



## AVERTISSEMENT

Ce document est le fruit d'un long travail approuvé par le jury de soutenance et mis à disposition de l'ensemble de la communauté universitaire élargie.

Il est soumis à la propriété intellectuelle de l'auteur. Ceci implique une obligation de citation et de référencement lors de l'utilisation de ce document.

D'autre part, toute contrefaçon, plagiat, reproduction illicite encourt une poursuite pénale.

Contact : [ddoc-theses-contact@univ-lorraine.fr](mailto:ddoc-theses-contact@univ-lorraine.fr)

## LIENS

Code de la Propriété Intellectuelle. articles L 122. 4

Code de la Propriété Intellectuelle. articles L 335.2- L 335.10

[http://www.cfcopies.com/V2/leg/leg\\_droi.php](http://www.cfcopies.com/V2/leg/leg_droi.php)

<http://www.culture.gouv.fr/culture/infos-pratiques/droits/protection.htm>

---

FACULTE DES SCIENCES

U.F.R. Sciences et Techniques de la Matière et des Procédés

Ecole doctorale SESAMES

Thèse

présentée pour l'obtention du titre de

**Docteur de l'Université Henri Poincaré, Nancy-I**

en Chimie Informatique et Théorique

par **Nicolas 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**

Soutenue publiquement le 22 octobre 2001

Membres du jury :

Président :	M. Jean-Louis RIVAIL	Professeur, Université Henri Poincaré Nancy I
Rapporteurs :	M. Daniel BORGIS	Professeur, Université Evry-Val d'Essonne
	M. Jean-Pierre DAUDEY	Directeur de Recherches CNRS, Université Paul Sabatier Toulouse III
Examineurs :	M. Xavier ASSFELD	Maître de Conférences, Université Henri Poincaré Nancy I
	M. Hervé TOULHOAT	Docteur, Institut Français du Pétrole, Rueil-Malmaison

---

Laboratoire de Chimie Théorique, Unité Mixte de Recherche CNRS-UHP 7565

Faculté des Sciences - 54500 Vandœuvre-lès-Nancy



*Aux deux étoiles qui illuminent ma vie, Karine et Alhena*



## Remerciements

Ce travail a été mené à bien au sein du laboratoire de chimie théorique de Nancy. Tout d'abord, je veux remercier Monsieur Rivail pour la grande confiance et l'intérêt constant qu'il m'a témoignés depuis le DEA. Sa participation au jury en était une preuve supplémentaire et je lui en suis très reconnaissant.

Je ne peux que clamer mon immense gratitude à Xavier Assfeld qui m'a guidé et accompagné durant ces quatre années en acceptant de diriger ces recherches. Nos discussions furent nombreuses, ses suggestions innombrables, le volume de mes questions étant au moins aussi considérable. J'espère très sincèrement que nos chemins se croiseront à nouveau et, qui sait, que nous travaillerons ensemble encore une fois.

Mes remerciements s'adressent également à Messieurs Borgis et Daudey qui m'ont fait l'honneur de rapporter les travaux présentés dans ce mémoire.

Monsieur Toulhoat a accepté d'examiner et de juger ce travail et je l'en remercie. J'espère qu'il y a trouvé un nouvel outil qui sera bénéfique à tous les membres du Groupe de Recherches « Dynamique moléculaire quantique appliquée à la catalyse ».

Au laboratoire, János Ángyán a permis le développement d'une partie de la méthode présentée ici et a fourni le programme de calcul du potentiel de Madelung, MadEwa. Merci pour le temps qu'il m'a accordé.

Un grand merci à tous les autres membres du laboratoire, passés ou présents, et à toutes les personnes qui se sont déplacées le jour de ma soutenance. Je pense en particulier à Stéphanie Chalmet qui m'a accompagné depuis le DEA ; MC Jérôme Golebiowski (aka Dr Dre) ; Yannick Jeanvoine ; Walid Harb qui a récemment repris le flambeau LSCF/MM pour une étude difficile (courage !).



# Sommaire

<b>1</b>	<b>Introduction</b>	<b>11</b>
1.1	Structure des zéolithes . . . . .	11
1.2	Méthodes usuelles . . . . .	13
1.2.1	Les méthodes « périodiques » . . . . .	14
1.2.2	Les méthodes « cluster » . . . . .	16
1.2.3	Inconvénients des méthodes actuelles – Nos solutions . . . . .	18
<b>2</b>	<b>QM</b>	<b>25</b>
2.1	Méthodes quantiques . . . . .	26
2.1.1	Expression de la fonction d’onde multiélectronique . . . . .	26
2.1.2	Équations de Hartree-Fock . . . . .	27
2.1.3	Équations de Roothaan . . . . .	28
2.1.4	La méthode de résolution auto-cohérente SCF . . . . .	30
2.1.5	Méthodes de la fonctionnelle de la densité . . . . .	31
2.1.6	Les méthodes post Hartree-Fock . . . . .	32
2.2	La méthode LSCF . . . . .	34
2.2.1	Optimisation de la fonction d’onde . . . . .	34
2.2.2	Modification des équations de Roothaan . . . . .	38
2.2.3	Gradients de l’énergie LSCF . . . . .	40
2.3	Etats ionisés ; etats excités . . . . .	42
2.3.1	Quelques repères méthodologiques . . . . .	43
2.3.2	Description des états ionisés de coeur par la méthode LSCF . . . . .	45
2.3.3	Description des états excités par la méthode LSCF . . . . .	45
2.3.4	Application à l’ionisation de molécules . . . . .	47
2.3.5	Application aux états excités . . . . .	49



2.4	Conclusion . . . . .	50
<b>3</b>	<b>QM/MM</b>	<b>55</b>
3.1	Méthodes QM/MM . . . . .	55
3.1.1	Mécanique moléculaire . . . . .	57
3.1.2	Les méthodes de type ONIOM . . . . .	58
3.1.3	Les méthodes de type soluté-solvant . . . . .	59
3.1.4	Les méthodes de type macromoléculaire . . . . .	59
3.2	LSCF/MM . . . . .	61
3.2.1	Les orbitales de liaison strictement localisées . . . . .	61
3.2.2	Le champ de forces hybride LSCF/MM . . . . .	72
3.2.3	Potentiel de liaison frontière . . . . .	73
3.2.4	Modification des charges à la frontière . . . . .	81
3.2.5	Protocole de simulation . . . . .	89
3.3	Etats ionisés de la crambine . . . . .	91
3.4	Conclusion . . . . .	93
<b>4</b>	<b>QM/MM/Madelung</b>	<b>99</b>
4.1	Interactions QM - potentiel externe . . . . .	100
4.1.1	Théorie . . . . .	101
4.1.2	Mise en oeuvre et validation . . . . .	104
4.2	LSCF/MM/Madelung . . . . .	111
4.2.1	Etude préliminaire . . . . .	112
4.2.2	Molécule modèle . . . . .	114
4.2.3	Paramétrisation O-Si . . . . .	119
4.3	Validation sur la chabazite . . . . .	120
4.4	Conclusion . . . . .	124
<b>5</b>	<b>Conclusion générale</b>	<b>127</b>
	<b>Annexe</b>	<b>131</b>
<b>A</b>	<b>Modifications du logiciel Gaussian98</b>	<b>131</b>
<b>B</b>	<b>Quelques exemples de fichiers de données LSCF</b>	<b>139</b>

## Avant-propos

L'étude de la réactivité catalysée par des zéolithes nécessite un modèle théorique élaboré qui prend en compte certaines caractéristiques physiques et chimiques de ces systèmes. Leur taille constitue le problème majeur à résoudre : un traitement quantique n'est possible qu'en adoptant certaines approximations qui concernent la structure de la zéolithe.

Les nombreux modèles proposés peuvent se classer en deux catégories. La première tire parti de la symétrie de translation présente dans les solides : ces systèmes sont abordés entièrement de façon quantique mais la périodicité introduite reste artificielle, notamment si des défauts chimiques doivent être considérés. La deuxième catégorie de méthodes ne s'intéresse qu'au centre réactif : seul un petit fragment du solide est abordé quantiquement, le reste est alors négligé ou pris en compte approximativement.

L'objectif de notre travail a été de mettre au point un modèle qui allie les avantages des deux approches. Nommé LSCF/MM/Madelung, il a nécessité le développement d'une méthode quantique originale, couplée ensuite à un champ de forces classique, l'ensemble ressentant alors les effets électrostatiques du potentiel de Madelung du solide.

Ce mémoire débutera avec un chapitre introductif présentant les méthodes évoquées ci-dessus en insistant sur leurs avantages et inconvénients respectifs. Nous introduirons alors notre modèle et justifierons son intérêt.

Le deuxième chapitre exposera les principes de la méthode quantique LSCF ainsi qu'une application à l'étude des états ionisés et excités de petites molécules.

Le troisième chapitre développera et validera le couplage entre la méthode LSCF et un champ de forces classique. Une attention particulière sera portée à la description de la frontière entre les deux mondes quantique et classique.

Le quatrième chapitre sera consacré à la mise au point d'un formalisme permettant de traiter les interactions entre un système moléculaire explicite et un potentiel externe. Son application au cas des zéolithes sera expliquée et le champ de forces hybride LSCF/MM/Madelung sera ainsi achevé.

Une conclusion générale terminera ce mémoire. Elle sera l'occasion de présenter un certain nombre

de pistes d'application des différentes composantes du logiciel développé tout au long de ces trois années de thèse.

En annexe seront récapitulées toutes les modifications apportées au logiciel de chimie quantique Gaussian98 qui a servi de support aux développements informatiques.

# Chapitre 1

## Introduction

Le travail exposé dans ce mémoire concerne essentiellement le développement d'un nouvel outil de modélisation des réactions chimiques catalysées par les zéolithes. Ce sujet est issu de la participation du laboratoire de chimie théorique de Nancy au groupe de recherche européen « Dynamique moléculaire quantique appliquée à la catalyse » et doit permettre une approche originale qui couple le savoir-faire nancéien dans les méthodes hybrides QM/MM et les récentes avancées dans la compréhension de la chimie des zéolithes dues à l'activité du GdR.

### 1.1 Structure des zéolithes

Les zéolithes sont des matériaux très intéressants largement exploités par les industriels. Ce sont des solides iono-covalents de formule générale  $M_{x/n}[(AlO_2)_x(SiO_2)_y].wH_2O$  où  $n$  représente la valence du cation  $M$ . Les structures zéolithiques sont formées de tétraèdres de  $[SiO_4]^{4-}$  et de  $[AlO_4]^{5-}$  (appelés structures primaires) qui partagent un atome d'oxygène avec leurs voisins. L'assemblage de ces tétraèdres donne lieu à la construction d'unités secondaires en forme de cages ou d'anneaux de tailles variées (figure 1.1). La diversité des possibilités d'assemblage de ces structures permet de construire de très nombreuses structures différentes par leurs formes et leurs tailles. Ainsi, si les premières zéolithes étaient d'origine naturelle, l'industrie en a maintenant synthétisé de nombreuses autres, dédiées à diverses tâches.

Cette « adaptabilité » de la structure des matériaux zéolithiques est une première raison de l'intérêt qu'on leur porte. Elle permet d'utiliser ces systèmes comme échangeur d'ions de tailles adaptées aux cages et canaux présents dans la zéolithe considérée. Citons par exemple le traitement des déchets radioactifs (strontium et césium), le traitement de l'eau (ammoniac). Les zéolithes sont aussi utilisées comme tamis moléculaire, afin de séparer certains ions. On peut moduler la taille des cages en changeant de

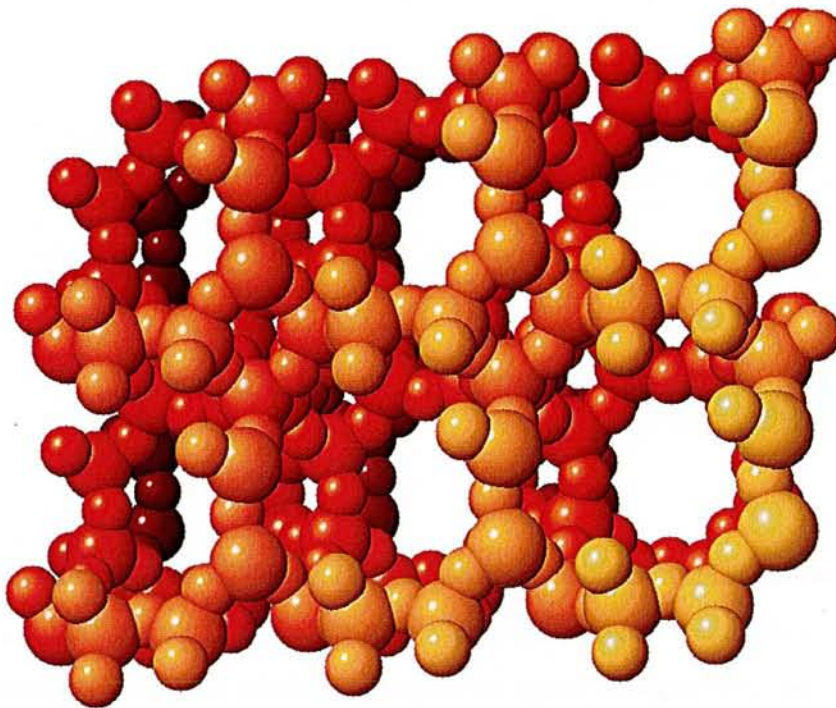


FIG. 1.1 – Anneaux caractéristiques de la structure de la chabazite

cation M.

L'assemblage des tétraèdres de  $\text{SiO}_4$  permet de conserver une charge électrique nulle de l'ensemble tridimensionnel. En revanche, lorsque l'on substitue un groupement  $[\text{AlO}_4]^{5-}$  à un groupement  $[\text{SiO}_4]^{4-}$ , un déséquilibre des charges apparaît. Pour conserver la neutralité électrique, une charge positive doit compenser cet excès négatif. Elle provient soit de cations (par exemple  $\text{Na}^+$ ) soit d'un proton qui se fixe sur un oxygène du site tétraédrique substitué. Il aura alors un caractère acide au sens de Brønsted (figure 1.2). Cette dernière particularité nous intéressera particulièrement : l'acidité module l'activité catalytique de la zéolithe. En effet, de par leur structure présentant de nombreux sites catalytiques ainsi qu'une grande surface interne, les zéolithes sont utilisées comme catalyseurs. Le principe général de la catalyse par les zéolithes est le suivant : la taille des pores de la zéolithe est sélective du réactif et/ou du produit de la réaction envisagée et l'acidité du site catalytique amorce ou non la réaction. L'application industrielle majeure de ce pouvoir catalytique des zéolithes concerne les réactions de craquage des hydrocarbures, c'est-à-dire la conversion des grosses chaînes carbonées en plus petites (pour produire de l'essence par exemple).

La relation acidité de Brønsted - pouvoir catalytique a fait l'objet de nombreuses études et de développements théoriques importants, mais la compréhension des différents phénomènes chimiques et

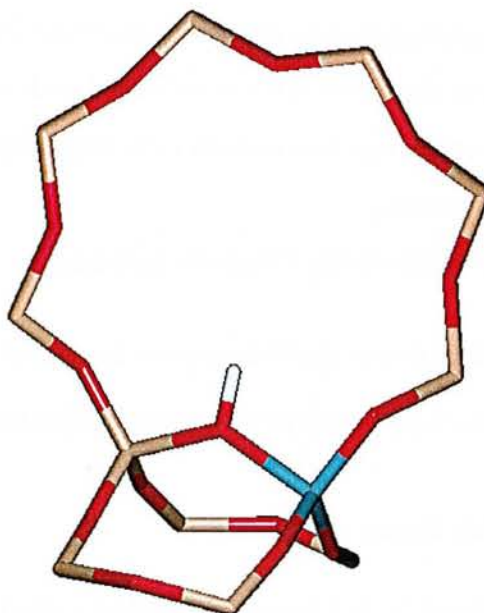


FIG. 1.2 – Site acide dans une zéolithe. Les codes de couleur sont : silicium jaune, aluminium bleu, oxygène rouge et hydrogène blanc.

physiques prenant place dans les zéolithes est loin d'être achevée. Notre démarche a pour but de proposer une approche originale qui doit apporter de nouveaux renseignements sur les processus responsables de l'activité catalytique au sein des zéolithes.

## 1.2 Les méthodes usuelles de simulation des zéolithes

Mais avant d'aborder les principes de notre modèle des zéolithes, nous allons rappeler les différentes techniques qui sont utilisées à l'heure actuelle dans ce domaine particulier de la chimie quantique. Il existe bien sûr des approches basées sur la mécanique classique, mais comme nous le verrons dans le chapitre suivant, elles ne sont pas bien adaptées aux études de réactivité. Nous citerons quand même les travaux de Henson et collaborateurs qui ont mis au point un champ de forces dédié aux zéolithes [1] ainsi que ceux du groupe de Sauer [2, 3]. Le programme GULP, mis au point par Gale [4], permet d'utiliser un potentiel classique défini par l'utilisateur et de tirer parti de la symétrie de translation. Il a été utilisé dans l'étude de la structure de zéolithes de type SAPO (dans lequel les atomes de silicium sont remplacés par des atomes d'aluminium et les atomes de substitution sont des phosphores) [5]. Les propriétés structurales et dynamiques des zéolithes peuvent également être étudiées par dynamique moléculaire, nous

renvoyons le lecteur à l'article de revue de Demontis et Suffritti [6].

Les différentes approches quantiques de la modélisation des zéolithes ont été illustrées il y a quelques années par Sauer [7, 8] et ont été abordées sommairement dans l'article de revue introduisant la modélisation des zéolithes de van de Graaf, Lan Njo et Smirnov [9]. Depuis, différentes évolutions ont eu lieu et de nouvelles techniques ont émergé.

Celles-ci se divisent en deux selon le type d'approche proposé :

- une approche plus physique faisant appel à la symétrie de translation présente dans ces systèmes ;
- une approche plus chimique qui s'intéresse plus précisément au centre réactif.

### 1.2.1 Les méthodes « périodiques »

Les méthodes traditionnelles de la chimie quantique (que nous détaillerons dans le chapitre suivant) sont basées pour la plupart sur la résolution des équations de Hartree-Fock, et plus particulièrement sur leur expression dans une base de fonctions localisées sur les noyaux des atomes : les équations de Roothaan. Les solutions de ces équations sont appelées orbitales moléculaires et sont délocalisées dans tout le système étudié. Dans le cas de systèmes périodiques infinis, la résolution de ces équations devient impossible : les orbitales moléculaires se développeraient sur une base infinie d'orbitales atomiques. Cependant, l'utilisation de la symétrie permet de transformer ces équations et de les rendre à nouveau solubles.

