'atome Y ayant, quelque soit la nature de l'atome, une charge nucléaire

égale à +3. Cependant, ces écarts restent acceptables dans l'optique de MD.

FIG. 5.9 – Profil énergétique (en u.a.) le long des liaisons C-C, C-N et C-O (en Å) pour les molécules d'éthane, de méthylamine et de méthanol au niveau de théorie RHF, et pour le fragment $\text{CH}_3\text{-R}$, au niveau LSCF(WP)/+3 FCO et SCCO avec la base 6-311G**.



(a) Profil énergétique de la liaison C-C



(b) Profil énergétique de la liaison C-N

(c) Profil énergétique de la liaison C-O

Pour déterminer l'influence des atomes d'hydrogène MM du fragment R (FIG. 5.8), nous avons reporté les distances d'équilibre et les angles dans le cas d'une interaction électrostatique prise au niveau *mechanical embedding* (ME) (champ de forces MM3 [163–170]) et avec *electronic embedding* (EE) (champ de forces OPLS [154–160]). Comme reporté dans la table 5.4, l'introduction de l'environnement MM n'influence pas les conclusions du paragraphe précédent, et ceci quelque soit le type d'interaction. Cependant, nous voudrions souligner la forte influence de la charge classique de l'atome Y. Par exemple, dans le cas de la molécule de méthanol, le fait de prendre en compte la charge classique de l'atome MM, induit une diminution de 0,034 Å de la distance C-O, comparé au cas ME. Comme les charges natives du champ de forces OPLS ne préservent pas l'électronéutralité (excepté pour la molécule d'éthane par raison de symétrie), nous reportons les mêmes calculs avec les charges classiques obtenues par la procédure RESP. Les résultats montrent de légères modifications des distances d'équilibre et des angles. Nous pouvons remarquer que l'utilisation des charges RESP pour ces trois molécules permet de diminuer l'écart

5.2. Prise en compte des électrons de cœur et de valence

entre le niveau de théorie RHF et LSCF/+3/MM SCCO. Nous concluons que le schéma LSCF/+3/MM SCCO permet d'obtenir une bonne description géométrique de la frontière QM/MM, et ceci de manière indépendante du champ de forces (ME ou EE) et des charges classiques utilisées. Nous allons maintenant nous intéresser à des systèmes de tailles plus importantes.

Nucléoside cytidine Afin de vérifier la validité de l'approche SCCO sur des systèmes de taille plus importante et d'étudier la variation des paramètres géométriques, nous allons considérer la molécule de cytidine. La partition QM/MM est décrite sur la figure 5.6. Pour le premier cas (CUT1 : FIG. 5.6), la frontière QM/MM a été placée au niveau de la liaison C-N connectant le cycle ribose et la base nucléotidique cytosine. Le cycle ribose est alors traité au niveau QM. Le second cas (CUT2 : FIG. 5.6) correspond à un calcul QM/MM comportant deux coupures localisées sur les liaisons C-C et C-O de la partie ribose. Nous traitons alors la base nucléotidique au niveau QM, et le reste au niveau MM. Pour ces trois liaisons (C-C, C-N et C-O), la SLBO a été déterminée grâce au critère WP sur les systèmes modèles (éthane, méthylamine et méthanol) et appliquée à ce système en accord avec le principe de transférabilité.

Les résultats concernant le nucléoside cytidine sont regroupés dans la table 5.5. Nous pourrions faire les commentaires suivants : dans les deux cas, l'approche LSCF(WP)/+3/MM SCCO donne des distances de liaison raisonnables, comparées aux calculs purement quantiques, avec une déviation maximale de 0,050 Å. Cette déviation maximale est trouvée dans le cas difficile où l'atome X est connecté à deux atomes Y (CUT2).

Pour conclure sur cette première partie des résultats concernant les paramètres structuraux, nous pouvons remarquer que le traitement LSCF/+3/MM SCCO permet d'obtenir une précision identique à celle obtenue dans le cas de l'approche LSCF/+1/MM FBP [214] et LSCF/+3/MM FCO [215]. Nous allons donc maintenant nous tourner vers l'étude des performances de cette méthodologie pour des valeurs énergétiques, comme les affinités protoniques.

Tripeptide Ace-Lys-NMe Dans ce paragraphe, nous allons examiner les performances générales de l'approche SCCO par rapport à des valeurs énergétiques. Comme les affinités protoniques sont très sensibles à l'environnement et très délicates à décrire correctement [199], nous avons décidé d'étudier la forme protonée et neutre du tripeptide Ace-Lys-NMe. Nous avons considéré ce système dans une conformation présentant une chaîne latérale complètement étendue. Lors des calculs QM/MM, nous n'avons pas introduit de contrainte sur la position des atomes. La frontière QM/MM est localisée entre les atomes C_α et C_β (FIG. 5.7), et la chaîne latérale de la lysine est traitée au niveau QM. Pour les calculs LSCF/+1/MM, le FBP correspondant à une liaison $C(sp^3)-C(sp^3)$ a été utilisé [214]. La SLBO et la FCO ont été déterminées sur la molécule d'éthane grâce à la théorie ELMO, suivant la même méthodologie que précédemment (voir 5.2.1.0). Le RMSD de la partie QM (resp. MM), $RMSD_{QM}$ ($RMSD_{MM}$ resp.), et le RMSD total du système ont été calculés suivant la méthode de Kabsch [322] implantée dans le programme VMD [323]. Afin d'obtenir des valeurs de référence, nous avons aussi traité le système entier au niveau RHF/6-311G**.

Chapitre 5. Description Avancée de la Frontière QM/MM

TAB. 5.4 – Charges classiques (en électrons) du champ de forces correspondant, sélection de distances (en Å) et d'angles (en degrés) calculés pour les molécules CH₃-R (R = CH₃, NH₂, OH) en utilisant le niveau de théorie LSCF(WP)/+3/MM SCCO et la base 6-311G**. Les déviations par rapport aux valeurs RHF sont reportées entre parenthèses, avec les RMSD correspondants aux distances RMSD_d et aux angles RMSD_∠ (en Å et en degrés) pour la structure des fragments CH₃-Y.

Champ de forces	MM3	OPLS	
Charges classiques	—	OPLS	RESP
Ethane (R = CH ₃)			
q(C)	—	-0,1800	-0,0189
q(H)	—	0,0600	0,0063
d(C-C)	1,567(0,040)	1,589(0,042)	1,565(0,038)
d(C-H)	1,095(0,009)	1,095(0,009)	1,094(0,008)
∠(H-C-C)	114,8(3,6)	114,6(3,4)	114,6(3,4)
∠(H-C-H)	103,6(-4,1)	103,9(-3,8)	103,9(-3,8)
RMSD _d	0,029	0,030	0,027
RMSD _∠	3,9	3,6	3,6
Méthylamine (R = NH ₂)			
q(N)	—	-0,9000	-0,4446
q(H)	—	0,3500	0,2223
d(C-N)	1,462(0,008)	1,465(0,011)	1,482(0,028)
d(C-H) ^a	1,097(0,009)	1,105(0,017)	1,100(0,012)
∠(H ₁ -C-N)	116,8(2,2)	117,1(2,5)	117,2(2,6)
∠(H ₂ -C-N)	113,8(4,4)	113,6(4,2)	113,4(4,0)
∠(H ₁ -C-H ₂)	104,1(-3,9)	103,9(-4,1)	103,9(-4,1)
∠(H ₂ -C-H ₂)	103,5(-3,8)	103,2(-4,1)	103,5(-3,8)
RMSD _d	0,009	0,014	0,022
RMSD _∠	3,7	3,8	3,7
Méthanol (R = OH)			
q(O)	—	-0,6830	-0,3177
q(H)	—	0,4180	0,3177
d(C-O)	1,358(-0,041)	1,325(-0,074)	1,354(-0,045)
d(C-H) ^a	1,099(0,014)	1,107(0,022)	1,101(0,016)
∠(H ₁ -C-O)	110,4(3,0)	111,4(4,0)	110,9(3,5)
∠(H ₂ -C-O)	116,0(4,0)	116,2(4,2)	116,0(4,0)
∠(H ₁ -C-H ₂)	103,9(-4,4)	103,7(-4,6)	103,9(-4,4)
∠(H ₂ -C-H ₂)	105,2(-3,5)	104,2(-4,5)	104,6(-4,1)
RMSD _d	0,031	0,055	0,034
RMSD _∠	3,8	4,3	4,0

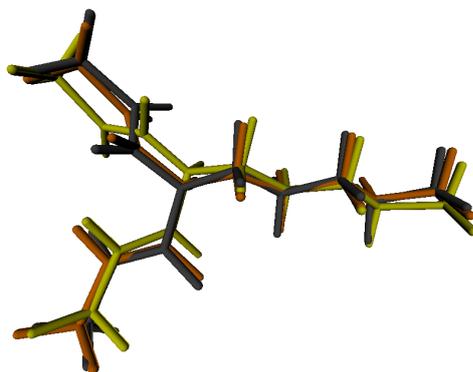
^aMoyenne sur le groupement méthyle

5.2. Prise en compte des électrons de cœur et de valence

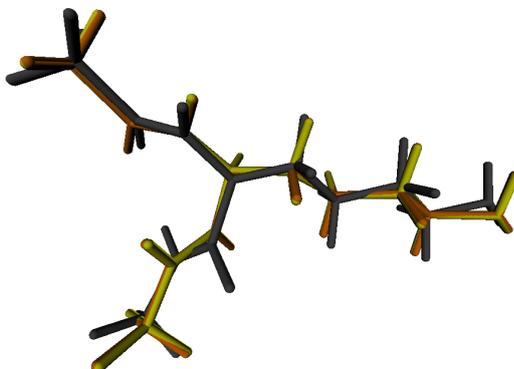
TAB. 5.5 – Distance d'équilibre X-Y de la molécule de cytidine au niveau RHF et LSCF(WP)/+3/MM SCCO avec la base 6-311G**. Deux localisations de la frontière QM/MM sont considérées, comme décrit sur la figure 5.6. Le champ de forces AMBER *ff99* a été utilisé et les charges classiques ont été déterminées grâce au protocole RESP. Les déviations par rapport aux valeurs RHF sont reportées entre parenthèses.

	RHF	LSCF(WP)/+3 SCCO
CUT1		
d(C-N)	1,473	1,488(0,015)
CUT2		
d(C-O)	1,389	1,369(-0,020)
d(C-C)	1,538	1,588(0,050)

FIG. 5.10 – Superposition des structures de la forme neutre et protonée du tripeptide Ace-Lys-NMe au niveau de théorie RHF (noir), LSCF(ELMO)/+3/MM FCO (jaune) et LSCF(ELMO)/+3/MM SCCO (orange).



(a) Forme protonée



(b) Forme neutre

Chapitre 5. Description Avancée de la Frontière QM/MM

Les résultats sont regroupés dans la table 5.6. Les trois méthodes donnent des résultats raisonnables pour les distances de liaison correspondant aux atomes frontières et des valeurs de RMSD_{QM} inférieures à 0,044 Å. La principale déviation entre les calculs purement quantiques et les structures QM/MM vient de la partie MM ($\text{RMSD} \simeq \text{RMSD}_{\text{MM}}$). Dans le cas de la forme protonée de la lysine, bien que la distance $C_\alpha\text{-}C_\beta$ obtenue avec la méthode LSCF(ELMO)/+3/MM SCCO soit la moins précise, toutes les méthodologies QM/MM donnent des valeurs similaires pour les RMSD, et donc des structures quasi-identiques (FIG. 5.10). Pour la forme neutre, la méthodologie LSCF/+3/MM ne permet pas d'obtenir des valeurs de RMSD de la même précision que celles obtenues dans le cas LSCF/+1/MM. Cependant, les méthodes FCO et SCCO permettent de réduire la déviation correspondant aux valeurs de ΔE (différence d'énergie entre la forme neutre et la forme protonée) de 1,4 et 1,0 kcal/mol, respectivement. Nous concluons que l'introduction de la description des électrons de cœur de l'atome Y par une FCO ou une SCCO n'affecte pas de façon significative les géométries de la partie QM et/ou de la partie MM, comparées à la méthodologie LSCF/+1/MM FBP. Elles permettent, aussi, de réduire les déviations énergétiques par rapport aux calculs totalement quantiques. Comme attendu, la prise en compte des électrons de cœur de l'atome Y par une FCO ou une SCCO conduit à des résultats similaires. Néanmoins, la méthodologie SCCO permet d'éviter la détermination de la FCO.

TAB. 5.6 – Données énergétiques et géométriques de la forme protonée (P) et neutre (N) de la lysine au niveau de théorie RHF, LSCF(ELMO)/+1/MM FBP, LSCF(ELMO)/+3/MM FCO et LSCF(ELMO)/+3/MM SCCO avec la base 6-311G**. Le champ de forces AMBER *ff99* est utilisé pour les calculs QM/MM et les charges de la partie MM sont déterminées grâce à la procédure RESP. Les distances et les valeurs de RMSD sont en Å, tandis que les énergies relatives $\Delta E = E_{\text{tot}}(\text{N}) - E_{\text{tot}}(\text{P})$ sont en kcal/mol.

	RHF	LSCF(ELMO)/+1/MM FBP	LSCF(ELMO)/+3/MM	
			FCO	SCCO
Forme protonée				
$d(C_\alpha\text{-}C_\beta)$	1,536	1,519(-0,017)	1,573(0,037)	1,589(0,053)
RMSD_{QM}	—	0,029	0,038	0,044
RMSD_{MM}	—	0,367	0,377	0,359
RMSD	—	0,374	0,393	0,362
Forme neutre				
$d(C_\alpha\text{-}C_\beta)$	1,526	1,520(-0,006)	1,573(0,047)	1,588(0,062)
RMSD_{QM}	—	0,028	0,035	0,037
RMSD_{MM}	—	0,571	1,154	1,140
RMSD	—	0,532	1,031	1,021
ΔE	235,4	239,0(3,6)	237,6(2,2)	238,0(2,6)

Conclusions

Dans cette section, nous avons introduit le concept d'orbitales de cœur auto-cohérentes [239] (SCCO) comme alternative à l'étude précédente utilisant des orbitales de cœur gelées [215] (FCO), dans le cadre de la méthode hybride LSCF/MM. La nouvelle méthodologie permet d'éviter la paramétrisation abusive de la frontière QM/MM et la détermination de la FCO. Nous avons montré que l'introduction explicite des électrons de cœur des atomes frontières par une SCCO conduit à des résultats de qualité comparable à ceux obtenus suivant l'approche FCO. Comparée aux calculs RHF, l'approche SCCO induit de faibles déviations géométriques et énergétiques sur, à la fois, les systèmes modèles et les molécules biologiques de taille plus importante (comme les polypeptides et les nucléosides). De plus, nous avons mis en évidence que plusieurs types de liaison, avec différentes polarités (C-C, C-N et C-O), peuvent être considérées, et correctement décrites. Ce dernier point est très important, car il n'est pas toujours possible de couper une liaison C-C, afin de diviser une macromolécule. Conformément à ces conclusions, dans le cadre des calculs LSCF/MM, le seul paramètre défini par l'utilisateur est la SLBO. Nous allons voir, dans la section suivante, que la méthodologie LSCF/MM permet aussi une partition du système au niveau de la liaison peptidique.

5.2.2 Electrons de valence : la liaison peptidique

Introduction

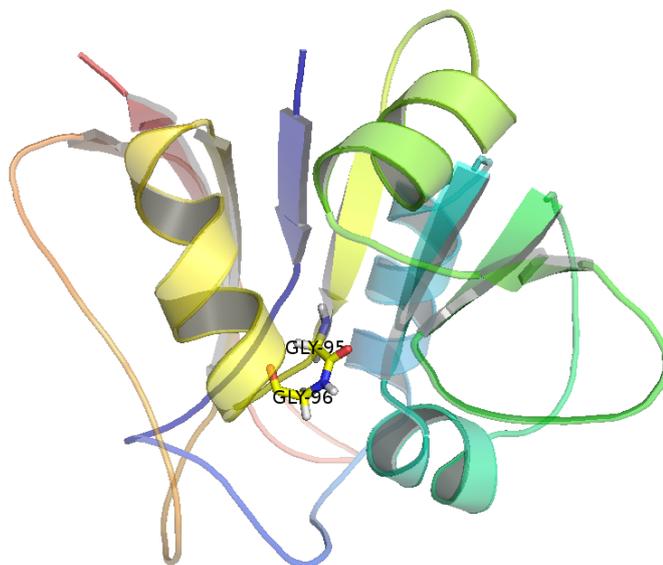
Dans la section précédente, nous avons montré que la méthode LSCF/MM peut être appliquée à différents types de liaisons (C-C, C-N et C-O). Bien que les systèmes envisagés ci-dessus présentent des liaisons de polarité différentes, nous n'avons envisagé que le cas de liaisons simples. Si nous nous intéressons plus particulièrement aux protéines et aux systèmes enzymatiques, nous rencontrons des liaisons à caractère d'ordre variable. La plus commune et la plus intéressante est, sans aucun doute, la liaison peptidique. Celle-ci est une liaison covalente entre un atome de carbone et d'azote qui assure la jonction entre deux acides aminés. Elle se forme par réaction de condensation du groupement carboxylique de l'un des acides aminés avec le groupement amine de l'autre, formant un groupement amide après élimination d'une molécule d'eau. Lors de l'étude d'un système biologique à l'aide d'une méthode hybride QM/MM, la première question que l'on se pose est de savoir où placer la ou les jonction(s) entre les deux parties QM et MM. En effet, elles doivent être placées de manière à ne pas introduire d'artefact sur la description de la partie QM, mais aussi de façon à limiter sa taille afin d'utiliser les outils théoriques les plus poussés lors de son étude.

De façon usuelle, la frontière QM/MM est placée entre deux atomes de carbones formant une liaison simple (cas de la liaison C_α - C_β [215, 241] ou C_α -C(=O) [236] dans les acides aminés). Ceci est motivé par le fait que la polarité de la liaison est faible, et très proche de celle d'une liaison C-H (utilisée par le schéma LA). Par exemple, lorsque nous avons étudié l'affinité protonique du tripeptide Ace-Lys-NMe (section 5.2.1), nous avons placé la frontière au niveau de la liaison C_α - C_β de la lysine. Comme nous le verrons dans le chapitre 7, lors de la détermination des spectres d'absorption UV-vis de polypeptides fonctionnalisés sur leur chaîne latérale par un groupement azobenzène, la frontière QM/MM

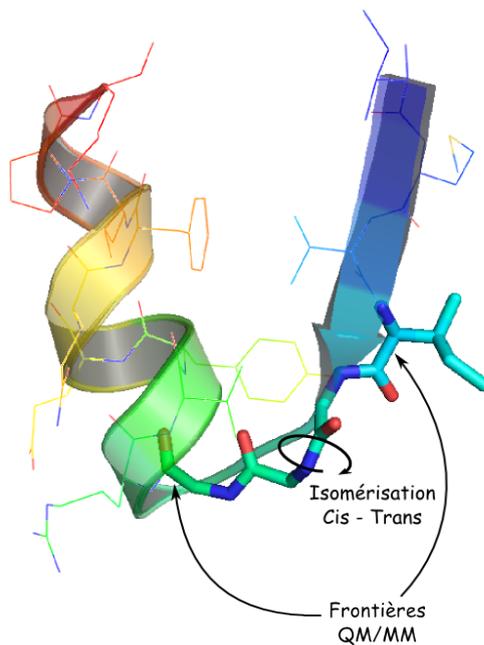
est aussi localisée entre les atomes C_α - C_β de l'acide aminé fonctionnalisé (FIG. 7.5). Dans ces cas, le squelette du polypeptide est traité au niveau MM, tandis que la chaîne latérale est traitée au niveau QM. L'étude de la Ref. [236] relève d'une situation où cette partition n'est pas possible. En effet, elle porte sur l'isomérisation *cis-trans* de la liaison peptidique entre deux glycines dans un fragment de 16 acides aminés (Ile 91 à Lys 106) de la dihydrofolate réductase (DHFR) de l'*Escherichia Coli* (entrée PDB 1DRE) [324]. Cette enzyme a la particularité de présenter une liaison peptidique *cis* entre les glycines 95 et 96 [325,326] (FIG. 5.11) Elle requiert donc d'inclure une partie du squelette du polypeptide dans la partie QM. Il en serait de même pour l'étude de la réaction d'hydrolyse de la liaison peptidique [327–330]. Cependant, dans certain cas, ces partitions peuvent être une limitation. Premièrement, la partie QM peut être trop petite (cas de la coupure C_α - C_β lors de l'étude d'un acide aminé ayant une chaîne latérale de petite taille, comme l'acide aspartique par exemple). Deuxièmement, elle peut être asymétrique (cas de la coupure C_α -C(=O)). Ces deux points peuvent induire une perturbation assez importante sur le processus chimique en cours d'étude. Deuxièmement, ces partitions peuvent impliquer une charge totale de la partie MM ne correspondant pas à un entier. Dans ce cas, les charges sont alors modifiées de façon plus ou moins arbitraire ou ajustées afin de préserver l'électroneutralité du système global (par la procédure RESP, par exemple). L'ensemble de ces situations est représenté sur la figure 5.12.

5.2. Prise en compte des électrons de cœur et de valence

FIG. 5.11 – Dihydrofolate réductase (DHFR) de l'*Escherichia Coli* (entrée PDB 1DRE)

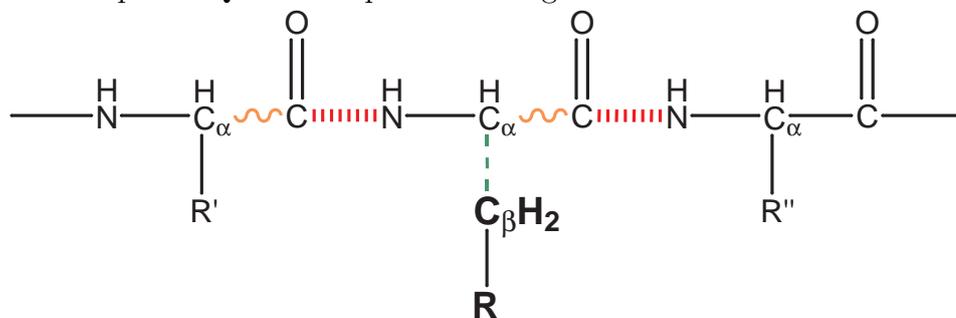


(a) Système entier : la liaison peptidique *cis* est localisée entre les résidus GLY-95 et GLY-96



(b) Système modèle (Ile 91 à Lys 106) : localisation des frontières QM/MM de type $C_{\alpha}-C(=O)$

FIG. 5.12 – Partition schématique d'un polypeptide. Les différentes localisations de la jonction QM/MM sont représentées : le cas de la liaison $C_\alpha-C_\beta$ est représenté en vert par une ligne discontinue, les liaisons $C_\alpha-C(=O)$ par un trait ondulé en orange, tandis que les liaisons peptidiques sont en bâtonnets rouges. Les atomes de la chaîne latérale appartenant à la partie QM sont représentés en gras.



Un choix plus naturel, mais beaucoup plus risqué, serait de délimiter, dans un système peptidique, les parties QM et MM au niveau de la liaison peptidique des acides aminés. Cependant, le caractère partiel de double liaison de cette liaison C-N ne peut pas être pris en compte par un schéma LA. D'ailleurs, Ferré et Olivucci ont reporté des instabilités de la fonction d'onde électronique dans ce cas [331]. Cependant, ce schéma de partition pourrait résoudre les contraintes précédentes. Dans cette section, nous allons montrer que la méthode LSCF permet de délimiter le système aux niveaux des liaisons peptidiques, sans paramètre additionnel (ni FBP [214], ni FCO [215]).

Comme dans le cas précédent, l'atome frontière Y requiert une attention spéciale, de par son caractère quanta-classique. En effet, cet atome participe aux interactions avec les atomes MM (liées et non-liées), mais aussi, comme il possède des fonctions de base, une charge nucléaire et participe à la liaison avec l'atome X en partageant des électrons. Il peut donc être aussi considéré comme un atome de la partie QM. Dans la plupart des cas, la liaison frontière est covalente ; chaque atome Y participe à la liaison frontière en amenant un électron. Ces atomes ont donc une charge nucléaire +1, que l'on additionne à sa charge classique. Notons que dans le cas d'une liaison dative, par exemple, l'atome Y peut participer par l'apport de 2 ou 0 électrons, selon qu'il soit donneur ou accepteur. Sa charge nucléaire est alors fixée à +2 ou 0, respectivement. Enfin, comme nous venons de le voir, si les électrons de cœur sont décrits explicitement, la charge nucléaire de Y est modifiée en conséquence. Dans ce cas, le FBP est nécessaire pour aboutir à une distance de liaison frontière correcte.

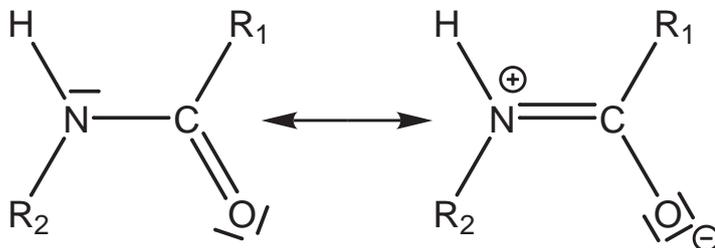
Partition

La méthode LSCF donne généralement de bons résultats lorsque la liaison frontière peut être représentée de façon correcte par une liaison localisée (dans le cas des liaisons C-C par exemple). *A contrario*, la liaison peptidique ne saurait être considérée comme locale. En effet, la liaison C-N possède un certain caractère de liaison double partiel, dû au transfert de charge de l'atome d'azote vers celui d'oxygène (FIG. 5.13). La résonance

5.2. Prise en compte des électrons de cœur et de valence

entre ces deux formes est responsable de la géométrie quasi-plane de la liaison peptidique (par conjugaison du système).

FIG. 5.13 – Résonance entre les deux structures de Lewis de la liaison peptidique.



Comme le nombre d'électrons entre les atomes C et N ne peut pas être modélisé par un nombre entier (supérieur à deux), et parce que ce nombre peut fluctuer en fonction de l'environnement (le poids de la forme zwitterionique peut, par exemple, augmenter si l'atome d'oxygène est engagé dans une liaison hydrogène), il est impossible de représenter cette liaison à partir d'une densité électronique gelée. Deux possibilités peuvent être considérées selon que la nature de l'atome frontière soit celui de carbone ou celui d'azote. Si l'atome frontière est l'atome de carbone, l'atome d'azote est alors traité au niveau quantique et peut partiellement délocaliser sa paire libre. En effet, plusieurs fonctions de base sont encore disponibles sur l'atome de carbone (certaines sont utilisées pour construire la SLBO), et il peut donc accepter la paire libre de l'azote. *A contrario*, si l'atome d'azote correspond à l'atome frontière, ce transfert de charge n'est pas possible car l'atome d'azote ne possède qu'un électron, qui est engagé dans la SLBO. Il serait donc intéressant de permettre une description de ses électrons de valence.

Dans la section précédente, nous avons montré que les électrons de cœur de l'atome Y pouvaient être pris en compte au niveau QM, dans le cadre d'une FCO, ou traités de manière équivalente à l'ensemble des autres électrons (SCCO). La paramétrisation de la frontière n'est alors plus nécessaire afin d'obtenir des distances de liaison correctes. La charge nucléaire de l'atome Y était alors augmentée de deux unités. Dans le cas considéré ici, les atomes de carbone et d'azote auront alors une charge de +3 [239]. Cependant, afin de permettre la description de la paire libre de l'azote quand ce dernier est l'atome frontière, nous allons donc, ici, considérer que la charge nucléaire est de +5, pour permettre l'introduction des deux électrons de valence représentant le doublet [240].

Afin d'illustrer la performance de cette méthode dans le cas usuel de la coupure $C_\alpha-C_\beta$ et dans le cas d'un coupure de la liaison peptidique, nous avons décidé d'étudier l'affinité protonique de l'histidine engagée dans un pentapeptide.

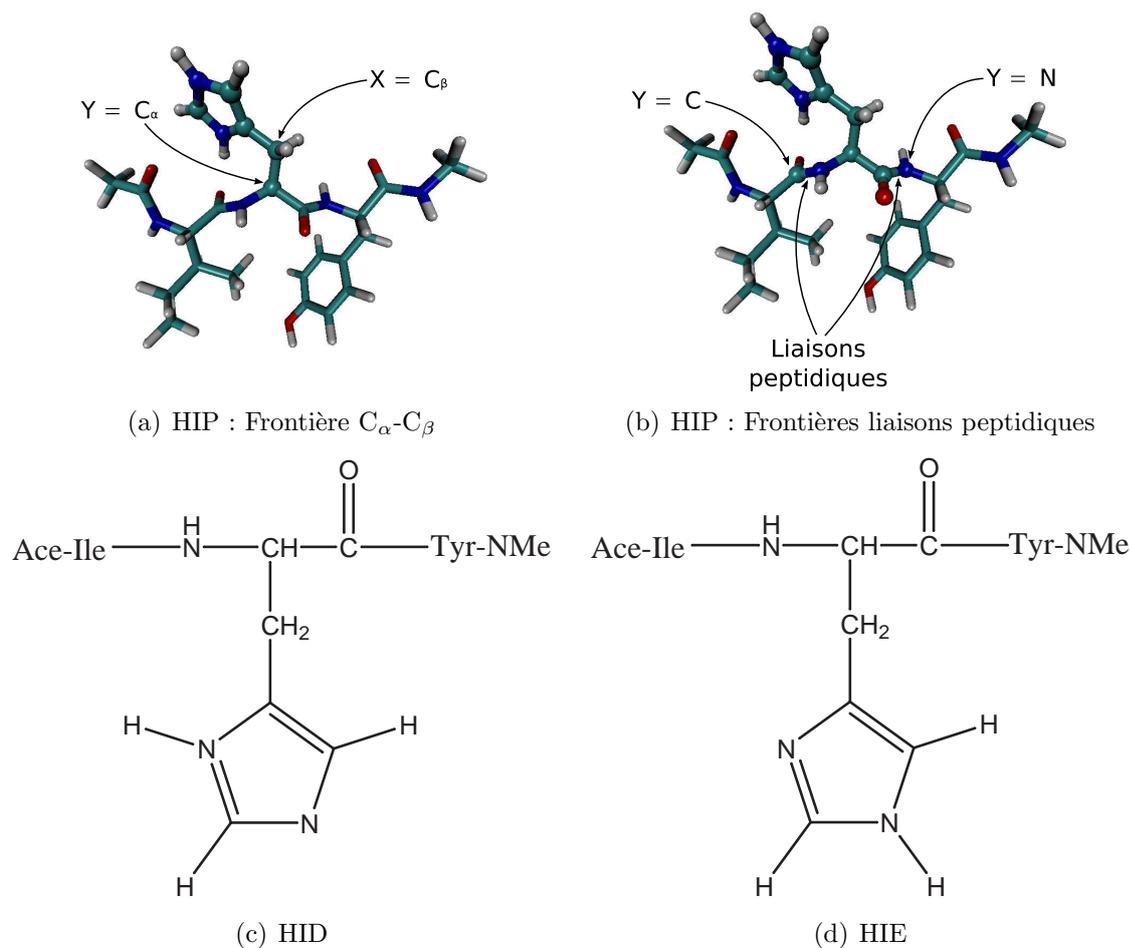
Calculs illustratifs : pentapeptide Ace-Ile-His-Tyr-NMe

Notre but est de comparer, par rapport à des calculs purement quantiques, plusieurs méthodologies QM/MM obtenues pour différentes localisations de la coupure QM/MM. Nous ne présenterons donc que des résultats obtenus sur des molécules dont la taille permet un traitement purement quantique mais avec une signification au niveau QM/MM. Nous

Chapitre 5. Description Avancée de la Frontière QM/MM

avons donc choisi un pentapeptide de séquence Ace-Ile-His-Tyr-NMe. Afin d'incorporer une base assez grande pour ne pas introduire d'artefact dûs à une base trop petite, nous avons fixé le niveau de théorie à RHF/6-311G**.

FIG. 5.14 – Partition QM/MM du pentapeptide Ace-Ile-His-Tyr-NMe en partie QM (boules et bâtons) et partie MM (bâtons fins). Les deux formes tautomères de l'acide aminé histidine (HID et HIE) sont présentées. La forme protonée est notée HIP.



Lorsque la frontière est située entre les atomes C_α et C_β de l'acide aminé histidine, le champ de forces CHARMM27 a été utilisé car il permet, par définition, de conserver une charge électrique entière pour la partie MM. Cependant, lorsque la frontière est située sur les deux liaisons peptidiques de l'histidine, nous utiliserons le champ de force AMBER *ff99*, car celui-ci est l'un des plus utilisés et qu'il permet aussi d'obtenir une charge entière pour la partie MM dans ce cas. Pour ce qui est de la SLBO, elle a été déterminée grâce au critère de localisation de WP (sur la molécule d'éthane dans le cas C_α - C_β , et sur la molécule de N-méthylacétamide dans le cas de la liaison peptidique).

Plusieurs types de calculs ont été envisagés. Tout d'abord, nous avons effectué des calculs purement quantiques au niveau RHF/6-311G**, afin d'obtenir une référence. Dans le

5.2. Prise en compte des électrons de cœur et de valence

cas de la coupure $C_\alpha-C_\beta$, deux méthodologies sont proposées. La première correspond au cas où l'atome frontière porte une charge +1 et qu'un FBP est utilisé [214]. Nous référerons par la suite à ces calculs par l'acronyme LSCF/+1/MM. Le second type de calcul correspond au cas où l'atome frontière porte une charge +3, et est référé par LSCF/+3/MM. Dans le cas de la frontière suivant la liaison peptidique, en plus des calculs LSCF/+1/MM et LSCF/+3/MM, nous avons ajouté la méthodologie LSCF/+5/MM correspondant à une charge +5 sur l'atome d'azote. Ceci revient à ajouter deux électrons de valence de l'atome d'azote au système QM quand ce dernier est considéré comme atome frontière.

TAB. 5.7 – **Frontière $C_\alpha-C_\beta$** . Distances des liaisons frontières (en Å) des trois états de protonation de l'histidine (HID, HIE et HIP, FIG. 5.14), engagée dans un pentapeptide au niveau de théorie RHF/6-311G**, suivant les méthodologies LSCF/+1/CHARMM27 et LSCF/+3/CHARMM27. Les RMSD par rapport à la géométrie de référence RHF/6-311G** sont donnés (en Å) pour le fragment QM (RMSD_{QM}), la partie MM (RMSD_{MM}), et le système entier (RMSD). Nous donnons aussi les affinités protoniques (ΔE en kcal/mol de l'acide aminé histidine, calculées aux mêmes niveaux de théorie ($\Delta E(\text{HID}) = E(\text{HID}) - E(\text{HIP})$ et $\Delta E(\text{HIE}) = E(\text{HIE}) - E(\text{HIP})$). Les déviations par rapport aux valeurs RHF sont reportées entre parenthèses.

	RHF	LSCF/+1/CHARMM27	LSCF/+3/CHARMM27
HIP			
d($C_\alpha-C_\beta$)	1,546	1,543(-0,003)	1,571(0,025)
RMSD _{QM}		0,055	0,186
RMSD _{MM}		0,857	0,861
RMSD		0,774	0,788
HID			
d($C_\alpha-C_\beta$)	1,543	1,541(-0,002)	1,562(0,019)
RMSD _{QM}		0,051	0,104
RMSD _{MM}		0,422	0,440
RMSD		0,390	0,405
HIE			
d($C_\alpha-C_\beta$)	1,532	1,536(0,004)	1,560(0,028)
RMSD _{QM}		0,037	0,218
RMSD _{MM}		0,951	0,960
RMSD		1,009	1,072
$\Delta E(\text{HID})$	257,5	256,0(1,5)	260,7(3,2)
$\Delta E(\text{HIE})$	258,5	260,5(2,0)	264,8(6,3)

Frontière $C_\alpha-C_\beta$ Les résultats concernant la frontière $C_\alpha-C_\beta$ (géométries et affinités protoniques) sont regroupés dans la table 5.7. Concernant les affinités protoniques, on peut remarquer que, par rapport aux calculs de référence RHF, les calculs QM/MM sont dans une marge d'erreur de quelques kcal/mol. L'erreur est un peu plus grande quand les électrons de cœur sont inclus. On peut penser que ces déviations (jusqu'à 6,3 kcal/mol)

sont très probablement dues à la proximité de la frontière QM/MM. Il est donc désirable de tester le deuxième type de coupure.

La comparaison des différentes géométries est effectuée par la valeur des RMSD, exprimés en Å. Les RMSD sont calculés suivant la méthode de Kabsch [322], implantée dans le logiciel VMD [323]. La première remarque est que la différence majeure entre les géométries QM et QM/MM provient de la partie MM, quelque soit le type de calculs LSCF/MM. Ceci est facilement compréhensible, car le champ de forces MM n'est pas paramétré pour reproduire les géométries en phase gazeuse, mais plutôt en phase condensée. Les déformations du squelette du pentapeptide induisent de légères modifications de la géométrie de la chaîne latérale de l'acide aminé histidine (décrit au niveau QM). Nous pouvons aussi noter que le RMSD du fragment MM est quasi-constant quelque soit le type de méthode LSCF/MM. Cependant, le RMSD du sous-système QM est plus important dans le cas de la méthode LSCF/+3/MM que dans celui de la méthode LSCF/+1/MM. Cette augmentation du RMSD pour la méthode LSCF/+3/MM est principalement due à la distance de la liaison frontière. Nous rappelons, que dans le cas de l'approche LSCF/+1, la description correcte de cette liaison requiert un potentiel frontière spécifique contenant cinq paramètres, afin d'avoir une distance de liaison correcte. On remarque que la différence pour l'approche LSCF/+1/MM est de quelques centièmes d'Ångström. Le passage au niveau LSCF/+3/MM conduit à un accord moins satisfaisant par rapport aux valeurs de référence RHF ($\simeq 2$ pm), restant cependant dans une marge d'erreur très acceptable. Cette légère déviation géométrique est certainement responsable de l'erreur au niveau des affinités protoniques.

Frontières liaisons peptidiques La table 5.8 regroupe les résultats concernant la liaison peptidique en tant que liaison frontière. Le schéma LSCF/+1/MM donne des valeurs d'affinités protoniques acceptables, par rapport aux autres méthodes QM/MM, bien que l'une d'entre elles soit moins proche (5,1 kcal/mol) de la valeur référence RHF (HID \rightarrow HIP), et ce contrairement au cas C_α - C_β . Plusieurs raisons peuvent être invoquées. Premièrement, la liaison peptidique ne peut pas être représentée de manière acceptable à l'aide uniquement d'une SLBO, comme mentionné précédemment. Deuxièmement, la charge nucléaire de l'atome Y (qui est égale à +1), n'est pas assez conséquente pour attirer le doublet de l'atome d'azote dans le cas de la liaison C_Y - N_X ce qui implique que les valeurs de $d(C_Y$ - $N_X)$ soient assez éloignées des résultats RHF et ce même en utilisant un FBP comme c'est le cas ici. Troisièmement, l'atome hybride d'azote de la liaison N_Y - C_X n'a aucun électron à partager. Par conséquent, dans deux des trois cas regroupés dans la table 5.8, la liaison peptidique ne possède pas une géométrie plane (valeur de $\tau(N_X$ - H - C_α - $C_Y)$ supérieure à 30°) car les densités électroniques décrivant les liaisons C_Y - N_X et N_Y - C_X ne sont pas assez importantes. Cependant, nous remarquons que les valeurs $d(N_Y$ - $C_X)$ dévient de 0,024 Å au maximum, ce qui reste très acceptable et est attribué à l'utilisation du FBP.

La méthode LSCF/+3 ne parvient pas à corriger tous les points mentionnés précédemment, bien que la charge nucléaire des atomes Y ait été augmentée. En effet, l'atome d'azote hybride n'a toujours pas d'électrons de valence afin de représenter la paire libre et le FBP est omis dans le cas de l'approche LSCF/+3. Par conséquent, les distances N_Y - C_X

5.2. Prise en compte des électrons de cœur et de valence

TAB. 5.8 – **Frontières liaisons peptidiques.** Distances des liaisons frontières (en Å) et valeurs de l'angle dièdre N_X -H- C_α - C_Y des trois états de protonation (HID, HIE, et HIP, FIG. 5.14) de l'acide aminé histidine engagé dans un pentapeptide Ace-Ile-His-Tyr-NMe, calculées au niveau de théorie RHF/6-311G**, LSCF(+1)/AMBER99, LSCF(+3)/AMBER99 et LSCF(+5)/AMBER99. Les valeurs des RMSD par rapport aux géométries de référence RHF/6-311G** sont données (en Å) pour le fragment QM (RMSD_{QM}), le fragment MM (RMSD_{MM}), et pour le système entier (RMSD). Nous donnons aussi les affinités protoniques (ΔE en kcal/mol) de l'acide aminé histidine, calculées aux mêmes niveaux de théorie ($\Delta E(\text{HID}) = E(\text{HID}) - E(\text{HIP})$ et $\Delta E(\text{HIE}) = E(\text{HIE}) - E(\text{HIP})$). Les déviations par rapport aux valeurs RHF sont reportées entre parenthèses.

	RHF	LSCF(+1)/AMBER99	LSCF(+3)/AMBER99	LSCF(+5)/AMBER99
HIP				
$d(C_Y-N_X)$	1,344	1,423(0,079)	1,381(0,037)	1,380(0,036)
$d(N_Y-C_X)$	1,339	1,319(-0,020)	1,416(0,077)	1,380(0,041)
$\tau(N_X-H-C_\alpha-C_Y)$	-10,5	-15,2	-14,2	-9,7
RMSD _{QM}		0,794	0,650	0,902
RMSD _{MM}		1,055	1,117	1,335
RMSD		1,033	1,335	1,021
HID				
$d(C_Y-N_X)$	1,351	1,436(0,085)	1,373(0,022)	1,373(0,022)
$d(N_Y-C_X)$	1,348	1,324(-0,024)	1,419(0,071)	1,380(0,032)
$\tau(N_X-H-C_\alpha-C_Y)$	-11,0	31,6	4,1	-2,5
RMSD _{QM}		0,197	0,268	0,238
RMSD _{MM}		0,838	0,813	1,136
RMSD		0,750	0,778	1,068
HIE				
$d(C_Y-N_X)$	1,345	1,437(0,092)	1,380(0,035)	1,380(0,035)
$d(N_Y-C_X)$	1,343	1,322(-0,021)	1,419(0,076)	1,379(0,036)
$\tau(N_X-H-C_\alpha-C_Y)$	14,5	33,5	13,6	-2,0
RMSD _{QM}		0,311	0,583	0,753
RMSD _{MM}		0,721	1,098	1,303
RMSD		0,784	1,183	1,474
$\Delta E(\text{HID})$	257,5	262,5(5,1)	267,0(9,5)	255,4(-2,1)
$\Delta E(\text{HIE})$	258,5	257,6(-0,9)	267,5(9,0)	258,5(0,0)

ont un écart important vis-à-vis des résultats RHF (0,071 Å au minimum), ce qui a pour conséquence d'augmenter la déviation entre les affinités protoniques obtenues au niveau RHF/6-311G** et grâce à la méthode LSCF/+3/AMBER99 (supérieures à 9,0 kcal/mol). Cependant, du fait de l'augmentation de deux unités de la charge nucléaire de C_Y , les valeurs $d(C_Y-N_X)$ sont beaucoup plus proches des valeurs RHF en comparaison aux résultats LSCF/+1/MM précédents. Ce dernier point permet aussi de restaurer le caractère quasi-plan de la liaison peptidique.

Nous pouvons remarquer plusieurs tendances très positives dues à l'utilisation de la méthodologie LSCF/+5. Ainsi, lorsque les deux électrons de valence sont ajoutés à l'atome d'azote hybride N_Y , la planarité de la liaison peptidique est conservée, ainsi que la qualité des résultats correspondant aux valeurs de $d(C_Y-N_X)$, mais aussi d'atteindre une description satisfaisante de la liaison N_Y-C_X . Pour preuve, les distances N_Y-C_X sont elles aussi en très bon accord avec les valeurs des calculs purement quantiques, tout comme les valeurs des affinités protoniques. Un autre point important à noter est que les deux liaisons frontières (C_Y-N_X et N_Y-C_X) deviennent de distance quasi égale lors de l'utilisation de l'approche LSCF/+5/MM.

Des conclusions équivalentes que celles effectuées dans le cas de la frontière entre les atomes C_α et C_β peuvent être faites dans le cas des valeurs des RMSD géométriques. La déviation la plus importante vient de la partie MM, induisant une modification de la partie traitée au niveau quantique. Une autre source d'erreurs entre les calculs purement quantiques et les calculs LSCF/MM provient de la liaison hydrogène entre l'atome d'hydrogène en position δ et l'atome d'oxygène de la liaison peptidique liant l'histidine à l'isoleucine. En effet, cette liaison existe dans le cas des calculs purement quantiques mais n'intervient pas dans les calculs QM/MM, quelque soit l'approche utilisée. Ceci est probablement dû à la description hybride de la liaison peptidique, et non particulièrement au schéma LSCF/+5/MM.

Conclusion

Dans cette section, nous avons montré que la méthode LSCF/MM peut être étendue avec succès à l'étude des polypeptides. La frontière QM/MM a été placée au niveau des liaisons peptidiques. Cette partition est particulièrement intéressante afin d'avoir :

- une partie QM de taille suffisante, limitant la perturbation de la frontière QM/MM.
- une description symétrique de l'acide aminé considéré.
- de contourner l'ajustement arbitraire des charges classiques du champ de forces.

Nous avons montré que le fait d'inclure des électrons de valence (LSCF/+5/MM) permet une description correcte de la liaison peptidique en tant que frontière QM/MM (distance de liaison, densité et géométrie plane), sans ajouter de paramètres.

Troisième partie

La Méthode du Champ Auto-Cohérent Local (LSCF) : Applications à l'Etude des Etats Electroniques de Systèmes Biomimétiques

6

Etats Ionisés et Excités de Cœur dans les Macromolécules

Publication

- **Core-ionized and core-excited states of macromolecules**
P.-F. Loos, X. Assfeld
Int. J. Quant. Chem., **2007**, *107*, 2343–2252

6.1 Introduction

Les états excités et ionisés de cœur permettent d’obtenir des informations très intéressantes sur la structure et la nature des molécules [332]. Bien que la majeure partie des études concerne des composés inorganiques, ou des petites molécules organiques, il a été récemment montré que les biomolécules pouvaient aussi être traitées [333–336]. Du point de vue des chimistes théoriciens, les états excités de cœur des macromolécules présentent plusieurs challenges :

- Pour décrire les états de cœur excités, nous devons éviter l’effondrement variationnel. De plus, nous devons être sûrs que les états excités obtenus sont bien orthogonaux à l’état fondamental. Cet axe de recherche est très fécond et plusieurs techniques différentes ont été proposées pour assurer ces deux points [337–341].
- La modélisation des macromolécules uniquement à l’aide de la mécanique quantique est encore, pour l’instant, hors de portée pour les chimistes théoriciens. Cependant, plusieurs méthodes hybrides QM/MM sont disponibles [65, 66, 70, 74, 183, 194, 199, 200, 202, 204–207, 214, 217, 219, 222, 225, 226] comme nous l’avons vu dans le chapitre 3 de la partie I.

Pris séparément, les deux problèmes sont correctement traités par les méthodes citées. Cependant, ils sont rarement considérés simultanément. Ainsi, plusieurs groupes ont proposé récemment des techniques combinant les méthodes QM/MM et TD-DFT [197, 342–350]. La méthode LSCF a été précédemment appliquée pour la détermination des états ionisés de cœur du polyéthylène, du polytétrachloroéthylène et de la cram-

bine [231]. Il convient de rappeler que dans le cadre de la méthode LSCF/MM, la jonction entre la partie QM et la partie MM est matérialisée par une SLBO doublement occupée, restant gelée lors de l'optimisation de la fonction d'onde. Ceci permet une description correcte du système hybride QM/MM et de traiter de façon rigoureuse le problème de la liaison pendante. Le « trou de cœur » est, quant à lui, représenté par une orbitale de cœur (CO pour *Core Orbital*) vide qui reste, elle aussi, gelée pendant le processus d'optimisation de la fonction d'onde du système pour éviter l'effondrement variationnel et assurer l'orthogonalité entre les états électroniques.

Dans ce chapitre, nous présentons tout d'abord une modification de la méthode LSCF permettant de traiter différemment le recouvrement entre orbitales occupées (SLBOs) et orbitales de cœur (COs) vacantes. La modification apportée est alors comparée à l'ancienne version de la méthode [231] sur un système simple (molécule de pentane). La convergence en fonction de la taille de la base est étudiée sur la molécule de méthane pour laquelle la valeur expérimentale n'est pas ambiguë, et permet d'utiliser une base de taille très importante (jusqu'à cc-pCV5Z). La validité de la méthode est établie grâce à des calculs sur des molécules organiques contenant au moins un atome de carbone, ainsi que sur des molécules contenant une liaison peptidique. De plus, nous appliquerons la méthode sur la molécule d'alanine dans différents environnements : nous considérerons le cas où la molécule est isolée (nous étudierons plusieurs conformations), engagée dans un tripeptide ou un pentadécapeptide en conformation hélice α et brin β . Ces deux derniers calculs seront réalisés grâce à la méthode LSCF/MM.

6.2 Théorie

La méthode LSCF a déjà été détaillée dans le chapitre 4, et nous ne rappellerons ici que les points nécessaires à la compréhension de la modification apportée. Par commodité, dans la suite de ce chapitre, nous noterons l'orbitale de cœur considérée CO (*Core Orbital*). Celle-ci peut être occupée (état fondamental) ou bien encore vacante (état excité ou ionisé). Pour étudier les états de cœur excités et ionisés, nous devons effectuer deux calculs distincts, et ce indépendamment du choix de la CO devant être étudiée. Le premier calcul consiste à calculer l'énergie de l'état fondamental, $E(\Psi)$ où Ψ représente le déterminant de cet état, tandis que le deuxième correspond au calcul de l'énergie de l'état électronique où la CO est vacante et gelée $E(\Psi')$ (afin d'éviter l'effondrement variationnel et d'assurer l'orthogonalité entre l'état fondamental et les états excités). Ici, nous noterons Ψ' le déterminant de l'état excité. L'énergie d'excitation ou d'ionisation E est calculée par la différence entre ces deux énergies.

$$E = E(\Psi') - E(\Psi) \quad (6.1)$$

Il existe deux méthodes possibles afin de définir la CO que nous appellerons interne et externe. La première méthode (interne) consiste à utiliser un critère de localisation interne (comme Pipek-Mezey ou Boys-Foster) appliqué sur la fonction d'onde optimisée de l'état fondamental. Le principe externe s'appuie sur le principe de transférabilité. La CO est donnée par l'utilisateur à partir d'un calcul atomique ou un critère de localisation sur une autre molécule par exemple. Dans ce cas, le calcul de l'état fondamental doit être

effectué avec la méthode LSCF à l'aide de cette CO occupée et gelée. Ceci est le seul moyen d'obtenir une orthogonalité stricte entre l'état fondamental et l'état excité. En effet, dans le premier cas (critère interne), la CO est obtenue par rotation dans l'espace des MOs occupées à l'aide d'une transformation unitaire (Section 4.5.2). Elle n'est donc plus fonction propre du Fockien de l'état fondamental mais reste orthogonale aux autres MOs du système. Son utilisation en tant que FO ne modifie donc en rien l'énergie de cet état. Cependant, lorsque cette CO est transférée d'un autre système, elle n'est pas forcément orthogonale aux autres MOs. Il est alors nécessaire d'effectuer un calcul additionnel pour déterminer l'énergie de l'état fondamental, prenant en compte cette contrainte. Nous allons montrer que la condition d'orthogonalité entre la CO et les autres MOs occupées permet d'assurer l'orthogonalité entre l'état fondamental et les états excités. Le déterminant de l'état fondamental $|\Psi\rangle$, construit à l'aide des spinorbitales optimisées $\{\phi_{CO}, \phi_i\}_{1 \leq i \leq n}$, est donné par la relation,

$$|\Psi\rangle = |\phi_{CO}\phi_2 \dots \phi_n\rangle \quad (6.2)$$

tandis que celui de l'état excité $|\Psi'\rangle$, construit à l'aide des spinorbitales $\{\phi'_i\}_{1 \leq i \leq n}$ optimisées pour l'état ionisé de cœur, se développe suivant :

$$|\Psi'\rangle = |\phi'_1\phi'_2 \dots \phi'_n\rangle \quad (6.3)$$

Dans le déterminant $|\Psi\rangle$ intervient la spinorbitale $|\phi_{CO}\rangle$ correspondant à la CO. Grâce aux spécificités de la méthode LSCF, on peut assurer que :

$$\forall i \in [2, \dots, n], \langle \phi_i | \phi_{CO} \rangle = 0 \quad (6.4)$$

$$\forall j \in [1, \dots, n], \langle \phi'_j | \phi_{CO} \rangle = 0 \quad (6.5)$$

Le recouvrement entre les deux déterminants est donné par :

$$S = \frac{1}{n!} \langle \Psi | \Psi' \rangle = \frac{1}{n!} \langle \phi_{CO}\phi_2 \dots \phi_n | \phi'_1\phi'_2 \dots \phi'_n \rangle \quad (6.6)$$

En développant cette relation en utilisant les règles de la Ref. [86] (p. 75) :

$$S = \frac{n(n-1)}{2n!} \sum_{ij}^{n!} (-1)^{p_i} (-1)^{p_j} \int \int \dots \int \mathcal{P}_i \{ \phi_{CO}\phi_2 \dots \phi_n \} \times \mathcal{P}_j \{ \phi'_1\phi'_2 \dots \phi'_n \} d\tau_1 d\tau_2 \dots d\tau_n \quad (6.7)$$

où \mathcal{P}_i est l'opérateur qui génère la $i^{\text{ème}}$ permutation dans l'ensemble $\{\phi_{CO}\phi_2 \dots \phi_n\}$ et où p_i est le nombre de transpositions nécessaire pour atteindre cette permutation à partir de l'ensemble initial. Or, en développant la relation (6.7), on remarque que le terme $\langle \phi'_j | \phi_{CO} \rangle$ est présent dans chaque terme, d'où $S = 0$, ce qui indique que l'état fondamental est bien orthogonal à l'état excité du fait de l'orthogonalité de ces spinorbitales avec la CO. Nous allons maintenant nous intéresser à la méthode d'orthogonalisation de ces FOs.