Ces méthodes utilisent le formalisme des orbitales de Bloch adaptées à la symétrie translationnelle présente dans un solide parfait. Les orbitales de Bloch (OB) sont des combinaisons linéaires de fonctions de base. Dans le cas d'un système de dimension 1 [10], celles-ci sont données par

$$OB_j(\vec{k}) = \frac{1}{\sqrt{n}} \sum_{m=-n'+1}^{n'} \exp(i\vec{k} \cdot \vec{r}_m) (\phi_j)_m \quad (1.1)$$

tel que  $\vec{k}$  est le vecteur d'onde caractéristique des propriétés de l'orbitale de Bloch par rapport à une translation d'un vecteur multiple de  $\vec{a}$  ( $a$  est la longueur de la maille),  $n' = n/2$  avec  $n$  le nombre de mailles constituant le système périodique,  $\vec{r}_m = m \cdot \vec{a}$  et  $(\phi_j)_m$  une fonction de base d'indice  $j$  présente dans la maille  $m$ . Dans l'espace réciproque des vecteurs  $\vec{k}$ , les orbitales de Bloch ont la propriété suivante :

$$OB_j(\vec{k}) = OB_j(\vec{k} + p \cdot \vec{a}^*) \quad (1.2)$$

où  $\vec{a}^* = \frac{2\pi}{a} \vec{a}$  est le vecteur qui engendre l'espace réciproque et  $p \in \mathbb{Z}$ . De la même façon, elles sont adaptées à la symétrie de translation de l'espace direct puisque la transformation  $\vec{r}_m^{\rightarrow'} = \vec{r}_m^{\rightarrow} + p \cdot \vec{a}$  engendre la même orbitale de Bloch à un facteur de module 1 près.

L'interaction entre les  $N$  orbitales de Bloch de même symétrie permet d'obtenir les états propres monoélectroniques du système : on parle alors d'orbitales cristallines (OC).

$$OC(\vec{k}) = \sum_{j=1}^N C_j(\vec{k}) OB_j(\vec{k}) \quad (1.3)$$

Ces dernières sont fonctions propres d'un opérateur décrivant le système pour un vecteur  $\vec{k}$  donné. Nous qualifierons ce type d'approche de « Hartree-Fock périodique »<sup>1</sup>. On la retrouve par exemple dans le programme Crystal [12] mis au point par l'équipe de Pisani à Turin. Ce dernier permet également de résoudre les équations de Kohn-Sham, expression de la théorie de la fonctionnelle de la densité (voir chapitre suivant) et qui a l'avantage de prendre en compte la corrélation électronique.

L'utilisation de fonctions de base localisées n'est pas la seule possibilité à la résolution des équations de Hartree-Fock ou de Kohn-Sham. Ces dernières années sont apparues plusieurs méthodes basées sur l'utilisation d'ondes planes comme fonctions de base. Elles présentent de nombreux avantages, comme par exemple leur grande commodité de calcul ou l'absence d'erreur de superposition de base.

La combinaison de leur utilisation avec d'autres techniques comme les pseudo-potentiels et les simulations de dynamique moléculaire *ab initio* de type Car-Parinello ont permis l'émergence de nombreux programmes. Particulièrement, nous citerons le programme VASP (pour Vienna Ab initio Simulation Package [13]) développé par Hafner et son groupe et mis à disposition des membres du GdR.

Pisani a répertorié récemment dans Theochem la majorité des programmes qui utilisent l'approche exposée ci-dessus [14].

Les études utilisant l'approche périodique appliquée aux zéolithes ont été nombreuses. Le transfert du proton acide entre différents sites a été étudié par Benco [15, 16]. Le groupe de Gale s'est particulièrement intéressé à l'adsorption du méthanol dans les zéolithes et à sa transformation [17, 18, 19]. La substitution d'atomes de silicium par des atomes de germanium et son influence sur l'activité catalytique de zéolithes riches en titane a été étudiée par Zicovich-Wilson et Corma [20]. Demuth *et al.* ont caractérisé les effets d'une substitution Si par Al dans la mordénite [21]. Ugliengo a étudié l'influence du rapport Si/Al sur la fréquence de vibration du groupement OH acide [22].

Au laboratoire, Y. Jeanvoine a utilisé le logiciel VASP pour étudier les sites acides des zéolithes

---

1. Il est évident que des calculs plus simples, basés sur des Hamiltoniens monoélectroniques effectifs, peuvent aussi utiliser cette technique. Le programme Yaehmop [11] en est un exemple pour la méthode Hückel étendue.



HSSZ-13 et HSAPO-34 [23]. Nous nous servons de ses résultats comme référence et outil de comparaison pour valider notre méthode.

### 1.2.2 Les méthodes « cluster »

Pour tenir compte de la périodicité des systèmes cristallins dans les calculs de la chimie quantique, nous avons vu qu'il est nécessaire de modifier (et de compliquer) les équations. Une alternative à cette approche consiste à ne s'intéresser qu'à une petite portion du solide, ce qui permet ainsi de garder une approche moléculaire de la réactivité dans les solides. Les équations développées pour les molécules peuvent s'adapter sans modification majeure. On appelle cette approche « cluster ».

Le principe est donc de modéliser le site réactif par un petit ensemble d'atomes représentatifs ; cet ensemble est saturé par des atomes monovalents (la plupart du temps des atomes d'hydrogène). De par sa simplicité de mise en oeuvre et sa rapidité, cette approche est largement répandue, comme en témoigne la littérature abondante qui lui est consacrée. Les études se distinguent à la fois par la taille du cluster étudié et la méthode quantique utilisée.

Plusieurs équipes sont intéressées dans les phénomènes d'échange ou d'arrachement de proton à une molécule organique en présence du site acide : les états de transition caractérisant ces phénomènes ainsi que les chemins réactionnels sont activement recherchés [24, 25, 26, 27, 28, 29]. Brand, Curtiss et Iron ont étudié les sites acides dans la zéolithe ZSM-5 en fonction de la taille du cluster considéré : leurs conclusions montrent une convergence très lente de l'affinité protonique avec la taille [30]. Nous reviendrons plus tard sur cette remarque. Harmsen, Bates et van Santen se sont intéressés à l'influence du palladium sur l'acidité zéolithique [31]. La force du site acide est souvent caractérisée par son interaction avec une molécule de monoxyde de carbone [32]. Kassab a exploré la stabilité relative des complexes formés par l'adsorption de composés carbonylés sur le site acide [33]. Ici encore, il semble que la taille du cluster joue un rôle non-négligeable dans l'interprétation des résultats. L'utilisation de clusters très larges devient nécessaire [34], nous verrons pourquoi par la suite. L'influence de la taille et de la forme des clusters de silicate a également été étudiée par Pereira, Catlow et Rice [35, 36] et renforce l'idée que l'approche cluster ne peut être envisagée qu'après une réflexion sur les effets de taille et de forme. Nous citerons encore les travaux de Lee et collaborateurs concernant l'affinité du lithium pour des fragments anioniques de zéolithes [37] ou encore ceux de Rice, Chakraborty et Bell sur la réduction catalytique de cations divalents dans la zéolithe ZSM-5 [38].

Cette brève revue montre bien le champ d'investigation très étendu tant du point de vue chimique que théorique proposé aux équipes qui abordent la réactivité dans les zéolithes. Cependant, comme il

a été mentionné, l'inconvénient majeur de l'approche cluster vient de la taille forcément limitée de ces modèles. En particulier, ils ne permettent pas de prendre en compte les effets d'environnement sur la réaction, la contribution majeure à ces effets provenant des interactions électrostatiques à longue portée ( $\sim \frac{1}{r}$ ). Plusieurs solutions ont été proposées afin de prendre en compte ces interactions.

La première solution consiste évidemment à augmenter la taille du cluster [39]. Cependant les ressources informatiques nécessaires deviennent vite prohibitives.

Une autre approche consiste à incorporer le potentiel de Madelung (nous le présenterons plus en détail dans la dernière partie de ce mémoire) d'une façon ou d'une autre dans le calcul effectué sur le cluster. Ángyán a proposé de modifier directement les équations SCF [40, 41]. Par exemple, ce sont les effets du cristal qui permettent de montrer que la structure la plus stable du complexe ammoniac - site acide est la forme ionique (c'est-à-dire après transfert d'un proton).

Une autre approche, dite « embedded cluster », est très populaire. Elle consiste à plonger le cluster quantique au sein d'un réseau de charges ponctuelles choisies de façon à reproduire le potentiel de Madelung du cristal parfait. Ces charges sont souvent calculées à partir d'une approche périodique, telle que décrite plus haut : elles sont ajustées au potentiel électrostatique dans l'approximation des charges ponctuelles [42, 43]. La méthode SCREEP consiste à séparer ces interactions électrostatiques en deux couches : la première (la plus proche du cluster) est décrite à l'aide de charges ponctuelles dérivées de calculs périodiques ; la deuxième (la plus éloignée) est représentée par une surface de charges qui remplacent les nombreuses charges qui seraient encore nécessaires à l'obtention du potentiel de Madelung exact [44, 45].

Ces dernières techniques peuvent être qualifiées de « embedding » électronique, c'est-à-dire que l'environnement n'a pas d'influence stérique sur le site réactif. Pour améliorer ce modèle, il peut être avantageux de lui ajouter d'autres contributions, qui ont moins de portée, mais qui peuvent jouer néanmoins un rôle important. On parle alors de méthode QM/MM. Ces méthodes seront particulièrement détaillées dans la troisième partie de ce mémoire. Brièvement, nous dirons que l'environnement du site réactif est toujours représenté par des charges ponctuelles mais dont les interactions avec les autres charges et le cluster quantique peuvent être de différentes natures. La contribution majeure à ces techniques dans le domaine des zéolithes est due à l'équipe de Sauer à Berlin. Il ont mis au point une méthode appelée QM-Pot qui couple la description quantique du cluster avec une description à base de différents potentiels empiriques interatomiques pour l'environnement [46] (notons que dans ce modèle, les charges de l'environnement ne sont pas visibles du système quantique). L'équipe de Sauer a appliqué leur approche à l'étude des sites acides dans diverses zéolithes [47, 48]. Très récemment, QM-Pot a été reformulé et

étendu à de nouveaux types de potentiels interatomiques spécifiques aux solides [49], puis appliqué à diverses études comme la mobilité du proton dans la chabazite, la faujasite ou la H-ZSM-5 [50]. D'autres méthodes QM/MM appliquées aux zéolithes ont été récemment publiées, par exemple celle de de Vries *et al.* [51].

### 1.2.3 Inconvénients des méthodes actuelles – Nos solutions

Plusieurs études comparatives des approches cluster et périodique ont été publiées [52, 53, 54]. Toutes convergent sur la nécessité d'incorporer les effets d'environnement dans le calcul quantique. L'approche simple cluster n'est donc pas assez efficace pour permettre une approche quantitative de la réactivité dans les zéolithes. Une première solution proposée consiste à agrandir la taille du système. Cependant, les interactions électrostatiques étant à très longue portée, une taille adaptée à une bonne prise en compte de celles-ci devient vite prohibitive et reste hors de portée des calculateurs les plus puissants. L'inclusion des effets d'environnement via le mécanisme d'embedding pose des problèmes de description : doit-on polariser la partie quantique par des charges externes ou se contenter comme Sauer d'une description géométrique de l'environnement ? Si oui, quel type de charges doit-on utiliser ? Quelles formes de potentiel doit-on utiliser ? Faut-il définir un champ de forces spécifique à ce type de simulation ?

L'autre solution proposée, le calcul quantique utilisant la périodicité du système, se heurte également rapidement au problème informatique. En effet, à l'heure actuelle, seuls les systèmes de maille unitaire suffisamment petite peuvent être abordés par ces techniques. De telles zéolithes ont rarement un intérêt industriel, elles servent alors de modèle théorique aux gros systèmes. Il ressort donc que ni l'approche « cluster » ni l'approche périodique ne permettent de traiter au niveau quantique des systèmes dont la maille unitaire contient plus de deux cents atomes.

D'autres défauts inhérents à l'approche périodique sont soulignés. Le premier a trait à l'inclusion des défauts cristallins dans les calculs, en particulier la substitution à un atome de silicium d'un atome d'aluminium. La périodicité de ces substitutions n'a évidemment aucun sens physique et constitue un artefact de la méthode. De plus, si les mailles sont de petite taille, les interactions entre défauts ne sont pas négligeables. Dans un système réel, la distribution des sites de substitution est aléatoire et rien n'indique que celle-ci soit périodique. Une solution consiste à augmenter la taille de la maille en définissant une super-maille afin d'éloigner les défauts. Mais on se heurte alors à nouveau au problème de la taille du système à simuler qui devient vite inabordable.

D'autre part, la réaction chimique étudiée aurait lieu simultanément dans chaque maille (ou super-

maille) en appliquant le traitement périodique, ce qui n'a également rien de réaliste.

Concernant l'approche cluster, outre les effets d'environnement négligés, nous devons également mentionner les problèmes de terminaison du système quantique. En effet, pour sélectionner un système réactif dans un solide, il est nécessaire de couper certaines liaisons chimiques. Il en résulte des valences libres qui ne sont pas représentatives du système dans le solide. La solution la plus communément adoptée est l'utilisation d'atomes d'hydrogène pour saturer ces liaisons. Ces atomes n'ayant aucune réalité physique, ils posent de nombreux problèmes que nous détaillerons dans le chapitre consacré aux méthodes QM/MM.

L'approche que nous proposons tire parti des avantages des deux types d'approches mentionnées. Désirant étudier des réactions chimiques au sein des zéolithes, nous devons utiliser la mécanique quantique pour rendre compte des liaisons qui se rompent et qui se créent dans le site réactionnel. Afin de décrire correctement l'effet ponctuel et aléatoire des défauts chimiques, le petit agrégat quantique (modélisant le site réactionnel) sera entouré d'un cluster d'atomes explicites décrits par la mécanique moléculaire. Ceci permet de diminuer les temps de calcul tout en conservant une représentation précise de l'environnement direct de la réaction. Enfin, les effets électrostatiques à longue distance seront pris en compte en plongeant le cluster QM/MM dans le potentiel de Madelung du reste du cristal supposé parfait. Nous appellerons notre approche QM/MM/Madelung. Elle peut être représentée schématiquement par la figure 1.3.

L'originalité de notre démarche provient essentiellement de deux points : la coupure entre parties QM et MM ne se fait pas par l'intermédiaire d'atomes d'hydrogène mais en utilisant des orbitales de liaison strictement localisées, gelées durant le calcul quantique ; l'incorporation du potentiel de Madelung dans le calcul QM/MM est possible grâce au développement d'une nouvelle méthode d'interaction entre un potentiel externe et une distribution de charges.

La mise au point de notre approche s'effectuera en trois grandes étapes :

1. Mise au point théorique d'une nouvelle méthode quantique appelée Local Self Consistent Field, qui permet de geler des spinorbitales prédéfinies durant un calcul SCF. Notre méthode sera validée et appliquée aux problèmes des ionisations et excitations des électrons de cœur.
2. Définition du champ de forces hybride LSCF/MM. Elle nécessite la construction d'un potentiel de liaison frontière et l'ajustement des charges classiques à la frontière. Quelques exemples d'application seront donnés.
3. Mise au point de la méthode LSCF/MM/Madelung. Après avoir développé le formalisme théorique, nous démontrerons son efficacité sur quelques exemples simples. Puis nous l'appliquerons

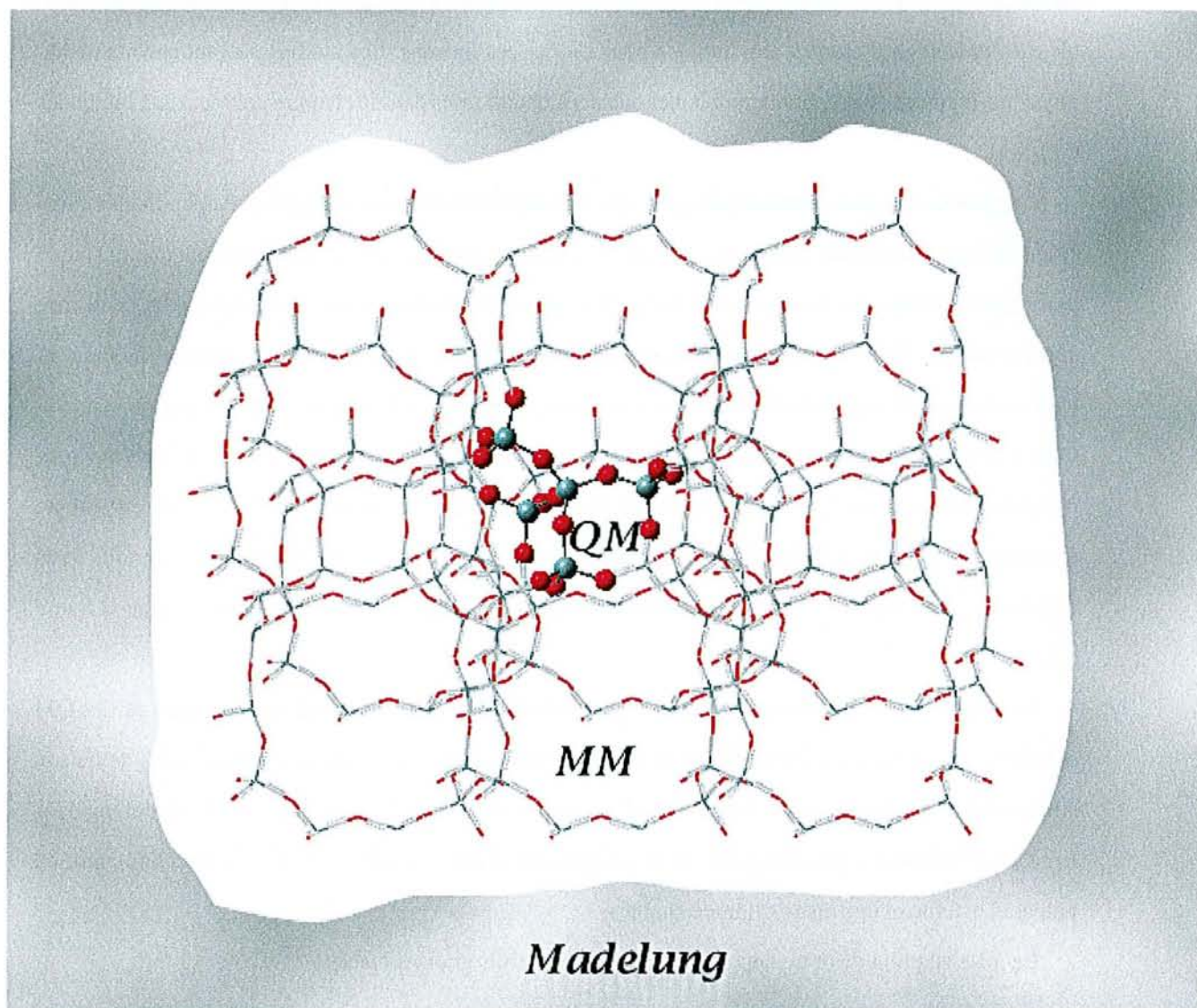


FIG. 1.3 – Représentation schématique de la méthode QM/MM/Madelung. La partie décrite par la mécanique quantique est représentée en « boules et bâtons », celle décrite par la mécanique moléculaire en « fils de fer » et le champ de Madelung par l'espace grisé.

à l'étude des sites acides dans la chabazite. Enfin, plusieurs pistes pour le futur seront données.

Tous les développements méthodologiques présentés dans ce mémoire ont été incorporés dans l'ensemble de programmes Gaussian98 [55]. La fin de ce mémoire sera consacrée à la présentation des ajouts et modifications qui ont été opérés dans ce programme. Une conclusion générale sera donnée et quelques perspectives concernant aussi bien la méthode LSCF que l'étude des zéolithes par LSCF/MM/Madelung seront avancées.

## Bibliographie

- [1] N.J. Henson, A.K. Cheetham, M. Stockenhuber et J.A. Lercher. *J. Chem. Soc., Faraday Trans.* **94**, 3759 (1998).
- [2] J.-R. Hill et J. Sauer. *J. Phys. Chem.* **99**, 9536 (1995).
- [3] K.-P. Schröder et J. Sauer. *J. Phys. Chem.* **100**, 11043 (1996).
- [4] J.D. Gale. *J. Chem. Soc., Faraday Trans.* **93**, 629 (1997).
- [5] G. Sastre, D.W. Lewis et C.R.A. Catlow. *J. Phys. Chem.* **100**, 6722 (1996).
- [6] P. Demontis et G.B. Suffritti. *Chem. Rev.* **97**, 2845 (1997).
- [7] J. Sauer. *Chem. Rev.* **89**, 199 (1989).
- [8] J. Sauer, P. Ugliengo, E. Garrone et V.R. Saunders. *Chem. Rev.* **94**, 2095 (1994).
- [9] B. van de Graaf, Swie Lan Njo et K.S. Smirnov. *Reviews in Computational Chemistry*, tome 14, chapitre Introduction to zeolite modeling, pages 137–223. Wiley-VCH, John Wiley and Sons, Inc., New York (2000).
- [10] C. Iung et E. Canadell. *Description orbitale de la structure électronique des solides I. De la molécule aux composés 1D*. Ediscience (1997).
- [11] G.A. Landrum. *YAEHMOP: Yet Another extended Hückel Molecular Orbital Package* (1997). YAEHMOP is freely available at URL: <http://overlap.chem.cornell.edu:8080/yaehmop.html>.
- [12] C. Roetti. *Lecture notes in chemistry*, chapitre Quantum-mechanical ab initio calculation of the properties of crystalline materials, page 125. Springer, Berlin (1996).
- [13] G. Kresse et J. Hafner. *Phys. Rev. B* **49**, 14251 (1994).
- [14] C. Pisani. *J. Mol. Struct. (Theochem)* **463**, 125 (1999).

- [15] L. Benco, T. Demuth, J. Hafner et F. Hutschka. *Chem. Phys. Lett.* **324**, 373 (2000).
- [16] L. Benco, T. Demuth, J. Hafner et F. Hutschka. *Chem. Phys. Lett.* **330**, 457 (2000).
- [17] R. Shah, J.D. Gale et M.C. Payne. *J. Phys. Chem.* **100**, 11688 (1996).
- [18] I. Stich, J.D. Gale, K. Terakura et M.C. Payne. *J. Am. Chem. Soc.* **121**, 3292 (1999).
- [19] J.D. Gale, R. Shah, M.C. Payne, I. Stich et K. Terakura. *Catal. Today* **50**, 525 (1999).
- [20] C.M. Zicovich-Wilson et A. Corma. *J. Phys. Chem. B* **104**, 4134 (2000).
- [21] T. Demuth, J. Hafner, L. Benco et H. Toulhoat. *J. Phys. Chem. B* **104**, 4593 (2000).
- [22] P. Ugliengo, B. Civalleri, C.M. Zicovich-Wilson et R. Dovesi. *Chem. Phys. Lett.* **318**, 247 (2000).
- [23] Y. Jeanvoine. *Modélisation ab initio des silicates, aluminosilicates et silicoaluminophosphates. Etude structurale et dynamique des sites acides de Bronsted dans HSSZ-13 et HSAPO-34*. Thèse de doctorat, Université Henri Poincaré, Nancy I (1997).
- [24] S.R. Blazzkowski, A.P.J. Jansen, M.A.C. Nascimento et R.A. van Santen. *J. Phys. Chem.* **98**, 12938 (1994).
- [25] S.R. Blazzkowski, M.A.C. Nascimento et R.A. van Santen. *J. Phys. Chem.* **100**, 3463 (1996).
- [26] P.M. Esteves, M.A.C. Nascimento et C.J.A. Mota. *J. Phys. Chem. B* **103**, 10417 (1999).
- [27] J.T. Fermann, C. Blanco et S. Auerbach. *J. Chem. Phys.* **112**, 6779 (2000).
- [28] J.T. Fermann et S. Auerbach. *J. Chem. Phys.* **112**, 6787 (2000).
- [29] J.A. Ryder, A.K. Chakraborty et A.T. Bell. *J. Phys. Chem. B* **104**, 6998 (2000).
- [30] H.V. Brand, L.A. Curtiss et L.E. Iton. *J. Phys. Chem.* **96**, 7725 (1992).
- [31] R. Harmsen, S. Bates et R.A. van Santen. *Faraday Discuss.* **106**, 443 (1997).
- [32] A. Khodakov, S.P. Bates, J. Dwyer, C.M. Windsor et N.A. Burton. *Phys. Chem. Chem. Phys.* **1**, 507 (1999).
- [33] E. Kassab, H. Jessri, M. Allavena et D. White. *J. Phys. Chem. A* **103**, 2766 (1999).
- [34] J. Sauer, U. Eichler, U. Meier, A. Schäfer, M. von Arnim et R. Ahlrichs. *Chem. Phys. Lett.* **308**, 147 (1999).
- [35] J.C.G. Pereira, C.R.A. Catlow et G.D. Price. *J. Phys. Chem. A* **103**, 3252 (1999).
- [36] J.C.G. Pereira, C.R.A. Catlow et G.D. Price. *J. Phys. Chem. A* **103**, 3268 (1999).
- [37] Y.-C. Lee, L.A. Curtiss, M.A. Ratner et D.F. Shriver. *Chem. Phys. Lett.* **321**, 463 (2000).
- [38] M.J. Rice, A.K. Chakraborty et A.T. Bell. *J. Phys. Chem. B* **104**, 9987 (2000).
- [39] S.A. Zygmunt, L.A. Curtiss, P. Zapol et L.E. Iton. *J. Phys. Chem. B* **104**, 1944 (2000).

- [40] M. Allavena, K. Seiti, E. Kassab, G. Ferenczy et J.G. Ángyán. *Chem. Phys. Lett.* **168**, 461 (1990).
- [41] G. Ferenczy, G.I. Czonka, G. Náray-szabó et J.G. Ángyán. *J. Comp. Chem.* **19**, 38 (1998).
- [42] S.P. Greatbanks, I.H. Hillier et P. Sherwood. *J. Comp. Chem.* **18**, 562 (1997).
- [43] P. Sherwood, A.H. de Vries, S.J. Collins, S.P. Greatbanks, N.A. Burton, M.A. Vincent et I.H. Hillier. *Faraday Discuss.* **106**, 79 (1997).
- [44] J.M. Vollmer, E.V. Stefanovich et T.N. Truong. *J. Phys. Chem. B* **103**, 9415 (1999).
- [45] J.M. Vollmer et T.N. Truong. *J. Phys. Chem. B* **104**, 6308 (2000).
- [46] U. Eichler, C.M. Kölmel et J. Sauer. *J. Comp. Chem.* **18**, 463 (1996).
- [47] U. Eichler, M. Brändle et J. Sauer. *J. Phys. Chem. B* **101**, 10035 (1997).
- [48] M. Brändle et J. Sauer. *J. Am. Chem. Soc.* **120**, 1556 (1998).
- [49] J. Sauer et M. Sierka. *J. Comp. Chem.* **26**, 1470 (2000).
- [50] M. Sierka et J. Sauer. *J. Phys. Chem. B* **105**, 1603 (2001).
- [51] A.H. de Vries, P. Sherwood, S.J. Collins, A.M. Rigby, M. Rigutto et G.J. Karner. *J. Phys. Chem. B* **103**, 6133 (1999).
- [52] M. Brändle, J. Sauer, R. Dovesi et N.M. Harrison. *J. Chem. Phys.* **109**, 10379 (1998).
- [53] P.E. Sinclair, A. de Vries, P. Sherwood, C.R. Catlow et R.A. van Santen. *J. Chem. Soc., Faraday Trans.* **94**, 3401 (1998).
- [54] J.-R. Hill, C.M. Freeman et B. Delley. *J. Phys. Chem. A* **103**, 3772 (1999).
- [55] M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, V.G. Zakrzewski, J.A. Montgomery, Jr., R.E. Stratmann, J.C. Burant, S. Dapprich, J.M. Millam, A.D. Daniels, K.N. Kudin, M.C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G.A. Petersson, P.Y. Ayala, Q. Cui, K. Morokuma, D.K. Malick, A.D. Rabuck, K. Raghavachari, J.B. Foresman, J. Cioslowski, J.V. Ortiz, A.G. Baboul, B.B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R.L. Martin, D.J. Fox, T. Keith, M.A. Al-Laham, C.Y. Peng, A. Nanayakkara, C. Gonzales, M. Challacombe, P.M.W. Gill, B. Johnson, W. Chen, M.W. Wong, J.L. Andres, C. Gonzales, M. Head-Gordon, E.S. Replogle et J.A. Pople. *Gaussian 98, Revision A.9* (1998). Gaussian, Inc., Pittsburgh PA.





## Chapitre 2

# Méthodologie et développements de la méthode Local Self Consistent Field *ab* *initio*

L'objectif de cette partie est d'abord de dresser un panorama succinct des différentes méthodes de chimie quantique utilisées durant cette étude. La majorité de celles-ci sont connues et développées depuis plus d'un demi-siècle et sont largement utilisées par la communauté des chimistes quanticiens. En particulier, nous exposerons les méthodes mono-déterminantales (Hartree-Fock et fonctionnelle de la densité) à travers les différentes approximations qui mènent à elles. Nous mentionnerons brièvement les méthodes post-Hartree-Fock que nous avons été amenés à utiliser.

Notre but est bien entendu de construire une méthode hybride à trois niveaux QM/MM/Madelung. La jonction entre les parties QM et MM est effectuée par l'intermédiaire d'orbitales de liaison strictement localisées. Ces orbitales, prédéfinies, ne doivent pas varier dans notre méthode. Pour ces raisons, nous avons développé une méthode quantique parfaitement adaptée à cette situation et nous l'appelons Local Self Consistent Field (LSCF).

Dans la deuxième partie, nous détaillerons les fondements théoriques de la méthode LSCF. Nous commencerons par donner la méthode de résolution itérative de ces équations en la comparant au processus standard SCF, puis nous verrons comment doivent être modifiées les équations de Roothaan afin de prendre en compte les particularités de la méthode LSCF, enfin nous obtiendrons une forme analytique des dérivées premières de l'énergie LSCF par rapport aux coordonnées nucléaires.

Finalement, nous montrerons qu'en tant que méthode quantique, le formalisme LSCF permet d'étu-

dier les états excités ou ionisés d'une molécule. Nous nous intéresserons alors particulièrement aux électrons de coeur.

## 2.1 Méthodes quantiques

Les états stationnaires d'un système moléculaire composé de  $N_a$  noyaux et  $n$  électrons sont déterminés en résolvant l'équation de Shrödinger indépendante du temps

$$\hat{H}\Psi = \mathbf{E}\Psi \quad (2.1)$$

où  $\Psi$  et  $\mathbf{E}$  sont respectivement les vecteurs et les valeurs propres de l'Hamiltonien  $\hat{H}$  du système. Ce dernier est un opérateur qui exprime toutes les interactions entre particules composant le système.

$$\hat{H} = \sum_{a=1}^{N_a} -\frac{1}{2M_a} \Delta_a + \sum_{i=1}^n -\frac{1}{2} \Delta_i - \sum_{a=1}^{N_a} \sum_{i=1}^n \frac{Z_a}{r_{ai}} + \sum_{a=1}^{N_a} \sum_{b>a}^{N_a} \frac{Z_a Z_b}{r_{ab}} + \sum_{i=1}^n \sum_{j>i}^n \frac{1}{r_{ij}} \quad (2.2)$$

Les deux premiers termes correspondent à l'énergie cinétique des noyaux et des électrons, le troisième terme à l'attraction noyaux-électrons et les deux derniers termes aux répulsions nucléaires et électroniques.

Dans le cadre de l'approximation de Born et Oppenheimer [1], le mouvement des noyaux est dé-couplé de celui des électrons. Il reste alors à étudier le comportement des électrons pour une géométrie donnée des noyaux. Celui-ci est décrit par une fonction d'onde électronique, état propre de l'Hamiltonien dans lequel le premier terme de (2.2) est négligé (par la suite appelé Hamiltonien électronique). Le potentiel électronique dans lequel se meuvent les noyaux est ainsi déterminé. Dans la suite, nous nous placerons toujours dans le cadre de cette approximation. Remarquons que l'avant-dernier terme de (2.2) est indépendant des coordonnées électroniques : il agit uniquement comme une correction à l'énergie électronique du système.

Les méthodes quantiques exposées ci-dessous sont relativement coûteuses en temps et en moyens informatiques. Elles ne peuvent être appliquées qu'à des systèmes moléculaires assez petits (d'une dizaine à une centaine d'atomes selon le niveau de calcul).

### 2.1.1 Expression de la fonction d'onde multiélectronique

Pour simplifier le problème à  $n$  corps exposé ci-dessus, deux autres approximations sont utilisées couramment. La première, appelée approximation orbitale, permet de décrire le comportement individuel

d'un électron  $i$  dans le champ des autres particules du système moléculaire. Notamment, il ressent le champ moyen des autres électrons. Il sera décrit par une fonction monoélectronique, la spinorbitale, dépendante de la position et du spin de l'électron  $\phi_i(r(i), s(i))$ .

L'approximation suivante consiste à négliger le couplage existant entre le mouvement de l'électron (moment cinétique orbitaire) et le moment cinétique de spin. Ainsi, une spinorbitale peut s'écrire comme le produit d'une fonction ne dépendant que de la position de l'électron  $\varphi_i(i)$  (orbitale atomique (OA) pour un atome, orbitale moléculaire (OM) dans le cas d'un système polyatomique) par une fonction de spin définie par

$$s(i) = \begin{cases} \alpha(i) & \text{pour un spin } \frac{1}{2} \\ \beta(i) & \text{pour un spin } -\frac{1}{2} \end{cases} \quad (2.3)$$

Ainsi,  $\phi_i(r(i), s(i)) = \phi_i(i) = \varphi_i(i) \cdot s(i)$ .

Nous cherchons maintenant une expression de la fonction d'onde électronique, compte tenu des approximations introduites ci-dessus. Les électrons sont de spin demi-entier et on les qualifie de fermions. Le principe d'exclusion de Pauli [2] qui en découle impose que lors de l'échange de deux électrons, la fonction d'onde électronique change de signe. Dans le cadre des approximations introduites au paragraphe précédent, la fonction d'onde électronique est exprimée habituellement sous la forme d'un déterminant de Slater [3] qui respecte cette antisymétrie (pour un jeu de spinorbitales orthonormées) :

$$\Psi = \frac{1}{\sqrt{n!}} \begin{vmatrix} \phi_1(1) & \phi_2(1) & \dots & \phi_n(1) \\ \phi_1(2) & \phi_2(2) & \dots & \phi_n(2) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_1(n) & \phi_2(n) & \dots & \phi_n(n) \end{vmatrix} \quad (2.4)$$

Pour simplifier l'écriture, nous utiliserons la notation  $\Psi = |\phi_1 \phi_2 \dots \phi_n|$ .

### 2.1.2 Équations de Hartree-Fock

L'Hamiltonien électronique s'écrit :

$$\hat{H} = \sum_{i=1}^n \left( -\frac{1}{2} \Delta_i + V_{Ni} \right) + \sum_{i=1}^n \sum_{j>i}^n \frac{1}{r_{ij}} = \sum_{i=1}^n \hat{H}^c(i) + \sum_{i=1}^n \sum_{j>i}^n \frac{1}{r_{ij}} \quad (2.5)$$

où  $\widehat{H}^c(i)$  représente l'énergie cinétique de l'électron  $i$  ainsi que l'attraction par l'ensemble des noyaux de la molécule. On l'appelle Hamiltonien de coeur et  $c'$  est un opérateur monoélectronique alors que la répulsion électronique s'exprime sous la forme d'un opérateur biélectronique. L'énergie du système représenté par le déterminant de Slater (2.4) est donnée par

$$E = \sum_i \langle \phi_i(1) | \widehat{H}^c(1) | \phi_i(1) \rangle + \sum_i \sum_{j>i} \left[ \langle \phi_i(1) \phi_j(2) | \frac{1}{r_{12}} | \phi_i(1) \phi_j(2) \rangle - \langle \phi_i(1) \phi_j(2) | \frac{1}{r_{12}} | \phi_i(2) \phi_j(1) \rangle \right] \quad (2.6)$$

Notre but est d'obtenir la meilleure énergie possible. Le théorème variationnel nous apprend que cette énergie doit correspondre à un extremum, c'est-à-dire qu'une petite variation d'une spinorbitale ne doit pas modifier sa valeur. C'est à partir de cette considération que Hartree et Fock ont mis au point leurs célèbres équations qui permettent la détermination des spinorbitales [4, 5] :

$$\widehat{F}\phi_i = \varepsilon_i \phi_i \quad (2.7)$$

où  $\widehat{F}$  est un opérateur monoélectronique appelé opérateur de Hartree-Fock qui se décompose en

$$\widehat{F}(1) = \widehat{H}^c(1) + \sum_i \left[ \widehat{J}_i(1) - \widehat{K}_i(1) \right] \quad (2.8)$$

Le premier terme est l'Hamiltonien de coeur, le deuxième un opérateur Coulombien et le troisième un opérateur d'échange, ces deux derniers étant définis par

$$\begin{aligned} \widehat{J}_i(1) \phi_j(1) &= \phi_j(1) \int_{(2)} \phi_i^*(2) \frac{1}{r_{12}} \phi_i(2) d\tau_2 \\ \widehat{K}_i(1) \phi_j(1) &= \phi_i(1) \int_{(2)} \phi_i^*(2) \frac{1}{r_{12}} \phi_j(2) d\tau_2 \end{aligned} \quad (2.9)$$

l'intégration se fait sur les coordonnées de l'électron 2.