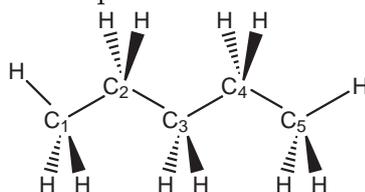
Lorsque toutes les FOs sont occupées, la méthode d'orthogonalisation de Löwdin peut être appliquée afin d'orthogonaliser ces fonctions, comme expliqué dans la partie 4.2.1, car elle correspond à des rotations dans l'espace occupé-occupé. Cependant, pour les états

excités ou ionisés, la transformation de Löwdin ne peut pas être utilisée car elle mélange les FOs occupées et vacantes. La fonction orthogonalisée résultante ne représente plus la fonction vacante souhaitée. La situation peut être dramatique lorsque la CO est, par exemple, proche spatialement d'une SLBO. Cependant, cette erreur est très faible dans la plupart des cas. En effet, le recouvrement entre ces orbitales est faible car la CO est principalement développée sur les AOs de cœur, tandis que la SLBO est principalement développée sur les AOs de valence. Un moyen d'éviter cette situation, c'est-à-dire la rotation entre FOs occupées et vacantes, est d'orthogonaliser l'espace des fonctions vacantes par la méthode de Löwdin (L_v pour Löwdin virtuelle), puis les fonctions occupées par rapport aux précédentes par la méthode de Gram-Schmidt²¹ (GS). Cette orthogonalisation par bloc, que nous noterons L_v+GS , est adoptée par la suite. De plus, dans cette étude, nous comparerons cette nouvelle procédure d'orthogonalisation avec celle qui correspondait à effectuer une transformation de Löwdin sur l'ensemble des FOs quelque soit leur nombre d'occupation, comme cela a déjà été fait [231].

6.3 Validation de la méthode : molécule de pentane

Pour vérifier l'importance de l'erreur induite par l'orthogonalisation globale de Löwdin qui était appliquée dans les calculs précédents [231], nous avons effectué des calculs sur la molécule de pentane à l'aide de la base 6-311++G**. Les géométries ont été optimisées au niveau B3LYP/6-311++G**. Deux types de CO (1s du carbone) ont été utilisés et obtenus à l'aide des critères de localisation Pipek-Mezey (PM) et Boys-Foster (BF). De plus, deux types de fonctionnelle d'échange-corrélation ont été testés : la fonctionnelle B3LYP [124, 128] et PBE0 [132]. Les cinq orbitales 1s des atomes de carbone ont été étudiées. Les énergies d'ionisation avec et sans SLBO entre les atomes C_1 et C_2 (FIG. 6.1) sont données dans la table 6.1.

FIG. 6.1 – Molécule de pentane et numérotation des atomes.



Différentes tendances peuvent être extraites de ces résultats. Premièrement, la différence entre l'ancien (Löwdin) et le nouveau schéma (L_v+GS) d'orthogonalisation n'est pas très importante (moins de 0,07 eV). Ceci est dû au fait que le recouvrement (dernière colonne de la table 6.1) entre la SLBO et la CO est faible. Ceci s'explique par la nature des MOs mises en jeu : la SLBO est plutôt développée sur les AOs de valence, tandis que la CO est développée principalement sur les AOs de cœur. Nous pouvons donc en déduire,

²¹Une troisième manière strictement équivalente serait d'effectuer une orthogonalisation de Gram-Schmidt sur l'ensemble des FOs en commençant par la CO

TAB. 6.1 – Energies d’ionisation de l’orbitale $1s$ du carbone (en eV) calculées au niveau de théorie DFT/6-311++G** pour la molécule de pentane. Les calculs sont effectués avec et sans SLBO entre les atomes C_1 et C_2 , avec l’ancienne (Löwdin) ou la nouvelle procédure (Lv+GS) d’orthogonalisation (voir la figure 6.1 pour la numérotation des atomes). Deux types de critères de localisation sont utilisés, Pipek-Mezey (PM) et Boys-Foster (BF). Les valeurs entre parenthèses correspondent à la différence d’énergie (en eV) entre les deux procédures d’orthogonalisation.

SLBO	Lv+GS						Löwdin						
	PM		BF		PM		BF		PM		BF		
Loc.	B3LYP	PBE0	B3LYP	PBE0	B3LYP	PBE0	B3LYP	PBE0	B3LYP	PBE0	B3LYP	PBE0	\langle SLBO CO \rangle
C_1	289,54	291,02	290,41	290,94	292,55	292,03	290,90(-0,04)	292,48(0,07)	291,99(0,09)	292,83(0,02)	292,31(-0,04)	290,37(0,01)	0,0055
C_2	290,26	290,98	290,41	291,63	292,35	291,85	291,60(0,03)	292,31(-0,04)	292,83(0,02)	292,31(-0,04)	290,92(0,01)	290,44(-0,01)	0,0024
C_3	290,32	290,86	290,29	290,38	290,93	290,36	290,38(0,00)	290,92(0,01)	290,37(0,01)	290,38(0,00)	291,01(-0,01)	290,44(-0,01)	0,0064
C_4	290,26	290,98	290,41	290,29	291,00	290,43	290,30(-0,01)	291,01(-0,01)	290,44(-0,01)	290,30(-0,01)	291,04(0,00)	290,42(0,00)	0,0048
C_5	289,54	291,02	290,41	289,56	291,04	290,42	289,56(0,00)	291,04(0,00)	290,42(0,00)	289,56(0,00)	291,04(0,00)	290,42(0,00)	0,0012

que bien qu’approchés, les résultats précédemment obtenus restent valides. L’autre point important est que la présence de la SLBO perturbe la fonction d’onde assez localement : l’erreur induite sur les énergies d’ionisation de l’orbitale $1s$ du carbone C_3 est inférieure à un dixième d’eV. D’autre part, nous pouvons remarquer que la méthode LSCF ne permet pas de calculer des énergies d’ionisation de cœur sur les atomes portant la SLBO : l’erreur est supérieure à 1,5 eV quelque soit la procédure d’orthogonalisation, ce qui n’est pas acceptable. Cette conclusion reste valable quelque soit le type de fonctionnelle utilisé (B3LYP ou PBE0), et quelque soit le critère de localisation choisi (BF ou PM).

Nous devons tout de même faire remarquer que les énergies d’ionisation sont très sensibles à la fonctionnelle et au critère de localisation utilisés. Elles peuvent aussi dépendre de la base d’AO. Ceci fait l’objet de la section 6.3.1. La seule manière de décider quelle méthodologie est la plus appropriée est de confronter nos résultats à des valeurs expérimentales ou à des résultats provenant d’autres approches théoriques. Ceci sera fait dans la section 6.3.2, où nous comparerons les énergies d’ionisation de cœur de petites molécules organiques contenant au moins un atome de carbone.

6.3.1 Influence de la base d’AO sur les énergies d’ionisation de cœur : molécule de méthane

L’influence de la base d’AO sur les énergies d’ionisation de cœur est déterminée par une étude de la convergence de l’énergie verticale d’ionisation de cœur $1s$ de la molécule de méthane. Cette molécule possède deux avantages majeurs :

- (i) elle ne contient qu’un atome de carbone, et donc la valeur expérimentale (290,84 eV) n’est pas ambiguë.
- (ii) De plus, le nombre d’atomes de cette molécule nous permet de tester les bases de façon exhaustive, y compris celles qui comportent un très grand nombre de fonctions gaussiennes.

L’ensemble des calculs est effectué avec la fonctionnelle PBE0. Pour chaque base de gaussiennes, la géométrie de la molécule est complètement optimisée. Le critère BF a été choisi pour localiser la CO. Les résultats sont rassemblés dans la table 6.2.

Nous pouvons remarquer une convergence rapide des énergies d’ionisation de cœur pour chaque famille de base (FIG. 6.2). Notons aussi que l’ajout de fonctions diffuses (famille aug-cc-pVXZ) n’influence pas les résultats. En effet, ces fonctions sont étendues loin du noyau, ce qui n’influence que peu les énergies d’ionisation. Tant que la base est assez étendue, la polarisation des fonctions de cœur n’améliore pas les résultats. Concernant les bases de Pople [351, 352] triple- ζ [353–356], celles-ci donnent des résultats très constants quelque soit le nombre de fonctions diffuses ou de polarisation (pour les atomes lourds ou légers). Les énergies d’ionisation obtenues avec la série 6-311G sont très proches de celles obtenues avec les bases de Dunning triple- ζ [357–364]. D’un point de vue pratique, les valeurs obtenues avec les bases de Pople 6-311G sont en bon accord avec la valeur expérimentale (290,84 eV) ce qui est certainement dû à une compensation d’erreurs fortuite (limitation de la taille de la base, corrections relativistes, erreur de *self-interaction* dans la fonctionnelle d’échange-corrélation [365], ...). Comme les fonctions diffuses et de polarisation pourraient être importantes dans certaines situations, autre que le méthane, nous

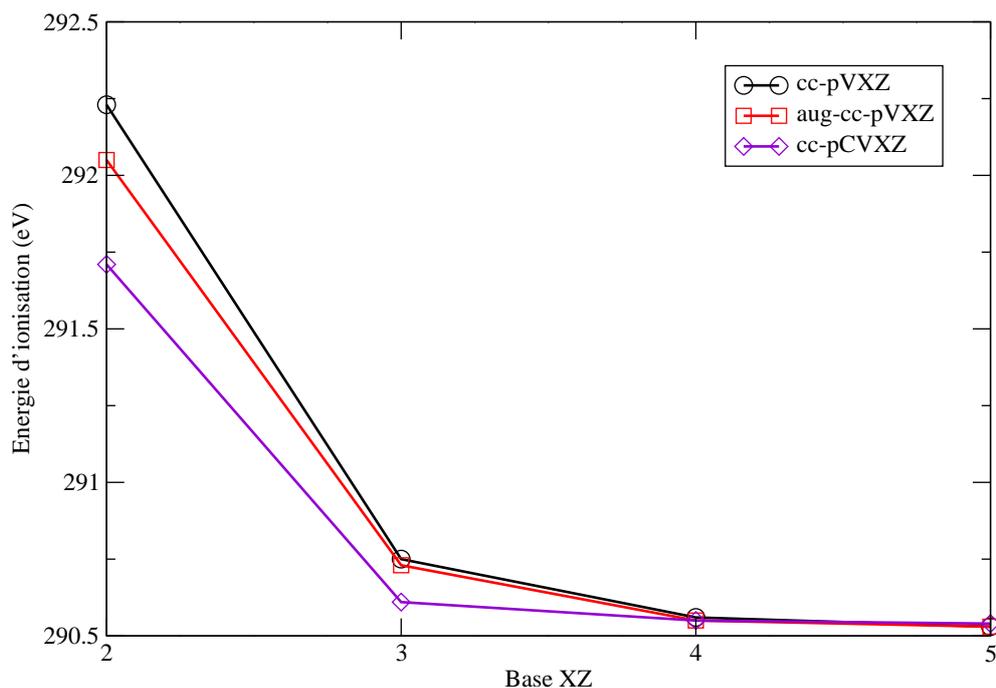
6.3. Validation de la méthode : molécule de pentane

TAB. 6.2 – Energies d’ionisation de l’orbitale 1s du carbone (en eV) de la molécule de méthane en fonction de la base calculée à l’aide de la fonctionnelle PBE0 et du critère de localisation Boys-Foster pour la CO.

Base	Energie d’ionisation (1s)
cc-pVXZ	
cc-pVDZ	292,23
cc-pVTZ	290,75
cc-pVQZ	290,56
cc-pV5Z	290,53
aug-cc-pVXZ	
aug-cc-pVDZ	292,05
aug-cc-pVTZ	290,73
aug-cc-pVQZ	290,55
aug-cc-pV5Z	290,53
cc-pCVXZ	
cc-pCVDZ	291,71
cc-pCVTZ	290,61
cc-pCVQZ	290,55
cc-pCV5Z	290,54
Pople triple- ζ	
6-311G	290,78
6-311G(d)	290,78
6-311G(d,p)	290,77
6-311G(2d,p)	290,75
6-311G(2d,2p)	290,69
6-311+G(d,p)	290,78
6-311++G(d,p)	290,78
Exp.	290,84

avons décidé de choisir la base 6-311++G** pour l'ensemble des calculs de ce chapitre.

FIG. 6.2 – Convergence des énergies d'ionisation de cœur de l'orbitale 1s (en eV) de la molécule de méthane en fonction de la base utilisée. L'entier X représente le nombre de ζ de la base de fonctions gaussiennes.



6.3.2 Molécules organiques

Les énergies d'ionisation de l'orbitale 1s du carbone d'une série de petites molécules organiques sont présentées dans la table 6.3. Les molécules sont réparties en deux groupes. Le premier groupe comprend les hydrocarbures, tandis que le second contient les molécules contenant des hétéroatomes (O et N). Toutes les géométries de l'état fondamental ont été optimisées au niveau B3LYP/6-311++G**. Seules les énergies verticales d'ionisation sont considérées : la géométrie du cation est exactement la même que celle de la forme neutre. Cette approximation est valide car le processus d'ionisation s'effectue très rapidement, sans que la géométrie n'ait eu le temps de se relaxer. Les énergies d'ionisation sont calculées à différents niveaux de théorie (B3LYP/6-311++G** et PBE0/6-311++G**) et avec différents critères de localisation (PM et BF). Les résultats obtenus sont comparés aux valeurs expérimentales de Jolly [366] et Saerthe [367].

6.3. Validation de la méthode : molécule de pentane

TAB. 6.3: Energies d'ionisation de l'orbitale 1s du carbone (en eV) calculées à différents niveaux de théorie et critères de localisation pour de petites molécules organiques. La première famille de molécules contient uniquement des hydrocarbures, tandis que la seconde contient des molécules avec des hétéroatomes (O ou N). Les valeurs expérimentales sont prises des références [366] et [367].

		Energies d'ionisation				Dev. Exp.		
		B3LYP		PBE0		B3LYP		PBE0
		PM	BF	BF	Exp.	PM	BF	BF
Méthane	C	290,91	291,45	290,78	290,84	0,07	0,61	0,07
Ethane	C	290,65	291,29	290,67	290,71	0,06	0,58	0,04
Ethylène	C	291,16	291,33	290,70	290,82	0,34	0,51	0,12
Ethyne	C	290,67	291,82	291,19	291,25	0,58	0,57	0,06
Propane	CH ₃	290,23	291,11	290,49				
	CH ₂	290,73	291,20	290,63				
	Moy.	290,48	291,15	290,56	290,57	0,09	0,58	0,01
Propène	CH ₂	290,26	290,73	290,11				
	CH	291,28	291,21	290,64				
	CH ₃	290,51	291,44	290,82				
	Moy.	290,68	291,13	290,52	290,68	0,00	0,45	0,16
Propyne	HC	290,48	290,82	290,21	290,37	0,11	0,45	0,16
	C	290,94	291,42	290,84	290,93	0,01	0,49	0,09
	CH ₃	289,98	292,32	291,70	291,76	1,78	0,56	0,06
Butane	CH ₃	289,74	291,05	290,44				
	CH ₂	290,52	291,03	290,46				
	Moy.	290,13	291,03	290,45	290,48	0,35	0,56	0,03
Pentane	C ₁	289,54	291,02	290,41				
	C ₂	290,26	290,98	290,41				
	C ₃	290,32	290,86	290,29				
	C ₄	290,26	290,98	290,41				
	C ₅	289,54	291,02	290,41				
	Moy.	289,99	290,97	290,39	290,42	0,43	0,55	0,03
RMSD ₁						0,60	0,54	0,09
Formaldéhyde	C	295,08	295,17	294,46	294,47	0,61	0,70	0,01
Acétone	CO				293,71	0,63	0,57	0,01
					293,88	0,46	0,40	0,18
	CH ₃				291,15	0,31	0,59	0,04
					291,23	0,39	0,51	0,12
Méthanol	C				292,3	0,32	0,81	0,14
					292,8	0,18	0,31	0,36
Ethanol	CH ₂	292,28	292,91	292,30	292,50	0,22	0,41	0,20

TAB. 6.3 – suite de la page précédente

		Energies d'ionisation				Dev. Exp.		
		B3LYP		PBE0		B3LYP		PBE0
		PM	BF	BF	Exp.	PM	BF	BF
Acide éthanoïque	CH ₃	290,56	291,32	290,69	291,10	0,54	0,22	0,41
	CH ₃	291,28	292,19	291,55	291,55	0,27	0,64	0,00
					291,6	0,32	0,59	0,05
Diméthyléther	COOH	295,92	295,96	295,35	295,38	0,54	0,58	0,03
					295,6	0,32	0,36	0,25
	CH ₃	292,25	292,83	292,17	292,17	0,08	0,66	0,00
Acide formique	C	296,36	296,42	295,75	292,55	0,30	0,28	0,38
					295,80	0,56	0,62	0,05
Méthylamine	CH ₃	291,54	292,19	291,55	295,90	0,46	0,52	0,15
					291,60	0,06	0,59	0,05
RMSD ₂						0,40	0,54	0,19
RMSD _{Tot}						0,49	0,54	0,16

En examinant les résultats de la table 6.3, nous pouvons remarquer que, par rapport à la fonctionnelle B3LYP, la fonctionnelle PBE0 permet d'obtenir des résultats beaucoup plus proches des données expérimentales. La déviation maximale par rapport à l'expérience est inférieure à 0,2 eV pour PBE0 et à 0,6 eV pour B3LYP pour le premier groupe de molécules ne contenant que des atomes de carbone et d'hydrogène. La déviation augmente à 0,4 eV pour PBE0 et à 0,8 eV pour B3LYP si l'on considère le deuxième groupe de molécules. Il faut aussi noter que les résultats obtenus avec le critère BF sont, en moyenne, de même qualité que ceux obtenus grâce au critère PM. Cependant, si nous examinons plus attentivement ces résultats, nous remarquons que le critère PM donne ponctuellement des résultats plus précis que BF (par exemple, la déviation pour la molécule de méthane est de 0,07 eV avec PM et 0,61 eV avec BF). Mais, de façon surprenante, le critère PM prédit une énergie d'ionisation très éloignée de l'expérience pour le carbone du groupement méthyle de la propyne. Les variations par rapport aux valeurs expérimentales sont comprises entre -1,78 et 0,63 eV. Le critère BF, quant à lui, montre un comportement beaucoup plus constant pour tous les types d'atomes de carbone : l'ensemble des variations théorie-expérience reste dans une gamme de 0,22-0,81 eV avec la fonctionnelle B3LYP et de -0,41 à 0,14 eV pour PBE0. Ce comportement semble être systématique. Dans le reste de cette étude, nous opterons donc pour le critère BF et la fonctionnelle PBE0.

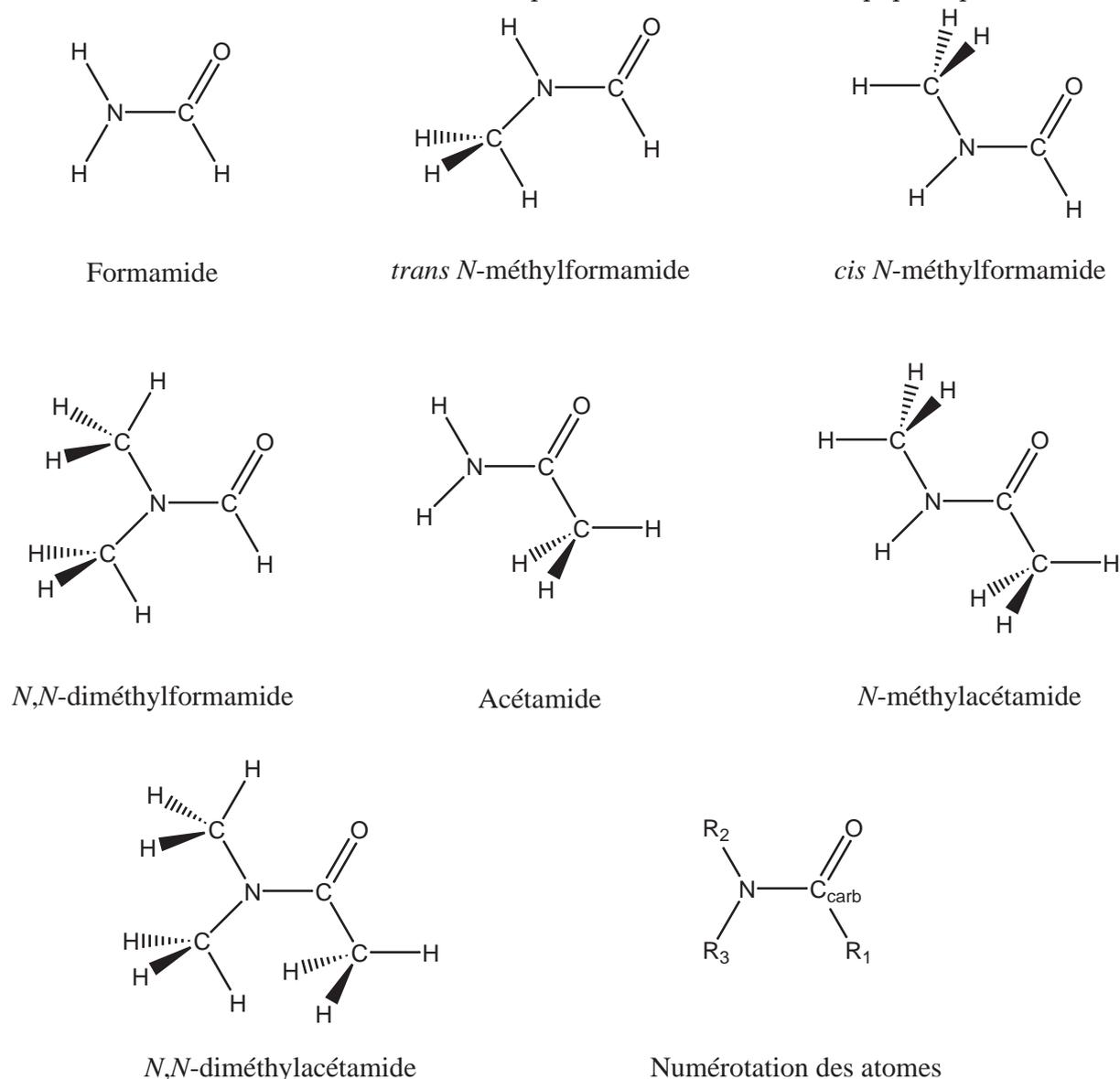
6.3.3 Molécules contenant une liaison peptidique

Avant d'étudier des systèmes macromoléculaires, nous nous sommes intéressés à des molécules biomimétiques simples pour lesquelles des résultats expérimentaux et théoriques étaient disponibles. Afin de tester la précision de notre méthode, nous avons donc calculé les énergies d'ionisation de l'orbitale 1s du carbone de plusieurs molécules contenant une liaison peptidique suivant la méthodologie précédemment définie. Les résultats sont

6.3. Validation de la méthode : molécule de pentane

présentés dans la table 6.4 et comparés aux valeurs expérimentales disponibles [366], ou avec d'autres méthodes théoriques [368–376]. Nous avons choisi les mêmes molécules que Chong et collaborateurs [376] avec l'ajout de l'isomère *cis* du N-méthylacétamide. Les molécules et les types d'atomes de carbone (C_{carb} , C_{R_1} , C_{R_2} et C_{R_3}) sont représentés sur la figure 6.3.

FIG. 6.3 – Molécules biomimétiques contenant une liaison peptidique.



Conformément à ces calculs, nous pouvons noter que notre méthode donne des résultats aussi précis que la méthode Δ KS de Chong *et al.*. La déviation maximale entre les deux méthodes est de 0,3 eV, ce qui est une marge d'erreur très acceptable. Même si le nombre de valeurs expérimentales est restreint, la différence la plus importante est de 0,2 eV pour

TAB. 6.4 – Energies d'ionisation de l'orbitale $1s$ du carbone (en eV) de molécules biomimétiques (Fig. 6.3 pour les types d'atomes de carbone et les conformations des molécules). Les calculs ont été effectués au niveau de théorie PBE0/6-311++G** à l'aide du critère de localisation Boys-Foster. Les valeurs de référence sont extraites de l'article de Chong *et al.* [368, 376]. Ces calculs sont réalisés grâce à la méthode Δ KS au niveau de théorie (PW86PW91)/cc-pCVTZ. Les valeurs expérimentales sont données entre parenthèses [366].

Atome de Carbone Molécules	C_{R_1}		C_{carb}		C_{R_2}		C_{R_3}	
	Ref.		Ref.		Ref.		Ref.	
Formamide			294,16 (294,45)	294,41				
Trans-N-Méthylformamide			293,64	293,92			292,50	292,49
Cis-N-méthylformamide			293,58	293,85	292,18	292,16		
N,N-diméthylformamide			293,25 (293,45)	293,52	291,92 (292,03)	291,86	292,27 (292,03)	292,24
Acétamide	291,20	291,19	293,87	295,17				
Trans-N-méthylacétamide	291,12	291,02	293,37	293,68	292,03	291,90		
Cis-N-méthylacétamide		291,03		293,73				292,29
N,N-diméthylacétamide	290,90	290,89	292,99	293,34	292,69	291,61	292,11	292,03

6.3. Validation de la méthode : molécule de pentane

notre méthode et de 0,3 eV pour la méthode Δ KS²². Nous pouvons donc conclure que les deux méthodes théoriques sont en bon accord avec les résultats expérimentaux.

6.3.4 La molécule d'alanine : influence de l'environnement biologique

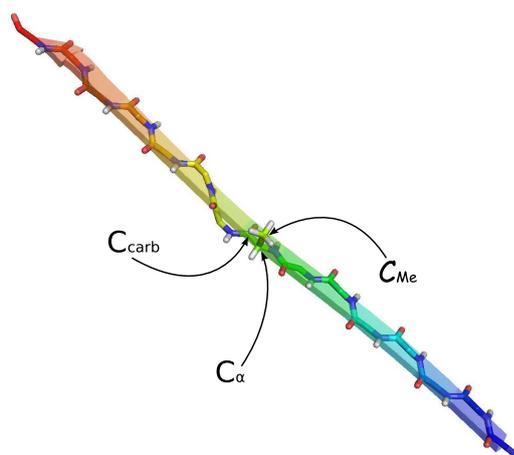
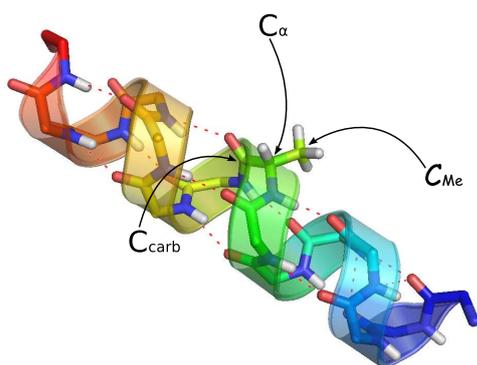
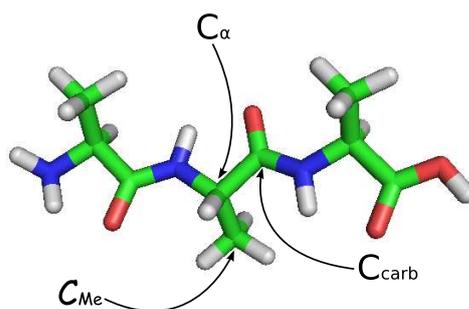
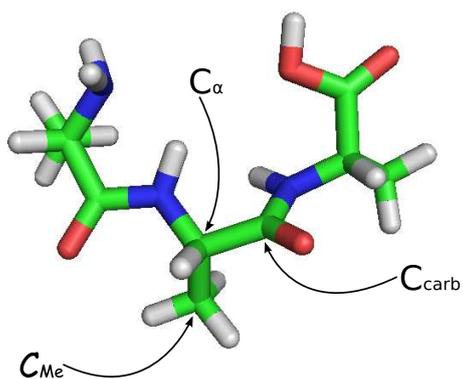
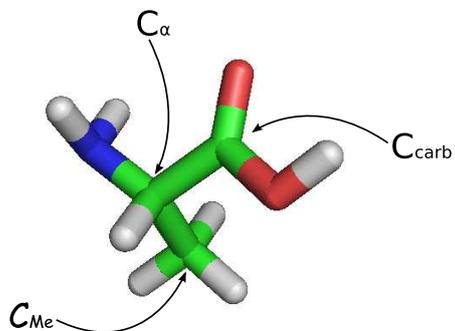
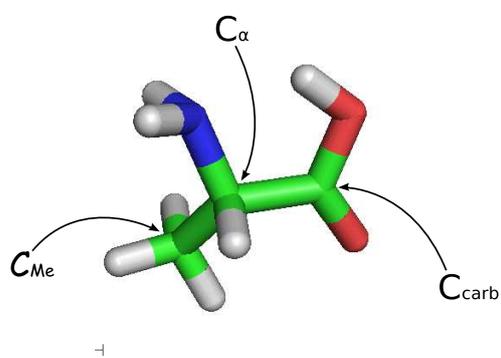
Dans une étude antérieure [231], Ferré *et al.* ont reporté les résultats de calculs QM/MM sur les énergies d'ionisation de l'orbitale 1s du carbone dans le cas d'un anion glutamate de la protéine crambine. Comme ce groupe fonctionnel est électriquement chargé, les charges ponctuelles classiques jouent un rôle de première importance. De la phase gazeuse à l'environnement de la protéine, nous avons déterminé un déplacement chimique de 4 eV. Ici, nous voudrions étudier l'influence d'une protéine modèle sur un acide aminé neutre. Nous avons choisi l'alanine comme acide aminé car des valeurs expérimentales sont disponibles pour cet acide aminé en phase gazeuse [380]. Dans cette situation, les calculs théoriques sont parfaitement adaptés à une comparaison directe avec les résultats expérimentaux.

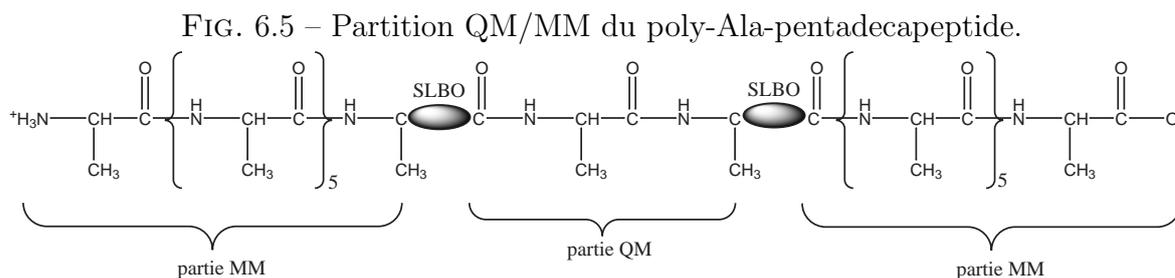
Dans un premier temps, nous avons calculé les énergies d'ionisation des deux structures les plus stables (FIG. 6.4). Deux conformères ont été trouvés à une énergie très proche (différence inférieure à 1 kcal.mol⁻¹). Puis, l'influence directe du milieu a été étudiée en considérant la molécule d'alanine engagée dans un tripeptide et dans un pentadécapeptide, tous les deux en conformation hélice α et brin β . Le calcul sur le pentadécapeptide a été effectué à l'aide de la méthode LSCF/MM (FIG. 6.5). Les géométries ont été complètement optimisées à leur niveau de théorie respectif, exception faite pour les calculs QM/MM où les positions des atomes du squelette du polypeptide ont été contraintes à rester fixes. L'ensemble des résultats est reporté dans la table 6.5.

²²La méthode Δ KS de Chong *et al.* est basée sur le modèle *unrestricted Generalized Transition State* (uGTS) de Williams *et al.* [377], généralisation du modèle *Transition State* (TS) de Slater [378]. Il consiste en un développement en série de Taylor de la différence d'énergie entre les deux états qui est fonction d'une variable continue λ . Les termes manquants sont déterminés grâce au théorème de Janak [379].

Chapitre 6. Etats Ionisés et Excités de Cœur dans les Macromolécules

FIG. 6.4 – Géométries des molécules d’alanine (deux conformations), des poly-Ala-tripeptides (conformations en hélice α et en brin β) et du poly-Ala-pentadécapeptide (conformations en hélice α et en brin β).





Pour la molécule d’alanine en phase gazeuse, nous remarquons que les énergies d’ionisation des trois atomes de carbone sont en très bon accord avec l’expérience, et ce pour chaque conformation. Il est donc impossible de se prononcer en faveur d’une conformation. Cependant, nous sommes tout à fait confiants sur la qualité des résultats pour ce système.

Pour le tripeptide, nous observons que les énergies d’ionisation sont très proches pour chaque conformation. Le seul changement majeur correspond à l’atome de carbone du groupement carbonyle pour lequel un déplacement de 1,5 eV est observé par rapport à la phase gazeuse. Ceci est facilement compréhensible sachant que l’environnement direct de cet atome de carbone est initialement un groupement carboxylique, puis un groupement amide. Par conséquent, cette valeur de l’énergie d’ionisation est très proche de celle calculée dans le cas de la liaison peptidique du composé N-méthylacétamide (TAB. 6.4).

En examinant les résultats obtenus dans le cas de la méthode LSCF/MM, nous pouvons conclure que l’environnement a un effet très faible. Ce résultat pouvait être anticipé car l’état fondamental est neutre, et donc faiblement polarisé par la partie classique. Ceci peut être vérifié en examinant la dernière colonne du tableau 6.5, où nous avons regroupé les énergies d’ionisation obtenues dans le cas de la méthode LSCF/MM avec les charges ponctuelles classiques mises à zéro (c’est-à-dire sans *electrostatic embedding*). Que ce soit en présence de charges classiques ou non, les résultats QM/MM sont très proches. La principale tendance que nous pouvons extraire de ces calculs est que l’*electrostatic embedding* augmente les énergies d’ionisation de 0,2-1,0 eV. Il est tout de même à noter que les liaisons hydrogène de l’hélice α n’induisent pas un déplacement significatif sur les énergies d’ionisation de l’orbitale 1s de l’atome de carbone du groupement carbonyle.

6.4 Conclusion

Dans le cadre de la méthode LSCF, nous avons développé une nouvelle approche pour calculer les états de cœur ionisés et excités de molécules organiques et de systèmes macromoléculaires. Ce nouveau schéma théorique permet de traiter de façon rigoureuse l’orthogonalité entre l’état fondamental et les états excités, grâce à une modification de la procédure d’orthogonalisation de la méthode LSCF originale [74, 214]. La validité de la méthode a été établie sur la molécule de pentane par comparaison entre l’ancienne (Löwdin) et la nouvelle procédure (Lv+GS). Concernant la description de l’orbitale de cœur, les bases de Pople triple- ζ permettent d’obtenir des résultats très satisfaisants. De plus, nous avons appliqué la méthode à un grand nombre de composés organiques,

TAB. 6.5 – Energies d'ionisation de l'orbitale $1s$ du carbone (en eV) de la molécule d'alanine isolée, engagée dans un tripeptide ou dans un pentadécapeptide en conformation hélice α et brin β . Voir FIG. 6.4 pour les géométries. Les calculs notés AMBER* sont effectués sans polarisation électrostatique de la fonction d'onde quantique par les charges ponctuelles classiques.

	Alanine			Tripeptide		Pentadécapeptide			
	PBE0		Exp.	PBE0		PBE0/AMBER		PBE0/AMBER*	
	Conf. 1	Conf. 2		Hélice α	Brin β	Hélice α	Brin β	Hélice α	Brin β
C_α	292,45	292,16	292,30	291,87	292,00	292,13	292,62	291,82	291,92
C_{carb}	294,60	294,92	295,30	293,69	293,87	293,78	294,63	293,06	293,64
C_{Me}	291,11	290,89	291,20	290,50	290,62	291,21	290,94	290,97	290,51

et particulièrement à un groupe de composés comportant une liaison peptidique. Les résultats ont été comparés aux références expérimentales, ainsi qu'à d'autres méthodes théoriques. Nous avons montré que la CO obtenue à l'aide du critère BF, associée avec la méthodologie PBE0/6-311++G**//B3LYP/6-311++G** permet d'obtenir des résultats précis.

Conformément à ces conclusions et à une étude préliminaire de la molécule d'alanine en tant que mono et tripeptide, nous avons calculé les énergies d'ionisation de l'orbitale $1s$ du carbone pour un pentadécapeptide en conformations hélice α et brin β dans le cadre de la méthode LSCF/MM. Nous concluons que les effets de l'environnement MM sur les énergies d'ionisation sont relativement faibles comparés aux études précédentes [231]. Cependant, l'*electrostatic embedding* doit être pris en compte car il augmente les énergies d'ionisation de 0,2-1,0 eV. Dans les travaux ultérieurs, la polarisation de la partie MM devra être prise en compte.

Comme perspective, nous sommes en ce moment en train d'étudier l'effet de l'erreur de self-interaction dans les fonctionnelles DFT [365]. En effet, dans le cas des électrons de cœur, ce phénomène peut avoir des effets très importants dû au caractère confinées de la densité électronique dans ces orbitales.

7

Spectroscopie UV-vis de Macromolécules

Publications

- **Theoretical investigation of the geometries and UV-vis spectra of Poly(L-glutamic acid) featuring a photochromic azobenzene side chain**
P.-F. Loos, J. Preat, A. D. Laurent, C. Michaux, D. Jacquemin, E. A. Perpète, X. Assfeld
J. Chem. Theor. Comput. **ASAP Article**
- **A TD-DFT investigation of UV spectra of pyranoïdic dyes : a NCM vs. PCM comparison**
J. Preat, P.-F. Loos, X. Assfeld, D. Jacquemin, E. A. Perpète
J. Mol. Struct. (THEOCHEM), **2007**, *808*, 85–91
- **DFT and TD-DFT investigation of IR and UV spectra of solvated molecules : Comparison of two SCRF continuum models**
J. Preat, P.-F. Loos ; X. Assfeld, D. Jacquemin, E. A. Perpète
Int. J. Quant. Chem., **2007**, *107*, 574–585

7.1 Chromisme

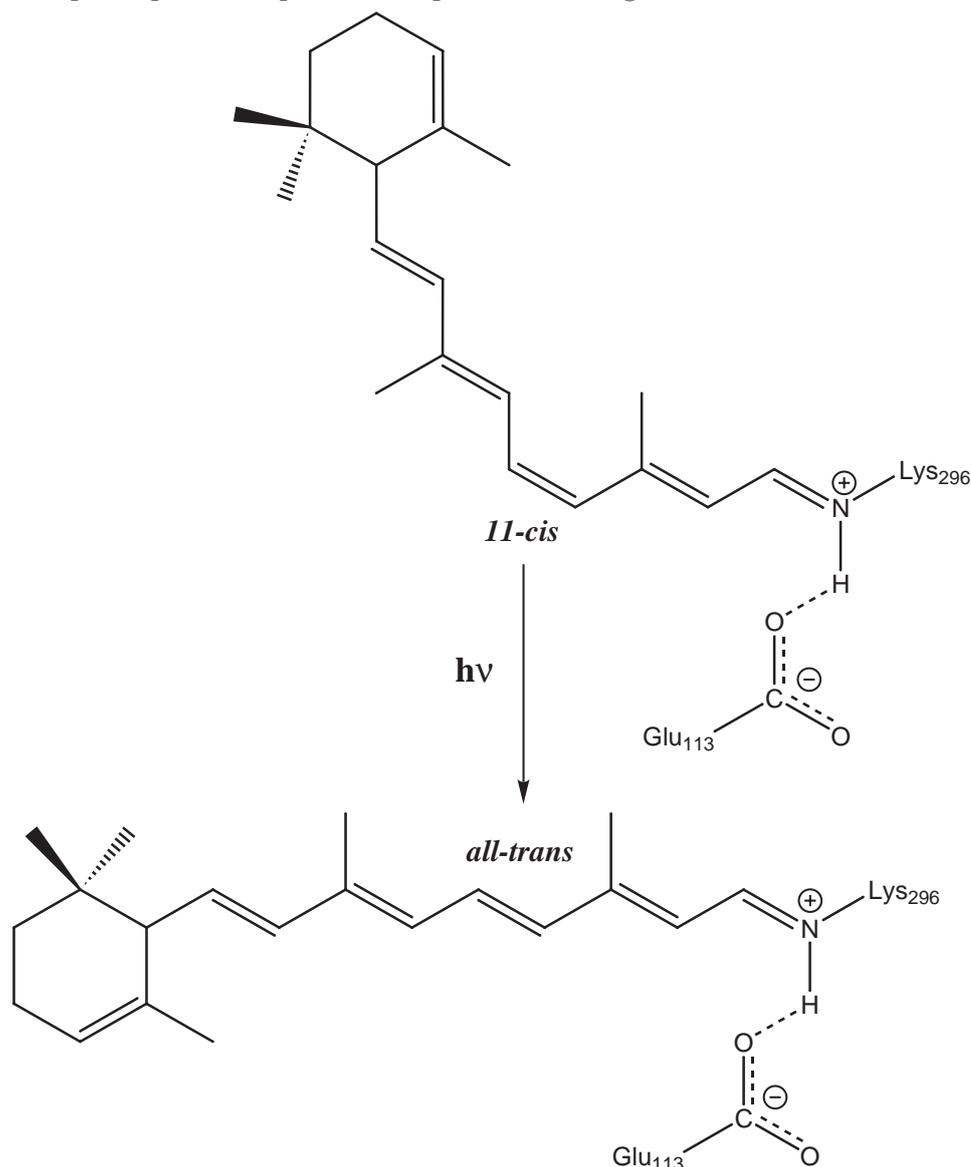
Le chromisme est un terme général qui correspond à un changement réversible de la couleur d'un composé vis-à-vis d'un stimulus externe par modification de sa distribution électronique. Il existe plusieurs types de chromisme attribués à différents phénomènes. Le photochromisme est la transformation réversible entre deux formes d'une même substance chimique (appelé chromophore) par absorption d'une onde électromagnétique, où les deux formes présentent des spectres d'absorption différents (variation de la couleur du composé par irradiation). Il est caractérisé par son rendement quantique qui correspond à l'efficacité du processus d'absorption, par la résistance à la fatigue de la substance chimique, ainsi

Chapitre 7. Spectroscopie UV-vis de Macromolécules

que par la nature des états photostationnaires.

Par exemple, la molécule de rétinale, polyène dérivé de la vitamine A, est le chromophore permettant la transduction de la lumière en signal visuel. Ces deux états photostationnaires sont les formes *11-cis* (rhodopsine) et *all-trans* (bathorhodopsin) qui correspondent à l'isomérisation de la double liaison C=C 11 de sa forme *cis* à sa forme *trans* (FIG. 7.1).

FIG. 7.1 – Photoisomérisation de la molécule de rétinale de la forme *11-cis* en *all-trans* dans la rhodopsine par absorption d'un photon d'énergie $h\nu$.

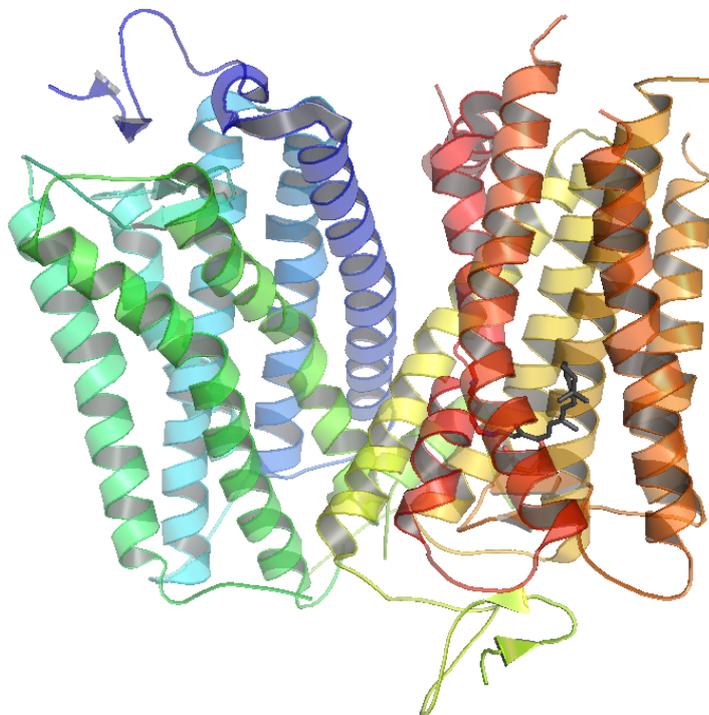


Le rétinale est présent dans les cellules photoréceptrices (bâtonnets) de la rétine qui initie les premières étapes du processus de la vision, et en particulier pour la vision nocturne. Il est le cofacteur de la rhodopsine appartenant à la famille des récepteurs couplés à une protéine G, qui est constituée d'une protéine (l'opsin) et d'un chromophore (le rétinale). Ce

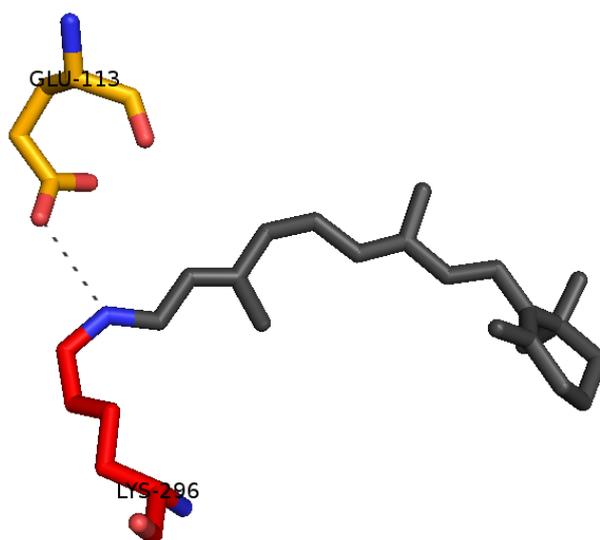
dernier est lié de façon covalente à la lysine 239 de l'opsin *via* une base de Schiff protonée, en interaction avec l'acide glutamique 113 servant de contre-ion (FIG. 7.1 et 7.2).

Chapitre 7. Spectroscopie UV-vis de Macromolécules

FIG. 7.2 – Structure obtenue par diffraction des rayons X (résolution 2,80 Å) de la rhodopsine bovine [381] (d'après l'entrée de la PDB 1F88). La molécule de rétinol (cofacteur) est représentée en gris foncé.



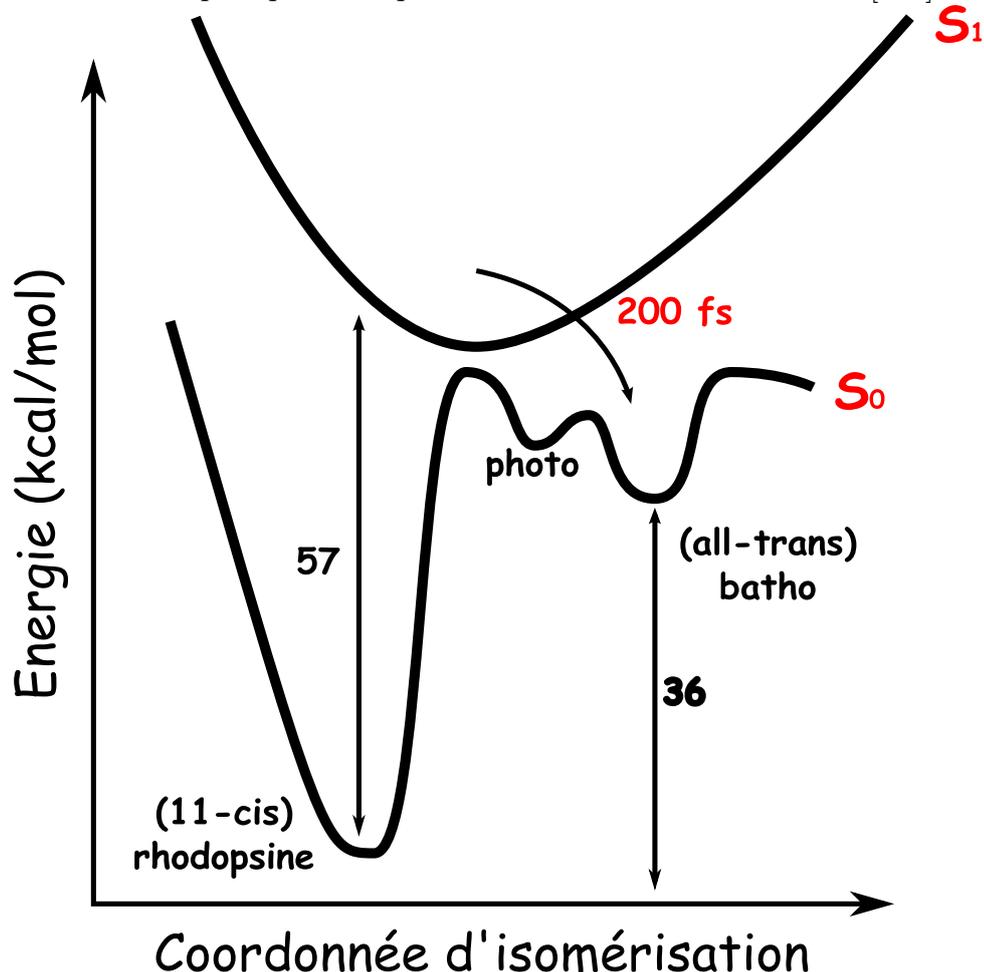
(a) Vue d'ensemble



(b) Site actif : rétinol, Lys 296 et Glu 113 (contre-ion)

D'un point de vue photochimique, le processus de photoisomérisation *cis-trans* correspond à une transition Frank-Condon de l'état fondamental singulet S_0 du rétinale vers son premier état excité singulet S_1 par absorption d'un photon à 498 nm (forme rhodopsine) [382,383]. Il s'en suit alors une relaxation en 200 fs (avec un rendement quantique de 0.67 [384]) vers la forme bathorhodopsine de l'état S_0 ²³. Ce processus permet de stocker une énergie de 36 kcal/mol [387], ce qui correspond au deux tiers de l'énergie du photon absorbé (57 kcal/mol). Ce processus photochimique est représenté sur la figure 7.3.

FIG. 7.3 – Premières étapes du processus photochimique initiant la vision : photoisomérisation de la forme 11-*cis* en *all-trans* par absorption d'un photon, puis formation de l'intermédiaire bathorhodopsin (batho) en 200 fs. L'intermédiaire photorhodopsine (photo) est détectable par spectroscopie Raman femtoseconde-stimulée [386].



La solvatochromie illustre la capacité d'une substance chimique à modifier sa couleur du fait d'un changement de polarité du solvant. Lorsque cette solvatochromie implique un déplacement vers les faibles longueurs d'onde en accord avec une augmentation de la

²³La forme intermédiaire (photorhodopsine) est très difficile à observer [385,386].

constante diélectrique du solvant, on parle alors de solvatochromie positive ou effet hypsochrome. Lorsque ce déplacement s'effectue vers les longueurs d'onde plus importante, il est alors nommé solvatochromie négative ou effet bathochrome. La solvatochromie a une influence importante sur l'absorption et l'émission, ainsi que sur les propriétés vibrationnelles [388, 389]. En effet, l'interaction soluté-solvant implique une stabilisation différente de l'état fondamental vis-à-vis de ces états excités, conséquence de la différence des moments dipolaires dans ces états. Il en résulte une modification de l'écart énergétique (variation de la position, de l'intensité et de la forme du pic d'absorption ou d'émission).

Par exemple, le composé 4,4'-bis(diméthylamino)fuchsonne présente un effet bathochromique (solvatochromie positive) car il est orange dans un solvant non polaire, tel que le toluène ($\epsilon = 2$), et devient rouge dans l'acétone ($\epsilon = 21$), ainsi que rouge-violet dans le méthanol ($\epsilon = 33$). Un exemple de solvatochromie négative (effet hypsochromique) provient du composé 2-(4'-hydroxystyryl)-N-méthyl-quinolinium bétaine, qui est une encre bleue dans le chloroforme ($\epsilon = 5$) et devient rouge-sang dans l'eau ($\epsilon = 78$).

Il est aussi important de noter que l'environnement biologique peut avoir un effet très important sur les longueurs d'onde d'absorption et d'émission d'un chromophore. Ceci est principalement dû au champ électrique non-uniforme créé par la protéine, ainsi que les interactions spécifiques du chromophore avec les acides aminés voisins et les molécules d'eau présentes dans le site actif. Pour illustrer ce phénomène, revenons au cas de la molécule de rétinol : sa longueur d'onde d'absorption dans le méthanol est de 445 nm. Elle subit un déplacement de 53 nm vers les grandes longueurs d'onde lorsque le rétinol est placé à l'intérieur de la rhodopsin (498 nm) [390]. Cet effet est appelé *opsin shift*.

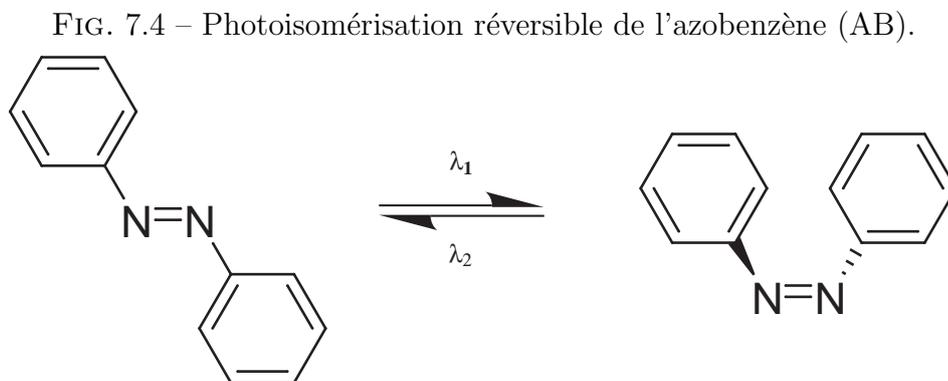
Dans la suite de ce chapitre, nous allons étudier ces différents facteurs (solvatochromie et influence du milieu biologique) sur les longueurs d'onde d'absorption UV-vis des formes *trans* et *cis* du chromophore azobenzène (AB).

7.2 Contexte et Problématique

Durant les vingt dernières années, l'intérêt pour les composés dérivés de l'AB a incroyablement augmenté, dû principalement à leurs importantes gammes d'applications industrielles. En particulier, les composés basés sur des motifs AB représentent 60 à 70 % de la production mondiale de colorants en absorption [391]. De plus, ils ont récemment été découverts comme étant des composés prometteurs en tant que médias périphériques de stockage [392–395] ou en tant que moteurs moléculaires [396, 397]. Ces procédés exploitent la capacité des colorants AB à se photoisomériser de façon réversible de la forme *trans* à la forme *cis* (FIG. 7.4). Dans les systèmes biologiques, l'isomérisation photoréversible d'une molécule attachée sur un système macromoléculaire, comme le pigment visuel rhodopsine [197, 342, 343, 383, 398–406], induit des changements conformationnels de la macromolécule, qui en réponse, conduit à une réponse physiologique de la protéine. Suivant cette idée, Pieroni et collaborateurs ont synthétisé des Poly(acide L-glutamique) contenant des chaînes latérales contenant un chromophore de type AB [407–409] (FIG. 7.5). Ces structures peuvent exister sous forme désordonnée (*random coil*) ou en structures ordonnées (comme les hélices α ou les feuilletts β), similaire aux systèmes biologiques. Ce biopolymère photochromique a été préparé à partir de Poly(acide L-glutamique) de haut

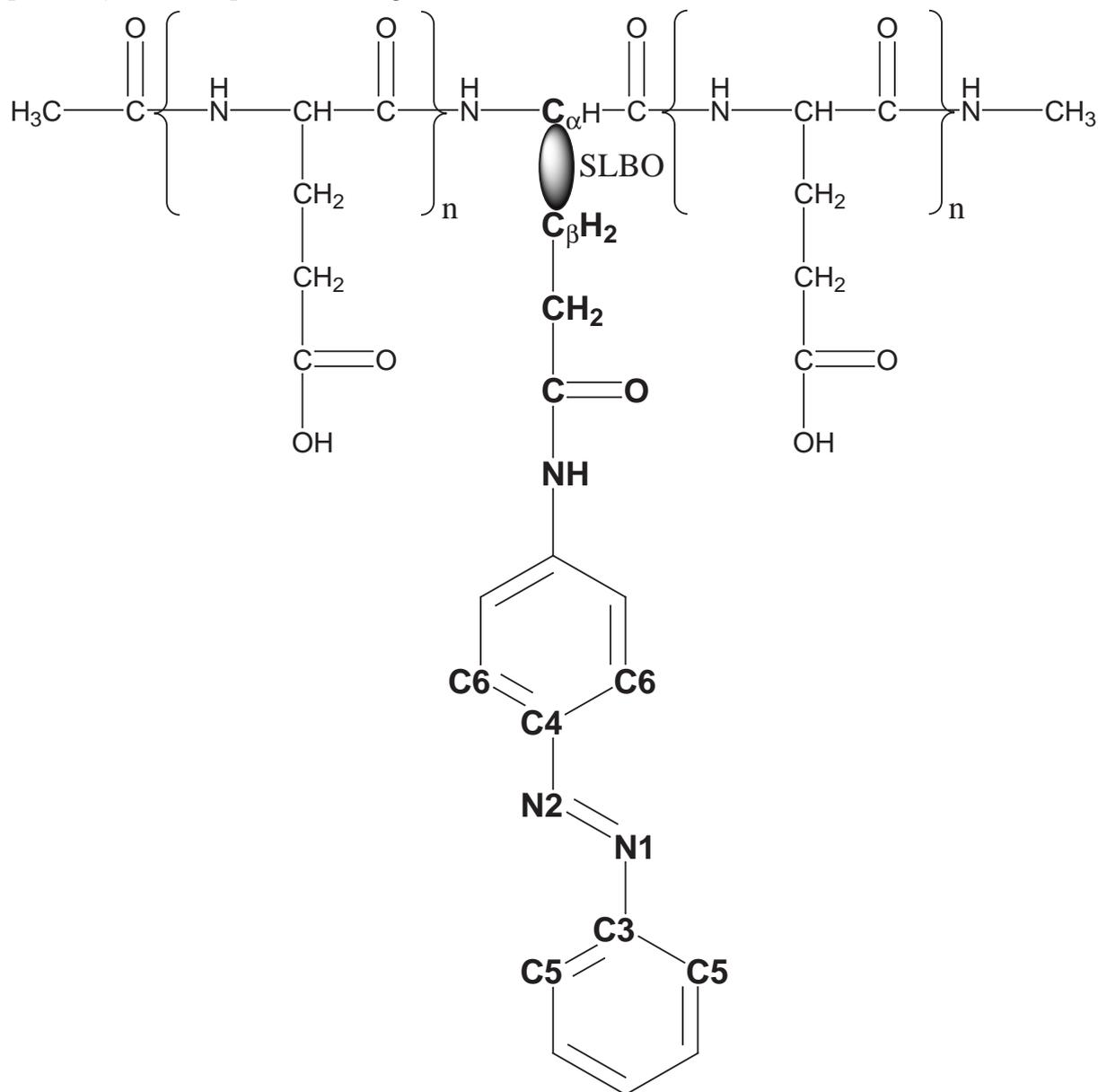
7.2. Contexte et Problématique

poids moléculaire ($\bar{M}_v = 200\ 000$) et des échantillons contenant de 13 à 56 % molaire de groupe AB ont été étudiés à différents pH. Il a été montré que l'irradiation à 350 nm produit une isomérisation de la forme *trans* à la forme *cis* (transition $\pi \rightarrow \pi^*$), tandis que la réaction inverse est obtenue par irradiation à 450 nm (transition $n \rightarrow \pi^*$), ou en laissant l'échantillon dans le noir [407].



Chapitre 7. Spectroscopie UV-vis de Macromolécules

FIG. 7.5 – Partition QM/MM du Poly(acide L-glutamique) contenant une chaîne latérale AB. La frontière QM/MM est localisée entre les atomes C_α et C_β et les atomes de la partie QM sont représentés en gras.



Dans l'eau, ce processus réversible est accompagné par des changements importants de la structure secondaire, détectés par dichroïsme circulaire (DC), ainsi que par une modification drastique du spectre d'absorption du colorant AB. Selon les conditions expérimentales (pH et % molaire d'AB), le polypeptide présente le spectre de DC d'une structure désordonnée (pelote statistique), α ou β selon les pourcentages relatifs de chaque structure [410]. A pH acide (4,7 – 5,0), l'échantillon de la forme *trans* contenant 16 et 21 % molaire de groupes AB indique la présence d'un taux prédominant d'hélice α , tandis

que l'échantillon à 36 % molaire présente la courbe de DC caractéristique d'une structure β . Lorsque le pH augmente (pH 8), tous les échantillons subissent une transition vers des formes désordonnées. Dans l'eau, l'irradiation produit l'isomérisation des chaînes latérales AB, et un effet remarquable sur les bandes de DC, qui sont aussi influencées par le pH et/ou le pourcentage de groupements AB. Par exemple, l'échantillon à 16 % molaire ne montre pas de variations de son spectre de DC, tandis que l'échantillon à 36 % molaire subit une transition β -désordonnée à pH 6,5. Pour l'échantillon à 21 % molaire, le pourcentage en structures α est augmenté par irradiation en dessous de pH 6,3.

Avant d'étudier les effets des conditions expérimentales (pH et % molaire d'AB), ainsi que l'aspect dynamique de ce processus conduisant à la modification de la conformation du polypeptide, il est nécessaire de comprendre les effets de la structure du polypeptide sur le spectre d'absorption UV-vis des deux isomères de l'AB (notés TAB et CAB).

Afin d'obtenir des spectres UV-vis de systèmes macromoléculaires de bonne qualité, nous avons besoin d'assurer deux points majeurs :

- La modélisation de macromolécules à l'aide de mécanique quantique uniquement est encore hors de portée des ressources informatiques actuelles.
- Pour obtenir un spectre UV-vis de bonne qualité, nous devons décrire de façon précise les états excités du chromophore.

Des propositions ont été faites pour résoudre ces problèmes un à un. Par exemple, le problème de la taille du système peut être contourné en utilisant une méthode QM/MM [65, 66, 183, 204–207, 216, 222, 225, 226]. Pour le deuxième problème, les calculs TD-DFT [133] permettent d'obtenir une détermination précise des énergies d'absorption associées à ces états excités pour un large panel de chromophores [388, 389, 411–414], et spécialement dans le cas des composés AB [415–417].

Nous remarquerons qu'un nombre restreint d'études combinant les calculs QM/MM et TD-DFT ont été précédemment publiées, comme les travaux de Elstner et collaborateurs [342, 343] et ceux de Vreven et collaborateurs [197], qui ont comparé plusieurs approches méthodologiques pour la détermination des longueurs d'onde d'absorption du chromophore rétinol. Rothlisberger *et al.* ont, quant à eux, reporté des calculs TD-DFT:MM [347, 349] basés sur des simulations Car-Parrinello, sur de l'acétone [346] et des aminocoumarines [345, 350] solvatées, ainsi que des résultats de calculs TD-DFT/MM sur le chromophore rétinol [344, 348]. Bien que Elstner *et al.* aient montré que les calculs TD-DFT ne donnaient pas de bons résultats dans le cas de la molécule de rétinol protonée et des polyènes neutres [342, 343], du fait de l'approximation locale de la fonctionnelle d'échange-corrélation [138], nous voudrions souligner les travaux de Crecca et Roitberg sur l'étude du mécanisme d'isomérisation de l'azobenzène et ses dérivés disubstitués [415]. En utilisant des calculs DFT et TD-DFT, ils ont étudié le chemin d'isomérisation de plusieurs dérivés de l'AB. Les différentes barrières et la topologie des surfaces des différents états électroniques sont en bon accord avec les résultats expérimentaux, ainsi qu'avec d'autres méthodes théoriques.

La TD-DFT n'est pas la seule méthode disponible pour obtenir une description qualitative des états excités. Les méthodes multiconfigurationnelles MCSCF [418, 419] (*Multi-configuration Self-Consistent Field*), et en particulier la méthode CASSCF [420] (*Complete Active Space Self-Consistent Field*), permettent d'obtenir une excellente description des états excités. Lorsque les fonctions d'onde CASSCF de l'état fondamental et des premiers

états excités sont utilisées comme déterminants de référence dans la théorie des perturbations au second ordre [421] (CASPT2 pour (*Complete Active Space Perturbation Theory 2*)), il est couramment reconnu que les résultats au niveau de la topologie des surfaces d'énergie potentielle, ainsi que les différentes barrières énergétiques sont en très bon accord avec l'expérience. Par exemple, Olivucci *et al.* ont développé, dans le cas de la bactériorhodopsin [197,342,343,383,398–406], un protocole CASPT2//CASSCF/AMBER permettant de déterminer les énergies des maxima d'absorption et de fluorescence à quelques kcal/mol des valeurs expérimentales. Cette méthodologie a également été utilisée avec succès lors de l'étude des propriétés d'émission de la protéine GFP (*Green Fluorescent Protein*) [422]. Récemment, des calculs CASPT2//CASSCF/6-31G* ont été reportés pour les premiers états singulets $n \rightarrow \pi^*$ et $\pi \rightarrow \pi^*$ de l'azobenzène afin d'élucider son mécanisme de photoisomérisation. Cependant, ces calculs sont très demandeurs en mémoire et en temps de calcul (14 électrons dans 12 orbitales), ce qui limite leur utilisation.

Dans ce chapitre, nous reportons donc le calcul des spectres UV-vis des formes TAB et CAB pour différentes structures secondaires et supersecondaires (motifs) de Poly(acide L-glutamique). Premièrement, nous étudierons les géométries et les spectres UV-vis des deux isomères en phase gazeuse et dans l'éthanol. Les résultats seront comparés aux valeurs expérimentales pour valider le schéma théorique. Ils serviront aussi de valeurs de référence afin de faire ressortir les effets spécifiques de l'environnement du polypeptide sur la géométrie et les longueurs d'onde d'absorption de la partie AB. En particulier, nous étudierons différentes structures du polypeptides contenant des hélices α et des brins β , tout comme d'autres structures (épingle à cheveux β , feuillet β , motif β - α - β et α - β - α).

7.3 Méthodologie

Les calculs QM/MM sont effectués grâce au formalisme LSCF [70, 74, 214, 215, 231, 232, 234, 237, 238]. La SLBO a été obtenue par un calcul préliminaire sur la molécule d'éthane [235]. Dans le cas présent, la frontière QM/MM est placée entre les atomes C_α et C_β de l'acide aminé sur lequel l'unité AB est greffée (FIG. 7.5).

La partie MM est décrite à l'aide du champ de forces AMBER *ff99* [144, 147] dans lequel nous avons considéré la forme protonée de l'acide glutamique dans le but d'éviter la polarisation artificielle due à un trop grand nombre de charges négatives. Nous avons donc choisi de modéliser l'effet des contre-ions par la forme protonée des chaînes glutamiques latérales. Cette stratégie a été employée dans plusieurs études dans le cadre des systèmes biologiques [423–425] pour rendre compte de la situation où le contre-ion est intimement lié à la partie chargée négativement. La charge classique de l'atome frontière C_α a été fixée à 0,0365 e au lieu de 0.0145 e pour assurer l'électronneutralité globale du système. Les paramètres van der Waals des atomes QM ont été choisis comme étant égaux aux valeurs définies pour les types d'atomes correspondants dans le champ de forces. De plus, les acides aminés N- et C-terminaux sont protégés par des groupements acétyle et N-méthylacétamide, respectivement.

Concernant les calculs QM, la fonctionnelle B3LYP [128] combinée avec la base 6-311G** a été utilisée pour les optimisations de géométrie du chromophore. Après une optimisation de géométrie de l'ensemble du système, nous avons effectué des calculs TD-

DFT pour évaluer les spectres UV-vis à l'aide de la fonctionnelle PBE0 [132, 426] et la base 6-311+G**. Des études préliminaires ont montré que pour la majorité des colorants organiques [388, 389, 411–414, 416, 417], ce niveau de théorie fournit des résultats fiables pour les géométries, ainsi que pour les spectres UV-vis. Pour prendre en compte les effets du solvant sur les systèmes modèles, le modèle de Continuum Polarisable (PCM) a été utilisé [42]. Pour les calculs TD-DFT utilisant le champ de réaction auto-cohérent comme modèle de solvatation, la méthode PCM où le solvant n'est pas en équilibre thermodynamique avec la solution (PCM non-équilibre) a été sélectionnée [427].

7.4 Résultats

7.4.1 Azobenzène en phase gazeuse et dans l'éthanol

La table 7.1 regroupe les paramètres structuraux et les spectres UV-vis des formes TAB et CAB en phase gazeuse et dans l'éthanol. Comme expliqué dans une étude précédente [416], une comparaison directe de nos résultats par rapport aux données expérimentales est très difficile car les structures de diffraction des rayons X (DRX) sont déterminées en phase solide, tandis que les mesures de diffraction d'électron en phase gazeuse (DEG) sont effectuées à des températures relativement élevées. Néanmoins, concernant l'isomère TAB en phase gazeuse, les résultats B3LYP donnent une déviation maximale de 0,013 Å pour les distances de liaison et de 1,4° pour les angles de valence si l'on considère les expériences les plus récentes de DRX et DEG. De plus, les mesures DEG prédisent une géométrie parfaitement plane de la forme TAB en phase gazeuse. Pour la forme CAB, des mesures DRX à basse température, ou des expériences de DEG ne sont pas disponibles dans la littérature. Cependant, la structure DRX, obtenue à haute température, de la Ref. [428] est en bon accord avec les résultats DFT, avec une déviation maximale de 0,013 Å, 2,2° et 2,3° pour les distances de liaison, les angles et les dièdres, respectivement.

Expérimentalement, les spectres d'absorption UV-vis présentent une bande de faible amplitude ($n \rightarrow \pi^*$) entre 380 et 520 nm, qui est une transition interdite par symétrie dans le cas de l'isomère *trans*. Une transition $\pi \rightarrow \pi^*$ autour de 330 nm pour la forme TAB, et 275 nm pour la forme CAB [429] est également présente.

Pour le spectre UV-vis de la forme TAB, les énergies d'absorption TD-PBE0 sont en bon accord avec les valeurs expérimentales. Bien que les énergies d'excitation soient sous-estimées de 20-40 nm en phase gazeuse et dans l'éthanol, les déplacements hypsochromiques/bathochromiques des transitions $n \rightarrow \pi^*/\pi \rightarrow \pi^*$ sont bien reproduits par les calculs théoriques. Expérimentalement, on note un faible déplacement entre +3 et -7 nm pour la transition $n \rightarrow \pi^*$, à comparer à un déplacement de -3 nm prédit par les calculs TD-PBE0. Pour l'excitation $\pi \rightarrow \pi^*$, nous prédisons un déplacement bathochromique de +17 nm, qui est très proche de +17 et +19 nm obtenu expérimentalement. Nous renvoyons le lecteur à la référence [416] pour plus de détails sur les géométries et les transitions UV-vis de la forme TAB.

Concernant le spectre d'absorption de la forme CAB, les déviations entre les résultats théoriques et expérimentaux sont du même ordre de grandeur que celles obtenues dans le cas de l'isomère *trans*. L'excitation $n \rightarrow \pi^*$ devient alors une transition permise de par le

caractère non-planaire de la forme CAB. De façon théorique, la prise en compte des effets de solvant induit un déplacement hypsochromique de -11 nm et un déplacement bathochromique de +13 nm pour les transitions $n \rightarrow \pi^*$ et $\pi \rightarrow \pi^*$, respectivement. Comparé au composé TAB, le solvant induit une augmentation du déplacement hypsochromique pour la transition $n \rightarrow \pi^*$, tandis que l'énergie d'excitation $\pi \rightarrow \pi^*$ reste inchangée, avec une nette décroissance de la force d'oscillateur. L'effet bathochromique du solvant sur la transition $\pi \rightarrow \pi^*$ est bien reproduite par la présente approche théorique. En effet, les résultats expérimentaux conduisent à un déplacement bathochromique de +16 nm. En revanche, le déplacement vers les grandes longueurs d'onde de 8-18 nm, expérimentalement observé dans le cas de la transition $n \rightarrow \pi^*$, n'est pas reproduit par l'approche TD-PBE0, qui reporte un déplacement vers les courtes longueurs d'onde de -11 nm. Cependant, le spectre UV-vis de la forme CAB est sujet à réserve : il est obtenu à relativement haute température (181-322 °C) par une différence de spectre. Ce spectre, obtenu par technique de photolyse flash, correspond à une différence d'énergie d'absorption entre une forme stable (TAB), préalablement déterminée, et une espèce métastable (CAB) [430].

La table 7.1 reporte aussi les modifications des paramètres géométriques et les longueurs d'onde d'absorption du processus de photoisomérisation du colorant AB (FIG. 7.4). Le changement de la distance C-N est bien reproduit, tandis que la variation des valeurs de $d(\text{N}=\text{N})$ et $\angle(\text{C}-\text{N}=\text{N})$ sont légèrement sous-estimées par l'approche théorique utilisée. La divergence majeure entre théorie et expérience vient de la sous-estimation de la variation de la valeur de $\tau(\text{C}-\text{C}-\text{N}=\text{N})$ par 10-20°. Cependant les angles dièdres sont très sensibles à l'effet de confinement dans le cristal. Néanmoins, les plus récentes expériences de DEG sur le composé TAB donnent une valeur de $\tau(\text{C}-\text{C}-\text{N}=\text{N})$ égale à 0,0°, correspondant à une déviation entre la théorie et l'expérience de seulement 0,2°. Pour les déplacements des longueurs d'onde d'absorption, les approches théoriques et expérimentales prédisent un déplacement vers les faibles longueurs d'onde pour les transitions $n \rightarrow \pi^*$ et $\pi \rightarrow \pi^*$, en phase gazeuse et dans l'éthanol. De plus, les résultats TD-PBE0 fournissent des déplacements énergétiques dans la gamme de ceux expérimentaux, sauf pour le déplacement de l'excitation $n \rightarrow \pi^*$ en phase gazeuse. Ceci pourrait être une fois encore attribué aux résultats expérimentaux de la valeur $\lambda^{n \rightarrow \pi^*}$ obtenue dans le cas de la forme CAB en phase gazeuse.

Conformément à des études théoriques ultérieures, les énergies d'excitation TD-DFT sont plus déplacées vers les grandes longueurs d'onde, comparées à d'autres approches théoriques comme SOPPA²⁴ et CC2 [431], conduisant à une différence plus importante entre les résultats théoriques et les longueurs d'onde expérimentales. Cependant, ces méthodes post Hartree-Fock sont très coûteuses en temps de calcul, et par conséquent très difficiles à utiliser pour des systèmes de taille plus importante, principalement lorsque les effets de solvant doivent être pris en compte. Dans ce cas, les fonctionnelles DFT, et notamment hybrides telles que B3LYP et PBE0, présentent un bon compromis entre précision et temps de calcul. De plus, la méthodologie précédemment employée a prouvé qu'elle permettait de reproduire les déplacements expérimentaux correspondants aux processus

²⁴La méthode SOPPA (*Second Order Polarization Propagator Approximation*) correspond à un développement perturbatif au second ordre du propagateur de polarisation grâce à une fonction de Green à deux particules. L'ordre zéro de ce développement correspondant à la méthode RPA (*Random Phase Approximation*) ou TDHF (*Time-Dependent Hartree-Fock*).

de solvatochromisme et de photoisomérisation pour le colorant AB. Nous allons donc utiliser cette méthodologie afin d'étudier l'effet de la structure secondaire d'un polypeptide sur le spectre d'absorption de la partie AB.

7.4.2 Poly(acide L-glutamique) avec une chaîne latérale AB

Comme la méthode LSCF/MM requiert une SLBO pour connecter la partie QM et MM, il est important de s'assurer que celle-ci ne perturbe pas les propriétés de la partie QM. L'influence de la SLBO est analysée dans le but de vérifier qu'aucun déplacement artificiel n'est induit sur les énergies d'excitation TD-DFT²⁵. Nous avons donc calculé les longueurs d'onde d'absorption de la partie QM, complétée avec un fragment méthyle à la place de la partie MM. Les calculs sont effectués au niveau QM, avec et sans SLBO (FIG. 7.5). Ces calculs préliminaires montrent que la SLBO crée une erreur de 0,09 et de 0,22 nm sur les transitions $n \rightarrow \pi^*$ et $\pi \rightarrow \pi^*$, respectivement. De plus, une précédente étude a démontré que l'introduction d'une SLBO n'induit que de très faibles variations de la géométrie QM [235].

Les paramètres géométriques des Poly(acide L-glutamique) contenant une chaîne latérale AB sont listés dans la table 7.2. Nous avons aussi reporté les géométries optimisées et le nombre d'acides aminés glutamiques correspondant aux différentes structures secondaires (voir FIG. 7.6 pour la forme TAB et FIG. 7.7 pour la forme CAB). Du fait de l'importance de la double liaison N=N correspondant au chromophore de l'AB, nous étudierons premièrement ces modifications. Les différents environnements MM impliquent un allongement systématique de 0,002-0,004 Å, comparé à la structure isolée de la forme TAB en phase gazeuse. Concernant la forme CAB greffée sur le Poly(acide L-glutamique), cette augmentation de la liaison N=N n'est pas systématique. En effet, pour la conformation en brin, en feuillet et en épingle à cheveux β , la valeur $d(N1=N2)$ est égale ou inférieure (0,001 Å) comparé à la valeur de la forme CAB isolée en phase gazeuse. Pour les autres structures secondaires, la partie MM induit une augmentation de la distance N=N par 0,002-0,005 Å. Des effets similaires sont induits par le solvant pour les deux isomères (TAB. 7.1). Les distances C-N sont aussi affectées par la structure du polypeptide, ainsi que les angles autour de l'unité N=N du chromophore, et ceci quelques soient les conformères considérés. Contrairement aux autres paramètres géométriques, où l'effet est très faible, les angles dièdres sont nettement plus affectés par la structure du polypeptide, conduisant à des formes non-planes pour la forme TAB. Par exemple, la conformation en hélice α , ainsi que le motif β - α - β induisent des valeurs de $\tau(N1=N2-C4-C5)$ de 20° environ. Pour le dérivé CAB, les structures contenant des motifs α (hélice α , motif α - β - α et β - α - β) conduisent à de larges modifications de la valeur de $\tau(N1=N2-C4-C6)$, mais aussi à un allongement de la distance de la double liaison N=N, comme précédemment mis en avant.

²⁵Dans cette partie, les calculs TD-DFT ont été effectués suivant l'approximation FC (*Frozen Core*), n'incluant pas les orbitales de cœur dans le calculs de énergies d'excitation. Nous noterons, de plus, que dans la méthodologie TD-LSCF/MM, les SLBOs ne sont pas prises en compte dans le calcul des énergies d'excitation, tout comme les MOs de cœur. Ceci est du au fait que les SLBOs ne sont pas fonctions propres du Fockien, et que donc, l'introduction de celles-ci dans le calcul TD-DFT impliquerait une modification drastique des équations TD-DFT.

TAB. 7.1 – Structures et spectres UV-vis de la molécule de TAB et CAB dans différents milieux, obtenus au niveau de théorie TD-PBE0/6-311+G**//B3LYP/6-311G**. Les modifications des paramètres structuraux et des spectres UV-vis suivant la photoisomérisation des deux isomères sont aussi reportées. Les distances sont en Å, tandis que les angles et les dièdres sont en degrés. $\lambda^{n \rightarrow \pi^*}$ et $\lambda^{\pi \rightarrow \pi^*}$ sont les longueurs d'onde (en nm) de la première transition $n \rightarrow \pi^*$ et $\pi \rightarrow \pi^*$, respectivement. Les forces d'oscillateurs sont données entre parenthèses.

Milieu	Méthode	d(N=N)	d(C-N)	$\angle(C-N=N)$	$\angle(G-C-N)$	$\tau(C-C-N=N)$	$\lambda^{n \rightarrow \pi^*}$	$\lambda^{\pi \rightarrow \pi^*}$	Ref.
TAB Gaz	TD-DFT//DFT	1,253	1,418	115,1	124,6, 115,5	0,0, 180,0	480(0,00)	327(0,78)	[416]
	DRX 1997 (82 K)	1,259	1,431	114,1	—	21,0	—	—	[432]
	DEG 2001 (407 K)	1,260	1,428	113,7	124,8	0,0, 180,0	—	—	[433]
Ethanol	Exp. 1981	—	—	—	—	—	444	303	[434]
	Exp. 1982	—	—	—	—	—	440(380)	301(21300)	[430]
	TD-DFT//DFT	1,255	1,419	115,5	124,7, 115,5	0,0, 180,0	477(0,00)	344(0,90)	Cette étude
CAB Gaz	Exp. 1953	—	—	—	—	—	443(510)	320(21300)	[435]
	Exp. 1961	—	—	—	—	—	437(510)	320(17300)	[436]
	TD-DFT//DFT	1,243	1,436	124,1	122,9, 116,5	51,0, -138,1	478(0,03)	292(0,08)	Cette étude
Ethanol	DRX 1971	1,253	1,449	121,9	122,5, 117,3	53,3	—	—	[428]
	Exp. 1982	—	—	—	—	—	425	265	[430]
	TD-DFT//DFT	1,246	1,436	124,2	122,7, 116,5	50,8, -138,0	467(0,05)	305(0,09)	Cette étude
TAB Gaz	Exp. 1953	—	—	—	—	—	433(1518)	281(5260)	[435]
	Exp. 1973	—	—	—	—	—	443(1514)	281(5248)	[437]
	TD-DFT//DFT	-0,010	0,018	9,0	-1,7, 1,0	51,0, 41,9	-2	-35	Cette étude
CAB Ethanol	DRX	-0,006	0,018	7,8	—	32,3	—	—	[428, 432]
	Exp.	—	—	—	—	—	-19 \rightarrow -15	-38 \rightarrow -35	[430, 434]
	TD-DFT//DFT	-0,009	0,017	8,7	-2,0, 1,0	50,8, 42,0	-10	-39	Cette étude
Ethanol	Exp.	—	—	—	—	—	-10 \rightarrow +6	-39	[435-437]

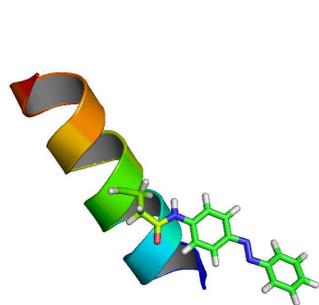
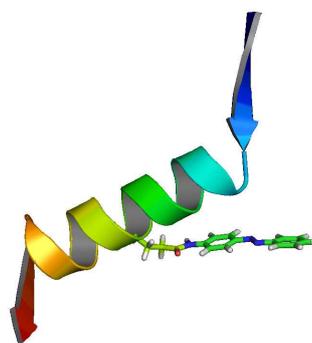
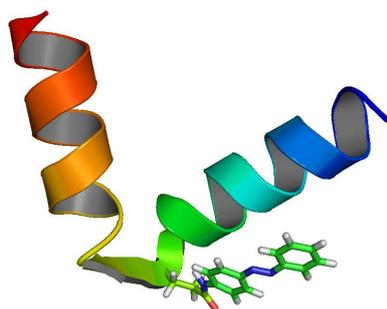
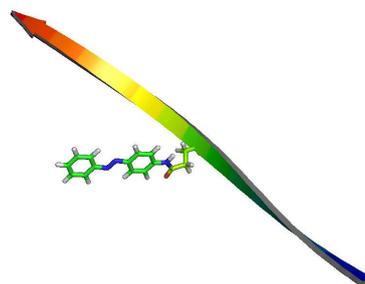
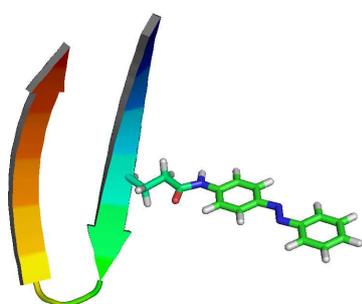
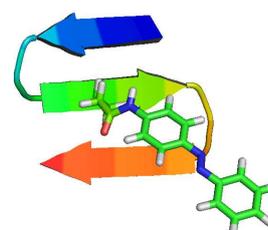
TAB. 7.2 – Paramètres structuraux des molécules TAB et CAB, isolées et greffées comme chaîne latérale de Poly(acide L-glutamique) obtenus au niveau de théorie B3LYP/6-311G**. La partie MM est décrite par le champ de forces AMBER *ff99*. Les distances sont données Å, tandis que les angles et les dièdres sont en degrés. Voir figure 7.5 pour la numérotation des atomes.

Structure	MM	d(N1=N2)	d(C3-N1)	d(N2-C4)	Z(C3-N1=N2)	Z(N1=N2-C4)	Z(C5-C3-N1)	Z(N2-C4-C6)	τ (C5-C3-N1=N2)	τ (N1=N2-C4-C6)	
TAB	—	1,253	1,418	1,418	115,1	115,1	124,6, 115,5	124,6, 115,5	0,0, 180,0	0,0, 180,0	
	Hélice- α	1,255	1,416	1,412	115,2	114,7	124,0, 116,1	124,4, 116,8	-23,5, 159,3	-12,2, 168,3	
	β - α - β	1,257	1,417	1,410	115,2	115,2	124,6, 115,6	124,7, 116,3	-7,8, 174,3	-3,2, -179,6	
	α - β - α	1,255	1,419	1,409	114,3	116,1	124,3, 116,0	125,9, 115,2	-8,0, 173,2	2,4, -175,6	
	Brin β	1,255	1,417	1,412	115,1	115,3	124,7, 115,5	125,2, 115,9	-1,3, 178,5	-0,9, 178,1	
	Epingle β	1,255	1,417	1,411	115,1	115,3	124,7, 115,6	125,1, 115,9	0,3, -179,7	0,2, -179,8	
	Feuillet β	1,255	1,415	1,415	115,1	115,7	124,3, 115,7	124,4, 116,4	17,2, -164,1	-18,3, 165,5	
	CAB	—	1,243	1,436	1,436	124,1	124,1	122,9, 116,5	122,9, 116,5	51,0, -138,1	51,0, -138,1
		Hélice α	1,245	1,433	1,430	124,4	125,1	122,0, 117,4	126,3, 114,6	-57,6, 131,4	-35,2, 154,0
		β - α - β	1,248	1,428	1,427	125,7	127,9	122,5, 117,1	129,2, 112,5	-56,8, 133,0	-20,6, 166,6
α - β - α		1,246	1,431	1,429	125,5	126,2	121,2, 118,2	127,9, 113,5	65,2, -124,4	21,0, -166,9	
Brin β		1,242	1,441	1,434	123,2	123,5	119,8, 119,4	122,1, 118,1	73,5, -116,7	56,4, -133,9	
Epingle β		1,242	1,437	1,437	122,8	123,5	119,7, 119,4	121,3, 118,6	64,8, -125,6	58,8, -133,1	
Feuillet β		1,243	1,431	1,436	123,5	125,0	122,9, 116,4	122,5, 117,4	51,9, -136,0	51,3, -139,4	

Chapitre 7. Spectroscopie UV-vis de Macromolécules

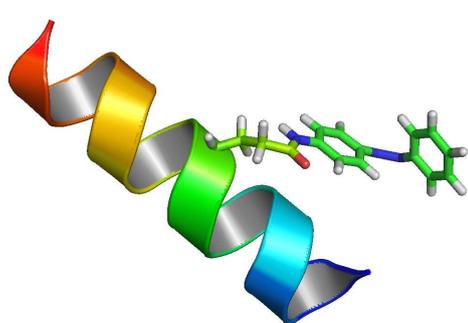
Les spectres UV-vis des Poly(acide L-glutamique) contenant des chaînes latérales AB sont reportés dans la table 7.3. Pour la forme TAB, un déplacement global vers les faibles longueurs d'onde (3-29 nm) et vers les grandes longueurs d'onde (23-36 nm) sont observés pour les transitions $n \rightarrow \pi^*$ et $\pi \rightarrow \pi^*$, respectivement, par rapport à la forme TAB isolée en phase gazeuse (valeurs $\Delta\lambda^{\text{AB}}$). Ceci est en accord avec les études précédentes sur les dérivés de la forme TAB [417]. Il faut noter que la transition $n \rightarrow \pi^*$ n'est plus interdite par symétrie, et possède d'ailleurs des forces d'oscillateur significatives (spécialement pour les conformations en hélice α et le motif β - α - β), bien que restant faibles comparées à celles de la transition $\pi \rightarrow \pi^*$. Le déplacement global de la bande $\pi \rightarrow \pi^*$ est aussi observé de la cas de la forme CAB (21-32 nm). Cependant, ce déplacement vers les faibles longueurs d'onde n'est pas systématique dans le cas de la bande $n \rightarrow \pi^*$: les conformères contenant principalement des motifs β (brin, épingle à cheveux et feuillet β) impliquent un déplacement vers les faibles longueurs d'onde (-20, -15 et -21 nm, respectivement), tandis que les conformations contenant des motifs β produisent un déplacement vers les longueurs d'onde plus importantes de 10-14 nm.

FIG. 7.6 – Géométries QM/MM optimisées des Poly(acide L-glutamique) contenant une chaîne photochromique TAB. La partie QM est représentée en bâtons, tandis que la partie MM est représentée en ruban. Nous reportons le nombre d'acides aminés de chaque structure entre parenthèses.

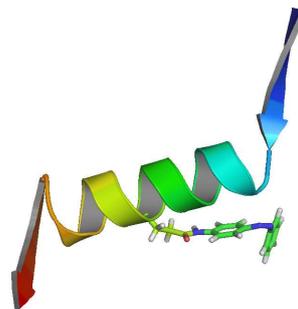
(a) Hélice α (15)(b) Motif β - α - β (24)(c) Motif α - β - α (29)(d) Brin β (15)(e) Epingle à cheveux β (15)(f) Feuillet β (16)

Chapitre 7. Spectroscopie UV-vis de Macromolécules

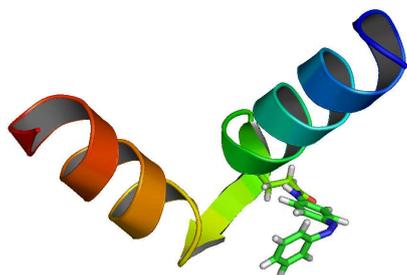
FIG. 7.7 – Géométries QM/MM optimisées des Poly(acide L-glutamique) contenant une chaîne photochromique CAB. La partie QM est représentée en bâtons, tandis que la partie MM est représentée en ruban. Nous reportons le nombre d'acides aminés de chaque structure entre parenthèses.



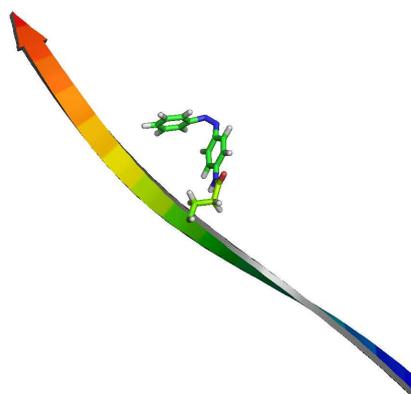
(a) Hélice α (15)



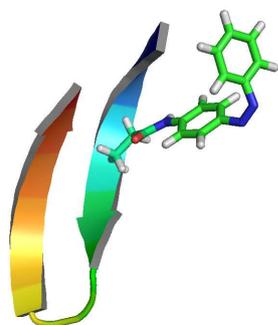
(b) Motif β - α - β (24)



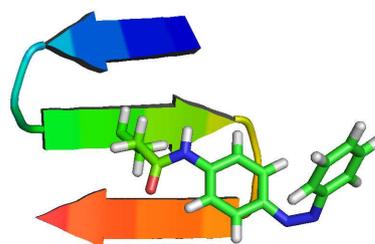
(c) Motif α - β - α (29)



(d) Brin β (15)



(e) Epingle à cheveux β (15)



(f) Feuillet β (16)

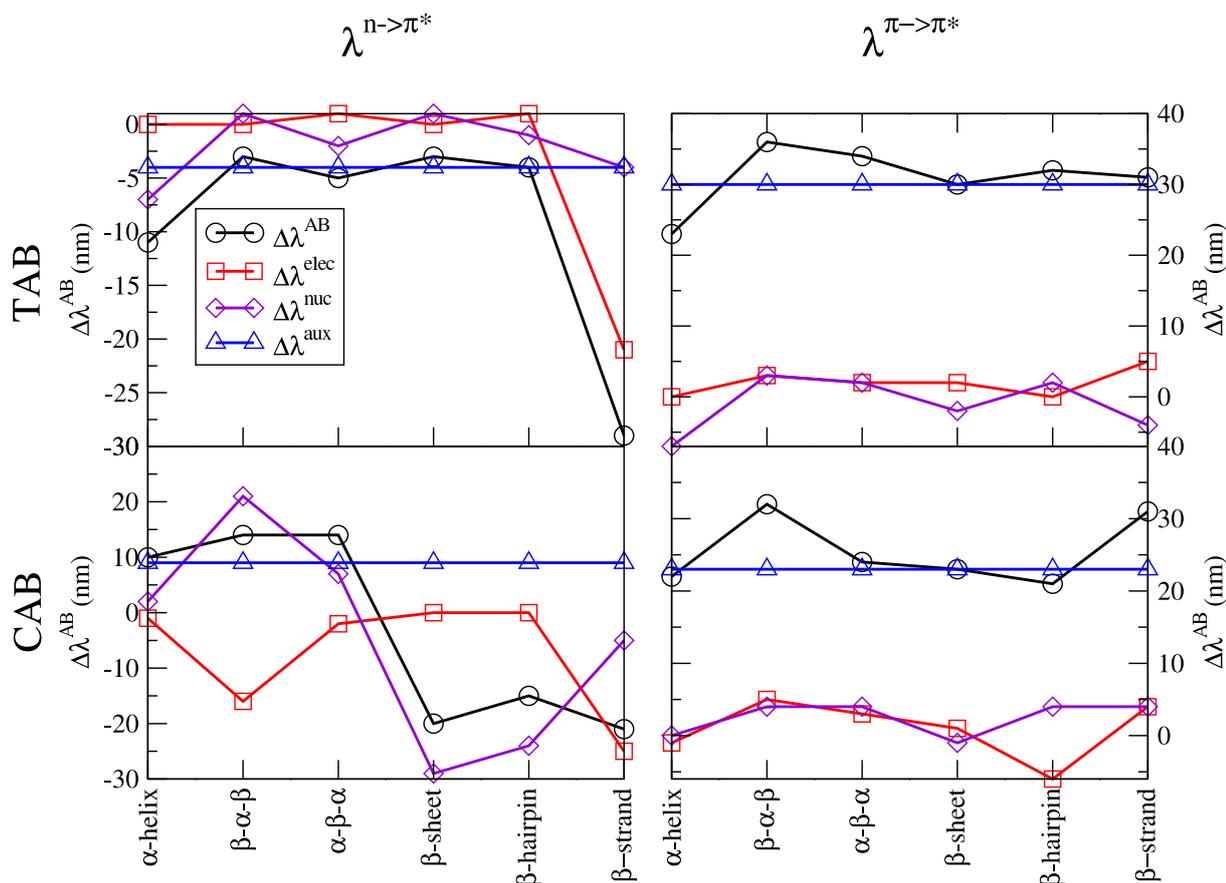
Pour clarifier l'origine physique de ces effets, nous proposons de décomposer cette valeur $\Delta\lambda^{\text{AB}}$ en trois composantes, regroupées dans la table 7.3 et visualisées sur la figure 7.8 :

- La polarisation de la fonction d'onde électronique due aux charges classiques du champ de forces ($\Delta\lambda^{\text{elec}}$). Cette quantité est déterminée par la différence entre les calculs TD-DFT/MM obtenus sans electronic embedding (EE) et le résultat du calcul TD-DFT/MM où les charges sont mises à zéro (avec la même géométrie).
- La modification géométrique induite par la partie MM sur la géométrie de la partie QM ($\Delta\lambda^{\text{nuc}}$) par rapport au système correspondant à la partie QM où l'on a remplacé la partie MM par un atome d'hydrogène (CPS pour *Capped Primary System*, voir FIG. 7.5). Ces déplacements énergétiques correspondent à la différence entre les résultats TD-DFT/MM où les charges sont mises à zéro et les valeurs des énergies TD-DFT obtenue sur la structure optimisée du CPS. Cette quantité correspond à l'effet de l'environnement MM sur la position des noyaux de la partie quantique.
- $\Delta\lambda^{\text{aux}}$ qui correspond au déplacement induit par le groupement auxochromique (incluant le groupement amide et une chaîne alkyle) utilisé pour greffer le photochrome AB sur l'extrémité de la chaîne latérale du polypeptide. Cette quantité, indépendante de la conformation du polypeptide, est définie comme la différence entre les énergies d'excitation entre le CPS et la partie AB.

Conformément à ces définitions, le déplacement global des énergies d'absorption allant de la forme isolée de la molécule d'AB vers le système entier est défini par la relation

$$\Delta\lambda^{\text{AB}} = \Delta\lambda^{\text{elec}} + \Delta\lambda^{\text{nuc}} + \Delta\lambda^{\text{aux}} \quad (7.1)$$

FIG. 7.8 – Déplacements énergétiques (en nm) entre les calculs TD-DFT/MM et l'AB en phase gazeuse ($\Delta\lambda^{AB}$). La valeur $\Delta\lambda^{AB}$ est décomposée en trois composantes venant de la polarisation électronique des charges MM ($\Delta\lambda^{elec}$), de la polarisation nucléaire de l'environnement MM ($\Delta\lambda^{nuc}$) et au déplacement auxochromique due à la substitution de la partie AB ($\Delta\lambda^{aux}$), tel que $\Delta\lambda^{AB} = \Delta\lambda^{elec} + \Delta\lambda^{nuc} + \Delta\lambda^{aux}$.



Pour la transition $\pi \rightarrow \pi^*$ des systèmes comportant la forme TAB, la composante majeure du déplacement vers les grandes longueurs d'onde est attribuée à la contribution auxochromique $\Delta\lambda^{aux}$. En effet, la polarisation due à l'environnement MM sur la fonction d'onde électronique et la position des noyaux a peu d'influence sur les valeurs de $\lambda^{\pi \rightarrow \pi^*}$. Cependant, un examen plus attentif des valeurs de $\Delta\lambda^{elec}$ montre que la polarisation des charges classiques implique un déplacement de 0 à 5 nm. Pour l'excitation $n \rightarrow \pi^*$, la situation est légèrement différente avec un déplacement auxochromique plus faible (-4 nm). Nous voudrions faire remarquer que l'effet de l'*electronic embedding* (EE) implique un large déplacement de -21 nm dans le cas de la structure secondaire en épingle à cheveux β . Cet effet correspond à la stabilisation de l'état fondamental comparée à l'état excité correspondant. Il provient d'une interaction spécifique entre un atome d'hydrogène du groupement carboxylique de la chaîne latérale et de la paire libre de l'unité N=N du chromophore de la forme TAB. Ceci met en lumière le rôle clé que peuvent jouer les chaînes glutamiques du polypeptide. De la même manière qu'une interaction spécifique

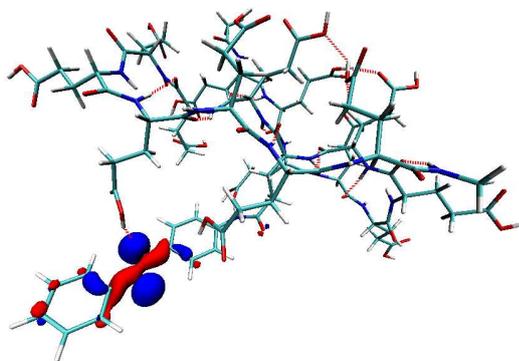
solvant-soluté, cette interaction intramoléculaire entre le chromophore et le polypeptide peut induire un fort déplacement des énergies d'excitation correspondantes, et tout spécialement pour la transition $n \rightarrow \pi^*$. Ceci s'explique par le caractère localisé des orbitales impliquées dans cette excitation : la MO n non-liante (FIG. 7.9a) et la MO π^* anti-liante (FIG. 7.9c) sont principalement localisées sur le chromophore du colorant azo. La MO n correspond aux paires libres de l'azote, tandis que l'orbitale π^* est principalement localisée sur la liaison diazo. La valeur de $\lambda^{\pi \rightarrow \pi^*}$ est moins sensible à cette interaction spécifique ($\Delta\lambda^{\text{elec}} = +5$ nm), puisque l'orbitale π est délocalisée sur l'ensemble de la molécule (FIG. 7.9b).

L'étude des excitations $\pi \rightarrow \pi^*$ des structures contenant le motif CAB conduit à des conclusions similaires au cas de l'isomère *trans*. Le déplacement vers les grandes longueurs d'onde de $\lambda^{\pi \rightarrow \pi^*}$ est principalement dû à l'effet de substitution sur la partie AB. Pour la transition $n \rightarrow \pi^*$, les structures contenant des motifs β montrent un déplacement vers les courtes longueurs d'onde. Le cas du feuillet β met en jeu le même phénomène que pour l'isomère TAB, tandis que les résultats concernant les conformations en brin et en épingle à cheveux β montrent une nette dépendance envers la modification de la géométrie de la partie AB, illustrée par les valeurs de $\Delta\lambda^{\text{nuc}}$. Comme mentionné précédemment, les deux conformères précédents possèdent les plus petites distances N=N et les plus importantes distances C-N (TAB. 7.2). L'effet, bien que moins prononcé, se retrouve dans la conformation en feuillet β , tandis que l'effet inverse est présent pour le motif β - α - β . L'augmentation de la distance N=N du chromophore se traduit par une valeur de +21 nm pour $\Delta\lambda^{\text{nuc}}$. Cependant, cette conformation MM n'induit pas un déplacement important vers les grandes longueurs d'onde, dû à l'effet opposé induit sur la valeur de $\Delta\lambda^{\text{elec}}$, conséquence de la proximité d'une chaîne latérale glutamique avec l'unité AB.

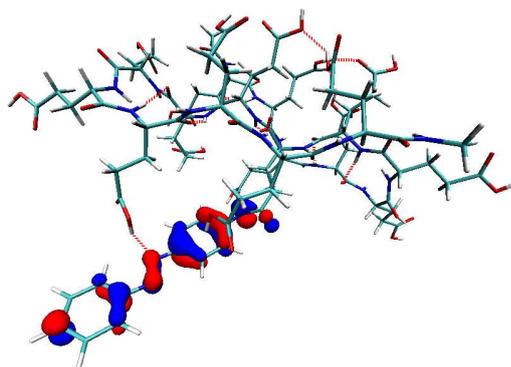
TAB. 7.3 – Spectres UV-vis de la molécule de TAB et CAB greffée comme chaîne latérale d'un Poly(acide L-glutamique) obtenus au niveau de théorie TD-PBE0/6-311+G**. La partie MM est décrite grâce au champ de forces AMBER ff99. $\lambda^{n \rightarrow \pi^*}$ et $\lambda^{\pi \rightarrow \pi^*}$ sont les longueurs d'onde (en nm) des premières transitions $n \rightarrow \pi^*$ et $\pi \rightarrow \pi^*$, respectivement. Les forces d'oscillateurs sont indiquées entre parenthèses. Nous reportons aussi les déplacements énergétiques (en nm) entre les valeurs TD-DFT/MM et la molécule d'AB en phase gazeuse ($(\Delta\lambda^{AB})$). La valeur $\Delta\lambda^{AB}$ est décomposée en trois composantes venant de la polarisation électrostatique due aux charges MM ($\Delta\lambda^{elec}$), de la polarisation nucléaire de l'environnement MM ($\Delta\lambda^{nuc}$) et du déplacement auxochromique due à la substitution de la partie AB ($\Delta\lambda^{aux}$), tel que $\Delta\lambda^{AB} = \Delta\lambda^{elec} + \Delta\lambda^{nuc} + \Delta\lambda^{aux}$.