La résolution de ces équations se fait par la méthode du champ auto-cohérent (Self Consistent Field ou SCF en anglais) qui est une méthode itérative rendue nécessaire lorsque l'on remarque que la définition de l'opérateur de Fock nécessite la connaissance préalable des orbitales moléculaires qu'il est sensé fournir. La procédure de résolution SCF sera détaillée plus loin.

### 2.1.3 Équations de Roothaan

Lorsque les orbitales moléculaires sont développées sur une base d'orbitales atomiques (OA) – on parle alors d'approximation LCAO (pour Linear Combination of Atomic Orbitals) – les équations de

Hartree-Fock peuvent également être exprimées dans cette base. Pour un système à couches fermées, c'est à Roothaan [6] que l'on doit cette expression. On considère alors que deux électrons de spins opposés peuvent être décrits par la même fonction d'espace, appelée orbitale moléculaire (OM). On parle alors de formalisme restreint ou RHF.

Considérons un jeu de  $N$  fonctions de base  $\{\dots \chi_\mu \dots \chi_\nu \dots\}$ , les OM se développent en combinaisons linéaires de ces fonctions dont les coefficients sont notés  $c_{\mu i}$  :

$$\varphi_i = \sum_{\mu=1}^N c_{\mu i} \chi_\mu \quad (2.10)$$

Ré-exprimant l'énergie électronique (2.6) dans cette base (en remarquant que les différents opérateurs n'agissent que sur la partie d'espace, on trouve que l'intégration sur le spin donne 1), la différenciant par rapport à un coefficient  $c_{\mu i}$ , on peut définir l'expression d'un élément de l'opérateur de Hartree-Fock dans la base des OA :

$$F_{\mu\nu} = \langle \chi_\mu | \hat{F} | \chi_\nu \rangle = H_{\mu\nu}^c + \sum_{\lambda\sigma} P_{\lambda\sigma} \left[ (\mu\nu | \lambda\sigma) - \frac{1}{2} (\mu\sigma | \lambda\nu) \right] \quad (2.11)$$

où  $H_{\mu\nu}^c = \langle \chi_\mu | \hat{H}^c | \chi_\nu \rangle$ ,  $(\mu\nu | \lambda\sigma) = \int \int \chi_\mu(1) \chi_\nu(1) \frac{1}{r_{12}} \chi_\lambda(2) \chi_\sigma(2) d\tau_1 d\tau_2$  et  $P_{\lambda\sigma} = 2 \sum_i c_{\lambda i} c_{\sigma i}$ .

Il est plus simple d'écrire cette expression sous forme matricielle :

$$\mathbf{F} = \mathbf{H} + \mathbf{G} \quad (2.12)$$

avec  $\mathbf{F}$  la matrice de Fock,  $\mathbf{H}$  la matrice Hamiltonien de coeur et  $\mathbf{G}$  la matrice résultant de la contraction de la matrice densité  $\mathbf{P}$  avec les intégrales biélectroniques.

Nous ne verrons pas ici en détail la dérivation des équations de Roothaan [6], car elles seront reprises plus loin avec la méthode LSCF. Les équations de Roothaan sont données en introduisant des multiplicateurs de Lagrange  $\varepsilon_i$  par :

$$\sum_{\mu} F_{\nu\mu} c_{\mu i} = \sum_{\mu} S_{\nu\mu} c_{\mu i} \varepsilon_i \quad (2.13)$$

ou en notation matricielle

$$\mathbf{FC} = \mathbf{SCE} \quad (2.14)$$

La résolution de ces équations n'est pas facile à cause de la non-orthogonalité de la base qui s'explique par la présence de la matrice de recouvrement  $\mathbf{S}$  de cette base. En passant dans une base orthogonale, l'expression 2.14 se réduit alors à  $\mathbf{F}'\mathbf{C}' = \mathbf{C}'\mathbf{E}$ , problème aux valeurs propres, soluble très facilement. Ceci est rendu possible en utilisant une transformation unitaire  $\mathbf{X}$  telle que  $\mathbf{X}^\dagger\mathbf{S}\mathbf{X} = \mathbf{1}$ . Habituellement, l'orthogonalisation de Löwdin est utilisée :  $\mathbf{X} = \mathbf{S}^{-\frac{1}{2}}$ .

Pour les systèmes à couches ouvertes, Berthier [7] et Pople-Nesbet [8] ont développé simultanément un formalisme qui permet de décrire avec des fonctions d'espace différentes les électrons de spin opposés. On parle alors de formalisme non restreint ou UHF. Cependant son utilisation pose un problème : la fonction d'onde électronique n'est plus fonction propre de l'opérateur  $\widehat{S}^2$  (il y a contamination de spin). Schlegel [9] a proposé une méthode basée sur un projecteur afin d'éliminer le premier contaminant :

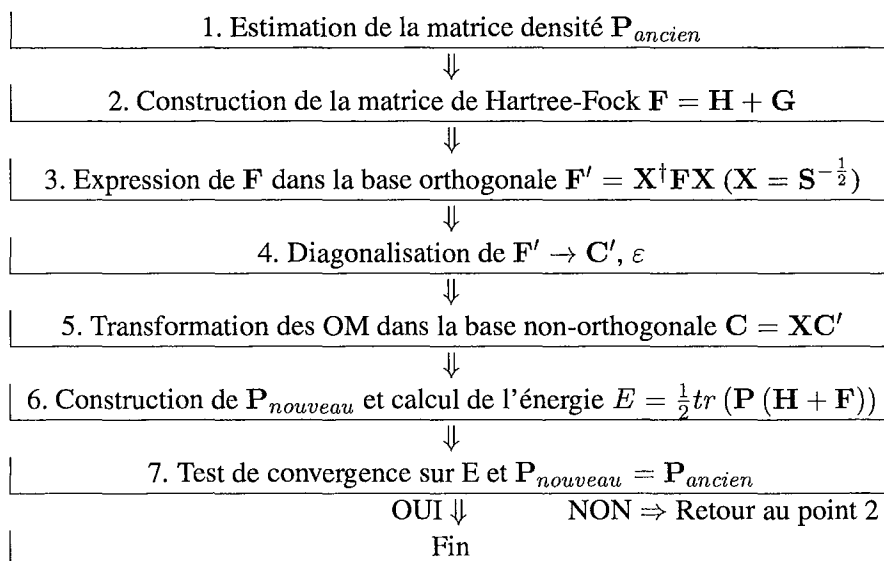
$$\widehat{A} = \frac{\widehat{S}^2 - (s+1)(s+2)}{\langle \Psi | \widehat{S}^2 | \Psi \rangle - (s+1)(s+2)} \quad (2.15)$$

avec  $s$  la valeur du spin recherchée de la fonction d'onde  $\Psi$ .

#### 2.1.4 La méthode de résolution auto-cohérente SCF

Comme il a été noté ci-dessus, la définition de l'opérateur de Fock nécessite de connaître au préalable les OM, qui sont pourtant les vecteurs propres de ce même opérateur. Pour résoudre ce dilemme, on utilise une méthode de résolution itérative, en partant d'un jeu d'OM approché. C'est la méthode SCF, adaptée aussi bien au formalisme RHF que UHF. De plus, la construction de l'expression de  $F_{\mu\nu}$  requiert le calcul d'un certain nombre d'intégrales mono- et bi-électroniques. Il existe un grand nombre de méthodes qui calculent efficacement ces intégrales et qui se sont améliorées au fil des années. La majorité de ces méthodes reposent sur l'utilisation de formules de récurrence qui permettent d'exprimer toutes les intégrales en fonction d'un petit jeu d'autres intégrales primaires. Les algorithmes en sont actuellement à un tel stade de développement que la détermination des intégrales n'est plus le processus limitant des programmes de chimie quantique. Nous citerons deux méthodes qui nous ont particulièrement intéressés : les polynômes de Rys [10] et la méthode Prism [11]. Rappelons également que les fonctions gaussiennes (plus précisément les contractions de gaussiennes) sont d'utilisation courante aujourd'hui comme fonctions de base. Le schéma algorithmique de la résolution SCF est donné dans la figure 2.1.

Le passage en base orthogonale de la matrice de Fock (pas 3) s'obtient généralement en utilisant la matrice  $\mathbf{X} = \mathbf{S}^{-\frac{1}{2}}$ . Souvent, les critères de convergence de la fonction d'onde (pas 7) portent à la fois sur l'énergie électronique et la matrice densité.

FIG. 2.1 – *Algorithme de résolution SCF des équations de Roothaan.*

Pour conclure sur la méthode Hartree-Fock, la différence entre l'énergie exacte du système et la limite de l'énergie Hartree-Fock (pour une base donnée) est appelée énergie de corrélation. Plusieurs méthodes, dites « post Hartree-Fock » permettent de prendre en compte une partie de la corrélation : soit en dépassant la description monodéterminantale de la fonction d'onde, soit en utilisant la méthode des perturbations (l'une n'excluant pas l'autre). Ces méthodes seront présentées succinctement dans la dernière partie de cette étude.

Une autre méthode de prise en compte de la corrélation électronique existe, tout en restant au même niveau de description monodéterminantale de la fonction d'onde. Elle est basée sur la théorie de la fonctionnelle de la densité et est exposée dans la section suivante.

### 2.1.5 Méthodes de la fonctionnelle de la densité

Durant cette thèse, nous avons également utilisé la théorie de la fonctionnelle de la densité (DFT). Elle a été introduite par Hohenberg et Kohn [12] qui ont établi le théorème suivant : la densité électronique  $\rho(r)$  d'un système dans son état fondamental détermine complètement (à une constante près) le potentiel externe  $v(r)$  (dans un système moléculaire isolé, le potentiel nucléaire). Autrement dit, l'Hamiltonien électronique ainsi que toutes les propriétés qui en découlent sont déterminés par la connaissance de cette densité électronique. On peut alors définir la fonctionnelle suivante :

$$E_v[\rho(r)] = \int v(r)\rho(r)dr + \mathbf{F}[\rho(r)] \quad (2.16)$$

La dépendance de  $\mathbf{F}$  en  $\rho(r)$  n'est pas connue et il faut donc recourir à des approximations. Kohn et



Sham [13] ont alors étendu à la densité électronique le théorème variationnel : la densité exacte correspond au minimum de l'énergie. Celle-ci est décomposée en différents termes : énergie cinétique d'une densité  $\rho(r)$  d'électrons sans interactions, interaction coulombienne des électrons entre eux et avec les noyaux et énergie d'échange-corrélation. L'Hamiltonien monoélectronique correspondant est :

$$\hat{H}_{KS} = -\frac{1}{2}\Delta + v(r) + V_{ee}[\rho(r)] + V_{xc}[\rho(r)] \quad (2.17)$$

Les équations qui en résultent sont semblables aux équations de Hartree-Fock :

$$\hat{H}_{KS}\varphi_i = \varepsilon_i\varphi_i \quad (2.18)$$

en exprimant  $\rho(r) = \sum_i \varphi_i^*(r) \cdot \varphi_i(r)$ . Les pseudo orbitales moléculaires (encore appelées orbitales de Kohn-Sham) peuvent être calculées numériquement ou développées sur une base d'orbitales atomiques comme nous l'avons vu avec les équations de Roothaan.

On ne connaît pas la forme exacte du potentiel d'échange-corrélation  $V_{xc}[\rho(r)]$ . Plusieurs formes approchées existent :

- Les fonctionnelles locales ne sont caractérisées que par la densité électronique locale  $\rho(r)$ . Cette approximation ne prend pas en compte les hétérogénéités de la distribution électronique et n'est donc valable que dans les cas où celle-ci varie peu.
- Les fonctionnelles à correction de gradient qui corrigent le défaut des fonctionnelles locales en prenant en compte le gradient de la densité électronique.
- Les fonctionnelles hybrides traitent la partie échange de la fonctionnelle comme un mélange d'échange Hartree-Fock et d'échange DFT (la partie corrélation restant purement DFT). Nous citerons par exemple les fonctionnelles B3LYP, B1B96, MPW91PW91 ou G961LYP. Elles sont actuellement considérées comme les plus performantes parmi les plus utilisées.

La DFT permet de prendre en compte une partie de la corrélation électronique à un coût informatique raisonnable, ce qui permet d'aborder des systèmes de taille relativement importante par rapport aux autres méthodes que nous allons maintenant décrire.

### 2.1.6 Les méthodes post Hartree-Fock

Ces méthodes ont pour caractéristique de prendre comme point de départ la fonction d'onde Hartree-Fock. Elles se sont beaucoup développées ces dernières années et permettent d'atteindre une précision

très impressionnante. Nous nous limiterons cependant à décrire deux types de méthodes que nous avons utilisés durant ces travaux. Il s'agit des interactions de configurations et de la méthode Møller-Plesset.

**Interaction de configurations (CI).** La fonction d'onde multiélectronique du système est construite comme une combinaison linéaire de déterminants orthogonaux, obtenus à partir du déterminant de Slater solution des équations de Hartree-Fock  $\Psi_0$  :

$$\Psi = C_0 \Psi_0 + \sum_i \sum_a^{occ \text{ virt}} C_i^a \Psi_i^a + \sum_{i < j} \sum_{a < b}^{occ \text{ virt}} C_{ij}^{ab} \Psi_{ij}^{ab} + \dots \quad (2.19)$$

Les déterminants sont obtenus en remplaçant dans  $\Psi_0$  une ou plusieurs spinorbitales occupées  $\phi_i, \phi_j, \dots$  par une ou plusieurs spinorbitales virtuelles ( $\phi_a, \phi_b, \dots$ ). On appelle ces déterminants des configurations mono, di, ... excitées. Les coefficients de l'équation 2.19 sont déterminés en diagonalisant l'Hamiltonien électronique (équation 2.5) exprimé sur la base des configurations. Les solutions obtenues sont des solutions approchées (au sens variationnel) de l'état fondamental (pour la valeur propre la plus basse) et de différents états excités.

Plus la base d'orbitales atomiques est étendue, plus le nombre d'excitations possible est élevé. La réalisation de toutes ces excitations conduit à un calcul Full CI qui, si la base est complète, est alors exact. Dans la pratique, ces calculs ne sont réalisables que pour de tous petits systèmes en base très limitée. On préfère souvent tronquer l'espace des configurations en se limitant aux excitations doubles (CID), simples et doubles (CISD), etc ...

**Møller-Plesset (MP).** Ces méthodes constituent un cas particulier de la méthode générale des perturbations. Dans notre cas, la corrélation électronique est considérée comme représentant une perturbation de l'énergie Hartree-Fock. Dans l'approche Møller-Plesset, l'Hamiltonien électronique est alors décomposé en deux termes :

$$\hat{H} = \sum_i^n \hat{F}_i + \hat{V} \quad (2.20)$$

où  $\hat{F}_i$  est l'opérateur de Hartree-Fock et  $\hat{V}$  est l'opérateur perturbatif. On trouve l'énergie Hartree-Fock comme égale à l'énergie non-perturbée plus la correction au premier ordre. Dans la pratique, les méthodes Møller-Plesset à l'ordre 2 (MP2) ou à l'ordre 4 (MP4) sont les plus couramment employées.

Cette approche n'est cependant valable que si la fonction d'onde d'ordre 0 est déjà une bonne approximation de la fonction d'onde exacte.

Nous concluons ce paragraphe en notant que ces méthodes post-Hartree-Fock sont beaucoup plus avides de ressources et de temps informatiques. Pour le cas particulier des zéolithes, où les précédentes études montrent que la corrélation est importante, ces méthodes deviennent vite prohibitives. Pour cette raison, nous leur préférerons une approche du type fonctionnelle de la densité.

## 2.2 La méthode Local Self Consistent Field

La méthode LSCF [14, 15, 16] a pour objectif d'optimiser la fonction d'onde électronique, soumise à la contrainte suivante : certaines spinorbitales prédéfinies sont gelées et le restent durant tout le calcul SCF. Ces spinorbitales sont fixées par l'utilisateur et représentent une certaine densité électronique immuable que ressentent les autres électrons. Il faut noter que d'autres méthodes ont été développées dans le même but :

- Fink exprime la matrice de Fock dans une base qui permet de ne diagonaliser qu'une petite partie de celle-ci. C'est de cette approche que nous sommes le plus proche [17, 18].
- Payne propose une méthode Hartree-Fock locale dans laquelle chaque orbitale est exprimée sur une base de fonctions localisées dans une région particulière du système étudié [19].
- Dans le même esprit que Fink, Bagus exprime la matrice de Fock dans une base qui permet ensuite d'annuler certains éléments hors-diagonaux [20].
- L'article de revue d'Ángyán donne un panorama très complet de cette discipline [21].

Nous montrerons d'abord comment la méthode de résolution SCF des équations de Hartree-Fock doit être modifiée pour prendre en compte cette contrainte. Puis nous déterminerons les équations de Roothaan modifiées. Enfin, les dérivées premières de l'énergie par rapport aux coordonnées nucléaires seront évaluées afin de permettre l'optimisation de la géométrie moléculaire.