Structure MM	$\lambda^{n \rightarrow \pi^*}$	$\Delta\lambda^{AB}$			$\lambda^{\pi \rightarrow \pi^*}$	$\Delta\lambda^{AB}$						
		Tot.	$\Delta\lambda^{elec}$	$\Delta\lambda^{nuc}$		$\Delta\lambda^{aux}$	Tot.	$\Delta\lambda^{elec}$	$\Delta\lambda^{nuc}$	$\Delta\lambda^{aux}$		
TAB	Hélice α	469(0,04)	-11	0	-7	-4	350(0,94)	23	0	-7	30	
	β - α - β	477(0,01)	-3	0	1	-4	363(1,00)	36	3	3	30	
	α - β - α	475(0,00)	-5	1	-2	-4	361(1,02)	34	2	2	30	
	Brin β	477(0,00)	-3	0	1	-4	357(0,98)	30	2	-2	30	
	Epingle à cheveux β	476(0,00)	-4	1	-1	-4	359(0,87)	32	0	2	30	
	Feuillet β	451(0,00)	-29	-21	-4	-4	358(0,75)	31	5	-4	30	
	CAB	Hélice α	488(0,05)	10	-1	2	9	314(0,35)	22	-1	0	23
		β - α - β	492(0,06)	14	-16	21	9	324(0,47)	32	5	4	23
		α - β - α	492(0,04)	14	-2	7	9	316(0,50)	24	3	4	23
		Brin β	458(0,04)	-20	0	-29	9	315(0,12)	23	1	-1	23
Epingle à cheveux β		463(0,05)	-15	0	-24	9	313(0,15)	21	-6	4	23	
Feuillet β	457(0,08)	-21	-25	-5	9	323(0,18)	31	4	4	23		

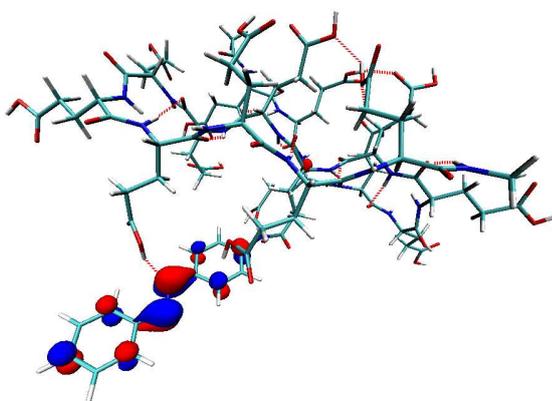
FIG. 7.9 – Isosurface ($0.05 \text{ u.a.}^{-3/2}$) des orbitales frontières impliquées dans les transitions $n \rightarrow \pi^*$ et $\pi \rightarrow \pi^*$ de la molécule TAB greffée sur le polypeptide en conformation en feuillet β .



(a) Orbitale n non-liante



(b) Orbitale π liante



(c) Orbitale π^* anti-liante

7.5 Conclusions

Grâce à une approche théorique combinant des calculs QM/MM et TD-DFT, nous avons reporté l'obtention de spectres UV-vis de Poly(acide L-glutamique) portant un motif AB sur l'une de leurs chaînes latérales. Les calculs ont été effectués au niveau de théorie TD-PBE0/6-311+G**//B3LYP/6-311G** en utilisant le champ de forces AMBER *ff99* pour traiter la partie MM. Les résultats sur l'unité AB, en phase gazeuse et en milieu éthanol, indiquent que cette approche théorique permet d'obtenir des géométries et des longueurs d'onde d'absorption en bon accord avec les valeurs expérimentales disponibles. En particulier, nous avons montré que les résultats TD-PBE0 reproduisent de façon convaincante les effets dus au solvatochromisme et au processus de photoisomérisation.

Dans l'optique d'étudier les effets de la structure du polypeptide, nous avons analysé l'effet des conformations régulièrement ordonnées comportant différents motifs α et β sur les spectres UV-vis de l'unité AB. Nous avons vu que les énergies d'absorption de la transition $\pi \rightarrow \pi^*$ sont globalement déplacées vers les grandes longueurs d'onde. Cet effet a été attribué à un déplacement auxochromique provenant principalement du groupement peptidique utilisé pour greffer la partie AB sur la chaîne latérale du polypeptide. De plus, la présence d'interactions spécifiques entre les chaînes glutamiques et la partie azobenzène, comme des liaisons hydrogène ou une modification importante de la partie QM, peut conduire à de larges déplacements de la transition $n \rightarrow \pi^*$ vers les courtes longueurs d'onde.

Conclusion Générale et Perspectives

Le but de cette thèse est de développer la méthode *Local Self-Consistent Field* (LSCF) du point de vue méthodologique, mais aussi de démontrer son applicabilité à de nombreux systèmes chimiques et d'inspiration biologique. Afin de présenter les travaux liés à ce projet, nous avons exposé les méthodes quantiques, classiques et mixtes dans les trois premiers chapitres de ce manuscrit. Puis, nous avons détaillé l'approche LSCF dans le chapitre 4 (modification du processus SCF, dérivées premières et secondes de l'énergie par rapport aux coordonnées nucléaires), ainsi que les points méthodologiques qui permettent son adaptabilité aux calculs QM/MM dans une large gamme de situations. Nous tenons à préciser que la détermination des équations conduisant aux dérivés secondes analytiques de l'énergie dans le cadre des théories HF et DFT, et plus particulièrement les modifications apportées aux équations *Coupled-Perturbed Hartree-Fock*, est un apport décisif dans l'optique des développements futurs de la méthodologie LSCF/MM. En effet, de par ses applications potentielles en réactivité (caractérisation des points de la surface d'énergie potentielle : minima, états de transition et intermédiaires réactionnels), l'évaluation par le biais d'une méthode analytique des fréquences et des modes normaux associés à une configuration nucléaire est un point crucial. Il en est de même pour les gradients de l'énergie MP2 : les systèmes chimiques comme les liaisons à deux centres et trois électrons (radicaux anions formés dans les protéines par addition d'un électron sur un pont disulfure) ne peuvent pas être traités correctement à l'aide de la méthode HF (manque de corrélation dynamique) ou des fonctionnelles hybrides (erreur de *self-interaction*). Afin d'obtenir une description satisfaisante, la corrélation électronique doit être incluse de façon rigoureuse : la méthode MP2 est alors un bon compromis entre précision et temps de calcul.

Parallèlement à ces développements théoriques, qui restent en cours d'implantation, nous avons regroupé dans le chapitre 5 plusieurs évolutions de la méthodologie LSCF/MM. Ils concernent la frontière QM/MM, et plus particulièrement le traitement de l'atome quanta-classique à l'interface des deux sous-systèmes. Dans un premier temps, nous avons proposé une ré-adaptation de l'orbitale localisée (SLBO) représentant l'interface QM/MM en accord avec les variations de densité électronique du système, conséquence de différents facteurs (effets de substituants, de solvant ou d'une charge ponctuelle). Cette méthode a été nommée OLSCF *Optimized Local Self-Consistent Field*.

Dans un deuxième temps, nous nous sommes attachés à trouver une solution afin d'éviter l'introduction des paramètres empiriques (FBP) dans le traitement de la frontière QM/MM. Dans cette optique, nous avons proposé d'introduire le traitement explicite des électrons de cœur de l'atome frontière Y à l'aide d'une orbitale de cœur gelée (FCO) ou auto-cohérente (SCCO). Cette dernière approche a l'avantage majeur d'éviter l'introduc-

tion de paramètres additionnels et la détermination de la FCO. En collaboration avec le Dr. Elise Dumont, la méthodologie basée sur l'utilisation de SCCO est utilisée pour l'étude de la capture électronique par des ponts disulfures dans des protéines [438–440], ou par des brins d'ADN [423–425, 441–450], et se révèle être particulièrement bien adaptée.

Enfin, nous avons étendu l'approche SCCO dans le but de traiter rigoureusement le cas de la liaison peptidique présente dans les systèmes enzymatiques ou les protéines. Afin d'atteindre cet objectif, nous avons proposé de prendre en compte certains électrons de valence de l'atome frontière quanto-classique lorsque celui-ci est un atome d'azote. Cette méthodologie permet une partition QM/MM systématique des systèmes peptidiques et a fourni des résultats très encourageants aux niveaux géométrique et énergétique. D'ailleurs, dans le cadre de la thèse d' Adèle D. Laurent supervisée par le Pr. Xavier Assfeld, cette méthode est en cours d'utilisation afin de déterminer les longueurs d'onde d'absorption de la famille des protéines fluorescentes [451–454].

Dans la troisième et dernière partie de cet exposé, nous avons regroupé deux exemples d'application de la méthode LSCF. L'étude des états ionisés ou excités de cœur est présentée dans le chapitre 6. Le couplage de la méthode LSCF purement quantique, qui permet de geler une orbitale quelque soit son occupation (occupée ou vacante) et sa nature (orbitale de cœur ou bien de valence), à la méthodologie LSCF/MM nous a conduit à une évaluation rigoureuse (par le biais d'une nouvelle procédure d'orthogonalisation) et qualitative des énergies d'ionisation de cœur de molécules organiques et biomimétiques. Nous avons choisi de présenter cette étude car il n'existe qu'un nombre restreint de méthodes capables de traiter à la fois la contrainte de taille du système et la représentation du « trou de cœur ». De plus, bien que ces spectroscopies d'électrons X soient actuellement appliquées à des molécules de taille restreinte, elles s'orientent ces dernières années vers des systèmes moléculaires de taille de plus en plus importante [333, 335, 336, 455, 456]. En collaboration avec le Pr. Carlo Adamo de l'ENSCP, cet axe s'oriente aussi vers l'étude de l'effet causé par l'erreur de *self-interaction* dans les fonctionnelles hybrides [365].

Le dernier chapitre (Chapitre 7) concerne la spectroscopie UV-visible de systèmes d'inspiration biologique, et plus particulièrement de la détermination qualitative des longueurs d'onde d'absorption du chromophore azobenzène en phase condensée et greffé sur des polypeptides de conformations variables. Dans ce chapitre, l'approche LSCF/MM et la TD-DFT sont couplées afin d'étudier les états excités de macromolécules. Nous nous sommes attachés à analyser l'influence des effets de solvant, ainsi que de la conformation du polypeptide sur les transitions UV-vis du chromophore azobenzène dans sa forme *cis* et *trans*. Cette partie est un premier pas au sein d'une étude beaucoup plus vaste. Elle concerne la conception d'architecture moléculaire capable de transformer l'énergie lumineuse en énergie mécanique par intégration d'un matériel inorganique au sein d'un composée d'inspiration biologique. En collaboration avec le Dr. David Ambrosek et le Pr. Chantal Daniel du Laboratoire de Chimie Quantique de Strasbourg et à l'aide de cette approche théorique combinant TD-DFT et MM, nous sommes en train d'étudier les valeurs des énergies d'absorption pour différents complexes du ruthénium intercalés dans un brin d'ADN [457].

Bien que la méthodologie LSCF/MM s'applique à des situations de plus en plus générales, de nombreux développements sont encore nécessaires. Le couplage de l'approche LSCF/MM et de la méthode PCM [42] (*Polarizable Continuum Model*) a été réalisé lors

de cette thèse et est applicable dans l'état actuel de notre version modifiée du logiciel Gaussian 03 [75]. Cependant, lorsque l'on souhaite traiter un système de grande taille, il n'est que très rarement possible et souhaitable d'insérer l'ensemble de la partie MM à l'intérieur du continuum diélectrique polarisable. Ceci est une conséquence directe de la demande en mémoire et temps de calcul de la détermination auto-cohérente des charges à la surface de la cavité. Il est alors nécessaire de développer une approche originale et rigoureuse afin de permettre une polarisation du continuum par les charges ponctuelles des atomes MM hors de la cavité. La thèse d'Adèle D. Laurent s'oriente principalement vers ce point.

Un autre axe prometteur est l'utilisation des champs de forces polarisables, comme AMOEBA [177–181] intégré dans le logiciel Tinker [76], au sein des méthodes mixtes QM/MM. Ceci a pour but de prendre en compte à la fois la polarisation de la fonction d'onde *ab initio* causée par l'environnement classique, mais aussi l'influence que peut engendrer la partie quantique sur les multipôles atomiques de la partie classique qui peut avoir un rôle important dans la compréhension des certains phénomènes ayant lieu dans les systèmes biologiques.

Par extension, le couplage des méthodes *ab initio* HF ou post-HF avec des méthodes demandant un coût de calcul inférieur (SE ou DFTB) semble une évolution logique et désirable des méthodes hybrides vers une plus grande précision. Ce point est particulièrement attractif de par l'augmentation actuelle de l'efficacité des méthodes SE et DFTB du point de vue algorithmique mais aussi de leur précision au sens chimique. Ce couplage entre méthodes *ab initio* et de type semi-empirique reste, cependant, un « challenge théorique » que se soit dans le cadre de la méthode LSCF ou des approches *Link-Atom*.

Dans ce manuscrit, nous avons détaillé les équations des gradients de l'énergie MP2 qui est la méthode la plus usuelle et la plus simple afin d'introduire un traitement rigoureux de la corrélation dynamique. Néanmoins, les modifications apportées aux équations CPHF qui sont associées aux spécificités de la méthode LSCF ouvrent la porte à la détermination des gradients de l'énergie dans le cas de nombreuses méthodes post-HF, telles que les théories *Coupled-Cluster* ou de configuration d'interaction. De plus, la détermination analytique des fréquences de vibration dans le cadre de la TD-DFT peut nous permettre de calculer les facteurs Frank-Condon dans le cas des phénomènes de fluorescence.

A

Modification des Equations CPHF pour la Méthode LSCF

Dans cette première annexe, nous détaillons la dérivation des équations *Coupled Perturbed Hartree-Fock* [252, 458–462] (CPHF) dans le cadre de la méthode LSCF. Ces équations permettent d’obtenir la matrice \mathbf{U}^x , matrice de réponse des orbitales à une perturbation x . Cette matrice est nécessaire à la connaissance des dérivées secondes de l’énergie par rapport aux positions nucléaires, dans le cas des méthodes HF et DFT (Section 4.3). Elle est aussi nécessaire dans le cas de certaines méthodes post-HF [463–466] (MP2 en particulier [252, 466–474]) afin de déterminer les dérivées premières de l’énergie par rapport aux coordonnées nucléaires (Annexe B). Elle permet d’exprimer la modification des coefficients des MOs lors d’une perturbation x à partir de leurs coefficients initiaux. L’équation ci-dessous rappelle la définition de cette matrice, et montre qu’elle permet de déterminer la dérivée par rapport à x d’une MO $|t\rangle$:

$$\begin{aligned} |t\rangle^x &= \left(\sum_{\mu} d_{\mu t} |\mu\rangle \right)^x = \sum_{\mu} d_{\mu t} |\mu\rangle^{(x)} + \sum_{\mu} d_{\mu t}^x |\mu\rangle = |t\rangle^{(x)} + \sum_{\mu} d_{\mu t}^x |\mu\rangle \\ &= |t\rangle^{(x)} + \sum_{\mu} \sum_{u}^{\text{all}} U_{ut}^x d_{\mu u} |\mu\rangle = |t\rangle^{(x)} + \sum_{u}^{\text{all}} U_{ut}^x |u\rangle \end{aligned} \quad (\text{A.1})$$

Dans l’ensemble de cette annexe, nous garderons en mémoire que la matrice \mathbf{U}^x est réelle. Nous rappelons ci-dessous la différence entre la dérivation par rapport à x et (x) que nous avons introduit dans la section 4.3. L’exposant x correspond à la dérivation d’une quantité suivant un déplacement nucléaire x , tandis que l’exposant (x) correspond seulement à la dérivation des intégrales sur la base des AOs (pas des coefficients) intervenant dans cette même quantité. Par exemple, la dérivée de \mathbf{F} et \mathbf{S} conduit à :

$$\begin{aligned} F_{\mu\nu}^x &= H_{\mu\nu}^{(x)} + \sum_{\lambda\sigma} P_{\lambda\sigma}^T G_{\mu\nu\lambda\sigma}^{(x)} + \sum_{\lambda\sigma} (P_{\lambda\sigma}^T)^x G_{\mu\nu\lambda\sigma} \\ &= F_{\mu\nu}^{(x)} + \sum_{\lambda\sigma} (P_{\lambda\sigma}^T)^x G_{\mu\nu\lambda\sigma} \\ S_{\mu\nu}^x &\equiv S_{\mu\nu}^{(x)} \end{aligned} \quad (\text{A.2})$$

Annexe A. Modification des Equations CPHF pour la Méthode LSCF

Dans le cas de la méthode LSCF, cette matrice ne peut pas être obtenue à partir de l'algorithme original [252], car certaines parties de la matrice de Fock, correspondant aux blocs SLBOs-SLBOs et SLBOs-MOs, ne sont pas diagonales, comme l'impose la méthode de résolution de la Ref. [252] (FIG. A.1). Dans la suite de cette section, nous allons donc détailler comment modifier les équations CPHF en conséquence.

FIG. A.1 – Structure de la matrice \mathbf{E} ou de la matrice de Fock dans la base des MOs.

$$\begin{array}{c}
 \vdots \\
 \langle i | \\
 \langle j | \\
 \vdots \\
 \langle a | \\
 \langle b | \\
 \vdots \\
 \langle l_P | \\
 \langle l_Q | \\
 \vdots
 \end{array}
 \begin{pmatrix}
 \cdots & |i\rangle & |j\rangle & \cdots & |a\rangle & |b\rangle & \cdots & |l_P\rangle & |l_Q\rangle & \cdots \\
 \ddots & \vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots & \\
 \cdots & \varepsilon_i & 0 & \cdots & 0 & 0 & \cdots & \varepsilon_{iP} & \varepsilon_{iQ} & \cdots \\
 \cdots & 0 & \varepsilon_j & \cdots & 0 & 0 & \cdots & \varepsilon_{jP} & \varepsilon_{jQ} & \cdots \\
 \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & & \vdots & \vdots & \\
 \cdots & 0 & 0 & \cdots & \varepsilon_a & 0 & \cdots & \varepsilon_{aP} & \varepsilon_{aQ} & \cdots \\
 \cdots & 0 & 0 & \cdots & 0 & \varepsilon_b & \cdots & \varepsilon_{bP} & \varepsilon_{bQ} & \cdots \\
 \vdots & \vdots & \vdots & & \vdots & \vdots & \ddots & \vdots & \vdots & \\
 \cdots & \varepsilon_{Pi} & \varepsilon_{Pj} & \cdots & \varepsilon_{Pa} & \varepsilon_{Pb} & \cdots & \varepsilon_{PP} & \varepsilon_{PQ} & \cdots \\
 \cdots & \varepsilon_{Qi} & \varepsilon_{Qj} & \cdots & \varepsilon_{Qa} & \varepsilon_{Qb} & \cdots & \varepsilon_{QP} & \varepsilon_{QQ} & \cdots \\
 \vdots & \vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots & \ddots
 \end{pmatrix}$$

A.1 Dérivation des équations CPHF dans la base des spinorbitales

A.1.1 Introduction de la matrice de réponse des orbitales

Commençons par séparer la matrice des coefficients de l'ensemble des MOs (\mathbf{D}) comme la somme des coefficients des MOs variationnellement optimisés (\mathbf{C}) et des coefficients des

A.1. Dérivation des équations CPHF dans la base des spinorbitales

SLBOs (\mathbf{L}) :

$$\begin{aligned}
 \mathbf{D} = \mathbf{C} + \mathbf{L} &= \left(\begin{array}{c|c} & \\ \hline \mathbf{C} & \mathbf{L} \\ \hline \end{array} \right) \\
 &= \left(\begin{array}{c|c} & \\ \hline \mathbf{C} & \mathbb{O} \\ \hline \end{array} \right) + \left(\begin{array}{c|c} & \\ \hline \mathbb{O} & \mathbf{L} \\ \hline \end{array} \right)
 \end{aligned} \tag{A.3}$$

où \mathbb{O} représente la matrice nulle. En accord avec cette décomposition, la matrice de réponse \mathbf{U}^x se développe sous la forme :

$$\mathbf{U}^x = (\mathbf{U}^Q)^x + (\mathbf{U}^G)^x \tag{A.4}$$

où $(\mathbf{U}^Q)^x$ et $(\mathbf{U}^G)^x$ représentent la matrice de réponse des orbitales de la partie optimisée variationnellement et de la partie gelée, respectivement. Ceci conduit au jeu d'équations suivant :

$$\mathbf{D}^x = \mathbf{D} \cdot \mathbf{U}^x \tag{A.5}$$

$$\mathbf{C}^x = \mathbf{D} \cdot (\mathbf{U}^Q)^x \tag{A.6}$$

$$\mathbf{L}^x = \mathbf{D} \cdot (\mathbf{U}^G)^x \tag{A.7}$$

La dérivation par rapport à x d'une MO $|t\rangle$ peut donc être explicitée suivant la relation :

$$\begin{aligned}
 |t\rangle^x &= \left(\sum_{\mu} d_{\mu t} |\mu\rangle \right)^x = \sum_{\mu} d_{\mu t} |\mu\rangle^{(x)} + \sum_{\mu} d_{\mu t}^x |\mu\rangle \\
 &= |t\rangle^{(x)} + \sum_{\mu} \sum_u^{\text{all}} U_{ut}^x d_{\mu u} |\mu\rangle = |t\rangle^{(x)} + \sum_u^{\text{all}} U_{ut}^x |u\rangle \\
 &= |t\rangle^{(x)} + \sum_p^{\text{all}} U_{tp}^x |p\rangle + \sum_P^{\text{all}} U_{tP}^x |P\rangle \\
 &= |t\rangle^{(x)} + \sum_p^{\text{all}} (U_{tp}^Q)^x |p\rangle + \sum_P^{\text{all}} (U_{tP}^G)^x |P\rangle
 \end{aligned} \tag{A.8}$$

Pour déterminer la dérivée d'une MO quelconque $|t\rangle^x$, il faut donc connaître les deux composantes de la matrice de réponse $(\mathbf{U}^Q)^x$ et $(\mathbf{U}^G)^x$. Grâce au caractère localisé des SLBOs, nous avons vu dans la section 4.5.4 que les dérivées par rapport à x des SLBOs

Annexe A. Modification des Equations CPHF pour la Méthode LSCF

(\mathbf{L}^x) peuvent être obtenues par différenciation numérique (EQ. (4.52)). La matrice $(\mathbf{U}^G)^x$ est obtenue par multiplication à gauche de la relation (A.7) par $(\mathbf{D})^{-1}$ ²⁶ :

$$(\mathbf{U}^G)^x = (\mathbf{D})^{-1} \cdot \mathbf{L}^x \quad (\text{A.9})$$

Cette expression permet de déterminer tous les coefficients $\{U_{tP}^x\}_{\substack{1 \leq t \leq N \\ 1 \leq P \leq L}}$ (FIG. A.2). Il nous reste donc à déterminer la matrice de réponse $(\mathbf{U}^Q)^x$, c'est-à-dire l'ensemble des coefficients $\{U_{tp}^x\}_{\substack{1 \leq t \leq N \\ 1 \leq p \leq N-L}}$, grâce aux équations CPHF.

FIG. A.2 – Structure de la matrice de réponse des orbitales \mathbf{U}^x .

$$\begin{array}{c}
 \vdots \\
 \langle i | \\
 \langle j | \\
 \vdots \\
 \langle a | \\
 \langle b | \\
 \vdots \\
 \vdots \\
 \langle l_P | \\
 \langle l_Q | \\
 \vdots
 \end{array}
 \left(
 \begin{array}{cccccc|cccc}
 \dots & |i\rangle & |j\rangle & \dots & |a\rangle & |b\rangle & \dots & \dots & |l_P\rangle & |l_Q\rangle & \dots \\
 \ddots & \vdots & \vdots & & \vdots & \vdots & & & \vdots & \vdots & \\
 \dots & U_{ii}^x & U_{ij}^x & \dots & U_{ia}^x & U_{ib}^x & \dots & \dots & U_{iP}^x & U_{iQ}^x & \dots \\
 \dots & U_{ji}^x & U_{jj}^x & \dots & U_{ja}^x & U_{jb}^x & \dots & \dots & U_{jP}^x & U_{jQ}^x & \dots \\
 \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & & & \vdots & \vdots & \\
 \dots & U_{ai}^x & U_{aj}^x & \dots & U_{aa}^x & U_{ab}^x & \dots & \dots & U_{aP}^x & U_{aQ}^x & \dots \\
 \dots & U_{bi}^x & U_{bj}^x & \dots & U_{ba}^x & U_{bb}^x & \dots & \dots & U_{bP}^x & U_{bQ}^x & \dots \\
 \vdots & \vdots & \vdots & & \vdots & \vdots & \ddots & & \vdots & \vdots & \\
 \vdots & \vdots & \vdots & & \vdots & \vdots & & \ddots & \vdots & \vdots & \\
 \dots & U_{Pi}^x & U_{Pj}^x & \dots & U_{Pa}^x & U_{Pb}^x & \dots & \dots & U_{PP}^x & U_{PQ}^x & \dots \\
 \dots & U_{Qi}^x & U_{Qj}^x & \dots & U_{Qa}^x & U_{Qb}^x & \dots & \dots & U_{QP}^x & U_{QQ}^x & \dots \\
 \vdots & \vdots & \vdots & & \vdots & \vdots & & & \vdots & \vdots & \ddots
 \end{array}
 \right)$$

$(\mathbf{U}^Q)^x$
 $(\mathbf{U}^G)^x$

A.1.2 Dérivation des équations séculaires et de la condition d'orthogonalité

Afin de dériver ces équations, nous allons partir des équations séculaires LSCF qui s'écrivent :

$$\mathbf{F} \cdot \mathbf{D} = \mathbf{S} \cdot \mathbf{D} \cdot \mathbf{E} \quad (\text{A.10})$$

Nous garderons en mémoire que la matrice \mathbf{E} n'est pas diagonale dans le cas de la méthode LSCF, contrairement au cas d'une résolution SCF standard. La dérivation de l'équation (A.10) par rapport à la quantité x conduit à :

$$\mathbf{F}^x \cdot \mathbf{D} + \mathbf{F} \cdot \mathbf{D}^x = \mathbf{S}^{(x)} \cdot \mathbf{D} \cdot \mathbf{E} + \mathbf{S} \cdot \mathbf{D}^x \cdot \mathbf{E} + \mathbf{S} \cdot \mathbf{D} \cdot \mathbf{E}^x \quad (\text{A.11})$$

²⁶Notons que \mathbf{D} est bien inversible. En effet, son déterminant est non nul car elle correspond à la matrice des coefficients des MOs. Cependant, la taille de cette matrice pouvant être importante, un algorithme itératif peut être utilisé afin de résoudre le système linéaire suivant : $\mathbf{L}^x = \mathbf{D} \cdot (\mathbf{U}^G)^x$.

A.1. Dérivation des équations CPHF dans la base des spinorbitales

En introduisant la matrice \mathbf{U}^x par le biais de la relation (A.5) :

$$\mathbf{F}^x \cdot \mathbf{D} + \mathbf{F} \cdot \mathbf{D} \cdot \mathbf{U}^x = \mathbf{S}^{(x)} \cdot \mathbf{D} \cdot \mathbf{E} + \mathbf{S} \cdot \mathbf{D} \cdot \mathbf{U}^x \cdot \mathbf{E} + \mathbf{S} \cdot \mathbf{D} \cdot \mathbf{E}^x \quad (\text{A.12})$$

La multiplication à gauche de l'équation précédente par \mathbf{D}^\dagger donne :

$$\begin{aligned} \mathbf{D}^\dagger \cdot \mathbf{F}^x \cdot \mathbf{D} + \mathbf{D}^\dagger \cdot \mathbf{F} \cdot \mathbf{D} \cdot \mathbf{U}^x &= \mathbf{D}^\dagger \cdot \mathbf{S}^x \cdot \mathbf{D} \cdot \mathbf{E} + \mathbf{D}^\dagger \cdot \mathbf{S} \cdot \mathbf{D} \cdot \mathbf{U}^x \cdot \mathbf{E} \\ &\quad + \mathbf{D}^\dagger \cdot \mathbf{S} \cdot \mathbf{D} \cdot \mathbf{E}^x \\ \mathbf{D}^\dagger \cdot \mathbf{F}^x \cdot \mathbf{D} + \mathbf{E} \cdot \mathbf{U}^x &= \mathbf{D}^\dagger \cdot \mathbf{S}^x \cdot \mathbf{D} \cdot \mathbf{E} + \mathbf{U}^x \cdot \mathbf{E} + \mathbf{E}^x \end{aligned} \quad (\text{A.13})$$

sachant que

$$\mathbf{D}^\dagger \cdot \mathbf{F} \cdot \mathbf{D} = \mathbf{E} \quad (\text{A.14})$$

$$\mathbf{D}^\dagger \cdot \mathbf{S} \cdot \mathbf{D} = \mathbb{I} \quad (\text{A.15})$$

où \mathbb{I} est la matrice identité. La condition d'orthogonalité des MOs donnée par (A.15) conduit à

$$\begin{aligned} (\mathbf{D}^\dagger)^x \cdot \mathbf{S} \cdot \mathbf{D} + \mathbf{D}^\dagger \cdot \mathbf{S}^{(x)} \cdot \mathbf{D} + \mathbf{D}^\dagger \cdot \mathbf{S} \cdot \mathbf{D}^x &= \mathbb{O} \\ (\mathbf{D}^x)^\dagger \cdot \mathbf{S} \cdot \mathbf{D} + \mathbf{D}^\dagger \cdot \mathbf{S}^{(x)} \cdot \mathbf{D} + \mathbf{D}^\dagger \cdot \mathbf{S} \cdot \mathbf{D}^x &= \mathbb{O} \\ (\mathbf{D} \cdot \mathbf{U}^x)^\dagger \cdot \mathbf{S} \cdot \mathbf{D} + \mathbf{D}^\dagger \cdot \mathbf{S}^{(x)} \cdot \mathbf{D} + \mathbf{D}^\dagger \cdot \mathbf{S} \cdot \mathbf{D} \cdot \mathbf{U}^x &= \mathbb{O} \\ (\mathbf{U}^x)^\dagger \cdot \mathbf{D}^\dagger \cdot \mathbf{S} \cdot \mathbf{D} + \mathbf{D}^\dagger \cdot \mathbf{S}^{(x)} \cdot \mathbf{D} + \mathbf{D}^\dagger \cdot \mathbf{S} \cdot \mathbf{D} \cdot \mathbf{U}^x &= \mathbb{O} \\ (\mathbf{U}^x)^\dagger + \mathbf{U}^x + \mathbf{D}^\dagger \cdot \mathbf{S}^{(x)} \cdot \mathbf{D} &= \mathbb{O} \end{aligned} \quad (\text{A.16})$$

En considérant que la matrice \mathbf{U}^x est réelle, on détermine la relation suivante :

$$U_{tu}^x + U_{ut}^x + S_{tu}^{(x)} = 0 \quad (\text{A.17})$$

Pour $t = u$, l'expression (A.17) donne :

$$U_{tt}^x = -\frac{1}{2} S_{tt}^{(x)} \quad (\text{A.18})$$

Connaissant les éléments $\{U_{tP}^x\}_{\substack{1 \leq t \leq N \\ 1 \leq P \leq L}}$ d'après la relation (A.9), l'équation (A.17) nous permet d'obtenir l'ensemble $\{U_{Pt}^x\}_{\substack{1 \leq P \leq L \\ 1 \leq t \leq N}}$. Il nous reste donc à déterminer le bloc MOs-MOs de la matrice \mathbf{U}^x qui correspond aux éléments $\{U_{pq}^x\}_{\substack{1 \leq p \leq N-L \\ 1 \leq q \leq N-L}}$ (FIG. A.2). A partir de la relation (A.13), on obtient la relation suivante :

$$F_{pq}^x + \sum_t^{\text{all}} \varepsilon_{pt} U_{tq}^x = \sum_t^{\text{all}} S_{pt}^{(x)} \varepsilon_{tq} + \sum_t^{\text{all}} U_{pt}^x \varepsilon_{tq} + \varepsilon_{pq}^x \delta_{pq} \quad (\text{A.19})$$

A.1.3 Dérivation de la matrice de Fock

En décomposant la somme sur l'ensemble des MOs (t) en une contribution sur les MOs variationnelles (r) et sur les SLBOs (R).

$$F_{pq}^x + \sum_r^{\text{all}} \varepsilon_{pr} U_{rq}^x + \sum_R \varepsilon_{pR} U_{Rq}^x = \sum_r^{\text{all}} S_{pr}^{(x)} \varepsilon_{rq} + \sum_R^{\text{all}} S_{pR}^{(x)} \varepsilon_{Rq} + \sum_r^{\text{all}} U_{pr}^x \varepsilon_{rq} + \sum_R^{\text{all}} U_{pR}^x \varepsilon_{Rq} + \varepsilon_{pq}^x \delta_{pq} \quad (\text{A.20})$$

Comme la matrice de Fock \mathbf{F} est diagonale dans le bloc MOs-MOs (FIG. A.1), $\varepsilon_{pq} = \delta_{pq} \varepsilon_p$:

$$F_{pq}^x + \varepsilon_p U_{pq}^x + \sum_R^{\text{all}} \varepsilon_{pR} U_{Rq}^x = S_{pq}^{(x)} \varepsilon_q + \sum_R^{\text{all}} S_{pR}^{(x)} \varepsilon_{Rq} + U_{pq}^x \varepsilon_q + \sum_R^{\text{all}} U_{pR}^x \varepsilon_{Rq} + \varepsilon_{pq}^x \delta_{pq} \quad (\text{A.21})$$

$$F_{pq}^x - S_{pq}^{(x)} \varepsilon_q + \sum_R^{\text{all}} \left[\varepsilon_{pR} U_{Rq}^x - S_{pR}^{(x)} \varepsilon_{Rq} - U_{pR}^x \varepsilon_{Rq} \right] = (\varepsilon_q - \varepsilon_p) U_{pq}^x + \varepsilon_{pq}^x \delta_{pq} \quad (\text{A.22})$$

Les éléments diagonaux de la relation précédente donne, en utilisant la relation (A.17),

$$\varepsilon_p^x = F_{pp}^x - S_{pp}^{(x)} \varepsilon_p + 2 \sum_R^{\text{all}} \varepsilon_{pR} U_{Rp}^x \quad (\text{A.23})$$

qui permet de déterminer les dérivées de valeurs propres du bloc MOs-MOs. Pour ce qui est des éléments non-diagonaux :

$$U_{pq}^x = \frac{1}{(\varepsilon_q - \varepsilon_p)} \left(F_{pq}^x - S_{pq}^{(x)} \varepsilon_q + \sum_R^{\text{all}} \left[\varepsilon_{pR} U_{Rq}^x + U_{Rp}^x \varepsilon_{Rq} \right] \right) \quad (\text{A.24})$$

L'équation (A.24) donne donc l'expression des éléments du bloc MOs-MOs de la matrice \mathbf{U}^x . Cependant, la matrice \mathbf{F}^x dépend elle aussi des éléments $\{U_{pq}^x\}_{\substack{1 \leq p \leq N-L \\ 1 \leq q \leq N-L}}$. Nous allons donc expliciter les éléments de la dérivée de la matrice de Fock \mathbf{F}^x . Il convient alors d'introduire l'expression des intégrales biélectroniques sur la base des MOs :

$$(tu|vw) = \sum_{\mu\nu\lambda\sigma} d_{\mu t} d_{\nu u} d_{\lambda v} d_{\sigma w} (\mu\nu|\lambda\sigma) \quad (\text{A.25})$$

Suivant cette définition, on peut en déduire la dérivée des intégrales ($pq|rs$),

$$(pq|rs)^x = (pq|rs)^{(x)} + \sum_t^{\text{all}} U_{tp}^x (tq|rs) + \sum_t^{\text{all}} U_{tq}^x (pt|rs) + \sum_t^{\text{all}} U_{tr}^x (pq|ts) + \sum_t^{\text{all}} U_{ts}^x (pq|rt) \quad (\text{A.26})$$

A.1. Dérivation des équations CPHF dans la base des spinorbitales

Ceci permet d'expliciter l'expression de F_{pq}^x . D'après (A.13), $F_{pq}^x = \sum_{\mu\nu} c_{\mu p} F_{\mu\nu}^x c_{\nu q}$, d'où :

$$\begin{aligned}
F_{pq}^x &= H_{pq}^{(x)} + G_{pq}^x \\
&= H_{pq}^{(x)} + \sum_i^{\text{occ}} [(pq|ii)^x - (pi|qi)^x] + \sum_I^{\text{occ}} [(pq|II)^x - (pI|qI)^x] \\
&= H_{pq}^{(x)} + \sum_i^{\text{occ}} [(pq|ii)^{(x)} - (pi|qi)^{(x)}] + \sum_I^{\text{occ}} [(pq|II)^{(x)} - (pI|qI)^{(x)}] \\
&\quad + \sum_i^{\text{occ}} \sum_t^{\text{all}} [U_{ti}^x(pq|ti) + U_{ti}^x(pq|it) - U_{ti}^x(pt|qi) - U_{ti}^x(pi|qt)] \\
&\quad + \sum_I^{\text{occ}} \sum_t^{\text{all}} [U_{tI}^x(pq|tI) + U_{tI}^x(pq|It) - U_{tI}^x(pt|qI) - U_{tI}^x(pI|qt)]
\end{aligned} \tag{A.27}$$

En se rappelant que,

$$\begin{aligned}
F_{pq}^{(x)} &= H_{pq}^{(x)} + G_{pq}^{(x)} \\
&= H_{pq}^{(x)} + \sum_i^{\text{occ}} [(pq|ii)^{(x)} - (pi|qi)^{(x)}] + \sum_I^{\text{occ}} [(pq|II)^{(x)} - (pI|qI)^{(x)}]
\end{aligned} \tag{A.28}$$

et en définissant la quantité A_{pqji} par l'expression²⁷,

$$A_{tuvw} = 2(tu|vw) - (tv|uw) - (tw|uv) \tag{A.29}$$

l'expression précédente conduit à :

$$F_{pq}^x = F_{pq}^{(x)} + \sum_i^{\text{occ}} \sum_t^{\text{all}} U_{ti}^x A_{pqt i} + \sum_I^{\text{occ}} \sum_t^{\text{all}} U_{tI}^x A_{pqt I} \tag{A.30}$$

Afin de simplifier l'expression ci-dessus, nous allons séparer la somme sur t en somme sur les orbitales occupées et vacantes pour les MOs et les SLBOs. Le deux derniers termes se décomposent donc suivant :

$$\begin{aligned}
\sum_i^{\text{occ}} \sum_t^{\text{all}} U_{ti}^x A_{pqt i} &= \sum_{ij}^{\text{occ}} U_{ji}^x A_{pqji} + \sum_i^{\text{occ}} \sum_a^{\text{virt}} U_{ai}^x A_{pqa i} \\
&\quad + \sum_{iJ}^{\text{occ}} U_{Ji}^x A_{pqJi} + \sum_i^{\text{occ}} \sum_A^{\text{virt}} U_{Ai}^x A_{pqAi} \\
\sum_I^{\text{occ}} \sum_t^{\text{all}} U_{tI}^x A_{pqt I} &= \sum_{Ij}^{\text{occ}} U_{jI}^x A_{pqjI} + \sum_I^{\text{occ}} \sum_a^{\text{virt}} U_{aI}^x A_{pqa I} \\
&\quad + \sum_{IJ}^{\text{occ}} U_{JI}^x A_{pqJI} + \sum_I^{\text{occ}} \sum_A^{\text{virt}} U_{AI}^x A_{pqAI}
\end{aligned} \tag{A.31}$$

²⁷Nous noterons la propriété suivante : $A_{tuvw} = A_{tuwv}$.

Annexe A. Modification des Equations CPHF pour la Méthode LSCF

Le partie correspondant au bloc occupées-occupées des MOs dans l'expression (A.31) peut encore être simplifié en effectuant la manipulation suivante :

$$\sum_{ij}^{\text{occ}} U_{ji}^x A_{pqji} = \frac{1}{2} \left[\sum_{ij}^{\text{occ}} U_{ji}^x A_{pqji} + \sum_{ij}^{\text{occ}} U_{ji}^x A_{pqji} \right] \quad (\text{A.32})$$

En échangeant i et j dans le second terme,

$$\begin{aligned} \sum_{ij}^{\text{occ}} U_{ji}^x A_{pqji} &= \frac{1}{2} \left[\sum_{ij}^{\text{occ}} U_{ji}^x A_{pqji} + \sum_{ji}^{\text{occ}} U_{ij}^x A_{pqij} \right] \\ &= \frac{1}{2} \sum_{ij}^{\text{occ}} [U_{ji}^x + U_{ij}^x] A_{pqji} = -\frac{1}{2} \sum_{ij}^{\text{occ}} S_{ij}^{(x)} A_{pqji} \end{aligned} \quad (\text{A.33})$$

car $A_{pqji} = A_{pqij}$. Suivant cette même dérivation, nous pouvons effectuer ces manipulations sur le bloc occupée-occupée des SLBOs :

$$\sum_{IJ}^{\text{occ}} U_{JI}^x A_{pqJI} = -\frac{1}{2} \sum_{IJ}^{\text{occ}} S_{IJ}^{(x)} A_{pqJI} \quad (\text{A.34})$$

La contribution des orbitales occupées du bloc SLBOs-MOs se contracte suivant :

$$\sum_{iJ}^{\text{occ}} U_{Ji}^x A_{pqJi} + \sum_{Ij}^{\text{occ}} U_{jI}^x A_{pqjI} = -\sum_{iJ}^{\text{occ}} S_{iJ}^{(x)} A_{pqJi} \quad (\text{A.35})$$

L'expression de la dérivée du Fockien est donc :

$$\begin{aligned} F_{pq}^x &= F_{pq}^{(x)} - \frac{1}{2} \sum_{ij}^{\text{occ}} S_{ij}^{(x)} A_{pqji} - \frac{1}{2} \sum_{IJ}^{\text{occ}} S_{IJ}^{(x)} A_{pqJI} - \sum_i^{\text{occ}} \sum_J^{\text{occ}} S_{iJ}^{(x)} A_{pqJi} \\ &+ \sum_a^{\text{virt}} \sum_J^{\text{occ}} U_{aJ}^x A_{pqJa} + \sum_I^{\text{occ}} \sum_A^{\text{virt}} U_{IA}^x A_{pqIA} + \sum_i^{\text{occ}} \sum_B^{\text{virt}} U_{iB}^x A_{pqBi} \\ &+ \sum_a^{\text{virt}} \sum_i^{\text{occ}} U_{ai}^x A_{pqai} \end{aligned} \quad (\text{A.36})$$

Nous remarquons donc que le seul bloc inconnue de la matrice \mathbf{U}^x dans l'expression de F_{pq}^x est le bloc occupée-virtuelle des MOs variationnelles (dernier terme de l'expression (A.36)), sachant que les éléments correspondants aux blocs MOs-SLBOs, SLBOs-MOs et SLBOs-SLBOs sont connus. L'introduction de (A.36) dans la relation (A.24) permet d'écrire :

$$U_{pq}^x = \frac{1}{(\varepsilon_q - \varepsilon_p)} Q_{pq}^x \quad (\text{A.37})$$

avec

$$Q_{pq}^x = B_{pq}^x + \sum_a^{\text{virt}} \sum_i^{\text{occ}} U_{ai}^x A_{pqai} \quad (\text{A.38})$$

A.2. Dérivation des équations CPHF pour un système à couches ouvertes

On définit la matrice \mathbf{B}^x par la relation (A.39). Cette matrice n'est constituée que de quantités connues.

$$\begin{aligned}
 B_{pq}^x &= F_{pq}^{(x)} - S_{pq}^{(x)} \varepsilon_q + \sum_R^{\text{all}} [\varepsilon_{pR} U_{Rq}^x + U_{Rp}^x \varepsilon_{Rq}] \\
 &\quad - \frac{1}{2} \sum_{ij}^{\text{occ}} S_{ij}^{(x)} A_{pqji} - \frac{1}{2} \sum_{IJ}^{\text{occ}} S_{IJ}^{(x)} A_{pqJI} - \sum_i^{\text{occ}} \sum_J^{\text{occ}} S_{iJ}^{(x)} A_{pqJi} \\
 &\quad + K_{pq}^x
 \end{aligned} \tag{A.39}$$

où la matrice \mathbf{K}^x correspond à :

$$K_{pq}^x = \sum_a^{\text{virt}} \sum_J^{\text{occ}} U_{aJ}^x A_{pq aJ} + \sum_I^{\text{occ}} \sum_A^{\text{virt}} U_{IA}^x A_{pq IA} + \sum_i^{\text{occ}} \sum_B^{\text{virt}} U_{iB}^x A_{pq iB} \tag{A.40}$$

Les équations *Coupled Perturbed Hartree-Fock* s'écrivent donc

$$\sum_i^{\text{occ}} \sum_a^{\text{virt}} [\delta_{ab} \delta_{ij} (\varepsilon_i - \varepsilon_a) - A_{aibj}] U_{ai}^x = B_{bj}^x \tag{A.41}$$

ou sous forme matricielle,

$$\mathbf{A}' \cdot (\mathbf{U}^Q)^x = \mathbf{B}^x \tag{A.42}$$

avec

$$A'_{aibj} = \delta_{ab} \delta_{ij} (\varepsilon_i - \varepsilon_a) - A_{aibj} \tag{A.43}$$

Ces équations permettent d'obtenir, par processus itératif [252], le bloc occupée-virtuelle des MOs variationnelles. La connaissance de ce bloc permet de déterminer l'ensemble des quantités faisant intervenir la dérivée des MOs. Par exemple, les éléments F_{pq}^x peuvent être directement obtenus grâce à la relation (A.36). Nous allons maintenant dériver ces équations dans le cas UHF, puis RHF.

A.2 Dérivation des équations CPHF pour un système à couches ouvertes

Dans le cas d'un système à couches ouvertes traité en formalisme UHF, l'ensemble des équations précédentes sont développées à partir des équations dans la base des spinorbitales. En séparant la somme sur toutes les MOs en une somme sur les MOs de spin α et de spin β :

$$\sum_t^{\text{all}} = \sum_{t^\alpha}^{\text{all}^\alpha} + \sum_{t^\beta}^{\text{all}^\beta} \tag{A.44}$$

Par exemple, dans ce cas la matrice $\mathbf{F}^{(x)}$ s'écrit :

$$F_{p^\alpha q^\alpha}^{(x)} = H_{p^\alpha q^\alpha}^{(x)} + \sum_{t^\alpha}^{\text{occ}^\alpha} [(p^\alpha q^\alpha | t^\alpha t^\alpha)^{(x)} - (p^\alpha t^\alpha | q^\alpha t^\alpha)^{(x)}] + \sum_{t^\beta}^{\text{occ}^\beta} (p^\alpha q^\alpha | t^\beta t^\beta)^{(x)} \tag{A.45}$$

Annexe A. Modification des Equations CPHF pour la Méthode LSCF

Les équations permettant de déterminer la matrice U^x , données par la relation (A.37), s'écrivent dans le cas des MOs de spin α :

$$U_{p^\alpha q^\alpha}^x = \frac{1}{(\varepsilon_q^\alpha - \varepsilon_p^\alpha)} Q_{p^\alpha q^\alpha}^x \quad (\text{A.46})$$

avec

$$Q_{p^\alpha q^\alpha}^x = B_{p^\alpha q^\alpha}^x + \sum_{a^\alpha} \sum_{i^\alpha}^{\text{virt}^\alpha \text{ occ}^\alpha} U_{a^\alpha i^\alpha}^x A_{p^\alpha q^\alpha a^\alpha i^\alpha} + \sum_{a^\beta} \sum_{i^\beta}^{\text{virt}^\beta \text{ occ}^\beta} U_{a^\beta i^\beta}^x A_{p^\alpha q^\alpha a^\beta i^\beta} \quad (\text{A.47})$$

$$Q_{p^\beta q^\beta}^x = B_{p^\beta q^\beta}^x + \sum_{a^\beta} \sum_{i^\beta}^{\text{virt}^\beta \text{ occ}^\beta} U_{a^\beta i^\beta}^x A_{p^\beta q^\beta a^\beta i^\beta} + \sum_{a^\alpha} \sum_{i^\alpha}^{\text{virt}^\alpha \text{ occ}^\alpha} U_{a^\alpha i^\alpha}^x A_{p^\beta q^\beta a^\alpha i^\alpha} \quad (\text{A.48})$$

et

$$A_{t^\alpha u^\alpha v^\alpha w^\alpha} = 2(t^\alpha u^\alpha | v^\alpha w^\alpha) - (t^\alpha v^\alpha | u^\alpha w^\alpha) - (t^\alpha w^\alpha | u^\alpha v^\alpha) \quad (\text{A.49})$$

$$A_{t^\beta u^\beta v^\beta w^\beta} = 2(t^\beta u^\beta | v^\beta w^\beta) - (t^\beta v^\beta | u^\beta w^\beta) - (t^\beta w^\beta | u^\beta v^\beta) \quad (\text{A.50})$$

$$A_{t^\alpha u^\alpha v^\beta w^\beta} = 2(t^\alpha u^\alpha | v^\beta w^\beta) \quad (\text{A.51})$$

$$A_{t^\beta u^\beta v^\alpha w^\alpha} = 2(t^\beta u^\beta | v^\alpha w^\alpha) \quad (\text{A.52})$$

La matrice \mathbf{B} (Eq. (A.39)) se développe suivant :

$$\begin{aligned} B_{p^\alpha q^\alpha}^x &= F_{p^\alpha q^\alpha}^{(x)} - S_{p^\alpha q^\alpha}^{(x)} \varepsilon_q^\alpha + \sum_{R^\alpha}^{\text{all}^\alpha} [\varepsilon_{pR}^\alpha U_{R^\alpha q^\alpha}^x + U_{R^\alpha p^\alpha}^x \varepsilon_{Rq}^\alpha] \\ &\quad - \sum_{i^\alpha j^\alpha}^{\text{occ}^\alpha} S_{i^\alpha j^\alpha}^{(x)} [(p^\alpha q^\alpha | i^\alpha j^\alpha) - (p^\alpha i^\alpha | q^\alpha j^\alpha)] \\ &\quad - \sum_{I^\alpha J^\alpha}^{\text{occ}^\alpha} S_{I^\alpha J^\alpha}^{(x)} [(p^\alpha q^\alpha | I^\alpha J^\alpha) - (p^\alpha I^\alpha | q^\alpha J^\alpha)] \\ &\quad - 2 \sum_{i^\alpha}^{\text{occ}^\alpha} \sum_{J^\alpha}^{\text{occ}^\alpha} S_{i^\alpha J^\alpha}^{(x)} [(p^\alpha q^\alpha | i^\alpha J^\alpha) - (p^\alpha i^\alpha | q^\alpha J^\alpha)] \\ &\quad - \sum_{i^\beta j^\beta}^{\text{occ}^\beta} S_{i^\beta j^\beta}^{(x)} (p^\alpha q^\alpha | i^\beta j^\beta) - \sum_{I^\beta J^\beta}^{\text{occ}^\beta} S_{I^\beta J^\beta}^{(x)} (p^\alpha q^\alpha | I^\beta J^\beta) \\ &\quad - 2 \sum_{i^\beta}^{\text{occ}^\beta} \sum_{J^\beta}^{\text{occ}^\beta} S_{i^\beta J^\beta}^{(x)} (p^\alpha q^\alpha | i^\beta J^\beta) + K_{p^\alpha q^\alpha}^x \end{aligned} \quad (\text{A.53})$$

A.3. Dérivation des équations CPHF pour un système à couches fermées

où nous avons posé

$$\begin{aligned}
K_{p^\alpha q^\alpha}^x &= \sum_{a^\alpha}^{\text{virt}^\alpha} \sum_{J^\alpha}^{\text{occ}^\alpha} U_{a^\alpha J^\alpha}^x [(p^\alpha q^\alpha | a^\alpha J^\alpha) - (p^\alpha a^\alpha | q^\alpha J^\alpha)] \\
&+ \sum_{I^\alpha}^{\text{occ}^\alpha} \sum_{A^\alpha}^{\text{virt}^\alpha} U_{I^\alpha A^\alpha}^x [(p^\alpha q^\alpha | I^\alpha A^\alpha) - (p^\alpha I^\alpha | q^\alpha A^\alpha)] \\
&+ \sum_{i^\alpha}^{\text{occ}^\alpha} \sum_{B^\alpha}^{\text{virt}^\alpha} U_{i^\alpha B^\alpha}^x [(p^\alpha q^\alpha | i^\alpha B^\alpha) - (p^\alpha i^\alpha | q^\alpha B^\alpha)] \\
&+ \sum_{a^\beta}^{\text{virt}^\beta} \sum_{J^\beta}^{\text{occ}^\beta} U_{a^\beta J^\beta}^x (p^\alpha q^\alpha | a^\beta J^\beta) \\
&+ \sum_{I^\beta}^{\text{occ}^\beta} \sum_{A^\beta}^{\text{virt}^\beta} U_{I^\beta A^\beta}^x (p^\alpha q^\alpha | I^\beta A^\beta) \\
&+ \sum_{i^\beta}^{\text{occ}^\beta} \sum_{B^\beta}^{\text{virt}^\beta} U_{i^\beta B^\beta}^x (p^\alpha q^\alpha | i^\beta B^\beta)
\end{aligned} \tag{A.54}$$

Des expressions analogues peuvent être déterminées pour $B_{p^\beta q^\beta}^x$, $K_{p^\beta q^\beta}^x$ et $F_{p^\beta q^\beta}^{(x)}$ en remplaçant α par β . Les équations CPHF s'écrivent donc :

$$\sum_{a^\alpha}^{\text{virt}^\alpha} \sum_{i^\alpha}^{\text{occ}^\alpha} \{ \delta_{ab} \delta_{ij} (\varepsilon_i^\alpha - \varepsilon_a^\alpha) - A_{a^\alpha i^\alpha b^\alpha j^\alpha} \} U_{a^\alpha i^\alpha}^x + \sum_{a^\beta}^{\text{virt}^\beta} \sum_{i^\beta}^{\text{occ}^\beta} \{ A_{a^\alpha i^\alpha b^\beta j^\beta} \} U_{a^\beta i^\beta}^x = B_{b^\alpha j^\alpha}^x \tag{A.55}$$

A.3 Dérivation des équations CPHF pour un système à couches fermées

Dans le cas d'un système à couche fermée, les équations CPHF peuvent être facilement dérivées du cas UHF. Dans ce cas,

$$\sum_t^{\text{all}} = \sum_{t^\alpha}^{\text{all}^\alpha} + \sum_{t^\beta}^{\text{all}^\beta} = 2 \sum_{t^\alpha}^{\text{all}^\alpha} \tag{A.56}$$

Par exemple,

$$F_{pq}^{(x)} = H_{pq}^{(x)} + \sum_t^{\text{occ}} [2(pq|tt)^{(x)} - (pt|qt)^{(x)}] \tag{A.57}$$

En utilisant ce principe, on en déduit les expressions de la matrice U^x ,

$$U_{pq}^x = \frac{1}{(\varepsilon_q - \varepsilon_p)} Q_{pq}^x \tag{A.58}$$

Annexe A. Modification des Equations CPHF pour la Méthode LSCF

avec

$$Q_{pq}^x = B_{pq}^x + \sum_a^{\text{virt}} \sum_i^{\text{occ}} U_{ai}^x A_{pqai} \quad (\text{A.59})$$

Dans le cas RHF, les éléments A_{tuvw} prennent la forme,

$$A_{tuvw} = 4(tu|vw) - (tv|uw) - (tw|uv) \quad (\text{A.60})$$

tandis que,

$$\begin{aligned} B_{pq}^x &= F_{pq}^{(x)} - S_{pq}^{(x)} \varepsilon_q + \sum_R^{\text{all}} [\varepsilon_{pR} U_{Rq}^x + U_{Rp}^x \varepsilon_{Rq}] \\ &\quad - \sum_{ij}^{\text{occ}} S_{ij}^{(x)} [2(pq|ij) - (pi|qj)] \\ &\quad - \sum_{IJ}^{\text{occ}} S_{IJ}^{(x)} [2(pq|IJ) - (pI|qJ)] \\ &\quad - 2 \sum_i^{\text{occ}} \sum_J^{\text{occ}} S_{iJ}^{(x)} [2(pq|iJ) - (pi|qJ)] + K_{pq}^x \end{aligned} \quad (\text{A.61})$$

avec

$$\begin{aligned} K_{pq}^x &= \sum_a^{\text{virt}} \sum_J^{\text{occ}} U_{aJ}^x [2(pq|aJ) - (pa|qJ)] \\ &\quad + \sum_I^{\text{occ}} \sum_A^{\text{virt}} U_{IA}^x [2(pq|IA) - (pI|qA)] + \sum_i^{\text{occ}} \sum_B^{\text{virt}} U_{iB}^x [2(pq|iB) - (pi|qB)] \end{aligned} \quad (\text{A.62})$$

Les équations CPHF se retrouvent donc sous la forme :

$$\sum_a^{\text{virt}} \sum_i^{\text{occ}} \{\delta_{ab} \delta_{ij} (\varepsilon_i - \varepsilon_a) - A_{aibj}\} U_{ai}^x = B_{bj}^x \quad (\text{A.63})$$

A.4 Contribution des atomes MM à la matrice de réponse

Dans le cas d'un système QM/MM, un opérateur additionnel monoélectronique est introduit afin de prendre en compte l'interaction électrostatique entre les électrons et les charges partielles de la partie MM (EQ. (3.7)). Comme il n'y a pas de fonctions de base sur les atomes de la partie MM, les équations CPHF peuvent être simplifiées pour les déplacements suivant ces atomes. Cependant, ceux-ci ont une contribution non-nulle dans la matrice de réponse des orbitales \mathbf{U}^x , et nous devons donc considérer explicitement ces degrés de liberté dans le processus CPHF [475]. En effet, le déplacement de l'un de ces atomes entraînent une modification des coefficients des atomes de la partie quantique.