### 2.2.1 Optimisation de la fonction d'onde

Dans l'approximation LCAO, les seules données nécessaires pour résoudre les équations de Hartree-Fock (ou de Kohn-Sham) sont les coefficients des orbitales gelées (OG) dans la base utilisée, et leur nombre d'occupation. Ces OG peuvent être de n'importe quelle forme, occupées ou non (nous verrons dans la dernière partie de ce chapitre l'intérêt qu'il y a à geler des spinorbitales vides). Comme il s'agit de spinorbitales, nous pouvons nous placer dans le formalisme restreint ou non-restreint des équations de Hartree-Fock. Pour plus de simplicité, nous utiliserons dans la suite le formalisme restreint : le nombre d'électrons est pair et il y a autant d'électrons de spin  $\alpha$  que de spin  $\beta$ .

Nous devons déterminer les orbitales moléculaires (OM), optimisées variationnellement, en fonction des OG. Afin de réduire le nombre d'intégrales à calculer, nous requérons que les OM soient orthogonales entre elles et également aux OG. La façon la plus directe de satisfaire cette condition consiste à exprimer les OM sur une base de fonctions déjà orthogonales aux OG. Pour cela, nous définissons une procédure en deux points.

Considérons la base de  $N$  fonctions non orthogonales  $\{\dots\chi_\mu\dots\chi_\nu\dots\}$ . Pour plus de commodité, la lettre  $\chi$  ne sera plus notée dans la suite. Soient  $L$  orbitales gelées  $|l_i\rangle$  qui s'expriment dans cette base par :

$$|l_i\rangle = \sum_{\mu} a_{\mu i} |\mu\rangle \quad (2.21)$$

et que l'on suppose mutuellement orthogonales (dans le cas contraire, elles sont orthogonalisées par la méthode de Löwdin [22]). Les  $N$  fonctions de base sont orthogonalisées aux OG en leur soustrayant leurs projections sur les OG. Soit  $|\tilde{\mu}\rangle$  une telle fonction :

$$|\tilde{\mu}\rangle = \sum_{\nu} \left[ 1 - \sum_{i=1}^L S_{\nu i}^2 \right]^{-\frac{1}{2}} \cdot \left[ \delta_{\mu\nu} - \sum_{i=1}^L \sum_{\eta} a_{\nu i} a_{\eta i} S_{\eta\mu} \right] |\nu\rangle \quad (2.22)$$

où  $S_{\nu i}$  exprime l'intégrale de recouvrement  $\langle\nu|l_i\rangle$ ,  $S_{\eta\mu}$  exprime  $\langle\eta|\mu\rangle$  et  $\delta_{\mu\nu}$  est le symbole de Kronecker.

Cette transformation s'exprime simplement par :

$$|\tilde{\mu}\rangle = \sum_{\nu} m_{\nu\tilde{\mu}} |\nu\rangle \quad (2.23)$$

ainsi, chaque colonne de la matrice  $\mathbf{M}$  correspond à une fonction de la nouvelle base  $\{|\tilde{\mu}\rangle\}$ .

Cependant, les  $L$  OG étant déjà exprimées sur la base des  $N$  OA, il existe alors  $L$  dépendances linéaires dans la base  $\{|\tilde{\mu}\rangle\}$ . Afin d'orthogonaliser entre elles les fonctions déjà orthogonales aux OG et d'éliminer ces dépendances linéaires, nous utilisons l'orthogonalisation canonique [22] qui s'écrit :

$$\mathbf{X} = \mathbf{U}\mathbf{s}^{-\frac{1}{2}} \quad (2.24)$$

où  $\mathbf{U}$  est la matrice unitaire des vecteurs propres du recouvrement des fonctions  $|\tilde{\mu}\rangle$  et  $\mathbf{s}$  la matrice des valeurs propres correspondantes. La matrice  $\mathbf{X}$  de dimension  $N \times (N - L)$  s'obtient en éliminant dans  $\mathbf{U}$  les  $L$  vecteurs propres correspondant aux valeurs propres nulles. Notons que cette transformation présente l'avantage d'orthogonaliser la base.

Les deux transformations précédentes peuvent être exprimées en une seule et permettent donc de déduire l'expression finale de la base orthogonale en fonction de la base originale :

$$\mathbf{B} = \mathbf{M} \cdot \mathbf{X} \quad (2.25)$$

où  $\mathbf{B}$  est une matrice de dimensions  $N \times (N - L)$ . Cette matrice est l'équivalente de la matrice  $\mathbf{S}^{-\frac{1}{2}}$  utilisée dans la procédure SCF habituelle (figure 2.1). Le remplacement de cette dernière par  $\mathbf{B}$  est la différence principale entre SCF et LSCF.

Une autre différence consécutive à l'utilisation des OG est la partition de la matrice densité qui sert à exprimer la matrice de Hartree-Fock (équation 2.11) :

$$F_{\mu\nu} = \langle \chi_\mu | \hat{F} | \chi_\nu \rangle = H_{\mu\nu}^c + \sum_{\lambda\sigma} P_{\lambda\sigma}^T \left[ (\mu\nu|\lambda\sigma) - \frac{1}{2} (\mu\sigma|\lambda\nu) \right] \quad (2.26)$$

avec  $\mathbf{P}^T = \mathbf{P}^Q + \mathbf{P}^G$  où  $\mathbf{P}^Q$  est issue des OM et  $\mathbf{P}^G$  provient des OG.

Grâce aux éléments exposés ci-dessus, il est maintenant possible d'effectuer des calculs LSCF. L'ensemble de programmes Gaussian98 [23], plus particulièrement le link 502, a été modifié en conséquence. Il faut noter que les dimensions de la matrice de Fock exprimée dans la base orthogonale sont réduites dans un calcul LSCF :  $\mathbf{F}' = \mathbf{B}^\dagger \mathbf{F} \mathbf{B}$  est de dimension  $(N - L) \times (N - L)$  au lieu de  $N \times N$ . La diagonalisation de cette matrice constituant un pas limitant, il est clair que la résolution LSCF nécessite moins de temps que la résolution SCF.

L'extension de LSCF au formalisme non-restreint est simple : deux matrices densité dues aux OG  $\alpha$  et  $\beta$  sont prises en compte et on construit deux matrices d'orthogonalisation  $\mathbf{B}^\alpha$  et  $\mathbf{B}^\beta$ . L'application de la méthode LSCF aux équations de Kohn-Sham est directe et ne nécessite aucune autre modification (en prenant garde cependant de prendre en compte la densité électronique due aux OG). De plus, l'utilisation de méthodes post-Hartree-Fock basées sur le déterminant de Slater LSCF est possible. Dans le cadre des méthodes de type Møller-Plesset, les OG ne doivent pas apparaître dans l'espace actif puisqu'elles ne sont pas fonctions propres de l'opérateur de Hartree-Fock.

Enfin, la méthode LSCF est exacte strictement (comparée à la méthode SCF) si on gèle des orbitales précédemment optimisées sur le même système ou des transformations unitaires de celles-ci (en prenant garde de ne pas mélanger les blocs d'orbitales occupées-occupées et virtuelles-virtuelles). Il est intéressant de noter qu'un jeu d'orbitales gelées définit une hypersurface d'énergie potentielle. Par exemple, considérons la déformation angulaire de la molécule d'eau en conservant la symétrie  $C_{2v}$  au

niveau HF/STO-3G. Pour la géométrie d'équilibre, l'OM occupée  $1b_1$  a pour expression dans la base des orbitales atomiques  $a.(O) + b.(H1 - H2)$  où  $(O)$  est l'orbitale  $2p_y$  de l'oxygène et  $(H1)$ ,  $(H2)$  sont les orbitales  $1s$  des atomes d'hydrogène, avec  $a$  et  $b$  valant respectivement 0,613 et 0,449). Quatre calculs sont effectués. Le premier correspond au calcul standard SCF où l'orbitale  $1b_1$  est ré-optimisée pour chaque valeur de l'angle  $H1-O-H2$ . Les trois autres sont des calculs LSCF où l'orbitale  $1b_1$  est gelée soit avec les coefficients correspondants à la géométrie d'équilibre, soit avec le coefficient  $a$  multiplié par 1,5 ou 0,5 (l'orbitale obtenue est bien sûr renormée). Les courbes d'énergie correspondant à un angle  $H1-O-H2$  variant entre  $60^\circ$  et  $180^\circ$  sont données dans la figure 2.2 et sont comparées à la courbe SCF<sup>1</sup>.

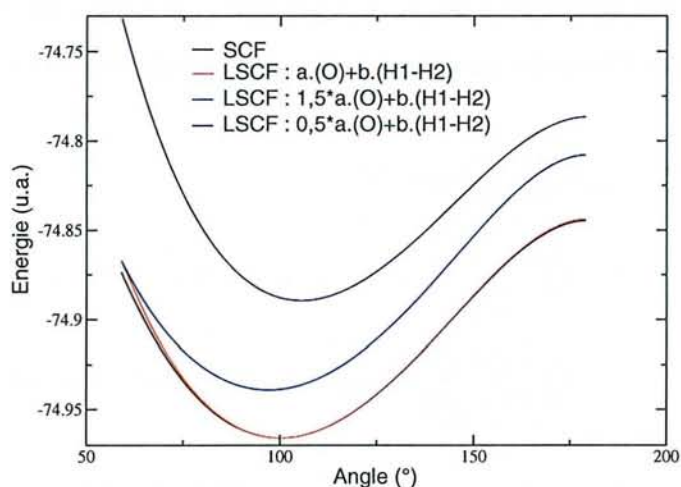


FIG. 2.2 – Variation de l'énergie en fonction de l'angle  $H-O-H$  de l'eau pour différentes OG de symétrie  $b_1$ .

Nous vérifions bien l'étroite dépendance de l'énergie LSCF envers la définition d'une orbitale gelée. Si l'orbitale gelée est l'OM optimisée pour l'angle d'équilibre (courbe LSCF), l'accord est bon : pour chaque angle, il suffit de renormer cette OG pour bien représenter l'OM optimisée correspondante. En revanche, si on modifie les poids relatifs des orbitales atomiques dans l'OG, les courbes sont très différentes. Les OG doivent donc être considérées comme des « paramètres » de la surface d'énergie potentielle au même titre que les positions des noyaux atomiques.

1. Pour la valeur de  $180^\circ$  de l'angle  $H1-O-H2$ , le groupe de symétrie est bien sûr  $D_{\infty h}$  et la symétrie de l'orbitale est  $\sigma_u^+$ .

### 2.2.2 Modification des équations de Roothaan

Nous dériverons ces équations pour un système à couches fermées à  $2n$  électrons. Dans une base de  $N$  fonctions réelles  $\{\dots \mu \dots \nu \dots\}$ , l'énergie électronique a la forme :

$$E = 2 \sum_i^n \sum_{\mu\nu}^N d_{\mu i} d_{\nu i} H_{\mu\nu} + \sum_{i,j}^n \sum_{\mu\nu\lambda\sigma}^N d_{\mu i} d_{\nu i} d_{\lambda j} d_{\sigma j} [2(\mu\nu|\lambda\sigma) - (\mu\sigma|\lambda\nu)] \quad (2.27)$$

où  $d_{\mu i}$  représente le coefficient de la fonction de base  $\mu$  dans l'orbitale  $i$  (gelée ou non). On les distinguera en notant  $a_{\mu l}$  ( $c_{\mu i}$ ) les coefficients des OA dans les OG (OM) occupées. Pour obtenir les équations de Roothaan modifiées, cette énergie est différenciée par rapport à ces coefficients. Au même titre que les positions nucléaires, les OG sont des paramètres de l'énergie. Il s'ensuit que  $\delta a_{\mu l} = 0, \forall \mu, l$  et :

$$\begin{aligned} \delta E &= 4 \sum_i^{n-L} \sum_{\mu\nu}^N \delta c_{\mu i} c_{\nu i} H_{\mu\nu} \\ &+ 4 \sum_{i,j}^{n-L} \sum_{\mu\nu\lambda\sigma}^N \delta c_{\mu i} c_{\nu i} c_{\lambda j} c_{\sigma j} [2(\mu\nu|\lambda\sigma) - (\mu\sigma|\lambda\nu)] \\ &+ 4 \sum_i^{n-L} \sum_l^L \sum_{\mu\nu\lambda\sigma}^N \delta c_{\mu i} c_{\nu i} a_{\lambda l} a_{\sigma l} [2(\mu\nu|\lambda\sigma) - (\mu\sigma|\lambda\nu)] \end{aligned} \quad (2.28)$$

En introduisant l'élément de matrice densité totale  $P_{\lambda\sigma}^T = \sum_j^n 2d_{\lambda j} d_{\sigma j}$ , l'équation 2.28 devient :

$$\delta E = 4 \sum_i^{n-L} \sum_{\mu} \delta c_{\mu i} \left( \sum_{\nu} c_{\nu i} H_{\mu\nu} + \sum_{\nu} c_{\nu i} \sum_{\lambda\sigma} P_{\lambda\sigma}^T \left[ (\mu\nu|\lambda\sigma) - \frac{1}{2} (\mu\sigma|\lambda\nu) \right] \right) \quad (2.29)$$

Nous pouvons maintenant utiliser la définition de la matrice de Hartree-Fock, dont un élément dans la base est donné par l'équation 2.26. Nous obtenons alors un jeu de  $(n - L) \times N$  équations :

$$\delta E = 4 \sum_i^{n-L} \sum_{\mu} \delta c_{\mu i} \sum_{\nu} c_{\nu i} F_{\mu\nu} \quad (2.30)$$

Les OM doivent être orthonormées :

$$\sum_{\mu\nu} c_{\mu i} S_{\mu\nu} c_{\nu j} = \delta_{ij} \quad (2.31)$$

et être aussi orthogonales à toutes les OG (occupées ou non) :

$$\sum_{\mu\nu} c_{\mu i} S_{\mu\nu} a_{\nu l} = 0 \quad (2.32)$$

Différenciant ces contraintes par rapport à  $c_{\mu i}$ , nous pouvons les introduire dans l'équation 2.30 grâce à la méthode des multiplicateurs de Lagrange :

$$\sum_i^{n-L} \sum_{\mu} \delta c_{\mu i} \left[ \sum_{\nu} c_{\nu i} F_{\mu\nu} - \sum_j^{n-L} \sum_{\zeta} \epsilon_{ij} c_{\zeta j} S_{\mu\zeta} - \sum_k^L \sum_{\sigma} \lambda_{ik} a_{\sigma k} S_{\mu\sigma} \right] \quad (2.33)$$

avec  $\frac{(n-L) \times (n-L+1)}{2}$  multiplicateurs  $\epsilon_{ij}$  correspondant aux équations 2.31 et  $(n-L) \times L$  multiplicateurs  $\lambda_{ik}$  correspondant aux équations 2.32. Le théorème variationnel indique que la valeur optimale des coefficients  $c_{\mu i}$  est obtenue pour  $\delta E = 0$ . On impose alors aux multiplicateurs d'annuler l'expression entre crochets dans 2.33 :

$$\sum_{\nu} c_{\nu i} \left[ F_{\mu\nu} - \sum_j^{n-L} S_{\mu\nu} \epsilon_{ij} \right] - \sum_{\nu} \sum_k^L S_{\mu\nu} a_{\nu k} \lambda_{ki} = 0 \quad (2.34)$$

Il y a donc un système homogène de  $(n-L) \times N$  équations à résoudre dont les inconnues sont les  $\frac{(n-L) \times (n-L+1)}{2} + (n-L) \times L$  multiplicateurs de Lagrange et les  $(n-L) \times N$  coefficients des orbitales moléculaires liés par  $n-L$  conditions de normations. On peut donc fixer arbitrairement  $\frac{(n-L) \times (n-L-1)}{2} + (n-L) \times L$  inconnues. Le premier terme correspond au nombre de multiplicateurs  $\epsilon_{ij}$  tels que  $i \neq j$ . Le deuxième correspond au bloc OM – OG, c'est-à-dire à tous les multiplicateurs de Lagrange qui expriment l'orthogonalité entre OM et OG. Nous n'utilisons pas la possibilité de mettre ces multiplicateurs de Lagrange à zéro afin de conserver la généralité des expressions. On pose donc  $\epsilon_{ij} = e_i \delta_{ij}$  et les équations 2.34 s'écrivent :

$$\sum_{\nu} c_{\nu i} [F_{\mu\nu} - S_{\mu\nu} \epsilon_i] = \sum_{\nu} \sum_k^L S_{\mu\nu} a_{\nu k} \lambda_{ki} \quad (2.35)$$

Ce sont les équations de Roothaan modifiées que nous résolvons à l'aide de la procédure LSCF exposée au paragraphe précédent. Les multiplicateurs de Lagrange sont déterminés par :

$$\epsilon_i = \sum_{\mu\nu} c_{\mu i}^* F_{\mu\nu} c_{\nu i} \quad (2.36)$$



et

$$\lambda_{ki} = \sum_{\mu\nu} a_{\mu l}^* F_{\mu\nu} c_{\nu i} \quad (2.37)$$

Les équations de Roothaan s'écrivent naturellement sous forme matricielle :

$$\mathbf{FC} - \mathbf{SCE} = \mathbf{SL}\Lambda \quad (2.38)$$

avec  $\mathbf{L}$  la matrice des OG exprimée dans la base originale.

Le passage dans la base orthogonale se fait grâce à la matrice  $\mathbf{B}$  (équation 2.25). En posant  $\mathbf{C} = \mathbf{BC}'$ , l'équation 2.38 devient :

$$\mathbf{FBC}' - \mathbf{SBC}'\mathbf{E} = \mathbf{SL}\Lambda \quad (2.39)$$

Multipliant à gauche par  $\mathbf{B}^\dagger$  et en posant  $\mathbf{F}' = \mathbf{B}^\dagger\mathbf{FB}$ , on obtient :

$$\mathbf{F}'\mathbf{C}' - \mathbf{C}'\mathbf{E} = \mathbf{B}^\dagger\mathbf{SL}\Lambda \quad (2.40)$$

Les OM et les OG sont orthogonales (équation 2.32), donc  $\mathbf{C}'^\dagger\mathbf{SL} = \mathbf{C}'^\dagger\mathbf{B}^\dagger\mathbf{SL} = 0$ . Donc nécessairement  $\mathbf{B}^\dagger\mathbf{SL} = 0$  et l'équation 2.40 se réduit à l'expression habituelle.

### 2.2.3 Gradients de l'énergie LSCF

Le calcul des dérivées premières de l'énergie électronique LSCF par rapport aux coordonnées nucléaires est nécessaire à l'optimisation de la géométrie du système étudié. Les forces agissant sur les noyaux sont égales à  $\vec{F} = -\overrightarrow{grad}E$ , ou encore,  $F_a$  étant la force agissant sur l'atome  $a$  à la position  $r_a$  :

$$F_a = -\frac{\partial E}{\partial r_a} \quad (2.41)$$

Pour déterminer l'expression à droite de l'expression 2.41, exprimons l'énergie électronique LSCF sous la forme suivante :

$$E = \frac{1}{2} \sum_{\mu\nu} P_{\nu\mu}^T (H_{\mu\nu} + F_{\mu\nu}) \quad (2.42)$$

Si l'on dérive cette expression par rapport aux coordonnées nucléaires [24], la tâche la plus difficile

consiste à déterminer la valeur du terme suivant :

$$\sum_{\mu\nu} \partial P_{\nu\mu}^T F_{\mu\nu} = \sum_{\mu\nu} \partial P_{\nu\mu}^Q F_{\mu\nu} + \sum_{\mu\nu} \partial P_{\nu\mu}^G F_{\mu\nu} \quad (2.43)$$

Par souci de concision, les dérivées d'une quantité  $X$  par rapport aux coordonnées  $a$  ( $\frac{\partial X}{\partial a}$ ) sont notées simplement  $\partial X$ .

Commençons par nous focaliser sur le premier terme de l'équation 2.43 dû aux OM :

$$\sum_{\mu\nu} \partial P_{\nu\mu}^Q F_{\mu\nu} = \sum_{\mu\nu} \sum_i n_i \partial c_{\mu i}^* F_{\mu\nu} c_{\nu i} + \sum_{\mu\nu} \sum_i n_i c_{\mu i}^* F_{\mu\nu} \partial c_{\nu i} \quad (2.44)$$

En utilisant les équations de Roothaan modifiées (équations 2.35) et les conditions d'orthonormalisation (équations 2.31 et 2.32), l'équation 2.44 devient :

$$\begin{aligned} \sum_{\mu\nu} \partial P_{\nu\mu}^Q F_{\mu\nu} &= - \sum_{\mu\nu} \sum_i n_i \epsilon_i c_{\mu i}^* c_{\nu i} \partial S_{\mu\nu} \\ &\quad - \sum_{\mu\nu} \sum_i \sum_l n_i c_{\mu i}^* a_{\nu l} \lambda_{li} \partial S_{\mu\nu} \\ &\quad - \sum_{\mu\nu} \sum_i \sum_l n_i \lambda_{li}^* a_{\mu l}^* c_{\nu i} \partial S_{\mu\nu} \\ &\quad - \sum_{\mu\nu} \sum_i \sum_l n_i c_{\mu i}^* S_{\mu\nu} \partial a_{\nu l} \lambda_{li} \\ &\quad - \sum_{\mu\nu} \sum_i \sum_l n_i \lambda_{li}^* \partial a_{\mu l}^* S_{\mu\nu} c_{\nu i} \end{aligned} \quad (2.45)$$

Utilisant les expressions des multiplicateurs de Lagrange (équations 2.36 et 2.37), il suit :

$$\begin{aligned} \sum_{\mu\nu} \partial P_{\nu\mu}^Q F_{\mu\nu} &= - \sum_{\mu\nu} \sum_{\sigma\zeta} \sum_i n_i c_{\sigma i}^* F_{\sigma\zeta} c_{\zeta i} c_{\mu i}^* c_{\nu i} \partial S_{\mu\nu} \\ &\quad - \sum_{\mu\nu} \sum_{\sigma\zeta} \sum_i \sum_l n_i c_{\mu i}^* a_{\nu l} a_{\sigma l}^* F_{\sigma\zeta} c_{\zeta i} \partial S_{\mu\nu} \\ &\quad - \sum_{\mu\nu} \sum_{\sigma\zeta} \sum_i \sum_l n_i c_{\sigma i}^* F_{\sigma\zeta} a_{\zeta l} a_{\mu l}^* c_{\nu i} \partial S_{\mu\nu} \\ &\quad - \sum_{\mu\nu} \sum_{\sigma\zeta} \sum_i \sum_l n_i c_{\mu i}^* S_{\mu\nu} \partial a_{\nu l} a_{\sigma l}^* F_{\sigma\zeta} c_{\zeta i} \\ &\quad - \sum_{\mu\nu} \sum_{\sigma\zeta} \sum_i \sum_l n_i c_{\sigma i}^* F_{\sigma\zeta} a_{\zeta l} \partial a_{\mu l}^* S_{\mu\nu} c_{\nu i} \end{aligned} \quad (2.46)$$

Maintenant, on peut définir une matrice  $W$  analogue à la matrice densité pondérée par l'énergie

utilisée habituellement :

$$\begin{aligned}
W_{\nu\mu} &= \sum_{\sigma\zeta} \sum_i n_i c_{\nu i} c_{\sigma i}^* F_{\sigma\zeta} c_{\zeta i} c_{\mu i}^* \\
&+ \sum_{\sigma\zeta} \sum_i \sum_l n_i a_{\nu l} a_{\sigma l}^* F_{\sigma\zeta} c_{\zeta i} c_{\mu i}^* \\
&+ \sum_{\sigma\zeta} \sum_i \sum_l n_i c_{\nu i} c_{\sigma i}^* F_{\sigma\zeta} a_{\zeta l} a_{\mu l}^*
\end{aligned} \tag{2.47}$$

que l'on introduit dans l'expression 2.46, puis réarrangeant les indices et sommant sur certains d'entre eux, on obtient l'expression :

$$\begin{aligned}
\sum_{\mu\nu} \partial P_{\nu\mu}^Q F_{\mu\nu} &= - \sum_{\mu\nu} W_{\nu\mu} \partial S_{\mu\nu} \\
&- \sum_{\mu\nu} \sum_l a_{\mu l}^* (FP^Q S)_{\mu\nu} \partial a_{\nu l} \\
&- \sum_{\mu\nu} \sum_l \partial a_{\mu l}^* (SP^Q F)_{\mu\nu} a_{\nu l}
\end{aligned} \tag{2.48}$$

Finalement, le terme 2.43 peut s'exprimer simplement sous la forme suivante :

$$\begin{aligned}
\sum_{\mu\nu} \partial P_{\nu\mu}^T F_{\mu\nu} &= - \sum_{\mu\nu} W_{\nu\mu} \partial S_{\mu\nu} \\
&- \sum_{\mu\nu} \sum_l \left[ n_l a_{\mu l}^* F_{\mu\nu} \partial a_{\nu l} - a_{\mu l}^* (FP^Q S)_{\mu\nu} \partial a_{\nu l} \right] \\
&- \sum_{\mu\nu} \sum_l \left[ n_l \partial a_{\mu l}^* F_{\mu\nu} a_{\nu l} - \partial a_{\mu l}^* (SP^Q F)_{\mu\nu} a_{\nu l} \right]
\end{aligned} \tag{2.49}$$

où  $n_l$  représente le nombre d'occupation de l'OG  $l$  (2 ou 0 en formalisme restreint). Cependant cette expression ne peut toujours pas être évaluée : les OG ne sont pas fonctions propres de l'opérateur de Hartree-Fock et il n'existe pas de moyen de déterminer les dérivées des coefficients des OG ( $\partial a_{\mu l}$ ).

Nous verrons dans l'application du formalisme LSCF aux méthodes QM/MM que l'utilisation d'orbitales strictement localisées permet de calculer ces dérivées.

### 2.3 Application de la méthode LSCF à la détermination des énergies d'excitation/ionisation.

La première application du formalisme quantique LSCF *ab initio* concerne le calcul des énergies d'excitation et d'ionisation de molécules de toutes tailles. En effet, si on gèle une spinorbitale occupée

de l'état fondamental d'un système moléculaire en la vidant, le calcul LSCF correspondant permet d'atteindre un état excité (si le nombre d'électrons est conservé) représenté sous la forme d'un déterminant de Slater orthogonal à l'état fondamental ou un état ionisé (si l'électron est enlevé).

Nous commencerons par rappeler les méthodes existantes en insistant particulièrement sur les méthodes mono-déterminantales appliquées aux ionisations/excitations des électrons de coeur.

Nous donnerons ensuite la justification mathématique de l'orthogonalité entre déterminants obtenus par la méthode LSCF.

Nous présenterons enfin quelques applications qui concernent la détermination quantitative des énergies d'ionisation de petites molécules.

### 2.3.1 Quelques repères méthodologiques

L'étude des états électroniques excités des molécules nécessite habituellement l'utilisation d'un formalisme post-Hartree-Fock, c'est-à-dire de dépasser la description mono-déterminantale du système. Nous citerons quelques méthodes à titre d'illustration, bien que nous ne nous soyons pas directement intéressés à elles pour ce travail.

- CIS pour Configuration Interactions Singles : seuls les déterminants issus d'une mono-excitation d'une spinorbitale de l'état fondamental sont considérés. L'Hamiltonien électronique est exprimé sur la base de ces configurations et diagonalisé.
- MCSCF pour Multi-Configurational SCF : les coefficients d'une combinaison de déterminants de Slater ainsi que les coefficients des orbitales atomiques dans les spinorbitales sont déterminés simultanément. Ce sont des calculs très lourds.
- Nous citerons encore les méthodes CCSD-EOM de Stanton [25] et TD-DFT [26], toutes deux également basées sur une description mono-déterminantale.

L'approche mono-déterminantale des états excités est en général trop pauvre pour obtenir des résultats quantitatifs. Cependant, dans certains cas comme les excitations ou ionisations des électrons de coeur (un électron est arraché de la couche la plus interne d'un atome), cette description est bien adaptée : en effet, le terme d'échange qui intervient entre l'électron de l'orbitale de départ et l'électron excité est négligeable (les états singulets et triplets qui résultent de l'excitation sont alors quasi-dégénérés). Ce sont ces cas qui nous intéresseront dans la suite. Il est important de rappeler que l'étude des états ionisés de coeur n'est pas triviale car un « trou » doit être créé dans la mer de Fermi. Généralement ce « trou » est comblé par la relaxation des électrons lors de l'optimisation de la fonction d'onde.

Mais avant d'exposer le principe de l'application de la méthode quantique LSCF aux états excités/ionisés, nous allons effectuer un bref récapitulatif des méthodes existantes, dédiées aux excitations/ionisations de coeur et qui utilisent une description mono-déterminantale de la fonction d'onde.

Nous pouvons commencer en rappelant le théorème de Koopmans [27] qui associe, au signe près, l'énergie des orbitales moléculaires issues d'un calcul Hartree-Fock, au potentiel d'ionisation de l'électron décrit par cette orbitale. Grâce à une bonne compensation des erreurs (les deux principales sont : la variation de l'énergie de corrélation est négligée ; les orbitales moléculaires ne se réadaptent pas lors du départ de l'électron), cette approximation fonctionne assez bien pour les électrons de valence. En revanche, l'accord est beaucoup plus délicat pour les potentiels d'ionisation des électrons internes.

L'analogie dite «  $Z+1$  » permet d'obtenir facilement des énergies d'ionisation de coeur : les niveaux d'énergie et les propriétés de valence d'un système ionisé de coeur sont très semblables à ceux du système contenant un proton supplémentaire (par exemple, l'analogie de CO avec un électron  $1s$  de C excité est  $\text{NO}^+$ ). Une récente application a concerné l'étude de la structure vibrationnelle d'hydrocarbures ionisés de coeur [28].

La méthode mono-déterminantale la plus utilisée est l'approche  $\Delta\text{SCF}$  de Bagus [29, 30, 31] qui autorise la relaxation des orbitales. Celle-ci nécessite deux calculs. Le premier, habituel, calcule l'énergie de l'état fondamental du système étudié. Pour le deuxième, une orbitale, identifiée comme étant celle qui décrit le mieux l'électron de coeur qui nous intéresse, est remplie avec un seul électron au lieu de deux (l'identification de cette orbitale peut être obtenue par une procédure de recouvrement maximum à chaque étape SCF). Cette méthode s'applique aussi bien à la description des états excités que ionisés. Cependant, le formalisme  $\Delta\text{SCF}$  ne peut satisfaire la condition d'orthogonalité entre l'état excité et l'état fondamental puisqu'il n'y a pas de référence au fondamental pendant le deuxième calcul. Ågren et collaborateurs ont repris ces idées et formulé la méthode  $\Delta\text{MCSCF}$  [32, 33]. Le deuxième calcul s'effectue lui-même en deux étapes : dans un premier temps, l'orbitale qui est occupée avec un seul électron est gelée ; dans un deuxième temps, toutes les orbitales sont relaxées par une optimisation de type Newton-Raphson. Enfin, récemment, le principe  $\Delta$  a été étendu à la fonctionnelle de la densité [34, 35] (on parlera alors de  $\Delta\text{KS}$  pour Kohn-Sham) : on a ainsi une méthode simple à mettre en oeuvre qui permet de relaxer les orbitales de l'état excité et de prendre en compte (au moins partiellement) la corrélation électronique. Les applications de ces méthodes  $\Delta$  ont été nombreuses et ont concerné aussi bien les excitations/ionisations des électrons de coeur que de valence. Notamment, elles ont permis d'aider à l'identification des spectres électroniques ainsi qu'à l'identification des structures moléculaires grâce aux déplacements chimiques des énergies induits par la présence de groupements caractéristiques

(par exemple, lors de l'adsorption de petites molécules sur une surface métallique) [36, 37, 38, 39, 40].

En 1969, Hunt et Goddard III ont proposé la technique IVO (pour Improved Virtual Orbitals [41]) qui permet de résoudre les équations Hartree-Fock et d'obtenir un jeu d'orbitales virtuelles qui soient de bonnes approximations des orbitales décrivant les états excités. Pour cela, ils considèrent l'opérateur  $\widehat{H}_i = \widehat{H}^c + \sum_{j \neq i}^n [\widehat{J}_j - \widehat{K}_j]$  construit pour chaque électron  $i$  : cet électron ressent le potentiel des  $n - 1$  autres électrons. Les spinorbitales virtuelles fonctions propres de cet opérateur représentent donc les niveaux excités de l'électron  $i$ .

Nous citerons enfin le formalisme développé par Chong [42], basé sur l'utilisation de la fonctionnelle de la densité et du modèle de l'état de transition généralisé non-restreint (uGTS) qui donne de très bons résultats aussi bien pour les ionisations/excitations de coeur que de valence [42, 43, 44, 45]. Ce modèle est la continuité de l'état de transition de Slater.

Les états excités de coeur sont des états très hauts en énergie. Il est donc illusoire de chercher à les représenter par des méthodes de type Interactions de Configurations pour de gros systèmes : il faudrait connaître les racines de l'Hamiltonien de hautes énergies avec une très grande précision, ce qui n'est pas possible à cause de la dégradation des résultats lorsqu'on monte en énergie.

### 2.3.2 Description des états ionisés de coeur par la méthode LSCF

La détermination du premier potentiel d'ionisation est triviale avec les outils standards de la chimie quantique. En revanche, il en va autrement lorsqu'on cherche à arracher un électron ne provenant pas de la plus haute orbitale moléculaire occupée (HOMO). En effet, la méthode variationnelle d'optimisation de la fonction d'onde permet d'obtenir l'état le plus stable, c'est-à-dire de plus basse énergie. Celui-ci ne correspond sûrement pas à un système comprenant une lacune électronique, créée par le départ d'un électron de coeur, dans ses couches internes : elle sera remplie par les électrons d'énergie supérieure.

Pour déterminer ces états ionisés de coeur avec la méthode LSCF, nous commençons par calculer l'état fondamental du système. Puis, nous gelons une spinorbitale optimisée occupée qui provient du déterminant de Slater représentant l'état fondamental et nous lui attribuons un nombre d'occupation nul. Enfin, un calcul LSCF est effectué sur le système ionisé. L'énergie d'ionisation est alors égale à la différence entre les deux calculs.

### 2.3.3 Description des états excités par la méthode LSCF

Par définition, les états excités d'une molécule sont les états propres de l'Hamiltonien du système, d'énergie supérieure à celle de l'état fondamental. Ils sont donc orthogonaux à l'état fondamental et

orthogonaux entre eux. Les méthodes de type  $\Delta$  présentées précédemment ne satisfont pas à cette condition. Nous avons choisi d'utiliser la méthode quantique LSCF afin de satisfaire à la contrainte d'orthogonalité entre l'état fondamental et les états excités. Chaque état excité est alors représenté par un déterminant de Slater orthogonal au déterminant de l'état fondamental. La contrainte d'orthogonalité est réalisée en gelant une spinorbitale de l'état fondamental que l'on vide préalablement. L'état excité obtenu est construit sur un jeu de spinorbitales, toutes orthogonales à la spinorbitale gelée. Pour prouver que cette technique conduit bien à un état orthogonal à l'état fondamental, considérons le déterminant de Slater de l'état  $\Psi$  d'un système de  $n$  électrons, construit à partir d'un jeu de spinorbitales  $\{\varphi_i\}$  :

$$\Psi = |\varphi_1 \varphi_2 \cdots \varphi_n| \quad (2.50)$$

Considérons un second déterminant  $\Psi'$  construit sur un jeu  $\{\phi_i\}$  de spinorbitales orthogonales à  $\varphi_k$ , appartenant à  $\Psi$ , qui est gelée par un calcul LSCF :

$$\Psi' = |\phi_1 \phi_2 \cdots \phi_n| \quad (2.51)$$

tel que :

$$\langle \varphi_k | \phi_i \rangle = 0, \forall i \in [1 \cdots n] \quad (2.52)$$

Le recouvrement entre les deux déterminants  $\Psi$  et  $\Psi'$  est donné par :

$$S = \langle \Psi | \Psi' \rangle = \frac{1}{n!} \left\langle \begin{vmatrix} \varphi_1(1) & \varphi_2(1) & \cdots & \varphi_n(1) \\ \varphi_1(2) & \varphi_2(2) & \cdots & \varphi_n(2) \\ \vdots & \vdots & \ddots & \vdots \\ \varphi_1(n) & \varphi_2(n) & \cdots & \varphi_n(n) \end{vmatrix} \middle| \begin{vmatrix} \phi_1(1) & \phi_2(1) & \cdots & \phi_n(1) \\ \phi_1(2) & \phi_2(2) & \cdots & \phi_n(2) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_1(n) & \phi_2(n) & \cdots & \phi_n(n) \end{vmatrix} \right\rangle \quad (2.53)$$

Développant cette expression et utilisant les règles énoncées page 75 de la référence [22], l'expression ci-dessus devient :

$$S = \frac{n(n-1)}{2} (n!)^{-1} \sum_i^{n!} \sum_j^{n!} (-1)^{p_i} (-1)^{p_j} \int d\tau_1 \cdots d\tau_n \times \wp_i \{\varphi_1(1) \cdots \varphi_n(n)\} \times \wp_j \{\phi_1(1) \cdots \phi_n(n)\} \quad (2.54)$$

tel que  $\wp_i$  est l'opérateur qui génère la  $i^{\text{ème}}$  permutation des électrons  $1 \cdots n$  et  $p_i$  est le nombre de

transpositions nécessaires pour obtenir cette permutation.