A.4. Contribution des atomes MM à la matrice de réponse

Dans le cas d'un déplacement nucléaire x suivant l'atome A de la partie MM (x_A^{MM}), l'élément $F_{pq}^{x_A^{\text{MM}}}$ se simplifie suivant :

$$F_{pq}^{(x_A^{\text{MM}})} \equiv H_{pq}^{(x_A^{\text{MM}})} \equiv \left\langle p \left| \left(\frac{q_A}{|\mathbf{R}_A - \mathbf{r}|} \right)^{(x_A^{\text{MM}})} \right| q \right\rangle \quad (\text{A.64})$$

car les dérivées suivant x_A^{MM} des intégrales mono et biélectroniques sont nulles. Comme il en est de même pour les dérivées de la matrice de recouvrement, l'expression de la matrice \mathbf{B}^x se simplifie sous la forme, dans le cas d'un système à couches fermées :

$$B_{pq}^{x_A^{\text{MM}}} = \left\langle p \left| \left(\frac{q_A}{|\mathbf{R}_A - \mathbf{r}|} \right)^{(x_A^{\text{MM}})} \right| q \right\rangle + \sum_R^{\text{all}} \left[\varepsilon_{pR} U_{Rq}^{x_A^{\text{MM}}} + U_{Rp}^{x_A^{\text{MM}}} \varepsilon_{Rq} \right] + K_{pq}^{x_A^{\text{MM}}} \quad (\text{A.65})$$

où $K_{pq}^{x_A^{\text{MM}}}$ est donnée par la relation (A.62). Les équations CPHF peuvent donc être résolues en utilisant la formule (A.63) en utilisant l'expression (A.65) pour la matrice \mathbf{B}^x .

B

Expression des gradients MP2 dans le cadre de la méthode LSCF

B.1 Introduction

La première expression analytique des gradients de l'énergie MP2 a été donnée par Pople *et al.* en 1979 [252]. Le développement d'un grand nombre d'algorithmes et de méthodologies permettant d'augmenter l'efficacité de ces calculs (méthode du *Z-vector*, algorithmes directs ou semi-directs, et approximation *Frozen Core*) [464, 468–473] ont conduit les méthodes de type MP2 [105] à être les plus fréquemment utilisées dans le but d'introduire la corrélation électronique dynamique de façon non-empirique. Dans le cas de la méthode LSCF, les FOs ne sont pas fonctions propres de la matrice de Fock, il est donc indispensable de prendre en compte cette information dans la dérivation des équations MP2. Ainsi, lors du calcul de l'énergie MP2, les FOs sont traitées de façon identique à une orbitale de cœur dans le cas de l'approximation *Frozen Core* (FC). Par conséquent, les diexcitations faisant intervenir les FOs ne sont pas prises en compte dans la correction au deuxième ordre de l'énergie électronique :

$$E^{(2)} = \frac{1}{4} \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{virt}} (ia||jb) T_{ij}^{ab} \quad (\text{B.1})$$

où les amplitudes sur la base des spinorbitales T_{ij}^{ab} sont définies par :

$$T_{ij}^{ab} = (ia||jb) / D_{ij}^{ab} \quad (\text{B.2})$$

avec

$$(ia||jb) = (ia|jb) - (ib|ja) \quad (\text{B.3})$$

$$D_{ij}^{ab} = \varepsilon_i + \varepsilon_j - \varepsilon_a - \varepsilon_b \quad (\text{B.4})$$

B.2 Dérivation des gradients MP2 dans la base des spinorbitales

B.2.1 Forme générale des gradients MP2

Le gradient de l'énergie totale MP2 ($E = E_{\text{HF}} + E^{(2)}$) est généralement écrit sous la forme :

$$E^x = \sum_{\mu\nu} P_{\mu\nu}^{\text{MP2}} H_{\mu\nu}^{(x)} + \sum_{\mu\nu} W_{\mu\nu}^{\text{MP2}} S_{\mu\nu}^{(x)} + \sum_{\mu\nu\lambda\sigma} \Gamma_{\mu\nu\lambda\sigma}^{\text{MP2}} (\mu\nu|\lambda\sigma)^{(x)} \quad (\text{B.5})$$

où $P_{\mu\nu}^{\text{MP2}}$, $W_{\mu\nu}^{\text{MP2}}$ et $\Gamma_{\mu\nu\lambda\sigma}^{\text{MP2}}$ sont respectivement la matrice densité à une particule, la matrice densité pondérée par l'énergie et la matrice densité à deux particules sur la base des AOs. Elles sont la somme des termes HF et d'une correction au second ordre ($P^{(2)}$, $W^{(2)}$ et $\Gamma_{\mu\nu\lambda\sigma}^{(2)}$). Pour des raisons d'implantation, la correction de la matrice densité à deux particules $\Gamma_{\mu\nu\lambda\sigma}^{(2)}$ s'exprime comme la somme d'un terme séparable et d'un terme non-séparable,

$$\Gamma_{\mu\nu\lambda\sigma}^{(2)} = \Gamma_{\mu\nu\lambda\sigma}^{\text{S}} + \Gamma_{\mu\nu\lambda\sigma}^{\text{NS}} \quad (\text{B.6})$$

B.2.2 Dérivée de l'énergie MP2

Dans le cadre de la méthode LSCF, nous allons détailler, dans cette partie, la dérivation des gradients MP2 dans la base des spinorbitales. Nous renvoyons le lecteur à l'article très détaillé de Aikens *et al.* [473] pour de plus amples informations sur le développement de ces équations dans le cas de l'approximation FC. La dérivée par rapport à x de la correction au deuxième ordre de l'énergie HF, $E^{(2)}$, s'écrit généralement sous la forme :

$$E^{(2)x} = \frac{1}{4} \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{virt}} [(ia||jb)^x T_{ij}^{ab} + (ia||jb) (T_{ij}^{ab})^x] \quad (\text{B.7})$$

En utilisant l'expression des dérivées des amplitudes,

$$\begin{aligned} (T_{ij}^{ab})^x &= (ia||jb)^x / D_{ij}^{ab} + (ia||jb) (1/D_{ij}^{ab})^x \\ &= (ia||jb)^x / D_{ij}^{ab} - (ia||jb) \frac{(D_{ij}^{ab})^x}{(D_{ij}^{ab})^2} \end{aligned} \quad (\text{B.8})$$

on en déduit la relation suivante :

$$E^{(2)x} = \frac{1}{2} \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{virt}} (ia||jb)^x T_{ij}^{ab} - \frac{1}{4} \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{virt}} (ia||jb)^2 \frac{(D_{ij}^{ab})^x}{(D_{ij}^{ab})^2} \quad (\text{B.9})$$

Introduction de la matrice de réponse

En substituant $(ia||jb)^x$ dans (B.9) par son expression dans (A.26), qui introduit la matrice de réponse des orbitales \mathbf{U}^x (Annexe A), et en utilisant l'expression de $(D_{ij}^{ab})^x$

B.2. Dérivation des gradients MP2 dans la base des spinorbitales

par le biais de l'équation (B.4) :

$$\begin{aligned}
 E^{(2)x} = & \frac{1}{2} \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{virt}} T_{ij}^{ab} \left[(ia||jb)^{(x)} + \sum_t^{\text{all}} U_{ti}^x (ta||jb) + \sum_t^{\text{all}} U_{ta}^x (it||jb) \right. \\
 & \left. + \sum_t^{\text{all}} U_{tj}^x (ia||tb) + \sum_t^{\text{all}} U_{tb}^x (ia||jt) \right] \\
 & - \frac{1}{4} \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{virt}} \frac{(ia||jb)}{D_{ij}^{ab}} T_{ij}^{ab} (\varepsilon_i^x + \varepsilon_j^x - \varepsilon_a^x - \varepsilon_b^x)
 \end{aligned} \tag{B.10}$$

Comme les plages de sommation sur i et j sont identiques, ainsi que celles sur a et b , l'équation (B.10) se simplifie comme suit :

$$\begin{aligned}
 E^{(2)x} = & \frac{1}{2} \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{virt}} T_{ij}^{ab} \left[(ia||jb)^{(x)} + 2 \sum_t^{\text{all}} U_{ti}^x (ta||jb) + 2 \sum_t^{\text{all}} U_{ta}^x (it||jb) \right] \\
 & - \frac{1}{2} \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{virt}} (T_{ij}^{ab})^2 (\varepsilon_i^x - \varepsilon_a^x)
 \end{aligned} \tag{B.11}$$

où nous avons utilisé l'expression (B.2) pour simplifier le dernier terme de (B.11). Nous allons maintenant scinder, comme dans l'annexe précédente, la somme sur t en une somme sur les MOs variationnelles et sur les SLBOs :

$$\begin{aligned}
 E^{(2)x} = & \frac{1}{2} \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{virt}} T_{ij}^{ab} \left[(ia||jb)^{(x)} + 2 \sum_p^{\text{all}} U_{pi}^x (pa||jb) + 2 \sum_p^{\text{all}} U_{pa}^x (ip||jb) \right] \\
 & + \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{virt}} \sum_P^{\text{all}} U_{Pi}^x T_{ij}^{ab} (Pa||jb) + \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{virt}} \sum_P^{\text{all}} U_{Pa}^x T_{ij}^{ab} (iP||jb) \\
 & - \frac{1}{2} \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{virt}} (T_{ij}^{ab})^2 (\varepsilon_i^x - \varepsilon_a^x)
 \end{aligned} \tag{B.12}$$

Les deux termes de la deuxième ligne de l'expression (B.12) correspondent à des termes connus car ils sont composés des blocs SLBOs-MOs de la matrice \mathbf{U}^x . Nous avons vu dans l'annexe précédente que ces deux blocs étaient facilement obtenus par différenciation numérique (EQ. (A.9)). Nous allons maintenant nous intéresser au bloc MOs-MOs, et plus particulièrement au bloc occupées-virtuelles. Afin d'isoler ces différents blocs, nous allons décomposer la somme sur l'ensemble des MOs (p) en une somme sur les MOs occupées

Annexe B. Expression des gradients MP2 dans le cadre de la méthode LSCF

(k) et virtuelles (c) :

$$\begin{aligned}
 E^{(2)x} = & \frac{1}{2} \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{virt}} T_{ij}^{ab} \left[(ia||jb)^{(x)} + 2 \sum_k^{\text{occ}} U_{ki}^x(ka||jb) + 2 \sum_c^{\text{virt}} U_{ci}^x(ca||jb) \right. \\
 & \left. + 2 \sum_k^{\text{occ}} U_{ka}^x(ik||jb) + 2 \sum_c^{\text{virt}} U_{ca}^x(ic||jb) \right] \\
 & + \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{virt}} \sum_P^{\text{all}} U_{Pi}^x T_{ij}^{ab}(Pa||jb) + \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{virt}} \sum_P^{\text{all}} U_{Pa}^x T_{ij}^{ab}(iP||jb) \\
 & - \frac{1}{2} \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{virt}} (T_{ij}^{ab})^2 (\varepsilon_i^x - \varepsilon_a^x)
 \end{aligned} \tag{B.13}$$

Termes de la matrice de réponse connus et singularités

Afin d'introduire le maximum de quantités connues, et d'éviter d'introduire des singularités dans les amplitudes, nous allons effectuer un ensemble de manipulations sur la plupart des termes de l'équation (B.13). Par exemple, commençons par le terme correspondant au bloc occupées-occupées des MOs. Nous allons diviser ce terme en deux parties égales et substituer U_{ki}^x à l'aide de l'expression (A.17) :

$$\begin{aligned}
 \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{virt}} \sum_k^{\text{occ}} T_{ij}^{ab} U_{ki}^x(ka||jb) = & \frac{1}{2} \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{virt}} \left[\sum_k^{\text{occ}} T_{ij}^{ab} U_{ki}^x(ka||jb) - \sum_k^{\text{occ}} T_{ij}^{ab} U_{ik}^x(ka||jb) \right] \\
 & - \frac{1}{2} \sum_{ijk}^{\text{occ}} \sum_{ab}^{\text{virt}} T_{ij}^{ab} S_{ki}^x(ka||jb)
 \end{aligned} \tag{B.14}$$

En décomposant la somme sur k suivant $\sum_k^{\text{occ}} U_{ki}^x = \sum_{k<i}^{\text{occ}} U_{ki}^x + U_{ii}^x + \sum_{k>i}^{\text{occ}} U_{ki}^x$ et en l'incorporant dans la relation (B.14), ceci conduit à :

$$\begin{aligned}
 \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{virt}} \sum_k^{\text{occ}} T_{ij}^{ab} U_{ki}^x(ka||jb) = & \frac{1}{2} \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{virt}} \left[\sum_{k<i}^{\text{occ}} T_{ij}^{ab} U_{ki}^x(ka||jb) + \sum_{k>i}^{\text{occ}} T_{ij}^{ab} U_{ki}^x(ka||jb) \right] \\
 & + \frac{1}{2} \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{virt}} T_{ij}^{ab} U_{ii}^x(ia||jb) \\
 & - \frac{1}{2} \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{virt}} \left[\sum_{k<i}^{\text{occ}} T_{ij}^{ab} U_{ik}^x(ka||jb) + \sum_{k>i}^{\text{occ}} T_{ij}^{ab} U_{ik}^x(ka||jb) \right] \\
 & - \frac{1}{2} \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{virt}} T_{ij}^{ab} U_{ii}^x(ia||jb) \\
 & - \frac{1}{2} \sum_{ijk}^{\text{occ}} \sum_{ab}^{\text{virt}} T_{ij}^{ab} S_{ki}^x(ka||jb)
 \end{aligned} \tag{B.15}$$

B.2. Dérivation des gradients MP2 dans la base des spinorbitales

Nous remarquons que les deux termes contenant U_{ii}^x s'annulent. De plus, en échangeant i et k dans les deux termes contenant la sommation $k < i$ (premier terme à l'intérieur les crochets) et en regroupant les termes comportant U_{ki}^x et U_{ik}^x :

$$\begin{aligned} \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{virt}} \sum_k^{\text{occ}} T_{ij}^{ab} U_{ki}^x(ka||jb) &= \frac{1}{2} \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{virt}} \sum_{k>i}^{\text{occ}} [T_{ij}^{ab} U_{ki}^x(ka||jb) - T_{kj}^{ab} U_{ki}^x(ia||jb)] \\ &- \frac{1}{2} \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{virt}} \sum_{k>i}^{\text{occ}} [T_{ij}^{ab} U_{ik}^x(ka||jb) - T_{kj}^{ab} U_{ik}^x(ia||jb)] \\ &- \frac{1}{2} \sum_{ijk}^{\text{occ}} \sum_{ab}^{\text{virt}} T_{ij}^{ab} S_{ki}^x(ka||jb) \end{aligned} \quad (\text{B.16})$$

En développant les amplitudes selon (B.2), on remarque que les deux termes multiplicatifs de U_{ki}^x sont identiques, et qu'il en est de même pour les termes multiplicatifs de U_{ik}^x . En combinant ces termes, on arrive à l'expression (B.17) :

$$\begin{aligned} \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{virt}} \sum_k^{\text{occ}} T_{ij}^{ab} U_{ki}^x(ka||jb) &= \frac{1}{2} \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{virt}} \sum_{k>i}^{\text{occ}} U_{ki}^x(ia||jb)(ka||jb) \left(\frac{1}{D_{ij}^{ab}} - \frac{1}{D_{kj}^{ab}} \right) \\ &- \frac{1}{2} \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{virt}} \sum_{k>i}^{\text{occ}} U_{ik}^x(ia||jb)(ka||jb) \left(\frac{1}{D_{ij}^{ab}} - \frac{1}{D_{kj}^{ab}} \right) \\ &- \frac{1}{2} \sum_{ijk}^{\text{occ}} \sum_{ab}^{\text{virt}} T_{ij}^{ab} S_{ki}^x(ka||jb) \end{aligned} \quad (\text{B.17})$$

En substituant U_{ki}^x et U_{ik}^x à l'aide de la relation (A.37) de l'annexe A, et en remarquant que,

$$\frac{1}{D_{ij}^{ab}} - \frac{1}{D_{kj}^{ab}} = \frac{D_{kj}^{ab} - D_{ij}^{ab}}{D_{ij}^{ab} D_{kj}^{ab}} = \frac{\varepsilon_k - \varepsilon_i}{D_{ij}^{ab} D_{kj}^{ab}} \quad (\text{B.18})$$

le premier terme de l'expression (B.17) peut être simplifié en (B.19) ce qui permet d'éliminer les différences des énergies des MOs, qui pourraient entraîner des divergences lorsque les orbitales sont dégénérées.

$$\frac{1}{2} \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{virt}} \sum_{k>i}^{\text{occ}} \frac{-Q_{ki}^x}{\varepsilon_k - \varepsilon_i} (ia||jb)(ka||jb) \frac{\varepsilon_k - \varepsilon_i}{D_{ij}^{ab} D_{kj}^{ab}} = -\frac{1}{2} \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{virt}} \sum_{k>i}^{\text{occ}} Q_{ki}^x T_{ij}^{ab} T_{kj}^{ab} \quad (\text{B.19})$$

Le second terme de (B.17) conduit à :

$$-\frac{1}{2} \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{virt}} \sum_{k>i}^{\text{occ}} \frac{Q_{ik}^x}{\varepsilon_k - \varepsilon_i} (ia||jb)(ka||jb) \frac{\varepsilon_k - \varepsilon_i}{D_{ij}^{ab} D_{kj}^{ab}} = -\frac{1}{2} \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{virt}} \sum_{k>i}^{\text{occ}} Q_{ik}^x T_{ij}^{ab} T_{kj}^{ab} \quad (\text{B.20})$$

Annexe B. Expression des gradients MP2 dans le cadre de la méthode LSCF

En substituant les expressions (B.19) et (B.20) dans (B.17), nous aboutissons à :

$$\begin{aligned} \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{virt}} \sum_k^{\text{occ}} T_{ij}^{ab} U_{ki}^x(ka||jb) &= - \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{virt}} \sum_{k>i}^{\text{occ}} \left(\frac{Q_{ki}^x + Q_{ik}^x}{2} \right) T_{ij}^{ab} T_{kj}^{ab} \\ &\quad - \frac{1}{2} \sum_{ijk}^{\text{occ}} \sum_{ab}^{\text{virt}} S_{ki}^{(x)} T_{ij}^{ab}(ka||jb) \end{aligned} \quad (\text{B.21})$$

En remplaçant ε_i^x par Q_{ii}^x dans le dernier terme de (B.13), et en le regroupant avec l'expression ci-dessus²⁸ :

$$\begin{aligned} &- \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{virt}} \sum_{k>i}^{\text{occ}} \left(\frac{Q_{ki}^x + Q_{ik}^x}{2} \right) T_{ij}^{ab} T_{kj}^{ab} - \frac{1}{2} \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{virt}} Q_{ii}^x T_{ij}^{ab} T_{ij}^{ab} \\ &= - \frac{1}{2} \sum_{ijk}^{\text{occ}} \sum_{ab}^{\text{virt}} \left(\frac{Q_{ki}^x + Q_{ik}^x}{2} \right) T_{ij}^{ab} T_{kj}^{ab} \end{aligned} \quad (\text{B.22})$$

De la même manière que le terme précédent mais en faisant attention au changement de signe, le bloc virtuelles-virtuelles de l'équation (B.13) peut être simplifié selon :

$$\begin{aligned} &\sum_{ij}^{\text{occ}} \sum_{ab}^{\text{virt}} \sum_{c>a}^{\text{virt}} \left(\frac{Q_{ca}^x + Q_{ac}^x}{2} \right) T_{ij}^{ab} T_{ij}^{cb} + \frac{1}{2} \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{virt}} Q_{aa}^x T_{ij}^{ab} T_{ij}^{ab} \\ &= \frac{1}{2} \sum_{ij}^{\text{occ}} \sum_{abc}^{\text{virt}} \left(\frac{Q_{ca}^x + Q_{ac}^x}{2} \right) T_{ij}^{ab} T_{ij}^{cb} \end{aligned} \quad (\text{B.23})$$

En remplaçant les expressions (B.22) et (B.23) dans (B.13) :

$$\begin{aligned} E^{(2)x} &= - \frac{1}{2} \sum_{ijk}^{\text{occ}} \sum_{ab}^{\text{virt}} \left(\frac{Q_{ki}^x + Q_{ik}^x}{2} \right) T_{ij}^{ab} T_{kj}^{ab} - \frac{1}{2} \sum_{ijk}^{\text{occ}} \sum_{ab}^{\text{virt}} S_{ki}^{(x)} T_{ij}^{ab}(ka||jb) \\ &\quad + \frac{1}{2} \sum_{ij}^{\text{occ}} \sum_{abc}^{\text{virt}} \left(\frac{Q_{ca}^x + Q_{ac}^x}{2} \right) T_{ij}^{ab} T_{ij}^{cb} - \frac{1}{2} \sum_{ij}^{\text{occ}} \sum_{abc}^{\text{virt}} S_{ca}^{(x)} T_{ij}^{ab}(ic||jb) \\ &\quad + \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{virt}} \sum_P^{\text{all}} U_{Pi}^x T_{ij}^{ab}(Pa||jb) + \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{virt}} \sum_P^{\text{all}} U_{Pa}^x T_{ij}^{ab}(iP||jb) \\ &\quad + \sum_{ij}^{\text{occ}} \sum_{abc}^{\text{virt}} U_{ci}^x(ca||jb) T_{ij}^{ab} + \sum_{ijk}^{\text{occ}} \sum_{ab}^{\text{virt}} U_{ka}^x(ik||jb) T_{ij}^{ab} + \frac{1}{2} \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{virt}} T_{ij}^{ab}(ia||jb)^{(x)} \end{aligned} \quad (\text{B.24})$$

Nous allons maintenant nous intéresser aux blocs faisant intervenir les SLBOs, c'est-à-dire aux deux termes de la troisième ligne de l'expression (B.24). Nous allons utiliser les mêmes manipulations que pour les deux blocs occupées-occupées et virtuelles-virtuelles,

²⁸On pourra vérifier la relation suivante : $Q_{pp}^x = \varepsilon_p^x$

B.2. Dérivation des gradients MP2 dans la base des spinorbitales

mais nous ne remplacerons pas les expressions de $U_{P_i}^x$ et $U_{P_a}^x$ car ce sont des quantités connues. Le terme contenant $U_{P_i}^x$ se simplifie selon,

$$\begin{aligned}
 \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{virt}} \sum_P^{\text{all}} U_{P_i}^x T_{ij}^{ab}(Pa||jb) &= \frac{1}{2} \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{virt}} \sum_P^{\text{all}} [U_{P_i}^x T_{ij}^{ab}(Pa||jb) + U_{P_i}^x T_{ij}^{ab}(Pa||jb)] \\
 &= \frac{1}{2} \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{virt}} \sum_P^{\text{all}} U_{P_i}^x T_{ij}^{ab}(Pa||jb) \\
 &\quad - \frac{1}{2} \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{virt}} \sum_P^{\text{all}} U_{iP}^x T_{ij}^{ab}(Pa||jb) \\
 &\quad - \frac{1}{2} \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{virt}} \sum_P^{\text{all}} S_{P_i}^{(x)} T_{ij}^{ab}(Pa||jb) \\
 &= \frac{1}{2} \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{virt}} \sum_P^{\text{all}} (U_{P_i}^x - U_{iP}^x) T_{ij}^{ab}(Pa||jb) \\
 &\quad - \frac{1}{2} \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{virt}} \sum_P^{\text{all}} S_{P_i}^{(x)} T_{ij}^{ab}(Pa||jb)
 \end{aligned} \tag{B.25}$$

tandis que le second terme faisant intervenir $U_{P_a}^x$ devient :

$$\begin{aligned}
 \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{virt}} \sum_P^{\text{all}} U_{P_a}^x T_{ij}^{ab}(iP||jb) &= \frac{1}{2} \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{virt}} \sum_P^{\text{all}} [U_{P_a}^x T_{ij}^{ab}(iP||jb) + U_{P_a}^x T_{ij}^{ab}(iP||jb)] \\
 &= \frac{1}{2} \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{virt}} \sum_P^{\text{all}} U_{P_a}^x T_{ij}^{ab}(iP||jb) \\
 &\quad - \frac{1}{2} \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{virt}} \sum_P^{\text{all}} U_{aP}^x T_{ij}^{ab}(iP||jb) \\
 &\quad - \frac{1}{2} \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{virt}} \sum_P^{\text{all}} S_{P_a}^{(x)} T_{ij}^{ab}(iP||jb) \\
 &= \frac{1}{2} \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{virt}} \sum_P^{\text{all}} (U_{P_a}^x - U_{aP}^x) T_{ij}^{ab}(iP||jb) \\
 &\quad - \frac{1}{2} \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{virt}} \sum_P^{\text{all}} S_{P_a}^{(x)} T_{ij}^{ab}(iP||jb)
 \end{aligned} \tag{B.26}$$

Avant de remplacer les expressions obtenues dans l'équation (B.24), nous allons identifier les différents termes correspondant à des parties de la matrice densité.

Identification des matrices densité

Chaque terme qui multiplie la dérivée d'un opérateur hamiltonien monoélectronique sera défini comme un terme correctif MP2 de la matrice densité ($P_{tu}^{(2)}$), tandis que chaque terme multipliant la dérivée de la matrice de recouvrement sera considéré comme un terme correctif MP2 de la matrice densité pondérée par l'énergie ($W_{tu}^{(2)} [I]$)²⁹. Suivant ces définitions, les termes correspondants aux blocs occupées-occupées et virtuelles-virtuelles deviennent,

$$-\frac{1}{2} \sum_{ijk}^{\text{occ}} \sum_{ab}^{\text{virt}} \left(\frac{Q_{ki}^x + Q_{ik}^x}{2} \right) T_{ij}^{ab} T_{kj}^{ab} = \sum_{ki}^{\text{occ}} \left(\frac{Q_{ki}^x + Q_{ik}^x}{2} \right) P_{ki}^{(2)} \quad (\text{B.27})$$

$$\frac{1}{2} \sum_{ij}^{\text{occ}} \sum_{abc}^{\text{virt}} \left(\frac{Q_{ca}^x + Q_{ac}^x}{2} \right) T_{ij}^{ab} T_{ij}^{cb} = \sum_{ca}^{\text{virt}} \left(\frac{Q_{ca}^x + Q_{ac}^x}{2} \right) P_{ca}^{(2)} \quad (\text{B.28})$$

$$-\frac{1}{2} \sum_{ijk}^{\text{occ}} \sum_{ab}^{\text{virt}} S_{ki}^{(x)} T_{ij}^{ab} (ka||jb) = \sum_{ki}^{\text{occ}} S_{ki}^{(x)} W_{ki}^{(2)} [I] \quad (\text{B.29})$$

$$-\frac{1}{2} \sum_{ij}^{\text{occ}} \sum_{abc}^{\text{virt}} S_{ca}^{(x)} T_{ij}^{ab} (ic||jb) = \sum_{ca}^{\text{virt}} S_{ca}^{(x)} W_{ca}^{(2)} [I] \quad (\text{B.30})$$

tandis que les blocs correspondant à la contribution des SLBOs sont transformés suivant,

$$\frac{1}{2} \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{virt}} \sum_P^{\text{all}} (U_{Pi}^x - U_{iP}^x) T_{ij}^{ab} (Pa||jb) = \sum_i^{\text{occ}} \sum_P^{\text{all}} (U_{Pi}^x - U_{iP}^x) P_{Pi}^{(2)} \quad (\text{B.31})$$

$$\frac{1}{2} \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{virt}} \sum_P^{\text{all}} (U_{Pa}^x - U_{aP}^x) T_{ij}^{ab} (iP||jb) = \sum_a^{\text{virt}} \sum_P^{\text{all}} (U_{Pa}^x - U_{aP}^x) P_{Pa}^{(2)} \quad (\text{B.32})$$

$$-\frac{1}{2} \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{virt}} \sum_P^{\text{all}} S_{Pi}^{(x)} T_{ij}^{ab} (Pa||jb) = \sum_i^{\text{occ}} \sum_P^{\text{all}} S_{Pi}^{(x)} W_{Pi}^{(2)} [I] \quad (\text{B.33})$$

$$-\frac{1}{2} \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{virt}} \sum_P^{\text{all}} S_{Pa}^{(x)} T_{ij}^{ab} (iP||jb) = \sum_a^{\text{virt}} \sum_P^{\text{all}} S_{Pa}^{(x)} W_{Pa}^{(2)} [I] \quad (\text{B.34})$$

Nous avons donc identifié les termes suivants :

- BLOC MOs-MOs :
- Matrice densité :

$$P_{ij}^{(2)} = -\frac{1}{2} \sum_k^{\text{occ}} \sum_{ab}^{\text{virt}} T_{ik}^{ab} T_{jk}^{ab} \quad (\text{B.35})$$

$$P_{ab}^{(2)} = \frac{1}{2} \sum_{ij}^{\text{occ}} \sum_c^{\text{virt}} T_{ij}^{ac} T_{ij}^{bc} \quad (\text{B.36})$$

²⁹Le terme correctif de la matrice densité pondérée par l'énergie $W_{tu}^{(2)}$ est annoté $[I]$ car nous en définirons d'autres dans la suite de cette annexe.

B.2. Dérivation des gradients MP2 dans la base des spinorbitales

– Matrice densité pondérée par l'énergie :

$$W_{ij}^{(2)} [I] = -\frac{1}{2} \sum_k^{\text{occ}} \sum_{ab}^{\text{virt}} T_{ik}^{ab} (ja||kb) \quad (\text{B.37})$$

$$W_{ab}^{(2)} [I] = -\frac{1}{2} \sum_{ij}^{\text{occ}} \sum_c^{\text{virt}} T_{ij}^{ac} (ib||jc) \quad (\text{B.38})$$

– BLOC SLBOs-MOs :

– Matrice densité :

$$P_{Pi}^{(2)} = \frac{1}{2} \sum_k^{\text{occ}} \sum_{ab}^{\text{virt}} T_{ik}^{ab} (Pa||kb) \quad (\text{B.39})$$

$$P_{Pa}^{(2)} = \frac{1}{2} \sum_{ij}^{\text{occ}} \sum_c^{\text{virt}} T_{ij}^{ac} (iP||jc) \quad (\text{B.40})$$

– Matrice densité pondérée par l'énergie :

$$W_{Pi}^{(2)} [I] = -\frac{1}{2} \sum_k^{\text{occ}} \sum_{ab}^{\text{virt}} T_{ik}^{ab} (Pa||kb) \quad (\text{B.41})$$

$$W_{Pa}^{(2)} [I] = -\frac{1}{2} \sum_{ij}^{\text{occ}} \sum_c^{\text{virt}} T_{ij}^{ac} (iP||jc) \quad (\text{B.42})$$

Lagrangien MP2

Le regroupement de l'ensemble de ces expressions dans la relation (B.24) permet d'obtenir l'expression (B.43) donnant la dérivée de la correction de l'énergie au second ordre :

$$\begin{aligned} E^{(2)x} &= \sum_{ij}^{\text{occ}} \left(\frac{Q_{ij}^x + Q_{ji}^x}{2} \right) P_{ij}^{(2)} + \sum_{ij}^{\text{occ}} S_{ij}^{(x)} W_{ij}^{(2)} [I] \\ &+ \sum_{ab}^{\text{virt}} \left(\frac{Q_{ab}^x + Q_{ba}^x}{2} \right) P_{ab}^{(2)} + \sum_{ab}^{\text{virt}} S_{ab}^{(x)} W_{ab}^{(2)} [I] \\ &+ \sum_i^{\text{occ}} \sum_P^{\text{all}} (U_{Pi}^x - U_{iP}^x) P_{Pi}^{(2)} + \sum_i^{\text{occ}} \sum_P^{\text{all}} S_{Pi}^{(x)} W_{Pi}^{(2)} [I] \\ &+ \sum_a^{\text{virt}} \sum_P^{\text{all}} (U_{Pa}^x - U_{aP}^x) P_{Pa}^{(2)} + \sum_a^{\text{virt}} \sum_P^{\text{all}} S_{Pa}^{(x)} W_{Pa}^{(2)} [I] \\ &+ \sum_{ij}^{\text{occ}} \sum_{abc}^{\text{virt}} U_{ci}^x (ca||jb) T_{ij}^{ab} + \sum_{ijk}^{\text{occ}} \sum_{ab}^{\text{virt}} U_{ka}^x (ik||jb) T_{ij}^{ab} \\ &+ \frac{1}{2} \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{virt}} T_{ij}^{ab} (ia||jb)^{(x)} \end{aligned} \quad (\text{B.43})$$

Annexe B. Expression des gradients MP2 dans le cadre de la méthode LSCF

A ce stade, les quantités qui restent inconnues sont les parties occupées-virtuelles et virtuelles-occupées de la matrice de réponse des orbitales (U_{ci}^x et U_{ka}^x , neuvième et dixième terme de l'expression (B.43)). Cependant les termes faisant intervenir les quantités Q_{ij}^x et Q_{ab}^x (premier et troisième termes de (B.43)) dépendent elles aussi de ces éléments du bloc occupées-virtuelles. A l'aide de la relation (A.38), nous allons donc substituer ces quantités afin d'isoler l'ensemble des termes qui dépendent de ce bloc.

$$\begin{aligned}
E^{(2)x} &= \sum_{ij}^{\text{occ}} \left(\frac{B_{ij}^x + B_{ji}^x}{2} \right) P_{ij}^{(2)} + \sum_{ij}^{\text{occ}} S_{ij}^{(x)} W_{ij}^{(2)} [I] \\
&+ \sum_{ab}^{\text{virt}} \left(\frac{B_{ab}^x + B_{ba}^x}{2} \right) P_{ab}^{(2)} + \sum_{ab}^{\text{virt}} S_{ab}^{(x)} W_{ab}^{(2)} [I] \\
&+ \sum_i^{\text{occ}} \sum_P^{\text{all}} (U_{Pi}^x - U_{iP}^x) P_{Pi}^{(2)} + \sum_i^{\text{occ}} \sum_P^{\text{all}} S_{Pi}^{(x)} W_{Pi}^{(2)} [I] \\
&+ \sum_a^{\text{virt}} \sum_P^{\text{all}} (U_{Pa}^x - U_{aP}^x) P_{Pa}^{(2)} + \sum_a^{\text{virt}} \sum_P^{\text{all}} S_{Pa}^{(x)} W_{Pa}^{(2)} [I] \\
&+ \sum_{ij}^{\text{occ}} \sum_c^{\text{virt}} \sum_k^{\text{occ}} U_{ck}^x A_{ijck} P_{ij}^{(2)} + \sum_{ab}^{\text{occ}} \sum_c^{\text{virt}} \sum_k^{\text{occ}} U_{ck}^x A_{abck} P_{ab}^{(2)} \\
&+ \sum_{ij}^{\text{occ}} \sum_{abc}^{\text{virt}} U_{ci}^x (ca||jb) T_{ij}^{ab} + \sum_{ijk}^{\text{occ}} \sum_{ab}^{\text{virt}} U_{ka}^x (ik||jb) T_{ij}^{ab} \\
&+ \frac{1}{2} \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{virt}} T_{ij}^{ab} (ia||jb)^{(x)}
\end{aligned} \tag{B.44}$$

où nous avons utilisé la propriété de symétrie suivante :

$$\sum_{pq}^{\text{all}} \sum_c^{\text{virt}} \sum_k^{\text{occ}} U_{ck}^x A_{pqck} = \sum_{pq}^{\text{all}} \sum_c^{\text{virt}} \sum_k^{\text{occ}} U_{ck}^x A_{qpck} \tag{B.45}$$

Nous allons maintenant effectuer quelques manipulations sur les deux nouveaux termes introduits dans (B.43) (onzième et douzième termes de l'expression précédente) afin de faire apparaître le terme occupées-virtuelles de la matrice densité pondérée par l'énergie $W_{ai}^{(2)}$ et de factoriser les termes contenant les éléments du type U_{ai}^x . Le douzième terme de l'expression devient :

$$\begin{aligned}
\sum_{ijk}^{\text{occ}} \sum_{ab}^{\text{virt}} U_{ka}^x (ik||jb) T_{ij}^{ab} &= - \sum_{ijk}^{\text{occ}} \sum_{ab}^{\text{virt}} U_{ak}^x (ik||jb) T_{ij}^{ab} - \sum_{ijk}^{\text{occ}} \sum_{ab}^{\text{virt}} S_{ak}^{(x)} (ik||jb) T_{ij}^{ab} \\
&= - \sum_{ijk}^{\text{occ}} \sum_{ab}^{\text{virt}} U_{ak}^x (ik||jb) T_{ij}^{ab} + \sum_k^{\text{occ}} \sum_a^{\text{virt}} S_{ak}^{(x)} W_{ak}^{(2)} [I]
\end{aligned} \tag{B.46}$$

avec

$$W_{ai}^{(2)} [I] = - \sum_{jk}^{\text{occ}} \sum_b^{\text{virt}} T_{jk}^{ab} (ij||bk) \tag{B.47}$$

B.2. Dérivation des gradients MP2 dans la base des spinorbitales

En remplaçant l'expression obtenue dans $E^{(2)x}$, nous obtenons

$$\begin{aligned}
E^{(2)x} = & \sum_{ij}^{\text{occ}} \left(\frac{B_{ij}^x + B_{ji}^x}{2} \right) P_{ij}^{(2)} + \sum_{ij}^{\text{occ}} S_{ij}^{(x)} W_{ij}^{(2)} [I] \\
& + \sum_{ab}^{\text{virt}} \left(\frac{B_{ab}^x + B_{ba}^x}{2} \right) P_{ab}^{(2)} + \sum_{ab}^{\text{virt}} S_{ab}^{(x)} W_{ab}^{(2)} [I] \\
& + \sum_i^{\text{occ}} \sum_P^{\text{all}} (U_{Pi}^x - U_{iP}^x) P_{Pi}^{(2)} + \sum_i^{\text{occ}} \sum_P^{\text{all}} S_{Pi}^{(x)} W_{Pi}^{(2)} [I] \\
& + \sum_a^{\text{virt}} \sum_P^{\text{all}} (U_{Pa}^x - U_{aP}^x) P_{Pa}^{(2)} + \sum_a^{\text{virt}} \sum_P^{\text{all}} S_{Pa}^{(x)} W_{Pa}^{(2)} [I] \\
& + \sum_k^{\text{occ}} \sum_a^{\text{virt}} S_{ak}^{(x)} W_{ak}^{(2)} [I] + \frac{1}{2} \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{virt}} T_{ij}^{ab} (ia||jb)^{(x)} \\
& + \sum_{ij}^{\text{occ}} \sum_c^{\text{virt}} \sum_k^{\text{occ}} U_{ck}^x A_{ijck} P_{ij}^{(2)} + \sum_{ab}^{\text{occ}} \sum_c^{\text{virt}} \sum_k^{\text{occ}} U_{ck}^x A_{abck} P_{ab}^{(2)} \\
& + \sum_{ij}^{\text{occ}} \sum_{abc}^{\text{virt}} U_{ci}^x (ca||jb) T_{ij}^{ab} - \sum_{ijk}^{\text{occ}} \sum_{ab}^{\text{virt}} U_{ak}^x (ik||jb) T_{ij}^{ab}
\end{aligned} \tag{B.48}$$

Puis, en factorisant les termes contenant U_{ck}^x ,

$$\begin{aligned}
E^{(2)x} = & \sum_{ij}^{\text{occ}} \left(\frac{B_{ij}^x + B_{ji}^x}{2} \right) P_{ij}^{(2)} + \sum_{ij}^{\text{occ}} S_{ij}^{(x)} W_{ij}^{(2)} [I] \\
& + \sum_{ab}^{\text{virt}} \left(\frac{B_{ab}^x + B_{ba}^x}{2} \right) P_{ab}^{(2)} + \sum_{ab}^{\text{virt}} S_{ab}^{(x)} W_{ab}^{(2)} [I] \\
& + \sum_i^{\text{occ}} \sum_P^{\text{all}} (U_{Pi}^x - U_{iP}^x) P_{Pi}^{(2)} + \sum_i^{\text{occ}} \sum_P^{\text{all}} S_{Pi}^{(x)} W_{Pi}^{(2)} [I] \\
& + \sum_a^{\text{virt}} \sum_P^{\text{all}} (U_{Pa}^x - U_{aP}^x) P_{Pa}^{(2)} + \sum_a^{\text{virt}} \sum_P^{\text{all}} S_{Pa}^{(x)} W_{Pa}^{(2)} [I] \\
& + \sum_k^{\text{occ}} \sum_a^{\text{virt}} S_{ak}^{(x)} W_{ak}^{(2)} [I] + \frac{1}{2} \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{virt}} T_{ij}^{ab} (ia||jb)^{(x)} \\
& + \sum_c^{\text{virt}} \sum_k^{\text{occ}} U_{ck}^x \left[\sum_{ij}^{\text{occ}} A_{ijck} P_{ij}^{(2)} + \sum_{ab}^{\text{virt}} A_{abck} P_{ab}^{(2)} \right. \\
& \left. + \sum_j^{\text{occ}} \sum_{ab}^{\text{virt}} (ca||jb) T_{kj}^{ab} - \sum_{ij}^{\text{occ}} \sum_b^{\text{virt}} (ik||jb) T_{ij}^{cb} \right]
\end{aligned} \tag{B.49}$$

Annexe B. Expression des gradients MP2 dans le cadre de la méthode LSCF

que l'on peut noter,

$$\begin{aligned}
E^{(2)x} = & \sum_{ij}^{\text{occ}} \left(\frac{B_{ij}^x + B_{ji}^x}{2} \right) P_{ij}^{(2)} + \sum_{ij}^{\text{occ}} S_{ij}^{(x)} W_{ij}^{(2)} [I] \\
& + \sum_{ab}^{\text{virt}} \left(\frac{B_{ab}^x + B_{ba}^x}{2} \right) P_{ab}^{(2)} + \sum_{ab}^{\text{virt}} S_{ab}^{(x)} W_{ab}^{(2)} [I] \\
& + \sum_i^{\text{occ}} \sum_P^{\text{all}} (U_{Pi}^x - U_{iP}^x) P_{Pi}^{(2)} + \sum_i^{\text{occ}} \sum_P^{\text{all}} S_{Pi}^{(x)} W_{Pi}^{(2)} [I] \\
& + \sum_a^{\text{virt}} \sum_P^{\text{all}} (U_{Pa}^x - U_{aP}^x) P_{Pa}^{(2)} + \sum_a^{\text{virt}} \sum_P^{\text{all}} S_{Pa}^{(x)} W_{Pa}^{(2)} [I] \\
& + \sum_a^{\text{virt}} \sum_i^{\text{occ}} U_{ai}^x L_{ai} + \sum_a^{\text{virt}} \sum_i^{\text{occ}} S_{ai}^{(x)} W_{ai}^{(2)} [I] \\
& + \frac{1}{2} \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{virt}} T_{ij}^{ab} (ia||jb)^{(x)}
\end{aligned} \tag{B.50}$$

en utilisant la définition du lagrangien MP2 :

$$L_{ai} = \sum_{jk}^{\text{occ}} A_{aijk} P_{jk}^{(2)} + \sum_{bc}^{\text{virt}} A_{aibc} P_{bc}^{(2)} + \sum_j^{\text{occ}} \sum_{bc}^{\text{virt}} (ab||jc) T_{ij}^{bc} - \sum_{jk}^{\text{occ}} \sum_b^{\text{virt}} (ij||bk) T_{jk}^{ab} \tag{B.51}$$

Le bloc occupée-virtuelle U_{ai}^x doit être déterminé en utilisant les équations CPHF (Annexe A). Une fois cette quantité connue, il nous est possible de déterminer la valeur de $E^{(2)x}$, sachant que les autres termes correspondent à des quantités connues.