Développant cette expression, le recouvrement entre la spinorbitale  $\varphi_k$  et une spinorbitale  $\phi_i$  apparaît toujours en facteur. Cette quantité étant nulle, il s'ensuit que l'expression 2.53 est nulle également. Ainsi, par construction, l'état  $\Psi'$  est bien orthogonal à l'état  $\Psi$ .

L'énergie d'excitation d'un électron est alors égale (comme pour la méthode  $\Delta$ SCF) à la différence :

$$\Delta E = E(\Psi') - E(\Psi) \quad (2.55)$$

Si maintenant nous gelons une spinorbitale  $\varphi_l$  de  $\Psi$  et une spinorbitale  $\phi_a$  de  $\Psi'$ , le déterminant  $\Psi''$  obtenu grâce à la méthode LSCF sera également orthogonal à  $\Psi$  et  $\Psi'$ . De cette manière, nous pouvons obtenir de proche en proche tout le spectre d'excitation d'un électron (pour un jeu de fonctions de base donné) uniquement à l'aide de calculs LSCF. En revanche, nous notons que les états excités issus de deux électrons décrits par des spinorbitales différentes dans l'état fondamental ne sont pas orthogonaux.

Comme indiqué précédemment, nous pouvons également utiliser une description post Hartree-Fock basée sur le déterminant de Slater d'un état excité.

Dans la suite de ce mémoire, nous utiliserons la méthode présentée ici à la détermination des énergies d'ionisation de petites molécules afin de valider notre approche. Puis, à la fin du prochain chapitre, nous la couplerons au formalisme hybride QM/MM présenté, afin d'étudier les processus d'ionisation de coeur dans les protéines.

### 2.3.4 Application à l'ionisation de molécules

Nous devons commencer par valider notre approche, c'est-à-dire déterminer avec quelle précision la méthode LSCF permet d'obtenir des énergies d'ionisation en bon accord avec les valeurs expérimentales et les valeurs théoriques obtenues avec les meilleures méthodes actuelles. Nous avons choisi pour ces dernières les résultats publiés par Triguero [34] pour les ionisations de coeur et ceux publiés par Shapley [35] concernant les ionisations des électrons de valence de petites molécules. Ces études utilisent le formalisme  $\Delta$ KS décrit précédemment, considéré comme la méthode la plus performante actuellement. De plus, des corrections relativistes sont incluses dans les résultats concernant les électrons de coeur (0,2 eV pour le carbone et 0,4 eV pour l'oxygène). Pour notre part, nous avons choisi de nous focaliser uniquement sur la méthode LSCF : aucune correction n'est appliquée aux résultats.

Toutes les valeurs reportées ci-dessous sont calculées sur la géométrie expérimentale. La base utilisée doit être suffisamment étendue afin de s'affranchir des effets de taille : nous avons choisi la base de qualité



quintuple- $\zeta$  cc-pV5Z de Dunning (choisie également par Triguero et Shapley). Trois niveaux de théorie seront considérés pour les calculs LSCF : Hartree-Fock (HF), Møller-Plesset au second ordre (MP2) et fonctionnelle de la densité B3PW91. Grâce à ces deux dernières méthodes, nous prendrons ainsi en compte une partie de la corrélation électronique.

Les énergies d'ionisation de cinq molécules sont données dans le tableau I.

TAB. I – *Energies d'ionisation théoriques et expérimentales, et déviations à l'expérience ( $\Delta E$ ) en eV. La colonne Ref reporte les résultats publiés [34, 35] (incluant des corrections relativistes). Des résultats expérimentaux sont donnés dans la dernière colonne [34, 35, 46].*

Molécule	Orbitale	HF	$\Delta E$	MP2	$\Delta E$	B3PW91	$\Delta E$	Ref	$\Delta E$	Exp
H <sub>2</sub>	1 $\sigma_g$	15,33	-0,65	16,25	0,27	16,44	0,46	16,40	0,42	15,98
HF	1 $\pi$	14,31	-1,88	16,92	0,73	16,21	0,02	16,64	0,45	16,19
	3 $\sigma$	18,26	-1,56	20,88	1,06	19,98	0,16	20,16	0,34	19,82
	2 $\sigma$	40,37	0,79	42,04	2,46	38,31	-1,27	37,55	-2,03	39,58
	1 $\sigma$	692,80	-1,40	693,78	-0,42	692,81	-1,39			694,20
CO	5 $\sigma$	13,23	-0,78	14,80	0,79	14,06	0,05	13,97	-0,04	14,01
	1 $\pi$	15,08	-1,83	18,06	1,15	17,11	0,20	17,47	0,56	16,91
	4 $\sigma$	19,00	-0,72	21,52	1,80	19,92	0,20	19,86	0,14	19,72
	3 $\sigma$	38,22		39,89		36,79				
	2 $\sigma$	296,31	0,11	297,18	0,98	295,95	-0,35	296,05	-0,15	296,20
H <sub>2</sub> O	1 $\sigma$	541,22	-1,28	542,41	-0,09	541,57	-0,93	542,24	-0,26	542,50
	1b <sub>1</sub>	10,97	-1,65	12,97	0,35	12,70	0,08	12,82	0,20	12,62
	3a <sub>1</sub>	13,26	-1,48	15,20	0,46	14,84	0,10	14,98	0,24	14,74
	1b <sub>2</sub>	17,44	-1,11	19,34	0,79	18,93	0,38	19,08	0,53	18,55
	2a <sub>1</sub>	34,05	1,85	34,53	2,33	32,44	0,24	31,80	-0,40	32,20
	1a <sub>1</sub>	538,90	-1,03	539,25	-0,68	538,86	-1,07	539,53	-0,40	539,93
CH <sub>4</sub>	1t <sub>2</sub>	13,31	-1,09	14,51	0,11	14,23	-0,17	14,19	-0,21	14,40
	2a <sub>1</sub>	24,15	1,25	24,34	1,44	22,64	-0,26	21,96	-0,94	22,90
	1a <sub>1</sub>	290,57	-0,29	290,31	-0,55	290,37	-0,49	290,65	-0,21	290,86

Une moyenne des déviations à l'expérience peut être calculée : les RMS (« Root Mean Square ») sont donnés dans le tableau II.

TAB. II – *Moyennes des déviations à l'expérience des énergies d'ionisation (en eV).*

Electrons	HF	MP2	B3PW91	Ref
Valence	1,35	1,28	0,42	0,70
Coeur	0,97	0,62	0,93	0,27

Concernant les ionisations des électrons de valence, la méthode LSCF basée sur la DFT et la fonc-

tionnelle hybride B3PW91 donne de très bons résultats, meilleurs que ceux publiés. Même pour un cas difficile comme l'orbitale  $2\sigma$  de HF, l'application de LSCF améliore la précision du résultat. L'utilisation des formalismes Hartree-Fock et de MP2 ne permet pas d'avoir de bons résultats : la première donne systématiquement une erreur supérieure à 1 eV ; quant à la seconde, plus les énergies d'ionisation sont grandes, plus l'erreur devient importante (le RMS résultant étant du même ordre que celui calculé pour HF).

Les énergies d'ionisation des électrons de coeur présentent des tendances différentes. Les formalismes HF et DFT ont des RMS voisins de l'ordre de 1 eV tandis que la méthode MP2 a un RMS de 0,62 eV relativement satisfaisant, cette dernière semble être la méthode de choix dans ce cas. Cependant, l'ionisation d'un électron 1s du carbone dans CO présente un comportement inverse : l'erreur HF est très petite (0,11 eV) tandis que l'erreur MP2 est importante (0,98 eV). Nous avons alors trouvé que la contamination de spin est très importante dans ce cas ( $\langle S^2 \rangle = 1,32$  au lieu de 0,75 pour un doublet). La méthode des projecteurs (équation 2.15) permet d'annihiler la contribution du premier contaminant. Nous obtenons alors  $\langle S^2 \rangle = 0,99$  et les énergies d'ionisation correspondantes sont 295,19 eV pour HF et 296,13 eV pour MP2 (soit des erreurs de 1,01 et -0,07 eV). Les RMS obtenus après correction de la contamination de spin sont 1,07 eV pour le formalisme HF et 0,44 eV pour la méthode MP2. Cette dernière s'impose donc comme la méthode de choix. La comparaison des déviations avec celles issues des résultats publiés est difficile : ces derniers comportent des corrections relativistes. Le procédé correspondant n'est pas très clairement expliqué ; cependant si on enlève 0,2 eV aux résultats portant sur l'ionisation de coeur du carbone et 0,4 eV à ceux concernant l'oxygène, le RMS monte alors à 0,58 eV, valeur supérieure à celle obtenue par la méthode LSCF basée sur MP2.

La méthode quantique LSCF *ab initio* semble donc être particulièrement bien adaptée pour la détermination des énergies d'ionisation : le cas des électrons de valence est traité préférentiellement par DFT et le cas des électrons de coeur par MP2.

### 2.3.5 Application aux états excités

En principe, l'approche LSCF des états excités, de n'importe quelle multiplicité, est possible. Il reste que la description monodéterminantale de ces états ne permet d'espérer que des résultats qualitatifs. Cependant, comme il a été mentionné précédemment, il est possible de construire un espace de configurations excitées basé sur un déterminant LSCF. Il est donc possible de réaliser des calculs de type interaction de configurations.

Nous avons testé cette possibilité en déterminant les énergies d'excitation de la molécule  $H_2$  qui

permettent de passer de l'état fondamental singulet à l'un des trois premiers états excités, eux aussi singulets. Bien évidemment, ce cas est le pire que l'on puisse imaginer : l'échange entre les deux électrons du système est toujours très grand. Les résultats, pour différentes méthodes basées sur LSCF (Hartree-Fock, DFT, MP2, CISD), sont donnés dans le tableau III et comparés aux valeurs obtenues par différentes méthodes conventionnelles (CIS, MCSCF). Les résultats donnés par la méthode MCSCF correspondent à ceux d'une interaction de configurations complète pour la base considérée (6-311++G\*\*).

TAB. III – Trois premières énergies d'excitation singulet→singulet (en eV) de  $H_2$ .

Méthodes basées sur LSCF						Méthodes usuelles	
HF	BLYP	B1LYP	B3LYP	MP2	CISD	CIS	MCSCF
12,55	12,06	12,32	12,30	13,17	13,30	13,66	13,52
15,56	15,35	15,55	15,57	16,16	16,31	17,64	17,42
22,80	21,81	22,16	22,12	22,95	23,06	25,51	25,30

De toute évidence, les résultats basés sur le formalisme LSCF sont mitigés. Même les valeurs correspondant à une interaction de configurations simples et doubles basée sur le déterminant de Slater adapté à l'état excité singulet sont relativement éloignées de celles données par les méthodes conventionnelles.

Nous avons choisi cet exemple car il nous permet d'estimer l'erreur maximale due à notre méthode. Il est évident que cette dernière donne des résultats beaucoup plus satisfaisants pour des états excités de coeur.

L'application de la méthode LSCF aux états excités n'est pas directe. En forçant une description monodéterminantale de ces états, la contamination de spin ne peut être évitée. Une solution future pourrait consister à construire un jeu de déterminants excités grâce à LSCF pour différents états et différentes multiplicités de spin et de construire l'Hamiltonien électronique dans la base de ces déterminants.

## 2.4 Conclusion

Nous avons défini dans ce chapitre la méthode quantique LSCF *ab initio* qui permet d'effectuer un calcul SCF, en ayant au préalable défini certaines spinorbitales (de n'importe quel type et forme) qui restent gelées durant le calcul. En particulier, nous avons exposé le principe de résolution des équations correspondantes, que nous avons formulées dans le cas d'un système à couches fermées. Les dérivées premières de l'énergie LSCF par rapport aux coordonnées nucléaires ont été calculées de façon analytique, mais ne peuvent actuellement pas être évaluées car les orbitales gelées ne sont pas vecteurs propres de l'opérateur de Fock.

Outre l'approche QM/MM que nous aborderons dans le chapitre suivant, le formalisme LSCF quantique peut être utilisé pour la détermination des énergies d'ionisation (et dans une moindre mesure, d'excitation) de n'importe quel électron d'un système moléculaire. Elle permet d'atteindre des états de même multiplicité que l'état fondamental sans crainte de catastrophe variationnelle. Cependant elle ne permet d'obtenir à l'heure actuelle que des énergies d'excitation/ionisation verticales. A cause de la description non-restreinte, nos résultats sont entachés de la contamination de spin.

Une autre piste que nous proposons pour le formalisme quantique LSCF mène aux méthodes à croissance linéaire [47, 48], pour lesquelles le temps de calcul croît linéairement avec la taille du système. La méthode de résolution des équations LSCF demande de passer dans une base orthogonale dont la taille (par rapport à celle de la base originale) est diminuée du nombre d'orbitales gelées. Si ce nombre est suffisamment grand, la diagonalisation de la matrice de Fock dans cette « petite » base devient aisée, même pour un grand système. C'est à partir de cette idée que la méthode LSCF à croissance linéaire pourra être développée.

## Bibliographie

- [1] M. Born et J.R. Oppenheimer. *Ann. Physik* **84**, 457 (1927).
- [2] W. Pauli Jr. *Z. Physik* **31**, 765 (1925).
- [3] J.C. Slater. *Phys. Rev.* **34**, 1293 (1929).
- [4] D.R. Hartree. *Proc. Cambridge Phil. Soc.* **24**, 89 (1928).
- [5] V. Fock. *Z. Physik* **61**, 126 (1930).
- [6] C.C.J. Roothaan. *Rev. Mod. Phys.* **23**, 69 (1951).
- [7] G. Berthier. *C.R. Acad. Sci. Paris* **238**, 91 (1954).
- [8] J.A. Pople et R.K. Nesbet. *J. Chem. Phys.* **22**, 571 (1954).
- [9] H.B. Schlegel. *J. Phys. Chem.* **92**, 3075 (1988).
- [10] M. Dupuis, J. Rys et H.F. King. *J. Chem. Phys.* **65**, 111 (1976).
- [11] P.M.W. Gill. *Advances in quantum chemistry*, tome 25, chapitre Molecular integrals over gaussian basis functions, pages 141–205. Academic Press, Inc. (1994).
- [12] P. Hohenberg et W. Kohn. *Phys. Rev. B* **136**, 864 (1964).

- [13] W. Kohn et L.J. Sham. *Phys. Rev. A* **140**, 1133 (1965).
- [14] X. Assfeld et J.-L. Rivail. *Chem. Phys. Lett.* **263**, 100 (1996).
- [15] X. Assfeld, N. Ferré et J.-L. Rivail. *ACS Symposium: Combined quantum mechanical and molecular mechanical methods*, chapitre The Local Self Consistent Field principles and applications to combined quantum mechanical - molecular mechanical computations on biomacromolecular systems, pages 234–249. 712. A.C.S., Washington, D.C. (1998).
- [16] N. Ferré, X. Assfeld et J.-L. Rivail. *Specific force field parameters determination for the hybrid ab initio QM/MM LSCF method*. Submitted to Journal of Computational Chemistry.
- [17] W.H. Fink. *J. Chem. Phys.* **57**, 1822 (1972).
- [18] W.J. Stevens et W.H. Fink. *Chem. Phys. Lett.* **139**, 15 (1987).
- [19] P.W. Payne. *J. Am. Chem. Soc.* **99**, 3787 (1977).
- [20] P.S. Bagus, K. Hermann et C.W. Bauschlicher, Jr. *J. Chem. Phys.* **80**, 4378 (1984).
- [21] J.G. Ángyán. *Theor. Chem. Acc.* **103**, 238 (2000).
- [22] A. Szabo et N.S. Ostlund. *Modern quantum chemistry*. McGraw-Hill, New York (1982).
- [23] M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, V.G. Zakrzewski, J.A. Montgomery, Jr., R.E. Stratmann, J.C. Burant, S. Dapprich, J.M. Millam, A.D. Daniels, K.N. Kudin, M.C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G.A. Petersson, P.Y. Ayala, Q. Cui, K. Morokuma, D.K. Malick, A.D. Rabuck, K. Raghavachari, J.B. Foresman, J. Cioslowski, J.V. Ortiz, A.G. Baboul, B.B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R.L. Martin, D.J. Fox, T. Keith, M.A. Al-Laham, C.Y. Peng, A. Nanayakkara, C. Gonzales, M. Challacombe, P.M.W. Gill, B. Johnson, W. Chen, M.W. Wong, J.L. Andres, C. Gonzales, M. Head-Gordon, E.S. Replogle et J.A. Pople. *Gaussian 98, Revision A.9* (1998). Gaussian, Inc., Pittsburgh PA.
- [24] J.A. Pople, R. Krishnan, H.B. Schlegel et J.S. Binkley. *Int. J. Quant. Chem.: Quant. Chem. Symp.* **13**, 225 (1979).
- [25] J.F. Stanton et R.J. Bartlett. *J. Chem. Phys.* **98**, 7029 (1993).
- [26] M.E. Casida, C. Jamorski, K.C. Casida et D.R. Salahub. *J. Chem. Phys.* **108**, 4439 (1998).
- [27] T. Koopmans. *Physica* **1**, 104 (1933).
- [28] T. Darrah Thomas, L. J. Saethre, S. L. Sorensen et S. Svensson. *J. Chem. Phys.* **109**, 1041 (1998).
- [29] P.S. Bagus. *Phys. Rev.* **139**, 619 (1965).
- [30] P.S. Bagus et H.F. Schaefer III. *J. Chem. Phys.* **55**, 1474 (1971).