Méthode du Z -vector

La méthode du Z -vector [464] est l'outil le plus utilisé pour déterminer la matrice de réponse des orbitales. En effet, à la vue des équations précédentes, il semble qu'il soit nécessaire de déterminer la matrice $(\mathbf{U}^Q)^x$, non seulement pour un déplacement suivant x pour un atome donné, mais bien pour l'ensemble des atomes du système suivant les directions x , y et z (*i.e.* $3N$). Cependant, la méthode du Z -vector permet de réduire de $3N$ à une seule le nombre de matrices à déterminer. Nous cherchons donc à déterminer la quantité suivante,

$$\sum_a^{\text{virt}} \sum_i^{\text{occ}} U_{ai}^x L_{ai} = \mathbf{L}^\dagger \cdot \mathbf{U}^x \tag{B.52}$$

à l'aide des équations CPHF que nous rappelons ci-dessous :

$$\mathbf{A}' \cdot (\mathbf{U}^Q)^x = \mathbf{B}^x \tag{B.53}$$

Par multiplication à gauche par $(\mathbf{A}')^{-1}$ des équations CPHF, ceci conduit à :

$$(\mathbf{U}^Q)^x = (\mathbf{A}')^{-1} \cdot \mathbf{B}^x \tag{B.54}$$

$$\mathbf{L}^\dagger \cdot \mathbf{U}^x = \mathbf{L}^\dagger \cdot (\mathbf{A}')^{-1} \cdot \mathbf{B}^x = \mathbf{Z}^\dagger \cdot \mathbf{B}^x \tag{B.55}$$

B.2. Dérivation des gradients MP2 dans la base des spinorbitales

avec

$$\mathbf{Z}^\dagger = \mathbf{L}^\dagger \cdot (\mathbf{A}')^{-1} \quad (\text{B.56})$$

\mathbf{Z} est alors déterminé par la résolution du système linéaire suivant :

$$(\mathbf{A}')^\dagger \cdot \mathbf{Z} = \mathbf{L} \quad (\text{B.57})$$

La connaissance de \mathbf{Z} permet alors d'évaluer la contribution de la matrice $(\mathbf{U}^Q)^x$ dans l'expression de $E^{(2)x}$ à l'aide de la relation :

$$\sum_a^{\text{virt}} \sum_i^{\text{occ}} U_{ai}^x L_{ai} = \sum_a^{\text{virt}} \sum_i^{\text{occ}} B_{ai}^x Z_{ai} \quad (\text{B.58})$$

La puissance de la méthode du *Z-vector* réside dans le fait que la matrice \mathbf{Z} , définie par la relation (B.56) n'est constituée que de quantités indépendantes du déplacement x . Ce vecteur est donc commun à l'ensemble des perturbations, qu'elles soient induites par un déplacement dipolaire, un champ électrique externe, etc. En accord avec l'expression de \mathbf{B} , on peut identifier \mathbf{Z} comme le terme correctif MP2 de la matrice densité :

$$\sum_a^{\text{virt}} \sum_i^{\text{occ}} B_{ai}^x Z_{ai} = \sum_a^{\text{virt}} \sum_i^{\text{occ}} B_{ai}^x P_{ai}^{(2)} \quad (\text{B.59})$$

En substituant cette relation dans (B.50) conduit à :

$$\begin{aligned} E^{(2)x} &= \sum_{ij}^{\text{occ}} \left(\frac{B_{ij}^x + B_{ji}^x}{2} \right) P_{ij}^{(2)} + \sum_{ij}^{\text{occ}} S_{ij}^{(x)} W_{ij}^{(2)} [I] \\ &+ \sum_{ab}^{\text{virt}} \left(\frac{B_{ab}^x + B_{ba}^x}{2} \right) P_{ab}^{(2)} + \sum_{ab}^{\text{virt}} S_{ab}^{(x)} W_{ab}^{(2)} [I] \\ &+ \sum_i^{\text{occ}} \sum_P^{\text{all}} (U_{Pi}^x - U_{iP}^x) P_{Pi}^{(2)} + \sum_i^{\text{occ}} \sum_P^{\text{all}} S_{Pi}^{(x)} W_{Pi}^{(2)} [I] \\ &+ \sum_a^{\text{virt}} \sum_P^{\text{all}} (U_{Pa}^x - U_{aP}^x) P_{Pa}^{(2)} + \sum_a^{\text{virt}} \sum_P^{\text{all}} S_{Pa}^{(x)} W_{Pa}^{(2)} [I] \\ &+ \sum_a^{\text{virt}} \sum_i^{\text{occ}} B_{ai}^x P_{ai}^{(2)} + \sum_a^{\text{virt}} \sum_i^{\text{occ}} S_{ai}^{(x)} W_{ai}^{(2)} [I] \\ &+ \frac{1}{2} \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{virt}} T_{ij}^{ab} (ia||jb)^{(x)} \end{aligned} \quad (\text{B.60})$$

Termes additionels de la matrice densité pondérée par l'énergie

Afin de démontrer qu'il existe des termes correctifs additionels de la matrice densité pondérée par l'énergie, nous allons remplacer l'expression des éléments B_{pq}^x par leur expression dans (A.39). De plus, nous utiliserons la relation (A.17) afin d'introduire la dérivée

Annexe B. Expression des gradients MP2 dans le cadre de la méthode LSCF

de la matrice de recouvrement du bloc SLBOs-MOs. Suivant ces points, nous obtenons :

$$\begin{aligned}
E^{(2)x} = & \sum_{ij}^{\text{occ}} \left\{ F_{ij}^{(x)} - S_{ij}^{(x)} \left(\frac{\varepsilon_i + \varepsilon_j}{2} \right) \right. \\
& + \frac{1}{2} \sum_P^{\text{all}} [(\varepsilon_{iP} U_{Pj}^x + U_{Pi}^x \varepsilon_{Pj}) + (\varepsilon_{jP} U_{Pi}^x + U_{Pj}^x \varepsilon_{Pi})] \\
& - \frac{1}{2} \sum_{kl}^{\text{occ}} S_{kl}^{(x)} A_{ijkl} - \frac{1}{2} \sum_{KL}^{\text{occ}} S_{KL}^{(x)} A_{ijLK} - \sum_{kL}^{\text{occ}} S_{kL}^{(x)} A_{ijLk} + K_{ij}^x \left. \right\} P_{ij}^{(2)} \\
& + \sum_{ab}^{\text{virt}} \left\{ F_{ab}^{(x)} - S_{ab}^{(x)} \left(\frac{\varepsilon_a + \varepsilon_b}{2} \right) \right. \\
& + \frac{1}{2} \sum_P^{\text{all}} [(\varepsilon_{aP} U_{Pb}^x + U_{Pa}^x \varepsilon_{Pb}) + (\varepsilon_{bP} U_{Pa}^x + U_{Pb}^x \varepsilon_{Pa})] \\
& - \frac{1}{2} \sum_{ij}^{\text{occ}} S_{ij}^{(x)} A_{abji} - \frac{1}{2} \sum_{IJ}^{\text{occ}} S_{IJ}^{(x)} A_{abJI} - \sum_{iJ}^{\text{occ}} S_{iJ}^{(x)} A_{abJi} + K_{ab}^x \left. \right\} P_{ab}^{(2)} \\
& + \sum_{ij}^{\text{occ}} S_{ij}^{(x)} W_{ij}^{(2)} [I] + \sum_{ab}^{\text{virt}} S_{ab}^{(x)} W_{ab}^{(2)} [I] \\
& + \sum_i^{\text{occ}} \sum_P^{\text{all}} (U_{Pi}^x - U_{iP}^x) P_{Pi}^{(2)} + \sum_i^{\text{occ}} \sum_P^{\text{all}} S_{Pi}^{(x)} W_{Pi}^{(2)} [I] \\
& + \sum_a^{\text{virt}} \sum_P^{\text{all}} (U_{Pa}^x - U_{aP}^x) P_{Pa}^{(2)} + \sum_a^{\text{virt}} \sum_P^{\text{all}} S_{Pa}^{(x)} W_{Pa}^{(2)} [I] \\
& + \sum_a^{\text{virt}} \sum_i^{\text{occ}} \left\{ F_{ai}^{(x)} - S_{ai}^{(x)} \varepsilon_i + \sum_P^{\text{all}} (\varepsilon_{aP} U_{Pi}^x + U_{Pa}^x \varepsilon_{Pi}) \right. \\
& - \frac{1}{2} \sum_{jk}^{\text{occ}} S_{jk}^{(x)} A_{aikj} - \frac{1}{2} \sum_{JK}^{\text{occ}} S_{JK}^{(x)} A_{aiKJ} - \sum_{jK}^{\text{occ}} S_{jK}^{(x)} A_{aiKj} + K_{ai}^x \left. \right\} P_{ai}^{(2)} \\
& + \sum_a^{\text{virt}} \sum_i^{\text{occ}} S_{ai}^{(x)} W_{ai}^{(2)} [I] + \frac{1}{2} \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{virt}} T_{ij}^{ab} (ia||jb)^{(x)}
\end{aligned} \tag{B.61}$$

B.2. Dérivation des gradients MP2 dans la base des spinorbitales

En utilisant la relation donnant $F_{pq}^{(x)}$ (Eq. (A.28)) et en isolant les termes multipliés par la dérivée de la matrice de recouvrement,

$$\begin{aligned}
E^{(2)x} = & \sum_{ij}^{\text{occ}} \left\{ H_{ij}^{(x)} + \sum_k^{\text{occ}} [(ij|kk)^{(x)} - (ik|jk)^{(x)}] \right\} P_{ij}^{(2)} \\
& + \sum_{ab}^{\text{virt}} \left\{ H_{ab}^{(x)} + \sum_k^{\text{occ}} [(ab|kk)^{(x)} - (ak|bk)^{(x)}] \right\} P_{ab}^{(2)} \\
& + \sum_a^{\text{virt}} \sum_i^{\text{occ}} \left\{ H_{ai}^{(x)} + \sum_k^{\text{occ}} [(ai|kk)^{(x)} - (ak|ik)^{(x)}] \right\} P_{ai}^{(2)} \\
& + \sum_{ij}^{\text{occ}} S_{ij}^{(x)} W_{ij}^{(2)} [I] + \sum_{ab}^{\text{virt}} S_{ab}^{(x)} W_{ab}^{(2)} [I] + \sum_a^{\text{virt}} \sum_i^{\text{occ}} S_{ai}^{(x)} W_{ai}^{(2)} [I] \\
& + \sum_{ij}^{\text{occ}} S_{ij}^{(x)} W_{ij}^{(2)} [II] + \sum_{ab}^{\text{virt}} S_{ab}^{(x)} W_{ab}^{(2)} [II] + \sum_a^{\text{virt}} \sum_i^{\text{occ}} S_{ai}^{(x)} W_{ai}^{(2)} [II] \\
& + \sum_{ij}^{\text{occ}} S_{ij}^{(x)} W_{ij}^{(2)} [III] + \frac{1}{2} \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{virt}} T_{ij}^{ab} (ia||jb)^{(x)} \\
& + \sum_{ij}^{\text{occ}} \left\{ \sum_P^{\text{all}} (\varepsilon_{iP} U_{Pj}^x + U_{Pi}^x \varepsilon_{Pj}) + K_{ij}^x \right\} P_{ij}^{(2)} \\
& + \sum_{ab}^{\text{virt}} \left\{ \sum_P^{\text{all}} (\varepsilon_{aP} U_{Pb}^x + U_{Pa}^x \varepsilon_{Pb}) + K_{ab}^x \right\} P_{ab}^{(2)} \\
& + \sum_a^{\text{virt}} \sum_i^{\text{occ}} \left\{ \sum_P^{\text{all}} (\varepsilon_{aP} U_{Pi}^x + U_{Pa}^x \varepsilon_{Pi}) + K_{ai}^x \right\} P_{ai}^{(2)} \\
& + \sum_i^{\text{occ}} \sum_P^{\text{all}} (U_{Pi}^x - U_{iP}^x) P_{Pi}^{(2)} + \sum_i^{\text{occ}} \sum_P^{\text{all}} S_{Pi}^{(x)} W_{Pi}^{(2)} [I] \\
& + \sum_a^{\text{virt}} \sum_P^{\text{all}} (U_{Pa}^x - U_{aP}^x) P_{Pa}^{(2)} + \sum_a^{\text{virt}} \sum_P^{\text{all}} S_{Pa}^{(x)} W_{Pa}^{(2)} [I] \\
& + \sum_{IJ}^{\text{occ}} S_{IJ}^{(x)} W_{IJ}^{(2)} [III] + \sum_{iJ}^{\text{occ}} S_{iJ}^{(x)} W_{iJ}^{(2)} [III]
\end{aligned} \tag{B.62}$$

où nous avons défini des nouveaux termes correctifs de la matrice densité pondérée par l'énergie :

$$W_{ij}^{(2)} [II] = - \left(\frac{\varepsilon_i + \varepsilon_j}{2} \right) P_{ij}^{(2)} \tag{B.63}$$

$$W_{ab}^{(2)} [II] = - \left(\frac{\varepsilon_a + \varepsilon_b}{2} \right) P_{ab}^{(2)} \tag{B.64}$$

$$W_{ai}^{(2)} [II] = -\varepsilon_i P_{ai}^{(2)} \tag{B.65}$$

Annexe B. Expression des gradients MP2 dans le cadre de la méthode LSCF

et

$$W_{ij}^{(2)} [III] = -\frac{1}{2} \sum_{pq}^{\text{all}} A_{pqji} P_{pq}^{(2)} \quad (\text{B.66})$$

$$W_{IJ}^{(2)} [III] = -\frac{1}{2} \sum_{pq}^{\text{all}} A_{pqJI} P_{pq}^{(2)} \quad (\text{B.67})$$

$$W_{iJ}^{(2)} [III] = -\sum_{pq}^{\text{all}} A_{pqJi} P_{pq}^{(2)} \quad (\text{B.68})$$

En plus des termes standards de la matrice densité pondérée par l'énergie, nous introduisons les termes dus aux blocs SLBOs-MOs suivant les relations :

$$\begin{aligned} \sum_{ij}^{\text{occ}} \sum_P^{\text{all}} (\varepsilon_{iP} U_{Pj}^x + U_{Pi}^x \varepsilon_{Pj}) P_{ij}^{(2)} &= \sum_{ij}^{\text{occ}} \sum_P^{\text{all}} \left(\varepsilon_{iP} U_{Pj}^x - S_{iP}^{(x)} \varepsilon_{Pj} - U_{iP}^x \varepsilon_{Pj} \right) P_{ij}^{(2)} \\ &= \sum_{ij}^{\text{occ}} \sum_P^{\text{all}} (\varepsilon_{iP} U_{Pj}^x - U_{iP}^x \varepsilon_{Pj}) P_{ij}^{(2)} \\ &+ \sum_i^{\text{occ}} \sum_P^{\text{all}} W_{Pi}^{(2)} [II] S_{Pi}^{(x)} \end{aligned} \quad (\text{B.69})$$

avec

$$W_{Pi}^{(2)} [II] = -\sum_j^{\text{occ}} \varepsilon_{Pj} P_{ij}^{(2)} \quad (\text{B.70})$$

Le second terme de l'expression précédente peut être simplifié en utilisant le fait que la matrice ε soit symétrique, et que $P_{ij}^{(2)} = P_{ji}^{(2)}$:

$$\begin{aligned} \sum_{ij}^{\text{occ}} \sum_P^{\text{all}} (\varepsilon_{iP} U_{Pj}^x - U_{iP}^x \varepsilon_{Pj}) P_{ij}^{(2)} &= \sum_{ij}^{\text{occ}} \sum_P^{\text{all}} \varepsilon_{iP} U_{Pj}^x P_{ij}^{(2)} - \sum_{ij}^{\text{occ}} \sum_P^{\text{all}} U_{iP}^x \varepsilon_{Pj} P_{ij}^{(2)} \\ &= \sum_{ij}^{\text{occ}} \sum_P^{\text{all}} \varepsilon_{iP} U_{Pj}^x P_{ij}^{(2)} - \sum_{ji}^{\text{occ}} \sum_P^{\text{all}} U_{jP}^x \varepsilon_{Pi} P_{ji}^{(2)} \\ &= \sum_{ij}^{\text{occ}} \sum_P^{\text{all}} \varepsilon_{Pi} (U_{Pj}^x - U_{jP}^x) P_{ij}^{(2)} \end{aligned} \quad (\text{B.71})$$

L'introduction de cette simplification dans l'équation (B.69) conduit à :

$$\begin{aligned} \sum_{ij}^{\text{occ}} \sum_P^{\text{all}} (\varepsilon_{iP} U_{Pj}^x + U_{Pi}^x \varepsilon_{Pj}) P_{ij}^{(2)} &= \sum_{ij}^{\text{occ}} \sum_P^{\text{all}} \left(\varepsilon_{iP} U_{Pj}^x - S_{iP}^{(x)} \varepsilon_{Pj} - U_{iP}^x \varepsilon_{Pj} \right) P_{ij}^{(2)} \\ &= \sum_{ij}^{\text{occ}} \sum_P^{\text{all}} \varepsilon_{Pi} (U_{Pj}^x - U_{jP}^x) P_{ij}^{(2)} \\ &+ \sum_i^{\text{occ}} \sum_P^{\text{all}} W_{Pi}^{(2)} [II] S_{Pi}^{(x)} \end{aligned} \quad (\text{B.72})$$

B.2. Dérivation des gradients MP2 dans la base des spinorbitales

En utilisant la même dérivation, le second terme faisant intervenir le bloc SLBOs-MOs donne :

$$\begin{aligned}
 \sum_{ab}^{\text{virt}} \sum_P^{\text{all}} (\varepsilon_{aP} U_{Pb}^x + U_{Pa}^x \varepsilon_{Pb}) P_{ab}^{(2)} &= \sum_{ab}^{\text{virt}} \sum_P^{\text{all}} \left(\varepsilon_{aP} U_{Pb}^x - S_{aP}^{(x)} \varepsilon_{Pb} - U_{aP}^x \varepsilon_{Pb} \right) P_{ab}^{(2)} \\
 &= \sum_{ab}^{\text{virt}} \sum_P^{\text{all}} \varepsilon_{Pa} (U_{Pb}^x - U_{Pb}^x) P_{ab}^{(2)} \\
 &\quad + \sum_a^{\text{virt}} \sum_P^{\text{all}} W_{Pa}^{(2)} [II] S_{Pa}^{(x)}
 \end{aligned} \tag{B.73}$$

avec

$$W_{Pa}^{(2)} [II] = - \sum_b^{\text{virt}} \varepsilon_{Pb} P_{ab}^{(2)} \tag{B.74}$$

Pour le bloc contenant la contribution du *Z-vector*,

$$\begin{aligned}
 &\sum_a^{\text{virt}} \sum_i^{\text{occ}} \sum_P^{\text{all}} (\varepsilon_{aP} U_{Pi}^x + U_{Pa}^x \varepsilon_{Pi}) P_{ai}^{(2)} \\
 &= \frac{1}{2} \left[\sum_a^{\text{virt}} \sum_i^{\text{occ}} \sum_P^{\text{all}} (\varepsilon_{aP} U_{Pi}^x + U_{Pa}^x \varepsilon_{Pi}) P_{ai}^{(2)} + \sum_a^{\text{virt}} \sum_i^{\text{occ}} \sum_P^{\text{all}} (\varepsilon_{aP} U_{Pi}^x + U_{Pa}^x \varepsilon_{Pi}) P_{ai}^{(2)} \right] \\
 &= \frac{1}{2} \sum_a^{\text{virt}} \sum_i^{\text{occ}} \sum_P^{\text{all}} \left(\varepsilon_{aP} U_{Pi}^x - S_{aP}^{(x)} \varepsilon_{Pi} - U_{aP}^x \varepsilon_{Pi} \right) P_{ai}^{(2)} \\
 &\quad + \frac{1}{2} \sum_a^{\text{virt}} \sum_i^{\text{occ}} \sum_P^{\text{all}} \left(-\varepsilon_{aP} U_{iP}^x - \varepsilon_{aP} S_{iP}^{(x)} + U_{Pa}^x \varepsilon_{Pi} \right) P_{ai}^{(2)} \\
 &= \frac{1}{2} \sum_a^{\text{virt}} \sum_i^{\text{occ}} \sum_P^{\text{all}} \varepsilon_{Pa} (U_{Pi}^x - U_{iP}^x) P_{ai}^{(2)} + \frac{1}{2} \sum_a^{\text{virt}} \sum_i^{\text{occ}} \sum_P^{\text{all}} \varepsilon_{Pi} (U_{Pa}^x - U_{aP}^x) P_{ai}^{(2)} \\
 &\quad + \sum_i^{\text{occ}} \sum_P^{\text{all}} W_{Pi}^{(2)} [III] S_{Pi}^{(x)} + \sum_a^{\text{virt}} \sum_P^{\text{all}} W_{Pa}^{(2)} [III] S_{Pa}^{(x)}
 \end{aligned} \tag{B.75}$$

avec

$$W_{Pi}^{(2)} [III] = -\frac{1}{2} \sum_a^{\text{virt}} \varepsilon_{Pa} P_{ai}^{(2)} \tag{B.76}$$

$$W_{Pa}^{(2)} [III] = -\frac{1}{2} \sum_i^{\text{occ}} \varepsilon_{Pi} P_{ai}^{(2)} \tag{B.77}$$

Annexe B. Expression des gradients MP2 dans le cadre de la méthode LSCF

En introduisant l'ensemble de ces relations dans l'expression (B.62) de $E^{(2)x}$, cela nous conduit à la relation (B.78) :

$$\begin{aligned}
E^{(2)x} = & \sum_{ij}^{\text{occ}} \left\{ H_{ij}^{(x)} + \sum_k^{\text{occ}} [(ij|kk)^{(x)} - (ik|jk)^{(x)}] \right\} P_{ij}^{(2)} \\
& + \sum_{ab}^{\text{virt}} \left\{ H_{ab}^{(x)} + \sum_k^{\text{occ}} [(ab|kk)^{(x)} - (ak|bk)^{(x)}] \right\} P_{ab}^{(2)} \\
& + \sum_a^{\text{virt}} \sum_i^{\text{occ}} \left\{ H_{ai}^{(x)} + \sum_k^{\text{occ}} [(ai|kk)^{(x)} - (ak|ik)^{(x)}] \right\} P_{ai}^{(2)} \\
& + \sum_{ij}^{\text{occ}} S_{ij}^{(x)} W_{ij}^{(2)} [I] + \sum_{ab}^{\text{virt}} S_{ab}^{(x)} W_{ab}^{(2)} [I] + \sum_a^{\text{virt}} \sum_i^{\text{occ}} S_{ai}^{(x)} W_{ai}^{(2)} [I] \\
& + \sum_{ij}^{\text{occ}} S_{ij}^{(x)} W_{ij}^{(2)} [II] + \sum_{ab}^{\text{virt}} S_{ab}^{(x)} W_{ab}^{(2)} [II] + \sum_a^{\text{virt}} \sum_i^{\text{occ}} S_{ai}^{(x)} W_{ai}^{(2)} [II] \\
& + \sum_{ij}^{\text{occ}} S_{ij}^{(x)} W_{ij}^{(2)} [III] + \frac{1}{2} \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{virt}} T_{ij}^{ab} (ia||jb)^{(x)} \\
& + \sum_{ij}^{\text{occ}} \left\{ \left[\sum_P^{\text{all}} \varepsilon_{Pi} (U_{Pj}^x - U_{jP}^x) \right] + K_{ij}^x \right\} P_{ij}^{(2)} \\
& + \sum_{ab}^{\text{virt}} \left\{ \left[\sum_P^{\text{all}} \varepsilon_{Pa} (U_{Pb}^x - U_{bP}^x) \right] + K_{ab}^x \right\} P_{ab}^{(2)} \\
& + \sum_a^{\text{virt}} \sum_i^{\text{occ}} \left\{ \left[\sum_P^{\text{all}} \frac{\varepsilon_{Pa}}{2} (U_{Pi}^x - U_{iP}^x) + \frac{\varepsilon_{Pi}}{2} (U_{Pa}^x - U_{aP}^x) \right] + K_{ai}^x \right\} P_{ai}^{(2)} \\
& + \sum_i^{\text{occ}} \sum_P^{\text{all}} (U_{Pi}^x - U_{iP}^x) P_{Pi}^{(2)} + \sum_a^{\text{virt}} \sum_P^{\text{all}} (U_{Pa}^x - U_{aP}^x) P_{Pa}^{(2)} \\
& + \sum_i^{\text{occ}} \sum_P^{\text{all}} S_{Pi}^{(x)} W_{Pi}^{(2)} [I] + \sum_i^{\text{occ}} \sum_P^{\text{all}} S_{Pi}^{(x)} W_{Pi}^{(2)} [II] + \sum_i^{\text{occ}} \sum_P^{\text{all}} S_{Pi}^{(x)} W_{Pi}^{(2)} [III] \\
& + \sum_a^{\text{virt}} \sum_P^{\text{all}} S_{Pa}^{(x)} W_{Pa}^{(2)} [I] + \sum_a^{\text{virt}} \sum_P^{\text{all}} S_{Pa}^{(x)} W_{Pa}^{(2)} [II] + \sum_a^{\text{virt}} \sum_P^{\text{all}} S_{Pa}^{(x)} W_{Pa}^{(2)} [III] \\
& + \sum_{IJ}^{\text{occ}} S_{IJ}^{(x)} W_{IJ}^{(2)} [III] + \sum_{iJ}^{\text{occ}} S_{iJ}^{(x)} W_{iJ}^{(2)} [III]
\end{aligned} \tag{B.78}$$

Les termes 1, 2, 3 et 11 qui contiennent la dérivé des intégrales biélectroniques contribuent à la matrice densité à deux particules. Plus particulièrement, les termes 1, 2 et 3 contribuent à la partie séparable ($\Gamma_{\mu\nu\lambda\sigma}^S$) de la matrice densité à deux particules, tandis que le onzième terme fait partie de la partie non-séparable ($\Gamma_{\mu\nu\lambda\sigma}^{\text{NS}}$). Nous présenterons leur formes explicites sur la base des AOs par la suite. Les onze premiers termes de l'expres-

B.2. Dérivation des gradients MP2 dans la base des spinorbitales

sion (B.78) correspondent à l'expression de $E^{(2)x}$ dans le cas d'un calcul SCF standard, tandis que les treize derniers représentent la correction lors d'un calcul LSCF. Avant de détailler la dérivation de ces équations dans le cas UHF, puis RHF, nous allons résumer les expressions des termes de la matrice densité à une particule et de la matrice densité pondérée par l'énergie :

– Matrice densité à une particule :

$$P_{ij}^{(2)} = -\frac{1}{2} \sum_k^{\text{occ}} \sum_{ab}^{\text{virt}} T_{ik}^{ab} T_{jk}^{ab} \quad (\text{B.79})$$

$$P_{ab}^{(2)} = \frac{1}{2} \sum_{ij}^{\text{occ}} \sum_c^{\text{virt}} T_{ij}^{ac} T_{ij}^{bc} \quad (\text{B.80})$$

$$P_{ai}^{(2)} = Z_{ai} \quad (\text{B.81})$$

$$P_{PQ}^{(2)} = 0 \quad (\text{B.82})$$

$$P_{Pi}^{(2)} = \frac{1}{2} \sum_k^{\text{occ}} \sum_{ab}^{\text{virt}} T_{ik}^{ab} (Pa||kb) \quad (\text{B.83})$$

$$P_{Pa}^{(2)} = \frac{1}{2} \sum_{ij}^{\text{occ}} \sum_c^{\text{virt}} T_{ij}^{ac} (iP||jc) \quad (\text{B.84})$$

– Matrice densité pondérée par l'énergie :

$$W_{ij}^{(2)} = -\frac{1}{2} \sum_k^{\text{occ}} \sum_{ab}^{\text{virt}} T_{ik}^{ab} (ja||kb) - \frac{1}{2} (\varepsilon_i + \varepsilon_j) P_{ij}^{(2)} - \frac{1}{2} \sum_{pq}^{\text{all}} A_{pqji} P_{ij}^{(2)} \quad (\text{B.85})$$

$$W_{ab}^{(2)} = -\frac{1}{2} \sum_{ij}^{\text{occ}} \sum_c^{\text{virt}} T_{ij}^{ac} (ib||jc) - \frac{1}{2} (\varepsilon_a + \varepsilon_b) P_{ab}^{(2)} \quad (\text{B.86})$$

$$W_{ai}^{(2)} = -\sum_{jk}^{\text{occ}} \sum_b^{\text{virt}} T_{jk}^{ab} (ij||bk) - \varepsilon_i P_{ai}^{(2)} \quad (\text{B.87})$$

$$W_{Pi}^{(2)} = -\frac{1}{2} \sum_k^{\text{occ}} \sum_{ab}^{\text{virt}} T_{ik}^{ab} (Pa||kb) - \sum_j^{\text{occ}} \varepsilon_{Pj} P_{ij}^{(2)} - \frac{1}{2} \sum_b^{\text{virt}} \varepsilon_{Pb} P_{ab}^{(2)} \quad (\text{B.88})$$

$$W_{Pa}^{(2)} = -\frac{1}{2} \sum_{ij}^{\text{occ}} \sum_c^{\text{virt}} T_{ij}^{ac} (iP||jc) - \sum_b^{\text{virt}} \varepsilon_{Pb} P_{ab}^{(2)} - \frac{1}{2} \sum_i^{\text{occ}} \varepsilon_{Pi} P_{ai}^{(2)} \quad (\text{B.89})$$

$$W_{IJ}^{(2)} = -\frac{1}{2} \sum_{pq}^{\text{all}} A_{pqJI} P_{pq}^{(2)} \quad (\text{B.90})$$

$$W_{iJ}^{(2)} = -\sum_{pq}^{\text{all}} A_{pqJi} P_{pq}^{(2)} \quad (\text{B.91})$$

B.3 Dérivation des gradients MP2 pour un système à couche ouverte

Dans cette section, nous allons développer en formalisme non-restreint les équations précédemment obtenues sur la base des spinorbitales. Nous expliciterons les expressions des matrices densités à une et deux particules, ainsi que celles de la matrice densité pondérée par l'énergie et le lagrangien MP2. L'expression des équations CPHF dans le cas d'un système à couche ouverte a été reportée dans la section A.2. Pour dériver les équations UHF, nous allons séparer les sommations sur les orbitales de spin α et β en partant des expressions sur la base des spinorbitales :

$$\sum_t^{\text{all}} = \sum_{t^\alpha}^{\text{all}^\alpha} + \sum_{t^\beta}^{\text{all}^\beta} \quad (\text{B.92})$$

B.3.1 Matrice densité à une particule

En décomposant les sommations, l'expression de $P_{ij}^{(2)}$ devient donc :

$$P_{ij}^{(2)} = P_{i^\alpha j^\alpha}^{(2)} + P_{i^\beta j^\beta}^{(2)} + P_{i^\alpha j^\beta}^{(2)} + P_{i^\beta j^\alpha}^{(2)} \quad (\text{B.93})$$

Les termes contenant à la fois des MOs de type α et β sont nuls de par l'intégration sur les variables de spin : $P_{i^\alpha j^\beta}^{(2)} = 0$ et $P_{i^\beta j^\alpha}^{(2)} = 0$. Il ne reste donc que les composantes $(\alpha\alpha)$ et $(\beta\beta)$ qui possèdent des expressions identiques et qui peuvent être déterminées en échangeant α et β . Pour l'expression de $P_{i^\alpha j^\alpha}^{(2)}$ par exemple :

$$\begin{aligned} P_{i^\alpha j^\alpha}^{(2)} = & - \sum_{k^\alpha}^{\text{occ}^\alpha} \sum_{ab^\alpha}^{\text{virt}^\alpha} \frac{[(i^\alpha a^\alpha | k^\alpha b^\alpha) - (i^\alpha b^\alpha | k^\alpha a^\alpha)] (j^\alpha a^\alpha | k^\alpha b^\alpha)}{D_{i^\alpha k^\alpha}^{a^\alpha b^\alpha} D_{j^\alpha k^\alpha}^{a^\alpha b^\alpha}} \\ & - \sum_{k^\beta}^{\text{occ}^\beta} \sum_{a^\alpha}^{\text{virt}^\alpha} \sum_{b^\beta}^{\text{virt}^\beta} \frac{[(i^\alpha a^\alpha | k^\beta b^\beta)] [(j^\alpha a^\alpha | k^\beta b^\beta)]}{D_{i^\alpha k^\beta}^{a^\alpha b^\beta} D_{j^\alpha k^\beta}^{a^\alpha b^\beta}} \end{aligned} \quad (\text{B.94})$$

Une relation identique peut être déterminée pour $P_{i^\beta j^\beta}^{(2)}$. La correction au deuxième ordre de la matrice densité du bloc virtuelles-virtuelles est,

$$\begin{aligned} P_{a^\alpha b^\alpha}^{(2)} = & - \sum_{i^\alpha j^\alpha}^{\text{occ}^\alpha} \sum_{c^\alpha}^{\text{virt}^\alpha} \frac{[(i^\alpha a^\alpha | j^\alpha c^\alpha) - (i^\alpha c^\alpha | j^\alpha a^\alpha)] (i^\alpha b^\alpha | j^\alpha c^\alpha)}{D_{i^\alpha j^\alpha}^{a^\alpha c^\alpha} D_{i^\alpha j^\alpha}^{b^\alpha c^\alpha}} \\ & - \sum_{i^\alpha}^{\text{occ}^\alpha} \sum_{j^\beta}^{\text{virt}^\beta} \sum_{c^\beta}^{\text{virt}^\beta} \frac{[(i^\alpha a^\alpha | j^\beta c^\beta)] [(i^\alpha b^\alpha | j^\beta c^\beta)]}{D_{i^\alpha j^\beta}^{a^\alpha c^\beta} D_{i^\alpha j^\beta}^{b^\alpha c^\beta}} \end{aligned} \quad (\text{B.95})$$

tandis que la correction contenant la contribution du *Z-vector*,

$$P_{a^\alpha i^\alpha}^{(2)} = Z_{a^\alpha i^\alpha} \quad (\text{B.96})$$

B.3. Dérivation des gradients MP2 pour un système à couche ouverte

où $Z_{a^\alpha i^\alpha}$ est déterminé en utilisant les équations CPHF dans le cas UHF (Eq. (A.55)). Nous détaillerons ci-dessous la forme explicite du lagrangien MP2 qui intervient dans ces équations. Pour les corrections du bloc SLBOs-MOs :

$$P_{P^\alpha i^\alpha}^{(2)} = \sum_{k^\alpha} \sum_{a^\alpha b^\alpha}^{\text{occ}^\alpha \text{virt}^\alpha} \frac{(P^\alpha a^\alpha | k^\alpha b^\alpha) - (P^\alpha b^\alpha | k^\alpha a^\alpha)}{D_{P^\alpha k^\alpha}^{a^\alpha b^\alpha}} (i^\alpha a^\alpha | k^\alpha b^\alpha) + \sum_{k^\beta} \sum_{a^\alpha} \sum_{b^\beta}^{\text{occ}^\beta \text{virt}^\alpha \text{virt}^\beta} \frac{(P^\alpha a^\alpha | k^\beta b^\beta)}{D_{P^\alpha k^\beta}^{a^\alpha b^\beta}} (i^\alpha a^\alpha | k^\beta b^\beta) \quad (\text{B.97})$$

$$P_{P^\alpha a^\alpha}^{(2)} = \sum_{i^\alpha j^\alpha} \sum_{c^\alpha}^{\text{occ}^\alpha \text{virt}^\alpha} \frac{(i^\alpha P^\alpha | j^\alpha c^\alpha) - (i^\alpha c^\alpha | j^\alpha P^\alpha)}{D_{i^\alpha j^\alpha}^{P^\alpha c^\alpha}} (i^\alpha a^\alpha | j^\alpha c^\alpha) + \sum_{i^\alpha} \sum_{j^\beta} \sum_{c^\beta}^{\text{occ}^\alpha \text{virt}^\beta \text{virt}^\beta} \frac{(i^\alpha P^\alpha | j^\beta c^\beta)}{D_{i^\alpha j^\beta}^{P^\alpha c^\beta}} (i^\alpha a^\alpha | j^\beta c^\beta) \quad (\text{B.98})$$

B.3.2 Matrice densité pondérée par l'énergie

Les différents termes de la matrice densité pondérée par l'énergie sont déterminés suivant les expressions :

$$W_{i^\alpha j^\alpha}^{(2)} [I] = - \sum_{k^\alpha} \sum_{a^\alpha b^\alpha}^{\text{occ}^\alpha \text{virt}^\alpha} \frac{(i^\alpha a^\alpha | k^\alpha b^\alpha) - (i^\alpha b^\alpha | k^\alpha a^\alpha)}{D_{i^\alpha k^\alpha}^{a^\alpha b^\alpha}} (j^\alpha a^\alpha | k^\alpha b^\alpha) - \sum_{k^\beta} \sum_{a^\alpha} \sum_{b^\beta}^{\text{occ}^\beta \text{virt}^\alpha \text{virt}^\beta} \frac{(i^\alpha a^\alpha | k^\beta b^\beta)}{D_{i^\alpha k^\beta}^{a^\alpha b^\beta}} (j^\alpha a^\alpha | k^\beta b^\beta) \quad (\text{B.99})$$

$$W_{a^\alpha b^\alpha}^{(2)} [I] = - \sum_{i^\alpha j^\alpha} \sum_{c^\alpha}^{\text{occ}^\alpha \text{virt}^\alpha} \frac{(j^\alpha a^\alpha | j^\alpha c^\alpha) - (j^\alpha c^\alpha | j^\alpha a^\alpha)}{D_{i^\alpha j^\alpha}^{a^\alpha c^\alpha}} (i^\alpha b^\alpha | j^\alpha c^\alpha) - \sum_{i^\alpha} \sum_{j^\beta} \sum_{c^\beta}^{\text{occ}^\alpha \text{virt}^\beta \text{virt}^\beta} \frac{(i^\alpha a^\alpha | j^\beta c^\beta)}{D_{i^\alpha j^\beta}^{a^\alpha c^\beta}} (i^\alpha b^\alpha | j^\beta c^\beta) \quad (\text{B.100})$$

$$W_{a^\alpha i^\alpha}^{(2)} [I] = -2 \sum_{j^\alpha k^\alpha} \sum_{b^\alpha}^{\text{occ}^\alpha \text{virt}^\alpha} \frac{(j^\alpha a^\alpha | k^\alpha b^\alpha) - (j^\alpha b^\alpha | k^\alpha a^\alpha)}{D_{j^\alpha k^\alpha}^{a^\alpha b^\alpha}} (i^\alpha j^\alpha | b^\alpha k^\alpha) - 2 \sum_{j^\alpha} \sum_{k^\beta} \sum_{b^\beta}^{\text{occ}^\alpha \text{virt}^\beta \text{virt}^\beta} \frac{(j^\alpha a^\alpha | k^\beta b^\beta)}{D_{j^\alpha k^\beta}^{a^\alpha b^\beta}} (i^\alpha j^\alpha | b^\beta k^\beta) \quad (\text{B.101})$$

Annexe B. Expression des gradients MP2 dans le cadre de la méthode LSCF

Pour le bloc SLBOs-MOs :

$$\begin{aligned}
 W_{P^\alpha i^\alpha}^{(2)} [I] &= - \sum_{k^\alpha}^{\text{occ}^\alpha} \sum_{a^\alpha b^\alpha}^{\text{virt}^\alpha} \frac{(i^\alpha a^\alpha | k^\alpha b^\alpha) - (i^\alpha b^\alpha | k^\alpha a^\alpha)}{D_{i^\alpha k^\alpha}^{a^\alpha b^\alpha}} (P^\alpha a^\alpha | k^\alpha b^\alpha) \\
 &\quad - \sum_{k^\beta}^{\text{occ}^\beta} \sum_{a^\alpha}^{\text{virt}^\alpha} \sum_{b^\beta}^{\text{virt}^\beta} \frac{(i^\alpha a^\alpha | k^\beta b^\beta)}{D_{i^\alpha k^\beta}^{a^\alpha b^\beta}} (P^\alpha a^\alpha | k^\beta b^\beta)
 \end{aligned} \tag{B.102}$$

$$\begin{aligned}
 W_{P^\alpha a^\alpha}^{(2)} [I] &= - \sum_{i^\alpha j^\alpha}^{\text{occ}^\alpha} \sum_{c^\alpha}^{\text{virt}^\alpha} \frac{(i^\alpha P^\alpha | j^\alpha c^\alpha) - (i^\alpha c^\alpha | j^\alpha P^\alpha)}{D_{i^\alpha j^\alpha}^{P^\alpha c^\alpha}} (i^\alpha a^\alpha | j^\alpha c^\alpha) \\
 &\quad - \sum_{i^\alpha}^{\text{occ}^\alpha} \sum_{j^\beta}^{\text{virt}^\beta} \sum_{c^\beta}^{\text{virt}^\beta} \frac{(i^\alpha P^\alpha | j^\beta c^\beta)}{D_{i^\alpha j^\beta}^{P^\alpha c^\beta}} (i^\alpha a^\alpha | j^\beta c^\beta)
 \end{aligned} \tag{B.103}$$

Pour ce qui est des contributions notées $[II]$ originaires des trois différents blocs MOs-MOs :

$$W_{i^\alpha j^\alpha}^{(2)} [II] = -\frac{1}{2} P_{i^\alpha j^\alpha}^{(2)} (\varepsilon_i^\alpha + \varepsilon_j^\alpha) \tag{B.104}$$

$$W_{a^\alpha b^\alpha}^{(2)} [II] = -\frac{1}{2} P_{a^\alpha b^\alpha}^{(2)} (\varepsilon_a^\alpha + \varepsilon_b^\alpha) \tag{B.105}$$

$$W_{a^\alpha i^\alpha}^{(2)} [II] = -P_{a^\alpha i^\alpha}^{(2)} \varepsilon_i^\alpha \tag{B.106}$$

Les expressions de $\mathbf{W}^{(2)} [II]$ et $\mathbf{W}^{(2)} [III]$ des blocs SLBOs-MOs sont :

$$W_{P^\alpha i^\alpha}^{(2)} [II] = - \sum_{j^\alpha}^{\text{occ}^\alpha} \varepsilon_{P^\alpha j^\alpha} P_{i^\alpha j^\alpha}^{(2)} \tag{B.107}$$

$$W_{P^\alpha a^\alpha}^{(2)} [II] = - \sum_{b^\alpha}^{\text{virt}^\alpha} \varepsilon_{P^\alpha b^\alpha} P_{a^\alpha b^\alpha}^{(2)} \tag{B.108}$$

$$W_{P^\alpha i^\alpha}^{(2)} [III] = -\frac{1}{2} \sum_{b^\alpha}^{\text{virt}^\alpha} \varepsilon_{P^\alpha b^\alpha} P_{a^\alpha b^\alpha}^{(2)} \tag{B.109}$$

$$W_{P^\alpha a^\alpha}^{(2)} [III] = -\frac{1}{2} \sum_{i^\alpha}^{\text{occ}^\alpha} \varepsilon_{P^\alpha i^\alpha} P_{a^\alpha i^\alpha}^{(2)} \tag{B.110}$$

B.3. Dérivation des gradients MP2 pour un système à couche ouverte

Enfin, les contributions provenant des blocs contenant les MOs et les SLBOs occupées sont obtenues suivant les relations :

$$W_{i^\alpha j^\alpha}^{(2)} [III] = - \sum_{p^\alpha q^\alpha}^{\text{all}^\alpha} P_{p^\alpha q^\alpha}^{(2)} [(p^\alpha q^\alpha | i^\alpha j^\alpha) - (p^\alpha i^\alpha | q^\alpha j^\alpha)] - \sum_{p^\beta q^\beta}^{\text{all}^\beta} P_{p^\beta q^\beta}^{(2)} (p^\beta q^\beta | i^\alpha j^\alpha) \quad (\text{B.111})$$

$$W_{I^\alpha J^\alpha}^{(2)} [III] = - \sum_{p^\alpha q^\alpha}^{\text{all}^\alpha} P_{p^\alpha q^\alpha}^{(2)} [(p^\alpha q^\alpha | I^\alpha J^\alpha) - (p^\alpha I^\alpha | q^\alpha J^\alpha)] - \sum_{p^\beta q^\beta}^{\text{all}^\beta} P_{p^\beta q^\beta}^{(2)} (p^\beta q^\beta | I^\alpha J^\alpha) \quad (\text{B.112})$$

$$W_{i^\alpha J^\alpha}^{(2)} [III] = -2 \sum_{p^\alpha q^\alpha}^{\text{all}^\alpha} P_{p^\alpha q^\alpha}^{(2)} [(p^\alpha q^\alpha | i^\alpha J^\alpha) - (p^\alpha i^\alpha | q^\alpha J^\alpha)] - 2 \sum_{p^\beta q^\beta}^{\text{all}^\beta} P_{p^\beta q^\beta}^{(2)} (p^\beta q^\beta | i^\alpha J^\alpha) \quad (\text{B.113})$$

B.3.3 Lagrangien MP2

Dans le cas d'un système à couche ouverte, la partie ($\alpha\alpha$) du lagrangien MP2 s'exprime sous la forme :

$$\begin{aligned} L_{a^\alpha i^\alpha} &= \sum_{j^\alpha k^\alpha}^{\text{occ}^\alpha} A_{a^\alpha i^\alpha j^\alpha k^\alpha} P_{j^\alpha k^\alpha}^{(2)} + \sum_{j^\beta k^\beta}^{\text{occ}^\beta} A_{a^\alpha i^\alpha j^\beta k^\beta} P_{j^\beta k^\beta}^{(2)} \\ &+ \sum_{b^\alpha c^\alpha}^{\text{virt}^\alpha} A_{a^\alpha i^\alpha b^\alpha c^\alpha} P_{b^\alpha c^\alpha}^{(2)} + \sum_{b^\beta c^\beta}^{\text{virt}^\beta} A_{a^\alpha i^\alpha b^\beta c^\beta} P_{b^\beta c^\beta}^{(2)} \\ &+ 2 \sum_{j^\alpha}^{\text{occ}^\alpha} \sum_{b^\alpha c^\alpha}^{\text{virt}^\alpha} \frac{(i^\alpha b^\alpha | j^\alpha c^\alpha) - (i^\alpha c^\alpha | j^\alpha b^\alpha)}{D_{i^\alpha j^\alpha}^{b^\alpha c^\alpha}} (a^\alpha b^\alpha | j^\alpha c^\alpha) \\ &+ 2 \sum_{j^\alpha}^{\text{occ}^\alpha} \sum_{b^\alpha}^{\text{virt}^\alpha} \sum_{c^\beta}^{\text{virt}^\beta} \frac{(i^\alpha b^\beta | j^\alpha c^\beta)}{D_{i^\alpha j^\alpha}^{b^\beta c^\beta}} (a^\alpha b^\alpha | j^\beta c^\beta) \\ &- 2 \sum_{j^\alpha k^\alpha}^{\text{occ}^\alpha} \sum_{b^\alpha}^{\text{virt}^\alpha} \frac{(j^\alpha a^\alpha | k^\alpha b^\alpha) - (j^\alpha b^\alpha | k^\alpha a^\alpha)}{D_{i^\alpha k^\alpha}^{a^\alpha b^\alpha}} (i^\alpha j^\alpha | b^\alpha k^\alpha) \\ &- 2 \sum_{j^\alpha}^{\text{occ}^\alpha} \sum_{k^\alpha}^{\text{occ}^\alpha} \sum_{b^\alpha}^{\text{virt}^\alpha} \frac{(j^\alpha a^\alpha | k^\beta b^\beta)}{D_{j^\alpha a^\beta}^{k^\alpha b^\beta}} (i^\alpha j^\alpha | b^\beta k^\beta) \end{aligned} \quad (\text{B.114})$$

Annexe B. Expression des gradients MP2 dans le cadre de la méthode LSCF

On peut alors déterminer les éléments $U_{a^\alpha i^\alpha}^x$ et $U_{a^\beta i^\beta}^x$ grâce aux équations CPHF (Eq. (A.55)) à l'aide de la méthode du *Z-vector* :

$$\sum_a^{\text{virt}} \sum_i^{\text{occ}} U_{ai}^x L_{ai} = \sum_{a^\alpha}^{\text{virt}^\alpha} \sum_{i^\alpha}^{\text{occ}^\alpha} U_{a^\alpha i^\alpha}^x L_{a^\alpha i^\alpha} + \sum_{a^\beta}^{\text{virt}^\beta} \sum_{i^\beta}^{\text{occ}^\beta} U_{a^\beta i^\beta}^x L_{a^\beta i^\beta} \quad (\text{B.115})$$

$$\sum_{a^\alpha}^{\text{virt}^\alpha} \sum_{i^\alpha}^{\text{occ}^\alpha} U_{a^\alpha i^\alpha}^x L_{a^\alpha i^\alpha} = \sum_{a^\alpha}^{\text{virt}^\alpha} \sum_{i^\alpha}^{\text{occ}^\alpha} B_{a^\alpha i^\alpha}^x Z_{a^\alpha i^\alpha} \quad (\text{B.116})$$

$$\sum_{a^\beta}^{\text{virt}^\beta} \sum_{i^\beta}^{\text{occ}^\beta} U_{a^\beta i^\beta}^x L_{a^\beta i^\beta} = \sum_{a^\beta}^{\text{virt}^\beta} \sum_{i^\beta}^{\text{occ}^\beta} B_{a^\beta i^\beta}^x Z_{a^\beta i^\beta} \quad (\text{B.117})$$

où les expressions de $B_{a^\alpha i^\alpha}^x$ et $B_{a^\beta i^\beta}^x$ sont fournies par l'équation (A.53).

B.3.4 Matrice densité à deux particules

Comme indiqué au début de cette annexe, la correction au deuxième ordre de la matrice densité à deux particules ($\Gamma_{\mu\nu\lambda\sigma}^{(2)}$) est décomposée en une contribution séparable ($\Gamma_{\mu\nu\lambda\sigma}^{\text{S}}$) et non-séparable ($\Gamma_{\mu\nu\lambda\sigma}^{\text{NS}}$) qui doivent être évaluées dans la base des AOs :

$$\Gamma_{\mu\nu\lambda\sigma}^{(2)} = \Gamma_{\mu\nu\lambda\sigma}^{\text{S}} + \Gamma_{\mu\nu\lambda\sigma}^{\text{NS}} \quad (\text{B.118})$$

La partie non-séparable exprimée sur la base des spinorbitales provient du onzième terme de la relation (B.78) :

$$\frac{1}{2} \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{virt}} T_{ij}^{ab} (ia||jb)^{(x)} \quad (\text{B.119})$$

En décomposant les sommations sur les MOs de spin α et β et en introduisant les coefficients des MOs sur la base des AOs, on arrive à la relation (B.120) :

$$\begin{aligned} \sum_{\mu\nu\lambda\sigma} \Gamma_{\mu\nu\lambda\sigma}^{\text{NS}} (\mu\nu|\lambda\sigma)^{(x)} &= \sum_{\mu\nu\lambda\sigma} \left\{ \sum_{i^\alpha j^\alpha}^{\text{occ}^\alpha} \sum_{a^\alpha b^\alpha}^{\text{virt}^\alpha} \frac{[(i^\alpha a^\alpha | j^\alpha b^\alpha) - (i^\alpha b^\alpha | j^\alpha a^\alpha)]}{\varepsilon_i^\alpha + \varepsilon_j^\alpha - \varepsilon_a^\alpha - \varepsilon_b^\alpha} c_{\mu i}^\alpha c_{\nu a}^\alpha c_{\lambda j}^\alpha c_{\sigma b}^\alpha \right. \\ &+ \sum_{i^\beta j^\beta}^{\text{occ}^\beta} \sum_{a^\beta b^\beta}^{\text{virt}^\beta} \frac{[(i^\beta a^\beta | j^\beta b^\beta) - (i^\beta b^\beta | j^\beta a^\beta)]}{\varepsilon_i^\beta + \varepsilon_j^\beta - \varepsilon_a^\beta - \varepsilon_b^\beta} c_{\mu i}^\beta c_{\nu a}^\beta c_{\lambda j}^\beta c_{\sigma b}^\beta \\ &+ 2 \sum_{i^\alpha}^{\text{occ}^\alpha} \sum_{j^\beta}^{\text{occ}^\beta} \sum_{a^\alpha}^{\text{virt}^\alpha} \sum_{b^\beta}^{\text{virt}^\beta} \frac{(i^\alpha a^\alpha | j^\beta b^\beta)}{\varepsilon_i^\alpha + \varepsilon_j^\beta - \varepsilon_a^\alpha - \varepsilon_b^\beta} c_{\mu i}^\alpha c_{\nu a}^\alpha c_{\lambda j}^\beta c_{\sigma b}^\beta \left. \right\} \\ &\times (\mu\nu|\lambda\sigma)^{(x)} \end{aligned} \quad (\text{B.120})$$

Le terme entre accolades représente donc la partie non-séparable $\Gamma_{\mu\nu\lambda\sigma}^{\text{NS}}$ de la matrice densité à deux particules. Nous tenons à faire remarquer que son expression est strictement identique au cas d'un calcul MP2 standard. Il ne demande donc pas de modification lors de l'implantation de la méthode LSCF. Intéressons nous maintenant à la partie séparable

B.3. Dérivation des gradients MP2 pour un système à couche ouverte

qui est composée, dans le cas de la dérivation sur la base des spinorbitales, des termes 1, 2 et 3 de l'expression (B.78). Il peut être factorisé sous la forme :

$$\sum_{pq}^{\text{all}} P_{pq}^{(2)} \sum_k^{\text{occ}} [(pq|kk)^{(x)} - (pk|qk)^{(x)}] \quad (\text{B.121})$$

En effectuant les mêmes opérations que sur la partie non-séparable, on obtient la relation (B.122) :

$$\begin{aligned} \sum_{\mu\nu\lambda\sigma} \Gamma_{\mu\nu\lambda\sigma}^S (\mu\nu|\lambda\sigma)^{(x)} &= \sum_{\mu\nu\lambda\sigma} \left[\left(\sum_{i^\alpha}^{\text{occ}^\alpha} c_{\lambda i}^\alpha c_{\sigma i}^\alpha \right) \left(\sum_{p^\alpha q^\alpha}^{\text{all}^\alpha} c_{\mu p}^\alpha c_{\nu q}^\alpha P_{p^\alpha q^\alpha}^{(2)} \right) \right. \\ &\quad \left. - \left(\sum_{i^\alpha}^{\text{occ}^\alpha} c_{\nu i}^\alpha c_{\sigma i}^\alpha \right) \left(\sum_{p^\alpha q^\alpha}^{\text{all}^\alpha} c_{\mu p}^\alpha c_{\lambda q}^\alpha P_{p^\alpha q^\alpha}^{(2)} \right) \right] (\mu\nu|\lambda\sigma)^{(x)} \\ &\quad + \sum_{\mu\nu\lambda\sigma} \left[\left(\sum_{i^\beta}^{\text{occ}^\beta} c_{\lambda i}^\beta c_{\sigma i}^\beta \right) \left(\sum_{p^\beta q^\beta}^{\text{all}^\beta} c_{\mu p}^\beta c_{\nu q}^\beta P_{p^\beta q^\beta}^{(2)} \right) \right. \\ &\quad \left. - \left(\sum_{i^\beta}^{\text{occ}^\beta} c_{\nu i}^\beta c_{\sigma i}^\beta \right) \left(\sum_{p^\beta q^\beta}^{\text{all}^\beta} c_{\mu p}^\beta c_{\lambda q}^\beta P_{p^\beta q^\beta}^{(2)} \right) \right] (\mu\nu|\lambda\sigma)^{(x)} \\ &\quad + \sum_{\mu\nu\lambda\sigma} \left[\left(\sum_{i^\beta}^{\text{occ}^\beta} c_{\lambda i}^\beta c_{\sigma i}^\beta \right) \left(\sum_{p^\alpha q^\alpha}^{\text{all}^\alpha} c_{\mu p}^\alpha c_{\nu q}^\alpha P_{p^\alpha q^\alpha}^{(2)} \right) \right] (\mu\nu|\lambda\sigma)^{(x)} \\ &\quad + \sum_{\mu\nu\lambda\sigma} \left[\left(\sum_{i^\alpha}^{\text{occ}^\alpha} c_{\lambda i}^\alpha c_{\sigma i}^\alpha \right) \left(\sum_{p^\beta q^\beta}^{\text{all}^\beta} c_{\mu p}^\beta c_{\nu q}^\beta P_{p^\beta q^\beta}^{(2)} \right) \right] (\mu\nu|\lambda\sigma)^{(x)} \end{aligned} \quad (\text{B.122})$$

En se rappelant les définitions des matrices densité α et β correspondant aux MOs variationnelles (Chapitre 4),

$$P_{\mu\nu}^{Q^\alpha} = \sum_{i^\alpha}^{\text{occ}^\alpha} c_{\mu i}^\alpha c_{\nu i}^\alpha \quad (\text{B.123})$$

$$P_{\mu\nu}^{Q^\beta} = \sum_{i^\beta}^{\text{occ}^\beta} c_{\mu i}^\beta c_{\nu i}^\beta \quad (\text{B.124})$$

et en introduisant les corrections MP2 des matrices densité sur la base des AOs,

$$P_{\mu\nu}^{Q^\alpha(2)} = \sum_{p^\alpha q^\alpha}^{\text{all}^\alpha} c_{\mu p}^\alpha c_{\nu q}^\alpha P_{p^\alpha q^\alpha}^{(2)} \quad (\text{B.125})$$

$$P_{\mu\nu}^{Q^\beta(2)} = \sum_{p^\beta q^\beta}^{\text{all}^\beta} c_{\mu p}^\beta c_{\nu q}^\beta P_{p^\beta q^\beta}^{(2)} \quad (\text{B.126})$$

$$(\text{B.127})$$

Annexe B. Expression des gradients MP2 dans le cadre de la méthode LSCF

la relation (B.122) devient :

$$\sum_{\mu\nu\lambda\sigma} \Gamma_{\mu\nu\lambda\sigma}^S (\mu\nu|\lambda\sigma)^{(x)} = \sum_{\mu\nu\lambda\sigma} \left[P_{\lambda\sigma}^{Q\alpha} P_{\mu\nu}^{Q\alpha(2)} - P_{\nu\sigma}^{Q\alpha} P_{\mu\lambda}^{Q\alpha(2)} + P_{\lambda\sigma}^{Q\beta} P_{\mu\nu}^{Q\beta(2)} - P_{\nu\sigma}^{Q\beta} P_{\mu\lambda}^{Q\beta(2)} + P_{\lambda\sigma}^{Q\beta} P_{\mu\nu}^{Q\alpha(2)} + P_{\lambda\sigma}^{Q\alpha} P_{\mu\nu}^{Q\beta(2)} \right] (\mu\nu|\lambda\sigma)^{(x)} \quad (\text{B.128})$$

Le terme entre crochets peut donc être identifié comme $\Gamma_{\mu\nu\lambda\sigma}^S$. En utilisant la matrice densité \mathbf{P}^Q à la place de $\mathbf{P}^T = \mathbf{P}^Q + \mathbf{P}^G$, le terme séparable de la matrice densité à deux particules reste inchangé lors d'un calcul LSCF. Nous allons maintenant pouvoir définir les matrices densité MP2 à une particule $P_{\mu\nu}^{\text{MP2}}$, pondérée par l'énergie $W_{\mu\nu}^{\text{MP2}}$ et à deux particules $\Gamma_{\mu\nu\lambda\sigma}^{\text{MP2}}$ sur la base des AOs en formalisme non-restreint :

$$P_{\mu\nu}^{\text{MP2}} = P_{\mu\nu}^{Q\alpha} + P_{\mu\nu}^{Q\beta} + P_{\mu\nu}^{Q\alpha(2)} + P_{\mu\nu}^{Q\beta(2)} \quad (\text{B.129})$$

$$W_{\mu\nu}^{\text{MP2}} = W_{\mu\nu}^{Q\alpha} + W_{\mu\nu}^{Q\beta} + W_{\mu\nu}^{Q\alpha(2)} + W_{\mu\nu}^{Q\beta(2)} \quad (\text{B.130})$$

$$\Gamma_{\mu\nu\lambda\sigma}^{\text{MP2}} = \Gamma_{\mu\nu\lambda\sigma}^Q + \Gamma_{\mu\nu\lambda\sigma}^{(2)} \quad (\text{B.131})$$

avec

$$W_{\mu\nu}^{Q\alpha} = \sum_{i^\alpha}^{\text{occ}^\alpha} \varepsilon_{i^\alpha} c_{\mu i}^\alpha c_{\nu i}^\alpha + \sum_{i^\beta}^{\text{occ}^\beta} \varepsilon_{i^\beta} c_{\mu i}^\beta c_{\nu i}^\beta \quad (\text{B.132})$$

$$\Gamma_{\mu\nu\lambda\sigma}^Q = \frac{1}{2} \left(P_{\mu\nu}^{Q\alpha} P_{\lambda\sigma}^{Q\alpha} - P_{\mu\lambda}^{Q\alpha} P_{\nu\sigma}^{Q\alpha} + P_{\mu\nu}^{Q\beta} P_{\lambda\sigma}^{Q\beta} - P_{\mu\lambda}^{Q\beta} P_{\nu\sigma}^{Q\beta} + P_{\mu\nu}^{Q\alpha} P_{\lambda\sigma}^{Q\alpha} + P_{\mu\nu}^{Q\beta} P_{\lambda\sigma}^{Q\beta} \right) \quad (\text{B.133})$$

La connaissance de ces trois matrices densité ($P_{\mu\nu}^{\text{MP2}}$, $W_{\mu\nu}^{\text{MP2}}$ et $\Gamma_{\mu\nu\lambda\sigma}^{\text{MP2}}$) permet de calculer les gradients de l'énergie MP2 données par l'équation (B.5).

B.4 Dérivation des gradients MP2 pour un système à couche fermée

L'expression des gradients MP2 étant maintenant connue dans le cas d'un système à couche ouverte, nous allons partir de ces expressions afin de déterminer leurs formes en formalisme restreint.

B.4.1 Matrice densité à une particule

L'expression de $P_{ij}^{(2)}$ est donc déterminée en sachant que $P_{i^\alpha j^\alpha}^{(2)} = P_{i^\beta j^\beta}^{(2)}$:

$$P_{ij}^{(2)} = P_{i^\alpha j^\alpha}^{(2)} + P_{i^\beta j^\beta}^{(2)} = 2P_{i^\alpha j^\alpha}^{(2)} \quad (\text{B.134})$$

B.4. Dérivation des gradients MP2 pour un système à couche fermée

Ceci nous permet de trouver l'expression des différentes composantes de la matrice densité à une particule :

$$P_{ij}^{(2)} = -2 \sum_k^{\text{occ}} \sum_{ab}^{\text{virt}} T_{ik}^{ab}(ja|kb) / D_{ik}^{ab} \quad (\text{B.135})$$

$$P_{ab}^{(2)} = 2 \sum_{ij}^{\text{occ}} \sum_c^{\text{virt}} T_{ij}^{ac}(ib|jc) / D_{ij}^{bc} \quad (\text{B.136})$$

$$P_{ai}^{(2)} = Z_{ai} \quad (\text{B.137})$$

où le terme Z_{ai} est déterminé grâce aux équations CPHF (Eq. (A.63)) et où les amplitudes sont définies, pour un système à couche fermée, par la relation :

$$T_{ij}^{ab} = \frac{[2(ia|jb) - (ib|ja)]}{D_{ij}^{ab}} \quad (\text{B.138})$$

Les termes des matrices densité à une particule pour le bloc SLBOs-MOs se développent selon :

$$P_{Pi}^{(2)} = 2 \sum_j^{\text{occ}} \sum_{ab}^{\text{virt}} T_{Pj}^{ab}(ia|jb) \quad (\text{B.139})$$

$$P_{Pa}^{(2)} = 2 \sum_{ij}^{\text{occ}} \sum_c^{\text{virt}} T_{ij}^{Pc}(ia|jc) \quad (\text{B.140})$$

B.4.2 Matrice densité pondérée par l'énergie

Concernant les termes de la matrice densité pondérée par l'énergie du bloc MOs-MOs ($\mathbf{W}^{(2)} [I]$ et $\mathbf{W}^{(2)} [II]$) :

$$W_{ij}^{(2)} [I] = -2 \sum_k^{\text{occ}} \sum_{ab}^{\text{virt}} T_{ik}^{ab}(ja|kb) \quad (\text{B.141})$$

$$W_{ab}^{(2)} [I] = -2 \sum_{ij}^{\text{occ}} \sum_c^{\text{virt}} T_{ij}^{ac}(ib|jc) \quad (\text{B.142})$$

$$W_{ai}^{(2)} [I] = -4 \sum_{jk}^{\text{occ}} \sum_b^{\text{virt}} T_{jk}^{ab}(ij|bk) \quad (\text{B.143})$$

$$W_{ij}^{(2)} [II] = -\frac{1}{2} P_{ij}^{(2)} (\varepsilon_i + \varepsilon_j) \quad (\text{B.144})$$

$$W_{ab}^{(2)} [II] = -\frac{1}{2} P_{ab}^{(2)} (\varepsilon_a + \varepsilon_b) \quad (\text{B.145})$$

$$W_{ai}^{(2)} [II] = -P_{ai}^{(2)} \varepsilon_i \quad (\text{B.146})$$

Annexe B. Expression des gradients MP2 dans le cadre de la méthode LSCF

Les expressions de $\mathbf{W}^{(2)} [I]$, $\mathbf{W}^{(2)} [II]$ et $\mathbf{W}^{(2)} [III]$ des blocs SLBOs-MOs sont :

$$W_{Pi}^{(2)} [I] = -2 \sum_k^{\text{occ}} \sum_{ab}^{\text{virt}} T_{Pk}^{ab} (ia|kb) \quad (\text{B.147})$$

$$W_{Pa}^{(2)} [I] = -2 \sum_{ij}^{\text{occ}} \sum_c^{\text{virt}} T_{ij}^{Pc} (ia|jc) \quad (\text{B.148})$$

$$W_{Pi}^{(2)} [II] = - \sum_j^{\text{occ}} \varepsilon_{Pj} P_{ij}^{(2)} \quad (\text{B.149})$$

$$W_{Pa}^{(2)} [II] = - \sum_b^{\text{virt}} \varepsilon_{Pb} P_{ab}^{(2)} \quad (\text{B.150})$$

$$W_{Pi}^{(2)} [III] = -\frac{1}{2} \sum_b^{\text{virt}} \varepsilon_{Pb} P_{ab}^{(2)} \quad (\text{B.151})$$

$$W_{Pa}^{(2)} [III] = -\frac{1}{2} \sum_i^{\text{occ}} \varepsilon_{Pi} P_{ai}^{(2)} \quad (\text{B.152})$$

Les termes notés $[III]$ de la matrice densité pondérée par l'énergie contenant les contributions des MOs et SLBOs occupées sont,

$$W_{ij}^{(2)} [III] = - \sum_{pq}^{\text{all}} P_{pq}^{(2)} [2(pq|ij) - (pi|qj)] \quad (\text{B.153})$$

$$W_{IJ}^{(2)} [III] = - \sum_{pq}^{\text{all}} P_{pq}^{(2)} [2(pq|IJ) - (pI|qJ)] \quad (\text{B.154})$$

$$W_{iJ}^{(2)} [III] = -2 \sum_{pq}^{\text{all}} P_{pq}^{(2)} [2(pq|iJ) - (pi|qJ)] \quad (\text{B.155})$$

Nous allons maintenant définir le lagrangien MP2 afin de déterminer la contribution du Z -vector par la résolution des équations CPHF en formalisme restreint (Eq. (A.63)).

B.4.3 Lagrangien MP2

Dans le cas d'un système à couche fermée, le lagrangien MP2 prend la forme :

$$L_{ai} = \sum_{jk}^{\text{occ}} A_{aijk} P_{jk}^{(2)} + \sum_{bc}^{\text{virt}} A_{aibc} P_{bc}^{(2)} + 4 \sum_j^{\text{occ}} \sum_{bc}^{\text{virt}} (ab|jc) T_{ij}^{bc} - 4 \sum_{jk}^{\text{occ}} \sum_b^{\text{virt}} (ij|bk) T_{jk}^{ab} \quad (\text{B.156})$$

sachant que

$$A_{pqrs} = 4(pq|rs) - (pr|qs) - (ps|qr) \quad (\text{B.157})$$

B.4.4 Matrice densité à deux particules

Ils nous restent à définir la matrice densité à deux particules pour un système à couche fermée. Sa partie non-séparable s'exprime sous la forme suivante :

$$\Gamma_{\mu\nu\lambda\sigma}^{\text{NS}} = 2 \sum_{ij} \sum_{ab} T_{ij}^{ab} c_{\mu i} c_{\nu a} c_{\lambda j} c_{\sigma b} \quad (\text{B.158})$$

Elle représente la transformation des amplitudes MP2 T_{ij}^{ab} de la base des MOs vers la base des AOs. La partie séparable se détermine à partir de son expression en formalisme non-restreint à l'aide de l'égalité des matrices densité à une particule \mathbf{P}^Q et $\mathbf{P}^{Q(2)}$ dans le cas du formalisme restreint ($P_{\lambda\sigma}^{Q\alpha} = P_{\lambda\sigma}^{Q\beta}$ et $P_{\mu\nu}^{Q\alpha(2)} = P_{\mu\nu}^{Q\beta(2)}$) :

$$\Gamma_{\mu\nu\lambda\sigma}^{\text{S}} = P_{\lambda\sigma}^Q P_{\mu\nu}^{Q(2)} - \frac{1}{2} P_{\nu\sigma}^Q P_{\mu\lambda}^{Q(2)} \quad (\text{B.159})$$

La partie séparable a une expression du même genre que la matrice densité à deux particules $\Gamma_{\mu\nu\lambda\sigma}^Q$. Nous allons donc définir les matrices densité MP2 qui, en formalisme restreint, suivent les relations :

$$P_{\mu\nu}^{\text{MP2}} = P_{\mu\nu}^Q + P_{\mu\nu}^{Q(2)} \quad (\text{B.160})$$

$$W_{\mu\nu}^{\text{MP2}} = 2W_{\mu\nu}^Q + W_{\mu\nu}^{Q(2)} \quad (\text{B.161})$$

$$\Gamma_{\mu\nu\lambda\sigma}^{\text{MP2}} = \Gamma_{\mu\nu\lambda\sigma}^Q + \Gamma_{\mu\nu\lambda\sigma}^{(2)} \quad (\text{B.162})$$

avec

$$P_{\mu\nu}^Q = 2 \sum_i^{\text{occ}} c_{\mu i} c_{\nu i} \quad (\text{B.163})$$

$$W_{\mu\nu}^Q = \sum_i^{\text{occ}} \varepsilon_i c_{\mu i} c_{\nu i} \quad (\text{B.164})$$

$$\Gamma_{\mu\nu\lambda\sigma}^Q = \frac{1}{2} P_{\mu\nu}^Q P_{\lambda\sigma}^Q - \frac{1}{4} P_{\mu\lambda}^Q P_{\nu\sigma}^Q \quad (\text{B.165})$$

Ceci clôt la détermination des gradients MP2 pour un système à couche fermée.

Annexe B. Expression des gradients MP2 dans le cadre de la méthode LSCF

Glossaire

Δ KS : Δ Kohn-Sham

AB : Azobenzène

ACA : Adjusted Connection Atom

ADN : Acide Désoxyribonucléique

AM1 : Austin Model 1

AMBER : Assisted Model Building with Energy Refinement

AMOEBA : Atomic Multipole Optimized Energetics for Biomolecular Applications

AO : Atomic Orbital

BF : Boys-Foster

BN : Bas Niveau

CAB : *Cis* Azobenzène

CASPT2 : Complete Active Space Perturbation Theory 2

CASSCF : Complete Active Space Self-Consistent Field

CC : Coupled-Cluster

CC2 : Coupled-Cluster 2

CHARMM : Chemistry at HARvard using Molecular Mechanics

CHELP : Charges from Electrostatic Potential

CHELPG : Charges from Electrostatic Potential, Grid Method

CI : Configuration Interaction

CIS : Configuration Interaction Single

CISD : Configuration Interaction Single Double

CO : Core Orbital

CPHF : Coupled Perturbed Hartree-Fock

D&C : Divide and Conquer

DC : Dichroïsme Circulaire

DEG : Diffraction des électrons en phase gazeuse

DFT : Density Functional Theory

DFTB : Density Functional based Tight Binding

DO : Dual Orbital

DRX : Diffraction des rayons X

EE : Electronic Embedding

EFP : Effective Fragment Potential

Glossaire

ELMO : Extremely Localized Molecular Orbitals

ER : Edmiston-Ruedenberg

EVB : Empirical Valence Bond

FBP : Frontier Bond Potential

FC : Frozen Core

FCO : Frozen Core Orbital

FO : Frozen Orbital

GB : Grande Base

GFP : Green Fluorescent Protein

GGA : Generalized Gradient Approximation

GHO : Generalized Hybrid Orbitals

GS : Gramm-Schmidt

GWA : GW Approximation

HF : Hartree-Fock

HK : Hohenberg-Kohn

HLA : Hydrogen Link Atom

HN : Haut Niveau

IMOMM : Integrated Molecular Orbital Molecular Mechanics

IMOMO : Integrated Molecular Orbital Molecular Orbital

IR : Infrarouge

KS : Kohn-Sham

L_v : Löwdin virtuelle

LA : Link Atom

LCAO : Linear Combination of Atomic Orbitals

LCC : Local Coupled-Cluster

LDA : Local Density Approximation

LH : Liaison Hydrogène

LJ : Lennard-Jones

LMO : Localized Molecular Orbital

LMP2 : Local Møller-Plesset 2

LO : Localized Orbital

LSCF : Local Self-Consistent Field

LSDA : Local Spin Density Approximation

MC : Monte-Carlo

MCSCF : Multiconfiguration Self-Consistent Field

MD : Molecular Dynamics

ME : Mechanical Embedding

MK : Merz-Kollman

MM : Molecular Mechanics

MMVB : Molecular Mechanics Valence Bond

MNDO : Modified Neglect of Diatomic Overlap
MO : Molecular Orbital
MP : Magnasco-Perico
MP3-MAIS : MP3-Method Adapted for Intermolecular Studies
MP3-PIF : MP3-Parametrizable Interaction Function
MPn : Møller-Plesset n

NEXAFS : Near Edge X-ray Absorption Fine Structure

OLSCF : Optimized Local Self-Consistent Field
ONIOM : Our own N-layer Integrated molecular Orbital and molecular Mechanics
OPLS : Optimized Potential for Liquid Simulation

PB : Petite Base
PB : Pseudobond
PBC : Periodic Boundary Conditions
PCILO : Perturbation Configuration Interaction using Localized Orbitals
PCM : Polarizable Continuum Model
PDB : Protein Data Bank
PM : Pipek-Mezey
PM3 : Parametric Model 3

QCP : Quantum Capping Potential
QM : Quantum Mechanics
QM/MM : Quantum Mechanics/Molecular Mechanics
QM/QM' : Quantum Mechanics/Quantum Mechanics'
QM:MM : Quantum Mechanics :Molecular Mechanics
QM:QM' : Quantum Mechanics :Quantum Mechanics'

RCD : Redistributed Charge and Dipole
ReaxFF : Reactive Force Field
RESP : Restreint ElectroStatic Potential
RHF : Restricted Hartree-Fock
RMSD : Root Mean Square Deviation
ROHF : Restricted Open-shell Hartree-Fock
RPA : Random Phase Approximation

SCCO : Self-Consistent Core Orbital
SCF : Self-Consistent Field
SCRf : Self-Consistent Reaction Field
SCSLBO : Self-Consistent Strictly Localized Bonding Orbital
SE : Semi-Empirique
SLABO : Strictly Localized Anti-Bonding Orbital
SLBO : Strictly Localized Bonding Orbital
SLO : Strictly Localized Orbital
SOPPA : Second Order Polarization Propagator Approximation
SPC : Simple Point Charge

Glossaire

SPC/E : Simple Point Charge/Extended

ST2 : Symmetrical Tetrad 2

TAB : *Trans* Azobenzène

TD-DFT : Time-Dependent Density Functional Theory

TD-HF : Time-Dependent Hartree-Fock

TDA : Tamm-Dancoff Approximation

TIP3P : Transferable intermolecular potential 3 Parameters

TIP4P : Transferable intermolecular potential 4 Parameters

TS : Transition State

UFF : Universal Force Field

uGTS : unrestricted Generalized Transition State

UHF : Unrestricted Hartree-Fock

UV-vis : Ultraviolet-visible

VN : von Niessen

WP : Weinstein-Pauncz

Bibliographie

- [1] A. C. T. van Duin, S. Dasgupta, F. Lorant, and W. A. Goddard III, *J. Phys. Chem. A*, **2001**, *105*, 9396.
- [2] A. Strachan, A. C. T. van Duin, D. Chakraborty, S. Dasgupta, and W. A. Goddard III, *Phys. Rev. Lett*, **2003**, *91*, 098301.
- [3] A. Strachan, E. M. Kober, A. C. T. van Duin, J. Oxgaard, and W. A. Goddard III, *J. Chem. Phys.*, **2005**, *122*, 054501.
- [4] K. D. Nielson, A. C. T. van Duin, J. Oxgaard, W.-Q. Deng, and W. A. Goddard III, *J. Phys. Chem. A*, **2005**, *109*, 493.
- [5] M. Buehler, A. C. T. van Duin, and W. A. Goddard III, *Phys. Rev. Lett.*, **2006**, *96*, 095505.
- [6] M. Buehler, H. Tang, A. C. T. van Duin, and W. A. Goddard III, *Phys. Rev. Lett*, **2007**, *99*, 165502.
- [7] A. Warshel and R. M. Weiss, *J. Am. Chem. Soc.*, **1980**, *102*, 6218.
- [8] J. Villà, J. Bentzien, A. González-Lafont, J. M. Lluch, J. Bertran, and A. Warshel, *J. Comput. Chem.*, **2000**, *21*, 607.
- [9] G. Hong, E. Rosta, and A. Warshel, *J. Phys. Chem. B*, **2000**, *110*, 19570.
- [10] G. Segal and J. A. Pople, *J. Chem. Phys.*, **1966**, *44*, 3289.
- [11] J. A. Pople, D. Beveridge, and P. Dobosh, *J. Chem. Phys.*, **1967**, *47*, 2026.
- [12] M. J. S. Dewar and W. Thiel, *J. Am. Chem. Soc.*, **1977**, *99*, 2338.
- [13] M. J. S. Dewar and W. Thiel, *J. Am. Chem. Soc.*, **1977**, *99*, 4899.
- [14] M. J. S. Dewar and H. S. Rzepa, *J. Am. Chem. Soc.*, **1978**, *100*, 777.
- [15] M. J. S. Dewar and M. L. McKee, *J. Comput. Chem.*, **1978**, *4*, 84.
- [16] L. P. Davis, R. M. Guidry, J. R. Williams, M. J. S. Dewar, and H. S. Rzepa, *J. Comput. Chem.*, **1981**, *2*, 433.
- [17] M. J. S. Dewar, M. L. McKee, and H. S. Rzepa, *J. Am. Chem. Soc.*, **1978**, *100*, 3607.
- [18] M. J. S. Dewar and E. F. Healy, *J. Comput. Chem.*, **1983**, *4*, 542.
- [19] M. J. S. Dewar, G. L. Grady, and J. J. P. Stewart, *J. Am. Chem. Soc.*, **1984**, *106*, 6771.
- [20] M. J. S. Dewar, G. L. Grady, K. M. Merz Jr., and J. J. P. Stewart, *Organometallics*, **1985**, *4*, 1964.

Bibliographie

- [21] M. J. S. Dewar and C. H. Reynolds, *J. Comput. Chem.*, **1986**, *2*, 140.
- [22] M. J. S. Dewar, E. G. Zoebisch, E. F. Healy, and J. J. P. Stewart, *J. Am. Chem. Soc.*, **1985**, *107*, 3902.
- [23] M. J. S. Dewar, C. Jie, and E. G. Zoebisch, *Organometallics*, **1988**, *7*, 513.
- [24] M. J. S. Dewar, , and K. M. Merz, *Organometallics*, **1988**, *7*, 522.
- [25] M. J. S. Dewar and C. Jie, *J. Mol. Struct.*, **1989**, *187*, 1.
- [26] M. J. S. Dewar and Y.-C. Yuan, *Inorg. Chem.*, **1990**, *3881*, 29.
- [27] M. J. S. Dewar and A. J. Holder, *Organometallics*, **1990**, *9*, 508.
- [28] E. Anders, R. Koch, and P. Freunsch, *J. Comput. Chem.*, **1993**, *14*, 1301.
- [29] J. J. P. Stewart, *J. Comput. Chem.*, **1989**, *10*, 209.
- [30] J. J. P. Stewart, *J. Comput. Chem.*, **1989**, *10*, 221.
- [31] D. Porezag, T. Frauenheim, T. Köhler, G. Seifert, and R. Kaschner, *Phys. Rev. B*, **1995**, *51*, 12947.
- [32] G. Seifert, D. Porezag, and T. Frauenheim, *Int. J. Quantum Chem.*, **1996**, *58*, 185.
- [33] M. Elstner, D. Porezag, G. Jungnickel, J. Elsner, M. Haugk, T. Frauenheim, S. Suhai, and G. Seifert, *Phys. Rev. B*, **1998**, *58*, 7260.
- [34] C. Köhler, G. Seifert, and T. Frauenheim, *Chem. Phys.*, **2005**, *23*, 309.
- [35] W. Han, M. Elstner, K. J. Jalkanen, T. Frauenheim, and S. Suhai, *Int. J. Quantum Chem.*, **2000**, *78*, 459.
- [36] Q. Cui, M. Elstner, T. Frauenheim, E. Kaxiras, and M. Karplus, *J. Phys. Chem. B*, **2001**, *105*, 569.
- [37] M. Elstner, P. Hobza, T. Frauenheim, S. Suhai, and E. Kaxiras, *J. Chem. Phys.*, **2001**, *114*, 5149.
- [38] T. A. Niehaus, S. Suhai, F. Della Sala, P. Lugli, M. Elstner, G. Seifert, and T. Frauenheim, *Phys. Rev. B*, **2001**, *63*, 085108.
- [39] T. A. Niehaus, D. Heringer, B. Torralva, and T. Frauenheim, *Eur. Phys. J. D*, **2005**, *35*, 467.
- [40] J. Tomasi and M. Persico, *Chem. Rev.*, **1994**, *94*, 2027.
- [41] C. J. Cramer and D. G. Truhlar, *Chem. Rev.*, **1999**, *99*, 2161.
- [42] J. Tomasi, B. Mennucci, and R. Cammi, *Chem. Rev.*, **2005**, *105*, 2999.
- [43] V. Dillet, D. Rinaldi, J. G. Ángyàn, and J.L. Rivail, *Chem. Phys. Lett.*, **1993**, *202*, 18.
- [44] V. Dillet, D. Rinaldi, and J.L. Rivail, *J. Phys. Chem.*, **1994**, *98*, 5043.
- [45] D. Rinaldi, A. Bouchy, J.L. Rivail, and V. Dillet, *Theor. Chem. Acc.*, **2004**, *120*, 2343.
- [46] D. Rinaldi, A. Bouchy, and J.L. Rivail, *Theor. Chem. Acc.*, **2006**, *116*, 664.
- [47] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, *J. Chem. Phys.*, **1983**, *79*, 926.

-
- [48] W. L. Jorgensen and J. D. Madura, *Mol. Phys.*, **1985**, *56*, 1381.
- [49] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, and J. Hermans; Reidel, Dordrecht, 1981; page 331; Intermolecular Forces.
- [50] H. J. C. Berendsen, J. R. Grigera, and T. P. Straatsma, *J. Phys. Chem.*, **1987**, *91*, 6269.
- [51] F. H. Stillinger and A. Rahman, *J. Chem. Phys.*, **1974**, *60*, 1545.
- [52] I. Tuñón, M. T. C. Martins-Costa, C. Millot, and M. F. Ruiz-López, *Chem. Phys. Lett.*, **1995**, *241*, 450.
- [53] I. Tuñón, M. T. C. Martins-Costa, C. Millot, and M. F. Ruiz-López, *J. Mol. Mod.*, **1995**, *1*, 196.
- [54] I. Tuñón, M. T. C. Martins-Costa, C. Millot, M. F. Ruiz-López, and J.-L. Rivail, *J. Comput. Chem.*, **1996**, *17*, 19.
- [55] I. Tuñón, M. T. C. Martins-Costa, C. Millot, and M. F. Ruiz-López, *J. Chem. Phys.*, **1997**, *106*, 3633.
- [56] M. Strnad, M. T. C. Martins-Costa, I. Tuñón C. Millot, M. F. Ruiz-López, and J. L. Rivail, *J. Chem. Phys.*, **1997**, *106*, 3643.
- [57] S. Chalmet and M. F. Ruiz-López, *J. Chem. Phys.*, **1999**, *111*, 1117.
- [58] S. Chalmet and M. F. Ruiz-López, *J. Chem. Phys.*, **2001**, *115*, 5220.
- [59] S. Chalmet and M. F. Ruiz-López, *J. Chem. Phys.*, **2006**, *124*, 194502.
- [60] M. I. Bernal-Uruchurtu, M. T. C. Martins-Costa, C. Millot, and M. F. Ruiz-López, *J. Comput. Chem.*, **2000**, *21*, 572.
- [61] G. Monard, M. I. Bernal-Uruchurtu, A. van der Vaart, K. M. Merz Jr., and M. F. Ruiz-López, *J. Phys. Chem. A*, **2005**, *109*, 3425.
- [62] W. Yang and T.-S. Lee, *J. Chem. Phys.*, **1995**, *103*, 5674.
- [63] S. L. Dixon and K. M. Merz Jr., *J. Chem. Phys.*, **1996**, *104*, 6643.
- [64] S. L. Dixon and K. M. Merz Jr., *J. Chem. Phys.*, **1997**, *107*, 879.
- [65] A. Warshell and M. Levitt, *J. Mol. Biol.*, **1976**, *103*, 227.
- [66] M. J. Field, P. A. Bash, and M. Karplus, *J. Comput. Chem.*, **1990**, *11*, 700.
- [67] S. Humbel, S. Sieber, and K. Morokuma, *J. Chem. Phys.*, **1996**, *105*, 1959.
- [68] M. Svensson, S. Humbel, and K. Morokuma, *J. Chem. Phys.*, **1996**, *105*, 3664.
- [69] T. Vreven, B. Mennucci, C. O. da Silva, K. Morokuma, and J. Tomasi, *J. Chem. Phys.*, **2001**, *115*, 62.
- [70] Y. Moreau, P.-F. Loos, and X. Assfeld, *Theor. Chem. Acc.*, **2004**, *112*, 228.
- [71] T. Vreven, K. S. Byun, I. Komáromi, S. Dapprich, J. A. Montgomery Jr., K. Morokuma, and M. J. Frisch, *J. Chem. Theor. Comput.*, **2006**, *117*, 815.
- [72] C. Leforestier, *J. Chem. Phys.*, **1978**, *68*, 4406.
- [73] R. Car and M. Parrinello, *Phys. Rev. Lett.*, **1985**, *55*, 2471.
- [74] X. Assfeld and J.-L. Rivail, *Chem. Phys. Lett.*, **1996**, *263*, 100.

Bibliographie

- [75] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, T. Vreven Jr., K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, and J. A. Pople; Gaussian 03, Revision B.05; Gaussian, Inc., Wallingford, CT, 2004.
- [76] J. W. Ponder; Tinker, version 4.2; Washington University : St. Louis, MO, 2004.
- [77] E. Schrödinger, *Ann. d. Physik*, **1925**, *79*, 361.
- [78] E. Schrödinger, *Ann. d. Physik*, **1925**, *79*, 489.
- [79] E. Schrödinger, *Ann. d. Physik*, **1925**, *79*, 734.
- [80] E. Schrödinger, *Ann. d. Physik*, **1925**, *80*, 437.
- [81] E. Schrödinger, *Ann. d. Physik*, **1925**, *81*, 109.
- [82] E. Schrödinger, *Die Naturwissenschaften*, **1926**, *14*, 664.
- [83] E. Schrödinger, *Phys. Rev.*, **1926**, *28*, 1049.
- [84] M. Born and J. R. Oppenheimer, *Ann. Physik.*, **1927**, *84*, 457.
- [85] J.-L. Rivail, *Eléments de Chimie Quantique à l'Usage des Chimistes*; EDP Sciences/CNRS Editions, 1999.
- [86] A. Szabo and N.S. Ostlund, *Modern Quantum Chemistry : Introduction to Advanced Structure Theory*; Dover publications Inc., Mineola, New-York, 1989.
- [87] R. McWeeny, *Methods of Molecular Quantum Mechanics*; Academic Press Inc., San Diego, 1989.
- [88] F. Jensen, *Introduction to Computational Chemistry*; John Wiley & Sons, Ltd., 1999.
- [89] C. J. Cramer, *Essentials of Computational Chemistry : Theories and Models*; John Wiley & Sons, Ltd., 2002.
- [90] T. Helgaker, P. Jørgensen, and J. Olsen, *Molecular Electronic-Structure Theory*; John Wiley & Sons, Ltd., 2000.
- [91] W. Pauli Jr., *Z. Physik*, **1925**, *31*, 765.
- [92] J. C. Slater, *Phys. Rev.*, **1929**, *29*, 1293.
- [93] D. R. Hartree, *Proc. Cambridge Phil. Soc.*, **1928**, *34*, 89.
- [94] V. Fock, *Z. Physik*, **1930**, *61*, 126.
- [95] J. A. Pople and R. K. Nesbet, *J. Chem. Phys.*, **1954**, *22*, 571.

-
- [96] G. Bertier, *Acad. Sci. Paris*, **1954**, 238, 91.
- [97] G. Bertier, *J. Chim. Phys.*, **1954**, 51, 363.
- [98] C. C. J. Roothaan, *Rev. Mod. Phys.*, **1951**, 23, 69.
- [99] P. O. Löwdin, *J. Chem. Phys.*, **1950**, 18, 365.
- [100] P. Carsky; John Wiley & Sons : Chichester, 1998; Vol. 1 of *The Encyclopedia of Computational Chemistry*; page 485.
- [101] W. Kutzelnigg, *J. Mol. Struct. (THEOCHEM)*, **1988**, 181, 33.
- [102] A. D. Mc Lean and G. S. Chandler, *J. Chem. Phys.*, **1980**, 72, 5639.
- [103] J. Gauss; John Wiley & Sons : Chichester, 1998; Vol. 1 of *The Encyclopedia of Computational Chemistry*; page 615.
- [104] R. J. Bartlett, *Modern Ideas in Coupled-Cluster Methods*; World Scientific : Singapore, 1997.
- [105] C. Møller and M. S. Plesset, *Phys. Rev.*, **1934**, 46, 618.
- [106] R. J. Bartlett, *Annu. Rev. Phys. Chem.*, **1981**, 32, 359.
- [107] P. Hohenberg and W. Kohn, *Phys. Rev. B*, **1964**, 136, 864.
- [108] R. G. Parr and W. Yang, *Density Functional Theory for Atoms and Molecules*; Oxford University Press, 1989.
- [109] R. M. Deizler and E. K. U. Gross, *Density Functional Theory*; Springer-Verlang : Heidelberg, Germany, 1995.
- [110] E. J. Baerends and O. V. Gritsenko, *J. Phys. Chem. A*, **1997**, 101, 5383.
- [111] W. Kohn and L. J. Sham, *Phys. Rev. A*, **1965**, 140, 1133.
- [112] S. F. Sousa, A. Fernandes, and M. J. Ramos, *J. Phys. Chem. A*, **2007**, 111, 10439.
- [113] L. H. Thomas, *Proc. Camb. Phil. Soc.*, **1927**, 23, 542.
- [114] E. Fermi, *Rend. Accad. Lince*, **1927**, 6, 602.
- [115] P. A. M. Dirac, *Proc. Camb. Phil. Soc.*, **1930**, 26, 376.
- [116] J. C. Slater, *Phys. Rev.*, **1951**, 81, 385.
- [117] D. M. Ceperley and B. J. Alder, *Phys. Rev. Lett.*, **1980**, 45, 566.
- [118] S. J. Vosko, L. Wilk, and M. Nusair, *Can. J. Phys.*, **1980**, 58, 1200.
- [119] J. P. Perdew and Y. Wang, *Phys. Rev. B*, **1986**, 33, 8800.
- [120] A. D. Becke, *Phys. Rev. A*, **1988**, 38, 3098.
- [121] J. P. Perdew, K. Burke, and M. Ernzerhof, *Phys. Rev. Lett.*, **1996**, 77, 3865.
- [122] J. P. Perdew, K. Burke, and M. Ernzerhof, *Phys. Rev. Lett.*, **1997**, 78, 1396.
- [123] J. P. Perdew, J. A. Chevary, S. H. Vosko, K. A. Jackson, M. R. Pederson, D. J. Singh, and C. Fiolhais, *Phys. Rev. B*, **1992**, 46, 6671.
- [124] C. Lee, W. Yang, and R. G. Parr, *Phys. Rev. B*, **1988**, 37, 785.
- [125] J. Tao, J. P. Perdew, V. N. Staroverov, and G. E. Scuseria, *Phys. Rev. Lett.*, **2003**, 91, 146401.

Bibliographie

- [126] L. Hedin, *Phys. Rev. Lett.*, **1965**, *139*, 796.
- [127] W. G. Aulbur, L. Jönsson, and J. W. Wilkins, *Solid State Physics*, **2000**, *54*, 1.
- [128] A. D. Becke, *J. Chem. Phys.*, **1993**, *98*, 5648.
- [129] J. P. Perdew and Y. Wang, *Phys. Rev. B*, **1992**, *45*, 13244.
- [130] J. P. Perdew, K. Burke, and Y. Wang, *Phys. Rev. B*, **1996**, *54*, 16533.
- [131] A. D. Becke, *J. Chem. Phys.*, **1993**, *98*, 1372.
- [132] C. Adamo and V. Barone, *J. Chem. Phys.*, **1999**, *110*, 6158.
- [133] E. Runge and E. K. U. Gross, *Phys. Rev. Lett.*, **1984**, *52*, 997.
- [134] R. van Leeuwen, *Phys. Rev. Lett.*, **1999**, *82*, 3863.
- [135] A. L. Fetter and J. D. Walecka, *Quantum Theory for Many-Particle Systems*; McGraw-Hill : New-York, 1971.
- [136] S. Hirata and M. Head-Gordon, *Chem. Phys. Lett.*, **1999**, *314*, 291.
- [137] J. del Bene, R. Ditchfield, and J. A. Pople, *J. Chem. Phys.*, **1971**, *55*, 2236.
- [138] A. Dreuw and M. Head-Gordon, *Chem. Rev.*, **2005**, *105*, 4009.
- [139] A. R. Leach, *Molecular Modelling : Principles and Applications*; Pearson Education EMA, 2001.
- [140] M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids*; Oxford University Press, 1989.
- [141] P. Kollman, R. Dixon, W. Cornell, T. Fox, C. Chipot, and A. Pohorille; 1997; Vol. 3 of *Computer Simulation of Biomolecular Systems*; page 83.
- [142] J. Aqvist, *J. Phys. Chem.*, **1990**, *94*, 8021.
- [143] W. S. Ross and C. C. Hardin, *J. Am. Chem. Soc.*, **1994**, *116*, 6070.
- [144] W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz Jr., D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. A. Kollman, *J. Am. Chem. Soc.*, **1995**, *117*, 5179.
- [145] T. E. Cheatham III, P. Cieplak, and P. A. Kollman, *J. Biomol. Struct. Dyn.*, **1999**, *16*, 845.
- [146] G. Moyna, H. J. Williams, R. J. Nachman, and A. I. Scott, *Biopolymers*, **1999**, *49*, 403.
- [147] J. Wang, P. Cieplak, and P. A. Kollman, *J. Comput. Chem.*, **2000**, *21*, 1049.
- [148] A. D. MacKerel Jr., C. L. Brooks III, L. Nilsson, B. Roux, Y. Won, and M. Karplus; John Wiley & Sons : Chichester, 1998; Vol. 1 of *The Encyclopedia of Computational Chemistry*; page 271.
- [149] N. Foloppe and A. D. MacKerell Jr., *J. Comput. Chem.*, **2000**, *21*, 86.
- [150] E. Neria, S. Fischer, and M. Karplus, *J. Chem. Phys.*, **1996**, *105*, 1902.
- [151] L. Nilsson and M. Karplus, *J. Comput. Chem.*, **1986**, *7*, 591.
- [152] B. R. Brooks, R. E. Bruccoleri, D. J. Olafson, D. J. States, S. Swaminathan, and M. Karplus, *J. Comput. Chem.*, **1983**, *4*, 187.

-
- [153] A. D. MacKerel Jr., D. Bashford, M. Bellott, R. L. Dunbrack Jr., J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchmir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, III W. E. Reiher, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiórkiewicz-Kuczera, D. Yin, and M. Karplus, *J. Phys. Chem. B*, **1998**, *102*, 3586.
- [154] W. L. Jorgensen, D. S. Maxwell, and J. Tirado-Rives, *J. Am. Chem. Soc.*, **1996**, *117*, 11225.
- [155] D. S. Maxwell, J. Tirado-Rives, and W. L. Jorgensen, *J. Comput. Chem.*, **1995**, *16*, 984.
- [156] W. L. Jorgensen and N. A. McDonald, *J. Mol. Struct. (THEOCHEM)*, **1998**, *424*, 145.
- [157] N. A. McDonald and W. L. Jorgensen, *J. Phys. Chem. B*, **1998**, *102*, 8049.
- [158] R. C. Rizzo and W. L. Jorgensen, *J. Am. Chem. Soc.*, **1999**, *121*, 4827.
- [159] M. L. P. Price, D. Ostrovsky, and W. L. Jorgensen, *J. Comput. Chem.*, **2001**, *22*, 1340.
- [160] G. A. Kaminski, R. A. Friesner, J. Tirado-Rives, and W. L. Jorgensen, *J. Phys. Chem. B*, **2001**, *105*, 6474.
- [161] M. T. Tribble N. L. Allinger, M. A. Miller, and D. H. Wertz, *J. Am. Chem. Soc.*, **1971**, *93*, 1637.
- [162] N. L. Allinger, *J. Am. Chem. Soc.*, **1977**, *99*, 3279.
- [163] N. L. Allinger, Y. H. Yuh, and J.-H. Lii, *J. Am. Chem. Soc.*, **1989**, *111*, 8551.
- [164] N. L. Allinger, H. J. Geise, W. Pyckhout, L. A. Paquette, and J. C. Gallucci, *J. Am. Chem. Soc.*, **1989**, *111*, 1106.
- [165] J.-H. Lii and N. L. Allinger, *J. Am. Chem. Soc.*, **1989**, *111*, 8567.
- [166] J.-H. Lii and N. L. Allinger, *J. Am. Chem. Soc.*, **1989**, *111*, 8576.
- [167] J.-H. Lii and N. L. Allinger, *J. Phys. Org. Chem.*, **1994**, *7*, 591.
- [168] N. L. Allinger, F. Li, and L. Yan, *J. Comput. Chem.*, **1990**, *11*, 848.
- [169] N. L. Allinger, F. Li, L. Yan, and J. C. Tai, *J. Comput. Chem.*, **1990**, *11*, 868.
- [170] J.-H. Lii and N. L. Allinger, *J. Comput. Chem.*, **1998**, *19*, 1001.
- [171] N. L. Allinger, K. Chen, and J.-H. Lii, *J. Comput. Chem.*, **1996**, *14*, 642.
- [172] N. Nevins, K. Chen, and N. L. Allinger, *J. Comput. Chem.*, **1996**, *14*, 669.
- [173] N. Nevins, K. Chen, and N. L. Allinger, *J. Comput. Chem.*, **1996**, *14*, 695.
- [174] N. L. Allinger, K. Chen, J. A. Katzeellenbogen, S. R. Wilson, and G. M. Anstead, *J. Comput. Chem.*, **1996**, *14*, 747.
- [175] S. L. Mayo, B. D. Olafson, and W. A. Goddard, *J. Phys. Chem.*, **1990**, *94*, 8897.
- [176] A. K. Rappé, C. J. Casewit, K. S. Colwell, W. A. Goddard III, and W. M. Skiff, *J. Am. Chem. Soc.*, **1992**, *114*, 10024.

Bibliographie

- [177] P. Ren and J. W. Ponder, *J. Comput. Chem.*, **2002**, *23*, 1497.
- [178] P. Ren and J. W. Ponder, *J. Phys. Chem. B*, **2003**, *107*, 5933.
- [179] J. W. Ponder and D. A. Case, *Adv. Prot. Chem.*, **2003**, *66*, 27.
- [180] A. Groosfield, P. Ren, and J. W. Ponder, *J. Am. Chem. Soc.*, **2003**, *125*, 15671.
- [181] P. Ren and J. W. Ponder, *J. Phys. Chem. B*, **2004**, *108*, 13427.
- [182] G. F. Smits and C. Altona, *Theor. Chem. Acc.*, **1985**, *67*, 461.
- [183] U. C. Singh and P. A. Kollman, *J. Comput. Chem.*, **1986**, *7*, 718.
- [184] F. Bernardi, M. Olivucci, and M. Robb, *J. Am. Chem. Soc.*, **1992**, *1606*, 227.
- [185] T. Ose, K. Watanabe, T. Mie, M. Honma, H. Watanabe, M. Yao, H. Oikawa, and I. Tanaka, *Nature*, **2003**, *422*, 185.
- [186] G. Pohnert, *ChemBioChem*, **2003**, *4*, 713.
- [187] T. Ose, K. Watanabe, M. Honma, H. Oikawa, and I. Tanaka, *Acta. Cryst. D*, **2004**, *60*, 1187.
- [188] C. R. W. Guimaraes, M. Udier-Blagovic, and W. L. Jorgensen, *J. Am. Chem. Soc.*, **2005**, *127*, 3577.
- [189] N. Ferré, A. Cembran, M. Garavelli, and M. Olivucci, *Theor. Chem. Acc.*, **2004**, *112*, 335.
- [190] B. H. Besler, K. M. Merz Jr., and P. A. Kollman, *J. Comput. Chem.*, **1990**, *11*, 431.
- [191] U. C. Singh and P. A. Kollman, *J. Comput. Chem.*, **1984**, *5*, 129.
- [192] L. E. Chirlian and M. M. Francl, *J. Comput. Chem.*, **1987**, *8*, 894.
- [193] C. M. Breneman and K. B. Wibber, *J. Comput. Chem.*, **1990**, *11*, 361.
- [194] F. Maseras and K. Morokuma, *J. Comput. Chem.*, **1995**, *16*, 1170.
- [195] M. Svensson, S. Humbel, D. J. Froese, T. Matsubara, S. Sieber, and K. Morokuma, *J. Phys. Chem.*, **1996**, *100*, 19357.
- [196] T. Vreven and K. Morokuma, O. Farkas, H. B. Schlegel, and M. J. Frisch, *J. Comput. Chem.*, **2003**, *24*, 760.
- [197] T. Vreven and K. Morokuma, *Theor. Chem. Acc.*, **2003**, *109*, 125.
- [198] H. P. Hratchian, P. V. Parandekar, K. Raghavavhari, M. J. Frisch, and T. Vreven, *J. Chem. Phys.*, **2008**, *128*, 034107.
- [199] H. Lin and D.G. Truhlar, *Theor. Chem. Acc.*, **2007**, *117*, 185.
- [200] S. Ranganathan and J. Gready, *J. Phys. Chem. B*, **1997**, *101*, 5614.
- [201] S. Dapprich, I. Komárino, K. S. Byun, K. Morokuma, and M. J. Frisch, *J. Mol. Struct. (THEOCHEM)*, **1999**, *461*, 1.
- [202] P. Cummins and J. Gready, *J. Phys. Chem. B*, **2000**, *104*, 4503.
- [203] T. Kerdcharoen and K. Morokuma, *Chem. Phys. Lett.*, **2002**, *355*, 257.
- [204] D. Das, K.P. Eurenus, E.M. Billings, P. Sherwood, D.C. Chatfield, M. Hodoseek, and B.R. Brooks, *J. Chem. Phys.*, **2002**, *117*, 10534.

-
- [205] Y. Zhang, T.-S. Lee, and W. Yang, *J. Chem. Phys.*, **1999**, *110*, 46.
- [206] I. Antes and W. Thiel, *J. Phys. Chem. A*, **1999**, *103*, 9290.
- [207] G. A. DiLabio, M. M. Hurley, and P. A. Christiansen, *J. Chem. Phys.*, **2002**, *166*, 9578.
- [208] Y. Zhang, *J. Chem. Phys.*, **2003**, *122*, 024114.
- [209] Y. Zhang, *Theor. Chem. Acc.*, **2006**, *116*, 43.
- [210] C. Xiao and Y. Zhang, *J. Chem. Phys.*, **2007**, *127*, 124102.
- [211] G. Ferenczy, J.-L. Rivail, P. R. Surján, and G. Náráy-Szabó, *J. Comput. Chem.*, **1992**, *13*, 830.
- [212] V. Théry, D. Rinaldi, J.-L. Rivail, B. Maigret, and G. Ferenczy, *J. Comput. Chem.*, **1994**, *15*, 269.
- [213] G. Monard, M. Loos, V. Théry, K. Baka, and J.-L. Rivail, *Int. J. Quantum Chem.*, **1996**, *58*, 153.
- [214] N. Ferré, X. Assfeld, and J.-L. Rivail, *J. Comput. Chem.*, **2002**, *23*, 610.
- [215] A. Fornili, P.-F. Loos, M. Sironi, and X. Assfeld, *Chem. Phys. Lett.*, **2006**, *427*, 236.
- [216] D. M. Philipp and R. A. Friesner, *J. Comput. Chem.*, **1999**, *20*, 1468.
- [217] R. B. Murphy, D. M. Philipp, and R. A. Friesner, *J. Comput. Chem.*, **2000**, *21*, 1442.
- [218] R. B. Murphy, D. M. Philipp, and R. A. Friesner, *Chem. Phys. Lett.*, **2000**, *321*, 113.
- [219] J. Gao, P. Amara, C. Alhambra, and M. J. Field, *J. Phys. Chem. A*, **1998**, *102*, 4714.
- [220] P. Amara, M. J. Field, C. Alhambra, and J. Gao, *Theor. Chem. Acc.*, **2000**, *104*, 336.
- [221] D. G. Truhlar, J. Gao, C. Alhambra, M. Garcia-Viloca, J. Corchado, M. L. Sánchez, and J. Villá, *Acc. Chem. Res.*, **2002**, *35*, 341.
- [222] J. Pu, J. Gao, and D.G. Truhlar, *J. Phys. Chem. A*, **2004**, *108*, 632.
- [223] J. Pu, J. Gao, and D.G. Truhlar, *J. Phys. Chem. A*, **2004**, *108*, 5454.
- [224] J. Jung, C. H. Choi, Y. Sugita, and S. Ten-no, *J. Chem. Phys.*, **2007**, *127*, 204102.
- [225] V. Kairys and J. H. Jensen, *J. Phys. Chem. A*, **2000**, *104*, 6656.
- [226] H. Lin and D.G. Truhlar, *J. Phys. Chem. A*, **2005**, *109*, 3991.
- [227] Y. Zhang, H. Lin, and D. G. Truhlar, *J. Chem. Theor. Comput.*, **2007**, *1378*, 3.
- [228] A. K. Rappé and W. A. Goddard III, *J. Phys. Chem.*, **1991**, *95*, 3358.
- [229] P. Amara and M. J. Field, *Theor. Chem. Acc.*, **2003**, *109*, 43.
- [230] L. G. Gorb, J.-L. Rivail, V. Théry, and D. Rinaldi, *Int. J. Quantum Chem.*, **1996**, *30*, 1525.
- [231] N. Ferré and X. Assfeld, *J. Chem. Phys.*, **2002**, *117*, 4119.

Bibliographie

- [232] N. Ferré and X. Assfeld, *J. Mol. Struct. (THEOCHEM)*, **2003**, 632, 83.
- [233] X. Assfeld, N. Ferré, and J.-L. Rivail, *Theor. Chem. Acc.*, **2004**, 111, 328.
- [234] Y. Moreau and X. Assfeld, *Lecture Series on Computer and Computational Sciences*, **2005**, 3, 1.
- [235] A. Fornili, Y. Moreau, M. Sironi, and X. Assfeld, *J. Comput. Chem.*, **2006**, 27, 515.
- [236] P.-F. Loos, X. Assfeld, and J.-L. Rivail, *Theor. Chem. Acc.*, **2007**, 118, 165.
- [237] P.-F. Loos and X. Assfeld, *J. Chem. Theor. Comput.*, **2007**, 3, 1047.
- [238] P.-F. Loos and X. Assfeld, *Int. J. Quantum Chem.*, **2007**, 107, 2243.
- [239] P.-F. Loos, A. Fornili, M. Sironi, and X. Assfeld, *Comp. Lett.*, **2007**, 4, 473.
- [240] P.-F. Loos and X. Assfeld, *AIP Conf. Proc.*, **2007**, 963, 308.
- [241] P.-F. Loos, J. Preat, A. D. Laurent, C. Michaux, D. Jacquemin, E. A. Perpète, and X. Assfeld, *J. Chem. Theor. Comput.*, **2008**, ASAP Article.
- [242] P. O. Löwdin, *Phys. Rev.*, **1955**, 97, 1474.
- [243] S. Huzinaga, *J. Chem. Phys.*, **1971**, 55, 5543.
- [244] N. Ferré; *Développements de la méthode LSCF ab initio et d'un champ de forces hybride QM/MM/Madelung pour l'étude de la réactivité dans les zéolithes*; PhD thesis, Laboratoire de Chimie Théorique, UMR CNRS-UHP 7565, Université Henri Poincaré, Nancy I, **2001**.
- [245] P. Pulay, *Mol. Phys.*, **1969**, 17, 197.
- [246] P. Pulay; Plenum, New-York, 1977; page 153; Applications of Electronic Structure Theory.
- [247] P. Pulay; John Wiley & Sons, 1987; Vol. 69 of *Advances in Chemical Physics : ab initio Methods in Quantum Chemistry (Part II)*; page 241.
- [248] Y. Yamaguchi, O. Osamura, J. D. Goddard, and H. F. Schaefer III, *A new Dimension to Quantum Mechanics. Analytical Derivative Methods in ab initio Molecular Electronic Structure Theory*; Oxford University Press, Oxford, 1994.
- [249] P. Pulay, *J. Chem. Phys.*, **1983**, 78, 5043.
- [250] Y. Moreau; *Développements d'une méthode théorique pour l'étude de la réactivité chimique des systèmes moléculaires complexes en solution*; PhD thesis, Equipe de Chimie et Biochimie Théoriques, UMR CNRS-UHP 7565, Université Henri Poincaré, Nancy I, **2005**.
- [251] J. Pu and D. G. Truhlar, *J. Chem. Theor. Comput.*, **2005**, 1, 54.
- [252] J. A. Pople, R. Krishnan, B. Schlegel, and J. S. Binkley, *Int. J. Quantum Chem.*, **1979**, 13, 225.
- [253] J. Pipek and P. Mezey, *J. Chem. Phys.*, **1989**, 90, 4916.
- [254] S. F. Boys, *Rev. Mod. Phys.*, **1960**, 32, 296.
- [255] J. M. Foster and S. F. Boys, *Rev. Mod. Phys.*, **1960**, 32, 300.
- [256] S. F. Boys; Academic, New-York, 1966; page 253; Quantum Theory of Atoms, Molecules and the Solid State.

-
- [257] C. Edmiston and K. Ruedenberg, *Rev. Mod. Phys.*, **1963**, *35*, 457.
- [258] C. Edmiston and K. Ruedenberg, *J. Chem. Phys.*, **1965**, *43*, S97.
- [259] C. Edmiston and K. Ruedenberg; Academic, New-York, 1966; page 263; Quantum Theory of Atoms, Molecules and the Solid State.
- [260] H. Weinstein and R. Pauncz, *Symp. Faraday Soc.*, **1968**, *2*, 23.
- [261] H. Weinstein, R. Pauncz, and M. Cohen, *Adv. Atomic. Molec. Phys.*, **1971**, *7*, 97.
- [262] V. Magnasco and A. Perico, *J. Chem. Phys.*, **1967**, *47*, 971.
- [263] V. Magnasco and A. Perico, *J. Chem. Phys.*, **1967**, *48*, 800.
- [264] A. Fornili; *Theory and Algorithms for the determination of Extremely Localized Molecular Orbitals and their application to QM/MM methods*; PhD thesis, Dipartimento di Chimica Fisica ed Elettrochimica, Università degli Studi di Milano, **2003**.
- [265] H. Stoll, G. Wagenblast, and H. Preuss, *Theor. Chem. Acc.*, **1980**, *57*, 169.
- [266] M. Couty, C.A. Bayse, and M.B. Hall, *Theor. Chem. Acc.*, **1997**, *97*, 96.
- [267] M. Sironi and A. Famulari, *Theor. Chem. Acc.*, **1999**, *103*, 417.
- [268] A. Fornili, M. Sironi, and M. Raimondi, *J. Mol. Struct. (THEOCHEM)*, **2003**, *632*, 157.
- [269] Z. Szekeres and P.R. Surján, *Chem. Phys. Lett.*, **2003**, *369*, 125.
- [270] J. W. Boughton and P. Pulay, *J. Comput. Chem.*, **1993**, *14*, 736.
- [271] S. Sæbø and P. Pulay, *Theor. Chem. Acc.*, **1986**, *69*, 357.
- [272] S. Sæbø and P. Pulay, *Ann. Rev. Phys. Chem.*, **1993**, *44*, 213.
- [273] S. Sæbø, W. Tong, and P. Pulay, *Ann. Rev. Phys. Chem.*, **1993**, *98*, 2170.
- [274] R. B. Murphy, M. D. Beachy, R. A. Friesner, and M. N. Ringnalda, *J. Chem. Phys.*, **1995**, *103*, 1481.
- [275] M. Schütz, G. Hetzer, and H. J. Werner, *J. Chem. Phys.*, **1999**, *111*, 5691.
- [276] G. Hetzer, M. Schütz, H. Stoll, and H.-J. Werner, *J. Chem. Phys.*, **2000**, *113*, 9443.
- [277] P. Y. Ayala and G. E. Scuseria, *J. Chem. Phys.*, **1999**, *110*, 3660.
- [278] P. E. Maslen and M. Head-Gordon, *Chem. Phys. Lett.*, **1999**, *283*, 102.
- [279] P. E. Maslen and M. Head-Gordon, *J. Chem. Phys.*, **1999**, *109*, 7093.
- [280] C. Hampel and H.-J. Werner, *J. Chem. Phys.*, **1996**, *104*, 6286.
- [281] M. Schütz and H.-J. Werner, *Chem. Phys. Lett.*, **2000**, *318*, 370.
- [282] M. Schütz and H.-J. Werner, *J. Chem. Phys.*, **2000**, *113*, 9986.
- [283] M. Schütz and H. J. Werner, *J. Chem. Phys.*, **2001**, *114*, 661.
- [284] M. Schütz, *J. Chem. Phys.*, **2002**, *116*, 8772.
- [285] M. Schütz, *Phys. Chem. Chem. Phys.*, **2002**, *4*, 3941.
- [286] R. S. Mulliken, *J. Chem. Phys.*, **1955**, *23*, 1833.
- [287] O. Maresca, A. Ionescu, A. Allouche, J. P. Aycard, M. Rajzmann, and F. Hutschka, *J. Mol. Struct. (THEOCHEM)*, **2003**, *119*, 620.

Bibliographie

- [288] W. von Niessen, *J. Chem. Phys.*, **1972**, *56*, 4290.
- [289] D. A. Kleier, T. A. Halgren, J. H. Hall, and W. N. Lipscomb, *J. Chem. Phys.*, **1974**, *61*, 3905.
- [290] J. Pipek, *Int. J. Quantum Chem.*, **1989**, *36*, 487.
- [291] T. Vreven and K. Morokuma, *J. Comput. Chem.*, **2000**, *21*, 1419.
- [292] Q. Cui, H. Guo, and M. Karplus, *J. Chem. Phys.*, **2002**, *117*, 5617.
- [293] V. Gogonea, L. M. Westerhoff, and K. M. Merz Jr., *J. Chem. Phys.*, **2000**, *113*, 5604.
- [294] N. Govind, Y. A. Wang, and E. A. Carter, *J. Chem. Phys.*, **1999**, *110*, 7677.
- [295] T. Kluner, N. Govind, Y. A. Wang, and E. A. Carter, *J. Chem. Phys.*, **2002**, *116*, 42.
- [296] T. M. Henderson, *J. Chem. Phys.*, **2006**, *125*, 014105.
- [297] P. Surján and I. Mayer, *Theor. Chem. Acc.*, **1981**, *59*, 603.
- [298] I. Mayer and P. Surján, *J. Chem. Phys.*, **1984**, *80*, 5649.
- [299] P. Surján, M. Révész, and I. Mayer, *J. Chem. Soc., Faraday Trans. 2*, **1981**, *77*, 1129.
- [300] P. Surján, *J. Mol. Struct. (THEOCHEM)*, **1988**, *169*, 95.
- [301] P. Surján, *Chem. Phys. Lett.*, **1985**, *386*, 117.
- [302] P. Surján, *Chem. Phys. Lett.*, **1982**, *483*, 92.
- [303] S. Diner, J.-P. Malrieu, and P. Clavery, *Theor. Chem. Acc.*, **1969**, *13*, 1.
- [304] J.-P. Malrieu, P. Clavery, and S. Diner, *Theor. Chem. Acc.*, **1969**, *13*, 18.
- [305] J.-P. Malrieu, *Semiempirical Methods in Electronic Structure Calculation*; G. A. Segal, Ed.; Plenum : New York, 1977.
- [306] V. Kvasnika, *Theor. Chem. Acc.*, **1974**, *61*, 34.
- [307] J. M. Cullen and M. C. Zerner, *Int. J. Quantum Chem.*, **1982**, *497*, 22.
- [308] A Boca, *Theor. Chem. Acc.*, **1982**, *179*, 61.
- [309] V. Barone, J. Douady, Y. Ellinger, and R. Subra, *J. Chem. Soc., Faraday Trans. 2*, **1989**, *1579*, 75.
- [310] F. Bartha, F. Bogar, and E. Kapuy, *J. Chem. Phys.*, **1990**, *38*, 215.
- [311] J. E. Subotnik, A. D. Dutoi, and M. Head-Gordon, *J. Chem. Phys.*, **2005**, *123*, 114108.
- [312] A. E. Reed, L. A. Curtiss, and F. Weinhold, *Chem. Rev.*, **1998**, *88*, 899.
- [313] J. J. W. McDouall, *Theor. Chem. Acc.*, **1992**, *83*, 339.
- [314] J. J. W. McDouall, *Theor. Chem. Acc.*, **1993**, *85*, 395.
- [315] T. Thorsteinsson and D. L. Cooper, *Mol. Phys.*, **1998**, *93*, 663.
- [316] P. W. Payne, *J. Am. Chem. Soc.*, **1977**, *99*, 3787.

-
- [317] N. Reuter, A. Dejaegere, B. Maigret, and M. Karplus, *J. Phys. Chem. A*, **2000**, *104*, 1720.
- [318] D. Bakowies and W. Thiel, *J. Phys. Chem.*, **1996**, *100*, 10580.
- [319] A. de Vries, P. Sherwood, S. Collins, A. Rigby, M. Rigutto, and G. Kramer, *J. Phys. Chem. B*, **1999**, *103*, 6133.
- [320] C. I. Bayly, W. D. Cornell, and P. A. Kollman, *J. Phys. Chem.*, **1993**, *97*, 10269.
- [321] W. D. Cornell, P. Cieplak, C. I. Bayly, and P. A. Kollman, *J. Am. Chem. Soc.*, **1993**, *115*, 9620.
- [322] W. Kabsch, *Acta Cryst.*, **1978**, *A34*, 827.
- [323] W. Humphrey, A. Dalke, and K. Schulten, *J. Molec. Graphics*, **1996**, *14*, 33.
- [324] M. R. Sawaya and J. Kraut, *Biochemistry*, **1997**, *36*, 586.
- [325] A. Jabs, M. S. Weiss, and R. Hilgenfeld, *J. Mol. Biol.*, **1999**, *286*, 291.
- [326] J. C. O'Neill Jr. A.-K. E. Svensson and C. R. Matthews, *J. Mol. Biol.*, **2003**, *326*, 569.
- [327] S. Antonczak, M. F. Ruiz-López, and J.-L. Rivail, *J. Mol. Mod.*, **1997**, *3*, 434.
- [328] S. Antonczak, M. F. Ruiz-López, and J.-L. Rivail, *J. Am. Chem. Soc.*, **1994**, *116*, 3912.
- [329] S. Antonczak, G. Monard, M. F. Ruiz-López, and J.-L. Rivail, *J. Am. Chem. Soc.*, **1998**, *120*, 8225.
- [330] S. Antonczak, G. Monard, M. F. Ruiz-López, and J.-L. Rivail, *J. Mol. Mod.*, **2000**, *6*, 527.
- [331] N. Ferré and M. Olivucci, *J. Mol. Struct. (THEOCHEM)*, **2003**, *632*, 71.
- [332] W. H. E. Schwarz, *Angew. Chem. Int. Ed.*, **1974**, *13*, 454.
- [333] E. Blomberg, P. M. Claesson, and J. C. Froberg, *Biomaterials*, **1998**, *19*, 371.
- [334] H. Fitzpatrick, P. F. Luckman, S. Eriksen, and K. Hammond, *J. Colloid. Interface Sci.*, **1992**, *149*, 1.
- [335] A. Lebugle, M. Subirade, and J. Gueguen, *Biochim. Biophys. Acta*, **195**, *107*, 1248.
- [336] M. L. Gordon, D. Tulumello, G. Cooper, A. P. Hitchcock, P. Glatzel, O. C. Mullins, S. P. Cramer, and U. Bergmann, *J. Phys. Chem. A*, **2003**, *107*, 8512.
- [337] V. Averbukh and L. S. Cederbaum, *J. Chem. Phys.*, **2006**, *125*, 094107.
- [338] J. Breidbach and L. S. Cederbaum, *J. Chem. Phys.*, **2007**, *126*, 034101.
- [339] D. P. Chong and Y. Takahata, *Chem. Phys. Lett.*, **2006**, *418*, 286.
- [340] Y. Takahata, A. K. Okamoto, and D. P. Chong, *Int. J. Quantum Chem.*, **2006**, *106*, 2581.
- [341] J. M. Herbert and J. V. Ortiz, *J. Phys. Chem. A*, **2000**, *104*, 11786.
- [342] M. Wanko, M. Garavelli, F. Bernardi T. A. Niehaus, T. Frauenheim, and M. Elstner, *J. Chem. Phys.*, **2004**, *120*, 1674.

Bibliographie

- [343] M. Wanko, M. Hoffmann, P. Strodel, A. Koslowski, W. Thiel, F. Neese, T. Frauenheim, and M. Elstner, *J. Phys. Chem. B*, **2005**, *109*, 3606.
- [344] U. F. Röhrig, L. Guidoni, and U. Rothlisberger, *Biochemistry*, **2002**, *41*, 10799.
- [345] M. Sulpizi, P. Carloni, J. Hutter, and U. Rothlisberger, *Phys. Chem. Chem. Phys.*, **2003**, *5*, 4798.
- [346] U. F. Röhrig, I. Frank, J. Hutter, A. Laio J. VandeVondele, and U. Rothlisberger, *ChemPhysChem*, **2003**, *4*, 1177.
- [347] M.-E. Moret, E. Tapavicza, L. Guidoni, U. F. Röhrig, M. Sulpizi, I. Tavernelli, and U. Rothlisberger, *Chimia*, **2005**, *59*, 493.
- [348] U. F. Röhrig, L. Guidoni, and U. Rothlisberger, *ChemPhysChem*, **2005**, *6*, 1836.
- [349] I. Tavernelli, U. F. Röhrig, and U. Rothlisberger, *Mol. Phys.*, **2005**, *103*, 963.
- [350] M. Sulpizi, U. F. Röhrig, J. Hutter, and U. Rothlisberger, *Int. J. Quantum Chem.*, **2005**, *101*, 671.
- [351] W. J. Hehre, R. F. Stewart, and J. A. Pople, *J. Chem. Phys.*, **1969**, *51*, 2657.
- [352] J. S. Binkley and J. A. Pople, *J. Am. Chem. Soc.*, **1980**, *102*, 939.
- [353] W. J. Hehre, J. S. Binkley, and J. A. Pople, *J. Chem. Phys.*, **1972**, *56*, 2257.
- [354] R. Krishnan, J. S. Binkley, R. Seeger, and J. A. Pople, *J. Chem. Phys.*, **1980**, *72*, 650.
- [355] M. J. Frisch, J. A. Pople, and J. S. Binkley, *J. Chem. Phys.*, **1984**, *80*, 3265.
- [356] M. M. Francl, W. J. Pietro, W. J. Hehre, J. S. Binkley, M. S. Gordon, D. J. DeFrees, and J. A. Pople, *J. Chem. Phys.*, **1982**, *77*, 3654.
- [357] T. H. Dunning Jr., *J. Chem. Phys.*, **1989**, *90*, 1007.
- [358] A. K. Wilson, T. van Mourik, and T. H. Dunning Jr., *J. Mol. Struct.*, **1996**, *388*, 301.
- [359] R. A. Kendall, T. H. Dunning Jr., and R. J. Harrison, *J. Chem. Phys.*, **1992**, *96*, 6796.
- [360] D. E. Woon and T. H. Dunning Jr., *J. Chem. Phys.*, **1995**, *103*, 4572.
- [361] A. K. Wilson and T. H. Dunning Jr., *J. Chem. Phys.*, **1997**, *106*, 8718.
- [362] D. Feller and K. A. Peterson, *J. Chem. Phys.*, **1998**, *108*, 154.
- [363] A. Halkier, T. Helgaker, P. Jørgensen, W. Klopper, H. Koch, J. Olsen, and A. K. Wilson, *Chem. Phys. Lett.*, **1998**, *286*, 243.
- [364] K. A. Petersoo, K. A. Wilson, D. E. Woon, and T. H. Dunning Jr., *Theor. Chem. Acc.*, **1997**, *97*, 251.
- [365] J. P. Perdew and A. Zunger, *Phys. Rev. B*, **1981**, *23*, 5048.
- [366] W. L. Jolly, K. D. Bomben, and C. J. Eyermann, *Atom. Data Nucl. Data Tables*, **1984**, *31*, 433.
- [367] L. J. Saethre, N. Berrah, J. D. Bazek, K. J. Borve, T. X. Carroll, E. Kukk, G. L. Gard, R. Winter, and T. D. Thomas, *J. Am. Chem. Soc.*, **2001**, *123*, 10729.

-
- [368] D. P. Chong, C.-H. Hu, and P. Duffy, *Chem. Phys. Lett.*, **1996**, *249*, 419.
- [369] M. Pulfer, C.-H. Hu, and D. P. Chong, *Chem. Phys.*, **1997**, *216*, 91.
- [370] C.-H. Hu and D. P. Chong, *Chem. Phys.*, **1997**, *216*, 99.
- [371] D. P. Chong and C.-H. Hu, *J. Chem. Phys.*, **1998**, *108*, 8950.
- [372] C. Bureau, D. P. Chong, G. Lecayon, and J. J. Delhalle, *J. Electron Spectrosc. Relat. Phenom.*, **1997**, *83*, 227.
- [373] G. Cavigliasso and D. P. Chong, *Can. J. Chem.*, **1999**, *77*, 24.
- [374] D. P. Chong and C. Bureau, *J. Electron Spectrosc. Relat. Phenom.*, **2000**, *106*, 1.
- [375] G. Cavigliasso and D. P. Chong, *J. Chem. Phys.*, **1999**, *111*, 9585.
- [376] D. P. Chong, P. Aplincourt, and C. Bureau, *J. Phys. Chem. A*, **2002**, *106*, 356.
- [377] A. R. Williams, R. A. deGroot, and C. B. Sommers, *J. Chem. Phys.*, **1975**, *63*, 628.
- [378] J. C. Slater, *Adv. Quantum. Chem.*, **1972**, *6*, 1.
- [379] J. F. Janak, *Phys. Rev. B*, **1978**, *18*, 7165.
- [380] I. Powis, E. E. Rennie, U. Hergenhahn, O. Kugeler, and R. Bussy-Socrate, *J. Phys. Chem. A*, **2003**, *107*, 25.
- [381] T. Okada, K. Palczewski, R. E. Stenkamp, and M. Miyano, *Science*, **2000**, *289*, 739.
- [382] H. Kandori, Y. Schichida, and Y. Yoshisawa, *Biochemistry (Moscow)*, **2001**, *66*, 1197.
- [383] R. A. Mathies and J. Lugtenburg, *Handbook of Biological Physics*; Vol. 3; D. G. Stavenga and W. J. de Grip and E. N. Pugh; Elsevier Science B. V., 2000.
- [384] J. B. Hurley, T. G. Ebrey, B. Honig, and M. Ottolenghi, *Nature*, **1977**, *270*, 540.
- [385] H. Nakamichi and T. Okada, *Angew. Chem. Int. Ed.*, **2006**, *45*, 4270.
- [386] P. Kukura, D. W. Mc Camant, S. Yoon, D. B. Wandschneider, and R. A. Mathies, *Science*, **2005**, *310*, 1006.
- [387] A. Cooper, *Nature*, **1979**, *282*, 531.
- [388] J. Preat, P.-F. Loos, X. Assfeld, D. Jacquemin, and E. A. Perpète, *Int. J. Quantum Chem.*, **2007**, *107*, 574.
- [389] J. Preat, P.-F. Loos, X. Assfeld, D. Jacquemin, and E. A. Perpète, *J. Mol. Struct. (THEOCHEM)*, **2007**, *808*, 85.
- [390] K. A. Freedman and R. S. Becker, *J. Am. Chem. Soc.*, **1986**, *108*, 1245.
- [391] H. Zollinger, *Color Chemistry, Syntheses, Properties and Applications of Organic Dyes and Pigments*; Wiley-VCH, Weinheim, 3 ed., 2003.
- [392] G. Sudesh Kumar and D. C. Neckers, *Chem. Rev.*, **1989**, *89*, 1915.
- [393] N. Tamai and H. Miyasaka, *Chem. Rev.*, **2000**, *100*, 1875.
- [394] A. Natansohn and P. Rochon, *Chem. Rev.*, **2002**, *102*, 4139.
- [395] Y. Y. Yu, M. Nakano, and T. Ikeda, *Nature*, **2003**, *425*, 145.
- [396] V. Balzani, A. Credi, F. M. Raymo, and J. F. Stoddart, *Angew. Chem. Int. Ed.*, **2000**, *29*, 3348.

Bibliographie

- [397] C. J. Qu, Q. C. Wang, and J. Ren H. Tian, *Org. Lett.*, **2004**, *6*, 2085.
- [398] R. Gonzales-Luque, M. Garavelli, F. Bernardi, M. Merchan, M. A. Robb, and M. Olivucci, *Proc. Nati. Acad. Sci. USA*, **2000**, *97*, 9379.
- [399] T. Andruniow, N. Ferré, and M. Olivucci, *Proc. Nati. Acad. Sci. USA*, **2004**, *101*, 17908.
- [400] A. Cembran, F. Bernardi, M. Olivucci, and M. Garavelli, *Proc. Nati. Acad. Sci. USA*, **2005**, *102*, 6255.
- [401] P. B. Coto, A. Sinicropi, N. Ferré, and M. Olivucci, *Proc. Nati. Acad. Sci. USA*, **2006**, *103*, 17154.
- [402] L. M. Frutos, T. Andruniow, F. Santoro, N. Ferré, and M. Olivucci, *Proc. Nati. Acad. Sci. USA*, **2007**, *104*, 7764.
- [403] A. Strambi, P. B. Coto, N. Ferré, and M. Olivucci, *Theor. Chem. Acc.*, **2007**, *118*, 185.
- [404] N. Ferré and M. Olivucci, *J. Am. Chem. Soc.*, **2003**, *125*, 6868.
- [405] T. Andruniow, S. Fantacci, F. De Angelis, N. Ferré, and M. Olivucci, *Angew. Chem. Int. Ed.*, **2005**, *44*, 6077.
- [406] F. Lumento, V. Zanirato, S. Fusi, E. Busi, L. Latterini, F. Elisei, A. Sinicropi, T. Andruniow, N. Ferré, R. Basosi, and M. Olivucci, *Angew. Chem. Int. Ed.*, **2007**, *46*, 414.
- [407] O. Pieroni, J. L. Houben, A. Fissi, P. Constantino, and F. Ciardelli, *J. Am. Chem. Soc.*, **1980**, *102*, 5913.
- [408] J. L. Houben, A. Fissi, D. Bacciola, N. Rosato, O. Pieroni, and F. Ciardelli, *Int. J. Biol. Macromol.*, **1983**, *5*, 94.
- [409] F. Ciardelli, O. Pieroni, A. Fissi, and J. L. Houben, *Biopolymers*, **1984**, *23*, 1423.
- [410] O. Pieroni, A. Fissi, N. Angelini, and F. Lenci, *Acc. Chem. Res.*, **2001**, *34*, 9.
- [411] J. Preat, D. Jacquemin, and E. A. Perpète, *Chem. Phys. Lett.*, **2005**, *415*, 20.
- [412] J. Preat, D. Jacquemin, V. Wathelet, J. M. André, and E. A. Perpète, *J. Phys. Chem. A*, **2006**, *110*, 8144.
- [413] D. Jacquemin, J. Preat, V. Wathelet, M. Fontaine, and E. A. Perpète, *J. Am. Chem. Soc.*, **2006**, *128*, 2072.
- [414] J. Preat, D. Jacquemin, V. Wathelet, J. M. André, and E. A. Perpète, *Chem. Phys.*, **2007**, *335*, 177.
- [415] C. R. Crecca and A. E. Roitberg, *J. Phys. Chem. A*, **2006**, *110*, 8188.
- [416] L. Briquet, D. P. Vercauteren, E. A. Perpète, and D. Jacquemin, *Chem. Phys. Lett.*, **2006**, *417*, 190.
- [417] L. Briquet, D. P. Vercauteren, J.-M. André, E. A. Perpète, and D. Jacquemin, *Chem. Phys. Lett.*, **2007**, *435*, 257.
- [418] R. Shepard ; John Wiley & Sons, 1987 ; Vol. 69 of *Advances in Chemical Physics : ab initio Methods in Quantum Chemistry (Part II)* ; page 63.

-
- [419] H.-J. Werner ; John Wiley & Sons, 1987 ; Vol. 69 of *Advances in Chemical Physics : ab initio Methods in Quantum Chemistry (Part II)* ; page 1.
- [420] B. O. Roos ; John Wiley & Sons, 1987 ; Vol. 69 of *Advances in Chemical Physics : ab initio Methods in Quantum Chemistry (Part II)* ; page 399.
- [421] B. O. Roos, K. Andersson, M. P. Fulscher, P.-A. Malmqvist, L. Serrano-Andres, K. Pierloot, and M. Merchan ; John Wiley & Sons, 1996 ; Vol. 93 of *Advances in Chemical Physics : New Methods in Computational Quantum Mechanics* ; page 219.
- [422] A. Sinicropi, T. Andruniow, N. Ferre, R. Basoci, and M. Olivucci, *J. Am. Chem. Soc.*, **2005**, *127*, 11534.
- [423] J. Gu, J. Wang, and H. F. Schaefer III, *J. Am. Chem. Soc.*, **2005**, *127*, 1053.
- [424] J. Gu, J. Wang, and H. F. Schaefer III, *J. Am. Chem. Soc.*, **2006**, *128*, 1250.
- [425] J. Gu, J. Wang, and J. Leszczynski, *J. Am. Chem. Soc.*, **2006**, *128*, 9322.
- [426] M. Ernzerhof and G. E. Scuseria, *J. Chem. Phys.*, **1999**, *110*, 5029.
- [427] M. Cossi and V. Barone, *J. Chem. Phys.*, **2001**, *115*, 4708.
- [428] A. Mostad and C. Romming, *Acta. Chem. Scand.*, **1971**, *25*, 3561.
- [429] T. Nagele, R. Hoche, W. Zinth, and J. Wachtveitl, *Chem. Phys. Lett.*, **1997**, *272*, 489.
- [430] J.-A. Andersson, R. Pettersson, and L. Tegner, *J. Photochem.*, **1982**, *20*, 17.
- [431] H. Fliegl, A. Kohn, C. Hattig, and R. Ahlrichs, *J. Am. Chem. Soc.*, **2003**, *125*, 9812.
- [432] J. Harada, K. Ogawa, and S. Tomoda, *Acta Cryst. B*, **1997**, *53*, 662.
- [433] T. Tsuji, H. Takashima, H. Takeuchi, T. Egawa, and S. Konaka, *J. Phys. Chem. A*, **2001**, *105*, 9347.
- [434] S. Millefiori and A. Millefiori, *J. Chem. Soc., Faraday Trans. II*, **1981**, *77*, 245.
- [435] P. P. Birnbaum, J. H. Linford, and D. W. G. Style, *Trans. Faraday Soc.*, **1953**, *49*, 735.
- [436] P. H. Gore and O. H. Wheeler, *J. Org. Chem.*, **1961**, *26*, 3295.
- [437] J. G. Grasselli, *CRC Atlas of Spectral Data and Physical Constants for Organic Compounds* ; CRC Press, A division of The Chemical Rubber Company, 18901, Cranwood Parkway, Cleveland, Ohio, 44128, 1973.
- [438] E. E. Dumont, P.-F. Loos, and X. Assfeld, *Chem. Phys. Lett.*, **2008**, *458*, 276.
- [439] E. E. Dumont, P.-F. Loos, and X. Assfeld, *J. Phys. Chem. B*, **2008**, *in press*.
- [440] E. E. Dumont, P.-F. Loos, A. D. Laurent, and X. Assfeld, *J. Chem. Theor. Comput.*, **2008**, *4*, 1171.
- [441] B. Boudaiffa, P. Cloutier, D. Hunting, M. A. Huels, and L. Sanche, *Science*, **2000**, *287*, 1658.
- [442] F. A. Hays, A. T. Teegarden, Z. J. R. Jones, M. Harms, D. Raup, J. Watson, E. Cavaliere, and P. S. Ho, *Proc. Nati. Acad. Sci. USA*, **2005**, *102*, 7157.
- [443] J. Simons, *Acc. Chem. Res.*, **2006**, *39*, 772.

Bibliographie

- [444] X. Li, M. D. Sevilla, and L. Sanche, *J. Am. Chem. Soc.*, **2003**, *125*, 13668.
- [445] X. Bao, J. Gu, J. Wang, and J. Leszczynski, *Proc. Natl. Acad. Sci. USA*, **2006**, *103*, 5658.
- [446] J. Berdys, I. Anusiewicz, P. Skurski, and J. Simons, *J. Am. Chem. Soc.*, **2004**, *126*, 6441.
- [447] I. Anusiewicz, J. Berdys, M. Sobczyk, P. Skurski, and J. Simons, *J. Phys. Chem. A*, **2004**, *108*, 11381.
- [448] J. Berdys, I. Anusiewicz, P. Skurski, and J. Simons, *J. Phys. Chem. A*, **2004**, *108*, 2999.
- [449] J. Berdys, P. Skurski, and J. Simons, *J. Phys. Chem. B*, **2004**, *108*, 5800.
- [450] R. Barrios, P. Skurski, and J. Simons, *J. Phys. Chem. B*, **2002**, *106*, 7991.
- [451] R. Nifosí, P. Amat, and V. Tozzini, *J. Comput. Chem.*, **2007**, *28*, 2366.
- [452] I. Demachy, J. Ridard, H. Laguitton-Pasquier, E. Durnerin, G. Vallverdu, P. Archirel, and B. Levy, *J. Phys. Chem. B*, **2005**, *109*, 24121.
- [453] Q. K. Timerghazin, H. J. Carlson, C. Liang, R. E. Campbell, and A. Brown, *J. Phys. Chem. B*, **2008**, *112*, 2533.
- [454] J. Gilmore and R. H. McKenzie, *J. Phys. Chem. A*, **2008**, *112*, 2162.
- [455] K. Kaznatcheyev, A. Osanna, C. Jacobsen, O. Plashkevych, O. Vahtras, H. Agren, V. Carravetta, and A. P. Hitchcock, *J. Phys. Chem. A*, **2002**, *106*, 3153.
- [456] Y. Zubavichus, A. Shaporenko, M. Grunze, and M. Zharnikov, *J. Phys. Chem. A*, **2005**, *109*, 6998.
- [457] M. Atsumi, L. González, and C. Daniel, *J. Photochem. Photobiol. A*, **2007**, *190*, 310.
- [458] R. M. Stevens, R. Pitzer, and W. N. Lipscomb, *J. Chem. Phys.*, **1963**, *38*, 550.
- [459] J. Gerratt and I. M. Mills, *J. Chem. Phys.*, **1968**, *49*, 1719.
- [460] J. Gerratt and I. M. Mills, *J. Chem. Phys.*, **1968**, *49*, 1730.
- [461] T. C. Caves and M. Karplus, *J. Chem. Phys.*, **1969**, *50*, 3649.
- [462] B. Paizs and I. Mayer, *Chem. Phys. Lett.*, **1994**, *220*, 97.
- [463] P. Jørgensen and J. Simons, *J. Chem. Phys.*, **1985**, *79*, 334.
- [464] N. C. Handy and H. F. Schaefer III, *J. Chem. Phys.*, **1984**, *81*, 5031.
- [465] J. E. Rice and R. D. Amos, *Chem. Phys. Lett.*, **1985**, *122*, 585.
- [466] N. C. Handy, R. D. Amos, J. F. Gaw, J. E. Rice, and E. D. Simandiras, *Chem. Phys. Lett.*, **1985**, *120*, 151.
- [467] P. Jørgensen and T. Helgaker, *J. Chem. Phys.*, **1988**, *89*, 1560.
- [468] M. Head-Gordon, J. A. Pople, and M. J. Frisch, *Chem. Phys. Lett.*, **1988**, *153*, 503.
- [469] M. J. Frisch, M. Head-Gordon, and J. A. Pople, *Chem. Phys. Lett.*, **1990**, *166*, 275.
- [470] M. J. Frisch, M. Head-Gordon, and J. A. Pople, *Chem. Phys. Lett.*, **1990**, *166*, 281.

-
- [471] T. J. Lee, S. C. Racine, J. E. Rice, and A. P. Rendell, *Mol. Phys.*, **1995**, *561*, 5031.
- [472] M. Head-Gordon, *Mol. Phys.*, **1999**, *96*, 673.
- [473] C. M. Aikens, S. P. Webb, R. L. Bell, G. D. Fletcher, M. W. Schmidt, and M. S. Gordon, *Theor. Chem. Acc.*, **2003**, *110*, 233.
- [474] S. Saebø, J. Baker, and K. Wolinski, *J. Chem. Phys.*, **2004**, *120*, 11423.
- [475] Q. Cui and M. Karplus, *J. Chem. Phys.*, **2000**, *112*, 1133.

Bibliographie

Table des figures

1.1	Représentation schématique de la relation entre les théories <i>Hartree-Fock</i> (HF), <i>Density Functional Theory</i> (DFT), <i>Time-Dependent Hartree-Fock</i> (TD-HF), <i>Time-Dependent Density Functional Theory</i> (TD-DFT), <i>Configuration Interaction Single</i> (CIS) et l'approximation <i>Tamm-Dancoff</i> (TDA) en TD-DFT.	22
3.1	Cluster d'eau — Partition QM/MM : la molécule d'eau traitée au niveau QM est représentée en boules et bâtons, tandis que la partie MM est représentée en bâtons épais.	28
3.2	Macrophomate synthase (MPS) [185–187] — Partition QM/MM : le site actif de l'enzyme est traité au niveau QM (bâtons fins). Il correspond à deux molécules d'eau (W1 et W2), deux acides aminés (Glu 185 et Asp 211), ainsi que deux oxygènes de la molécule de pyruvate. Ils forment un complexe octaédrique autour du magnésium (représenté en vert). Nous représentons l'état de transition hypothétique de la cycloaddition [4+2] (Diels-Alderase) entre le pyruvate (diénophile) et la pyrone (diène). L'existence de cet état de transition est sujet à controverse [188]. Le reste de l'enzyme est traité au niveau MM (<i>cartoon</i>). D'après l'entrée de la <i>Protein Data Bank</i> (PDB) 1IZC.	29
3.3	Différentes méthodologies QM/MM : Link-Atom, LSCF, GHO, RCD et EFP. Les atomes de la partie quantique sont notés Q et ceux de la partie classique C. Dans le cas de la méthode EFP, les atomes de la partie tampon sont notés B.	33
3.4	Principe d'un calcul ONIOM. Les calculs marqués par le symbole ■ sont effectués, tandis que l'énergie du calcul notée par ★ est extrapolée.	35
5.1	SLBO et SLABO de la liaison C-C centrale de la molécule de β -alanine obtenues par la méthode de PM au niveau B3LYP/6-311++G**. La localisation de la SLBO et de la SLABO a été effectuée sur l'espace des fonctions occupées et sur l'espace des fonctions virtuelles, respectivement.	58
5.2	Structure schématique de la matrice de recouvrement lors d'un calcul OLSCF.	59
5.3	Polarisation de la SLBO de la liaison C-C par les substituants électronégatifs X. Elle sera quantifiée par la relation (5.13).	64

Table des figures

5.4	Moitié de la population électronique (q , en électron) due à la SCSLBO de l'atome de carbone le plus proche de la charge classique (<i>i.e.</i> polarité de la liaison), fonction de la distance (d en Å) entre la charge classique positive et l'atome de carbone de la molécule d'éthane le plus proche de celle-ci. La charge ponctuelle est placée sur la ligne définie par les deux atomes de carbone.	65
5.5	Principe d'un calcul LSCF à l'aide d'une FCO d'après la Ref. [215].	68
5.6	Partition du nucléoside cytidine en partie QM (boules et bâtons) et MM (bâtons fins). Deux localisations de la frontière QM/MM sont considérées.	69
5.7	Partition du tripeptide Ace-Lys-NMe en partie QM (boules et bâtons) et MM (bâtons fins). La frontière QM/MM est localisée entre les atomes C_α et C_β de la lysine.	70
5.8	Numérotation des atomes et partition des fragments CH_3-R ($R = CH_3, NH_2, OH$). Les atomes d'hydrogène notés H_{MM} correspondent à ceux de la partie MM.	73
5.9	Profil énergétique (en u.a.) le long des liaisons C-C, C-N et C-O (en Å) pour les molécules d'éthane, de méthylamine et de méthanol au niveau de théorie RHF, et pour le fragment CH_3-R , au niveau LSCF(WP)/+3 FCO et SCCO avec la base 6-311G**	74
5.10	Superposition des structures de la forme neutre et protonée du tripeptide Ace-Lys-NMe au niveau de théorie RHF (noir), LSCF(ELMO)/+3/MM FCO (jaune) et LSCF(ELMO)/+3/MM SCCO (orange).	77
5.11	Dihydrofolate réductase (DHFR) de l' <i>Escherichia Coli</i> (entrée PDB 1DRE)	81
5.12	Partition schématique d'un polypeptide. Les différentes localisations de la jonction QM/MM sont représentées : le cas de la liaison $C_\alpha-C_\beta$ est représenté en vert par une ligne discontinue, les liaisons $C_\alpha-C(=O)$ par un trait ondulé en orange, tandis que les liaisons peptidiques sont en bâtonnets rouges. Les atomes de la chaîne latérale appartenant à la partie QM sont représentés en gras.	82
5.13	Résonance entre les deux structures de Lewis de la liaison peptidique.	83
5.14	Partition QM/MM du pentapeptide Ace-Ile-His-Tyr-NMe en partie QM (boules et bâtons) et partie MM (bâtons fins). Les deux formes tautomères de l'acide aminé histidine (HID et HIE) sont présentées. La forme protonée est notée HIP.	84
6.1	Molécule de pentane et numérotation des atomes.	94
6.2	Convergence des énergies d'ionisation de cœur de l'orbitale 1s (en eV) de la molécule de méthane en fonction de la base utilisée. L'entier X représente le nombre de ζ de la base de fonctions gaussiennes.	98
6.3	Molécules biomimétiques contenant une liaison peptidique.	101
6.4	Géométries des molécules d'alanine (deux conformations), des poly-Ala-tripeptides (conformations en hélice α et en brin β) et du poly-Ala-pentadécapeptide (conformations en hélice α et en brin β).	104
6.5	Partition QM/MM du poly-Ala-pentadécapeptide.	105

7.1	Photoisomérisation de la molécule de rétinol de la forme 11- <i>cis</i> en <i>all-trans</i> dans la rhodopsine par absorption d'un photon d'énergie $h\nu$	110
7.2	Structure obtenue par diffraction des rayons X (résolution 2,80 Å) de la rhodopsine bovine [381] (d'après l'entrée de la PDB 1F88). La molécule de rétinol (cofacteur) est représentée en gris foncé.	112
7.3	Premières étapes du processus photochimique initiant la vision : photoisomérisation de la forme 11- <i>cis</i> en <i>all-trans</i> par absorption d'un photon, puis formation de l'intermédiaire bathorhodopsin (batho) en 200 fs. L'intermédiaire photorhodopsine (photo) est détectable par spectroscopie Raman femtoseconde-stimulée [386].	113
7.4	Photoisomérisation réversible de l'azobenzène (AB).	115
7.5	Partition QM/MM du Poly(acide L-glutamique) contenant une chaîne latérale AB. La frontière QM/MM est localisée entre les atomes C_α et C_β et les atomes de la partie QM sont représentés en gras.	116
7.6	Géométries QM/MM optimisées des Poly(acide L-glutamique) contenant une chaîne photochromique TAB. La partie QM est représentée en bâtons, tandis que la partie MM est représentée en ruban. Nous reportons le nombre d'acides aminés de chaque structure entre parenthèses.	125
7.7	Géométries QM/MM optimisées des Poly(acide L-glutamique) contenant une chaîne photochromique CAB. La partie QM est représentée en bâtons, tandis que la partie MM est représentée en ruban. Nous reportons le nombre d'acides aminés de chaque structure entre parenthèses.	126
7.8	Déplacements énergétiques (en nm) entre les calculs TD-DFT/MM et l'AB en phase gazeuse ($\Delta\lambda^{AB}$). La valeur $\Delta\lambda^{AB}$ est décomposée en trois composantes venant de la polarisation électronique des charges MM ($\Delta\lambda^{elec}$), de la polarisation nucléaire de l'environnement MM ($\Delta\lambda^{nuc}$) et au déplacement auxochromique due à la substitution de la partie AB ($\Delta\lambda^{aux}$), tel que $\Delta\lambda^{AB} = \Delta\lambda^{elec} + \Delta\lambda^{nuc} + \Delta\lambda^{aux}$	128
7.9	Isosurface ($0.05 \text{ u.a.}^{-3/2}$) des orbitales frontières impliquées dans les transitions $n \rightarrow \pi^*$ et $\pi \rightarrow \pi^*$ de la molécule TAB greffée sur le polypeptide en conformation en feuillet β	131
A.1	Structure de la matrice \mathbf{E} ou de la matrice de Fock dans la base des MOs.	138
A.2	Structure de la matrice de réponse des orbitales \mathbf{U}^x	140

Table des figures

Liste des tableaux

4.1	Détails des interactions QM/MM. Les atomes de la partie quantique sont notés Q, et ceux de la partie MM sont notés C. Z_A et q_A représentent respectivement la charge nucléaire et la charge classique de l'atome A. $P_{\mu\nu}^T$ est l'élément $\mu\nu$ de la matrice densité totale du système.	52
5.1	Polarité de l'orbitale de liaison C-C (en électron) pour l'atome de carbone lié au substituant X dans la molécule CH_3CX_3	63
5.2	Polarité de l'orbitale de liaison C-C (en électron) pour l'atome de carbone porté par le groupement NH_2 de la molécule de β -alanine, due à la SCSLBO et à la SLBO(PM) ^a	64
5.3	Sélection de distances (en Å) et d'angles (en degrés) calculés pour $\text{CH}_3\text{-R}$ (R = CH_3 , NH_2 , OH) utilisant les niveaux de théorie RHF, LSCF(WP)/+1, LSCF(WP)/+3 FCO and LSCF(WP)/+3 SCCO avec la base 6-311G**. Nous reportons aussi la fréquence de vibration des liaisons X-Y (en cm^{-1}). Pour les calculs LSCF, les déviations par rapport aux valeurs RHF sont reportées entre parenthèses, avec les RMSD correspondant aux distances RMSD_d et aux angles RMSD_\angle (en Å et en degrés) pour la structure du fragment $\text{CH}_3\text{-Y}$	72
5.4	Charges classiques (en électrons) du champ de forces correspondant, sélection de distances (en Å) et d'angles (en degrés) calculés pour les molécules $\text{CH}_3\text{-R}$ (R = CH_3 , NH_2 , OH) en utilisant le niveau de théorie LSCF(WP)/+3/MM SCCO et la base 6-311G**. Les déviations par rapport aux valeurs RHF sont reportées entre parenthèses, avec les RMSD correspondants aux distances RMSD_d et aux angles RMSD_\angle (en Å et en degrés) pour la structure des fragments $\text{CH}_3\text{-Y}$	76
5.5	Distance d'équilibre X-Y de la molécule de cytidine au niveau RHF et LSCF(WP)/+3/MM SCCO avec la base 6-311G**. Deux localisations de la frontière QM/MM sont considérées, comme décrit sur la figure 5.6. Le champ de forces AMBER <i>ff99</i> a été utilisé et les charges classiques ont été déterminées grâce au protocole RESP. Les déviations par rapport aux valeurs RHF sont reportées entre parenthèses.	77

- 5.6 Données énergétiques et géométriques de la forme protonée (P) et neutre (N) de la lysine au niveau de théorie RHF, LSCF(ELMO)/+1/MM FBP, LSCF(ELMO)/+3/MM FCO et LSCF(ELMO)/+3/MM SCCO avec la base 6-311G**. Le champ de forces AMBER *ff99* est utilisé pour les calculs QM/MM et les charges de la partie MM sont déterminées grâce à la procédure RESP. Les distances et les valeurs de RMSD sont en Å, tandis que les énergies relatives $\Delta E = E_{\text{tot}}(\text{N}) - E_{\text{tot}}(\text{P})$ sont en kcal/mol. 78
- 5.7 **Frontière C $_{\alpha}$ -C $_{\beta}$.** Distances des liaisons frontières (en Å) des trois états de protonation de l'histidine (HID, HIE et HIP, FIG. 5.14), engagée dans un pentapeptide au niveau de théorie RHF/6-311G**, suivant les méthodologies LSCF/+1/CHARMM27 et LSCF/+3/CHARMM27. Les RMSD par rapport à la géométrie de référence RHF/6-311G** sont donnés (en Å) pour le fragment QM (RMSD_{QM}), la partie MM (RMSD_{MM}), et le système entier (RMSD). Nous donnons aussi les affinités protoniques (ΔE en kcal/mol) de l'acide aminé histidine, calculées aux mêmes niveaux de théorie ($\Delta E(\text{HID}) = E(\text{HID}) - E(\text{HIP})$ et $\Delta E(\text{HIE}) = E(\text{HIE}) - E(\text{HIP})$). Les déviations par rapport aux valeurs RHF sont reportées entre parenthèses. 85
- 5.8 **Frontières liaisons peptidiques.** Distances des liaisons frontières (en Å) et valeurs de l'angle dièdre N_X-H-C $_{\alpha}$ -C $_{\gamma}$ des trois états de protonation (HID, HIE, et HIP, FIG. 5.14) de l'acide aminé histidine engagé dans un pentapeptide Ace-Ile-His-Tyr-NMe, calculées au niveau de théorie RHF/6-311G**, LSCF(+1)/AMBER99, LSCF(+3)/AMBER99 et LSCF(+5)/AMBER99. Les valeurs des RMSD par rapport aux géométries de référence RHF/6-311G** sont données (en Å) pour le fragment QM (RMSD_{QM}), le fragment MM (RMSD_{MM}), et pour le système entier (RMSD). Nous donnons aussi les affinités protoniques (ΔE en kcal/mol) de l'acide aminé histidine, calculées aux mêmes niveaux de théorie ($\Delta E(\text{HID}) = E(\text{HID}) - E(\text{HIP})$ et $\Delta E(\text{HIE}) = E(\text{HIE}) - E(\text{HIP})$). Les déviations par rapport aux valeurs RHF sont reportées entre parenthèses. 87
- 6.1 Energies d'ionisation de l'orbitale 1s du carbone (en eV) calculées au niveau de théorie DFT/6-311++G** pour la molécule de pentane. Les calculs sont effectués avec et sans SLBO entre les atomes C $_1$ et C $_2$, avec l'ancienne (Löwdin) ou la nouvelle procédure (Lv+GS) d'orthogonalisation (voir la figure 6.1 pour la numérotation des atomes). Deux types de critères de localisation sont utilisés, Pipek-Mezey (PM) et Boys-Foster (BF). Les valeurs entre parenthèses correspondent à la différence d'énergie (en eV) entre les deux procédures d'orthogonalisation. 95
- 6.2 Energies d'ionisation de l'orbitale 1s du carbone (en eV) de la molécule de méthane en fonction de la base calculée à l'aide de la fonctionnelle PBE0 et du critère de localisation Boys-Foster pour la CO. 97

6.3	Energies d'ionisation de l'orbitale 1s du carbone (en eV) calculées à différents niveaux de théorie et critères de localisation pour de petites molécules organiques. La première famille de molécules contient uniquement des hydrocarbures, tandis que la seconde contient des molécules avec des hétéroatomes (O ou N). Les valeurs expérimentales sont prises des références [366] et [367].	99
6.4	Energies d'ionisation de l'orbitale 1s du carbone (en eV) de molécules biomimétiques (FIG. 6.3 pour les types d'atomes de carbone et les conformations des molécules). Les calculs ont été effectués au niveau de théorie PBE0/6-311++G** à l'aide du critère de localisation Boys-Foster. Les valeurs de référence sont extraites de l'article de Chong <i>et al.</i> [368, 376]. Ces calculs sont réalisés grâce à la méthode Δ KS au niveau de théorie (PW86PW91)/cc-pCVTZ. Les valeurs expérimentales sont données entre parenthèses [366].	102
6.5	Energies d'ionisation de l'orbitale 1s du carbone (en eV) de la molécule d'alanine isolée, engagée dans un tripeptide ou dans un pentadécapeptide en conformation hélice α et brin β . Voir FIG. 6.4 pour les géométries. Les calculs notés AMBER* sont effectués sans polarisation électrostatique de la fonction d'onde quantique par les charges ponctuelles classiques.	106
7.1	Structures et spectres UV-vis de la molécule de TAB et CAB dans différents milieux, obtenus au niveau de théorie TD-PBE0/6-311+G**//B3LYP/6-311G**. Les modifications des paramètres structuraux et des spectres UV-vis suivant la photoisomérisation des deux isomères sont aussi reportées. Les distances sont en Å, tandis que les angles et les dièdres sont en degrés. $\lambda^{n \rightarrow \pi^*}$ et $\lambda^{\pi \rightarrow \pi^*}$ sont les longueurs d'onde (en nm) de la première transition $n \rightarrow \pi^*$ et $\pi \rightarrow \pi^*$, respectivement. Les forces d'oscillateurs sont données entre parenthèses.	122
7.2	Paramètres structuraux des molécules TAB et CAB, isolées et greffées comme chaîne latérale de Poly(acide L-glutamique) obtenus au niveau de théorie B3LYP/6-311G**. La partie MM est décrite par le champ de forces AMBER <i>ff99</i> . Les distances sont données Å, tandis que les angles et les dièdres sont en degrés. Voir figure 7.5 pour la numérotation des atomes.	123
7.3	Spectres UV-vis de la molécule de TAB et CAB greffée comme chaîne latérale d'un Poly(acide L-glutamique) obtenus au niveau de théorie TD-PBE0/6-311+G**. La partie MM est décrite grâce au champ de forces AMBER <i>ff99</i> . $\lambda^{n \rightarrow \pi^*}$ et $\lambda^{\pi \rightarrow \pi^*}$ sont les longueurs d'onde (en nm) des premières transitions $n \rightarrow \pi^*$ et $\pi \rightarrow \pi^*$, respectivement. Les forces d'oscillateurs sont indiquées entre parenthèses. Nous reportons aussi les déplacements énergétiques (en nm) entre les valeurs TD-DFT/MM et la molécule d'AB en phase gazeuse ($\Delta\lambda^{AB}$). La valeur $\Delta\lambda^{AB}$ est décomposée en trois composantes venant de la polarisation électrostatique due aux charges MM ($\Delta\lambda^{elec}$), de la polarisation nucléaire de l'environnement MM ($\Delta\lambda^{nuc}$) et du déplacement auxochromique due à la substitution de la partie AB ($\Delta\lambda^{aux}$), tel que $\Delta\lambda^{AB} = \Delta\lambda^{elec} + \Delta\lambda^{nuc} + \Delta\lambda^{aux}$	130

Nancy-Université
Université
Henri Poincaré

Monsieur **LOOS Pierre-François**

DOCTORAT DE L'UNIVERSITÉ HENRI POINCARÉ, NANCY 1
en CHIMIE INFORMATIQUE & THEORIQUE

VU, APPROUVÉ ET PERMIS D'IMPRIMER

Nancy, le 26 juin 2008

Le Président de l'Université


J.P. FINANCE



UNIVERSITÉ HENRI POINCARÉ
Université Henri Poincaré, Nancy I
24-30 rue Lionnois - B.P. 3069 - 54013 NANCY Cédex
Tél. : 03 83 68 20 00 - Fax : 03 83 68 21 00

Résumé

Ces dix dernières années, les méthodes hybrides QM/MM combinant la mécanique quantique (QM) et la mécanique moléculaire (MM) se sont révélées particulièrement bien adaptées à l'étude des systèmes chimiques et biologiques. Elles ont permis la compréhension de nombreux phénomènes intervenant dans les macromolécules. Cependant, certaines situations requièrent une attention spéciale et/ou un niveau d'analyse et de théorie plus poussé. Ainsi, de nombreux développements sont nécessaires pour atteindre un niveau de précision suffisant à la compréhension du phénomène biochimique étudié. Afin d'atteindre cet objectif, plusieurs spécificités théoriques et méthodologiques sont en continuelle amélioration : (i) la jonction QM/MM est toujours un sujet délicat dans les méthodes mixtes et un axe de recherche très fécond. (ii) la complexité des mécanismes étudiés impose un traitement de la partie QM à des niveaux post-HF. Dans cette thèse, nous présenterons principalement la méthode *Local Self-Consistent Field* (LSCF), ainsi que ses derniers développements méthodologiques. Dans cette approche, la jonction entre les parties QM et MM est représentée à l'aide d'une orbitale localisée. Nous montrerons qu'elle permet de traiter de façon rigoureuse un grand nombre de situations faisant intervenir des systèmes biologiques, comme les ionisations de cœur et les spectres d'absorption UV-vis de polypeptides fonctionnalisés. De plus, nous présenterons quelques développements théoriques, comme la dérivation détaillée des équations *Coupled-Perturbed Hartree-Fock* dans le cadre de la méthode LSCF, ainsi que les principales étapes aboutissant à la détermination des gradients de l'énergie MP2.

Mots-clés: QM/MM, LSCF, liaison peptidique, ionisation de cœur, azobenzène, TD-DFT, CPHF, MP2