- [31] P.S. Bagus et H.F. Schaefer III. *J. Chem. Phys.* **56**, 224 (1972).
- [32] H.J.Aa. Jensen, P. Jorgensen et H. Ågren. *J. Chem. Phys.* **87**, 451 (1987).
- [33] P. Norman et H. Ågren. *J. Mol. Struct. (Theochem)* **401**, 107 (1997).
- [34] L. Triguero, O. Plashkevych, L.G.M. Pettersson et H. Ågren. *J. Electron Spectrosc. Rel. Phenom.* **104**, 195 (1999).
- [35] W.A. Shapley et D.P. Chong. *Int. J. Quant. Chem.* **81**, 34 (2001).
- [36] F. Parmigiani, G. Pacchini, F. Illas et P.S. Bagus. *J. Electron Spectrosc. Rel. Phenom.* **59**, 255 (1992).
- [37] K. Ueda, M. Okunishi, H. Chiba, K. Ohmori, Y. Sato, E. Shigemasa et N. Kosugi. *Chem. Phys. Lett.* **236**, 311 (1995).
- [38] F. Gel'mukhanov et H. Ågren. *J. Chem. Phys.* **103**, 5848 (1995).
- [39] P.S. Bagus, F. Illas, J. Casanovas et J.M. Jiménez-Mateos. *J. Electron Spectrosc. Rel. Phenom.* **83**, 151 (1997).
- [40] P.S. Bagus, F. Illas, G. Pacchioni et F. Parmigiani. *J. Electron Spectrosc. Rel. Phenom.* **100**, 215 (1999).
- [41] W.J. Hunt et W.A. Goddard III. *Chem. Phys. Lett.* **3**, 414 (1969).
- [42] D.P. Chong. *J. Chem. Phys.* **103**, 1842 (1995).
- [43] D.P. Chong, C.-H. Hu et P. Duffy. *Chem. Phys. Lett.* **249**, 491 (1996).
- [44] S. Kranias, C. Bureau, D.P. Chong, V. Brenner, I. George, P. Viel et G. Lecayon. *J. Phys. Chem. B* **101**, 10254 (1997).
- [45] O. Plashkevych, H. Ågren, L. Karlsson et L.G.M. Pettersson. *J. Electron Spectrosc. Rel. Phenom.* **106**, 51 (2000).
- [46] W.L. Jolly, K.D. Bomben et C.J. Eyermann. *Atomic Data and Nuclear Data Tables* **31**, 433 (1984).
- [47] W. Yang. *J. Mol. Struct. (Theochem)* **255**, 461 (1992).
- [48] A. Van der Vaart, V. Gogonea, S.L. Dixon et K.M. Merz Jr. *J. Comp. Chem.* **21**, 1494 (2000).



## Chapitre 3

# Application du formalisme LSCF *ab initio* aux méthodes QM/MM

La réactivité chimique ne peut pas être abordée dans les systèmes étendus uniquement avec les outils de la chimie quantique. Il existe d'autres moyens beaucoup moins coûteux mais basés sur une approche classique de la mécanique qui permettent le traitement théorique de milliers d'atomes ; cependant l'absence de prise en compte explicite des degrés de liberté électroniques ne permet pas de traiter correctement les problèmes de réactivité. La solution généralement mise en oeuvre ces dernières années consiste à coupler une approche quantique du centre réactionnel avec une approche classique de l'environnement de la réaction.

Nous commencerons par rappeler les principes sous-jacents à l'approche classique et que l'on nomme mécanique moléculaire. Puis nous exposerons les principes du couplage des méthodes quantiques et classiques, appelé champ de forces hybride QM/MM. En particulier, nous verrons quels sont les problèmes posés lorsque la séparation entre les deux mondes quantique et classique se fait à travers une liaison chimique. Nous montrerons que le formalisme LSCF s'applique naturellement à l'approche QM/MM à l'aide d'orbitales de liaison strictement localisées qui relient les sous-systèmes quantique et classique. Nous validerons notre méthode tant du point de vue énergétique que géométrique à l'aide d'exemples simples.

### 3.1 Méthodes hybrides QM/MM

L'étude de la réactivité dans les systèmes moléculaires de grande taille nécessite à la fois une description au niveau électronique du système et la capacité à traiter plusieurs centaines d'atomes. Le traitement



purement quantique de ces systèmes reste prohibitif, même si des algorithmes à croissance linéaire sont développés depuis quelques années [1, 2, 3]. La mécanique moléculaire ne peut prendre en compte les électrons. Il faut donc trouver une stratégie qui allie la capacité de traitement de la mécanique moléculaire et la description électronique de la mécanique quantique.

Les processus réactionnels chimiques ne concernent habituellement qu'un petit jeu d'atomes : cette propriété locale est à la base de la notion de transférabilité des propriétés chimiques. Le reste de la molécule est alors spectateur de la réaction ; il n'en est pas pour autant passif : la prise en compte de son influence sur le centre réactif est nécessaire. Cependant, une première solution serait de négliger l'effet de l'environnement chimique : on appelle cette approche « molécule modèle » ou, en anglais, « cluster ». Le choix d'un cluster assez étendu devrait permettre de prendre en compte la majorité des effets d'environnement. Cependant, certaines interactions, notamment électrostatiques, sont à longue portée et il faut alors définir un cluster quantique très étendu : on se retrouve face à l'impossibilité d'effectuer des calculs quantiques sur de tels systèmes. Il faut donc trouver une approximation, peu coûteuse, qui permette de prendre en compte les interactions entre les quelques atomes qui réagissent et leur environnement.

L'idée des méthodes hybrides QM/MM, développées depuis une vingtaine d'années, est de décrire à l'aide des méthodes de la mécanique quantique le centre réactionnel et de traiter l'environnement de la réaction à l'aide de la mécanique moléculaire. Les méthodes existantes diffèrent dans leurs façons de séparer les deux sous-systèmes et dans leurs prises en compte des interactions entre les deux sous-systèmes. L'Hamiltonien du système se décompose en trois termes :

$$\hat{H} = \hat{H}_{QM} + \hat{H}_{MM} + \hat{H}_{QM/MM} \quad (3.1)$$

où  $\hat{H}_{QM}$  et  $\hat{H}_{MM}$  sont les Hamiltoniens des sous-systèmes quantique et classique respectivement et  $\hat{H}_{QM/MM}$  est l'Hamiltonien d'interaction entre les deux sous-systèmes. C'est par ce dernier terme que se différencient le plus les différentes approches.

Nous commencerons par rappeler sommairement les principes de la mécanique moléculaire. Puis nous décrirons les diverses méthodes actuelles, en insistant particulièrement sur leurs différences et leurs domaines d'applicabilité. Parmi celles-ci, nous nous focaliserons sur les méthodes QM/MM qui permettent de traiter les macromolécules, c'est-à-dire les macrosystèmes pour lesquels la séparation entre les deux sous-systèmes se fait sur une ou plusieurs liaisons chimiques. Nous détaillerons alors la mise au point de notre champ de force hybride LSCF/MM *ab initio*.

### 3.1.1 Mécanique moléculaire

Les méthodes de mécanique moléculaire considèrent les atomes comme des points en interaction selon des lois issues de la mécanique newtonnienne. Ces interactions sont formulées de façon empirique et ajustées sur un jeu de molécules de référence. L'application de cet ensemble de lois à d'autres molécules qui ressemblent plus ou moins aux molécules de référence repose sur l'hypothèse souvent vérifiée de transférabilité des propriétés chimiques.

Cet ensemble d'interactions est appelé champ de forces. Il définit une surface de potentiel modèle sur laquelle se déplacent les atomes. Il existe une variété importante de ces champs de force, plus ou moins élaborés. Nous nous limiterons ici à la description des plus simples, qui seront utilisés dans cette étude (Amber [4], UFF [5], Dreiding [6], contenus dans Gaussian98 [7]). Ces trois champs de forces partagent la même représentation des potentiels empiriques qui les composent.

Les champs de forces classiques distinguent les interactions entre atomes liés et atomes non-liés. Ceci suppose la définition d'une connectivité qui recense toutes les liaisons chimiques présentes dans la molécule étudiée. La forme du champ de forces dépend donc directement de cette connectivité : si celle-ci varie au cours de l'étude (par exemple, au cours d'une réaction chimique), le champ de forces initial ne sera plus adapté à la poursuite de l'étude ; si l'on redéfinit le champ de forces en fonction de la nouvelle connectivité, on change alors de surface de potentiel et on ne peut pas comparer les résultats avant et après la rupture de la liaison. Ce dernier point constitue une faiblesse majeure de ces méthodes de mécanique moléculaire. Voyons maintenant la forme analytique des différentes composantes du champ de forces Amber. Elles sont de deux types différents, selon qu'elles mettent en jeu des interactions entre atomes liés ou non. On parlera de termes liés pour les interactions entre atomes connectés par une, deux ou trois liaisons (« condition 1-4 ») et de termes non-liés pour les autres interactions. Cependant, les interactions entre atomes séparés par trois liaisons sont comptées à la fois par le terme lié correspondant et les termes non-liés : un facteur est alors utilisé pour atténuer un des termes par rapport aux autres.

Souvent les termes liés sont représentés par des potentiels harmoniques.

- Terme de liaison (atomes connectés par une liaison)  $E_l = k_l(r - r_0)^2$  avec  $k_l$  constante de raideur de la liaison,  $r - r_0$  variation de longueur de liaison par rapport à une valeur de référence  $r_0$ .
- Terme angulaire (atomes connectés par deux liaisons)  $E_a = k_a(\theta - \theta_0)^2$  avec  $k_a$  constante de force de l'angle de valence,  $\theta - \theta_0$  variation d'angle par rapport à une valeur référence  $\theta_0$ .
- Terme de torsion (atomes connectés par trois liaisons)  $E_t = \frac{V_n}{2}[1 + \cos(n\phi - \gamma)]$  avec  $V_n$  la barrière de rotation d'ordre  $n$ ,  $\phi$  l'angle de torsion et  $\gamma$  le déphasage de celui-ci par rapport à l'origine.

Les termes non-liés sont de deux types : les interactions électrostatiques et celles de van der Waals.

- Terme électrostatique  $E_{ij}^q = \frac{q_i q_j}{\epsilon r_{ij}}$  avec  $q_i$  et  $q_j$  les charges ponctuelles définies sur les atomes  $i$  et  $j$ ,  $r_{ij}$  la distance entre ces deux atomes et  $\epsilon$  la constante diélectrique du milieu considéré.
- Terme de van der Waals  $E_{ij}^v = \frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^6}$  avec  $R_{ij}$  la distance entre les atomes  $i$  et  $j$ .

L'énergie totale d'un système est alors donnée par la somme de toutes ces contributions :

$$E_{MM} = \sum_l E_l + \sum_a E_a + \sum_t E_t + \sum_i \sum_{j>i} (E_{ij}^q + E_{ij}^v) \quad (3.2)$$

Dans des champs de forces plus élaborés, d'autres termes peuvent apparaître (termes liés croisés, liaisons hydrogène, . . . ) ; les termes liés peuvent être également de forme anharmonique ; des polarisabilités peuvent être introduites pour permettre la polarisation des atomes.

Du fait de sa simplicité, la mécanique moléculaire permet de traiter un nombre important d'atomes (plusieurs milliers) en des temps de calcul relativement courts. En revanche, les électrons n'étant pas traités explicitement et la définition de la connectivité du système étant requise, les processus électroniques comme les créations ou les ruptures de liaisons peuvent être difficilement modélisés par cette approche.

### 3.1.2 Les méthodes de type ONIOM

Les méthodes appelées ONIOM (également IMOMM, IMOMO) [8, 9], développées par Maseras et Morokuma reposent à la fois sur une partition géométrique du système et sur une hiérarchie des niveaux théoriques utilisés. Ce ne sont pas uniquement des méthodes QM/MM et c'est pourquoi nous les traitons à part. L'énergie du macrosystème est le résultat de la combinaison de trois autres énergies :

$$\mathbf{E} = \mathbf{E}_1 - \mathbf{E}_2 + \mathbf{E}_3 \quad (3.3)$$

où  $\mathbf{E}_1$  est l'énergie du sous-système réactionnel traité à un haut niveau de calcul quantique,  $\mathbf{E}_2$  est l'énergie de ce même sous-système traité à un bas niveau de calcul (quantique ou classique) et  $\mathbf{E}_3$  est l'énergie du macrosystème entier au même bas niveau de calcul. Immédiatement, on s'aperçoit que les interactions entre le centre réactionnel et son environnement ne sont prises en compte qu'au bas niveau de théorie. Particulièrement, si ce dernier est un champ de forces classique, la fonction d'onde du centre réactif n'est pas polarisée par l'environnement. Seul son effet stérique est pris en compte. Si celui-ci est prépondérant, alors ONIOM est la méthode de choix car elle est simple à mettre en oeuvre. En revanche, si les effets électroniques sont plus importants que les effets stériques, il faut alors soit utiliser

une méthode quantique comme bas niveau de calcul, soit utiliser une approche QM/MM incluant la polarisation de la partie quantique, comme celles que nous décrivons après.

Ces méthodes ont été appliquées à différentes études : réactions organiques [10], réactions organométalliques [11], modélisation de surfaces [12], . . . Woo, Cavallo et Ziegler ont également utilisé IMOMM (ONIOM avec la mécanique moléculaire comme bas niveau de théorie) pour des simulations dynamiques [13].

### 3.1.3 Les méthodes de type soluté-solvant

Lorsque la réaction chimique concerne un soluté de taille raisonnable en phase liquide, l'application du formalisme QM/MM est directe. Le soluté est traité quantiquement et le solvant classiquement. L'Hamiltonien d'interaction QM/MM est alors réduit à des termes non-liés, c'est-à-dire van der Waals et électrostatique. Les modèles développés sont plus ou moins élaborés. Il y a quelques années, Gao a publié une revue qui explore ces possibilités [14]. Le soluté peut être traité au niveau *ab initio*, semi-empirique ou DFT. Dès 1993, Wesolowski et Warshel proposaient de représenter le solvant comme une densité gelée dans un calcul DFT [15, 16]. A Nancy, sous l'impulsion de M.F. Ruiz-López, un programme de dynamique moléculaire DFT/MM a été mis au point et a servi à de nombreuses applications [17, 18, 19]. Les modèles de solvant explicites sont nombreux : rigides ou flexibles, polarisables ou non, à charge fluctuante [20]. Récemment, un modèle à trois couches a été développé qui combine une description DFT du soluté avec une description mixte du solvant MM + continuum (dans cette dernière, le solvant est considéré comme un milieu continu pourvu d'une constante diélectrique, polarisé par le soluté et le polarisant en retour) [21].

### 3.1.4 Les méthodes de type macromoléculaire

Les réactions chimiques se déroulant dans des milieux étendus comme des macromolécules, des solides ou sur des surfaces sont plus difficiles à traiter dans un formalisme QM/MM. En effet, la séparation entre les deux sous-systèmes nécessite de couper une ou plusieurs liaisons chimiques. La fragment quantique résultant de ces coupures est alors un radical dont les propriétés sont différentes de celles du fragment dans son environnement. Ce problème de liaisons frontières pendantes a donné lieu à de nombreux développements méthodologiques qui ont tenté, et tentent toujours, de décrire au mieux cette frontière entre les deux descriptions quantique et classique.

L'approche la plus simple, en apparence, consiste à saturer les liaisons que l'on coupe par des atomes monovalents : c'est la méthode du « link atom » (introduite initialement par Field, Bash et Karplus [22])

que l'on peut traduire en français par « atome de liaison » ou « atome de saturation ». L'atome monovalent le plus évident est l'atome d'hydrogène. Si la mise en oeuvre de ce schéma est relativement simple, plusieurs objections conceptuelles peuvent être formulées. L'introduction d'un atome supplémentaire augmente le nombre de degrés de liberté, même si l'atome est contraint à rester dans l'alignement de la liaison que l'on coupe. La surface d'énergie potentielle qui en résulte n'a plus les dimensions de celle concernant le système réel. De plus, la polarité de la liaison coupée peut être difficilement reproduite, surtout si celle-ci est largement polaire. Enfin, si les interactions entre l'atome de saturation et les atomes de l'environnement classique sont comptées, celles-ci n'ont aucun sens physique ; par contre, si elles ne sont pas comptées, la polarisation de la liaison frontière peut être complètement erronée. D'autres questions se posent également : l'atome de liaison doit-il être mobile ou doit-il être fixe ? Doit-il rester dans la direction de la liaison ? Certains schémas tentent de proposer une approche plus rationnelle comme la méthode SPLAM [23]. Malgré toutes ces difficultés conceptuelles, sa simplicité de mise en oeuvre fait du « link atom » la méthode la plus utilisée comme description de la frontière QM/MM et ce, dans de nombreuses applications variées [24, 25, 26, 27, 28, 29, 30, 31].

Des tentatives ont été faites pour corriger les défauts mentionnés en remplaçant l'atome d'hydrogène, soit par des « pseudo-halogènes » [32], soit par des groupements atomiques fantômes [33]. Antes et Thiel ont mis au point « l'atome de connection ajusté » [34] qui est un atome quantique auquel on ajoute des paramètres classiques supplémentaires afin de connecter les deux sous-systèmes. Dans le même esprit, Zhang, Lee et Yang ont développé l'approche de la « pseudo-liaison » [35, 36] : une liaison carbone-carbone est mimée par une liaison « pseudo-carbone » – carbone, l'atome de « pseudo-carbone » étant pourvu de sept électrons de valence et d'un potentiel de coeur effectif paramétré spécialement pour décrire la liaison coupée. De la même façon, Poteau et collaborateurs [37, 38] ont développé une approche de la frontière basée sur la paramétrisation de potentiels effectifs de groupe.

Une autre solution pour saturer les liaisons pendantes et décrire au mieux la frontière QM/MM, la première chronologiquement, a été proposée par Warshel et Levitt [39] : elle consistait à utiliser des orbitales hybrides de liaison pour décrire la liaison coupée. Rivail et collaborateurs ont repris ces idées et formulé la méthode « Local Self Consistent Field » (LSCF) au niveau semi-empirique [40, 41], puis au niveau *ab initio* [42, 43]. S'inspirant de LSCF, Gao a développé une méthode très voisine, dite « generalized hybrid orbital » [44, 45, 46]. Friesner a également repris les principes LSCF et a développé son propre champ de forces hybride [47, 48, 49]. Jensen et collaborateurs utilisent des orbitales localisées comme tampon entre le fragment quantique et l'environnement représenté par un potentiel effectif [50, 51]. A Nancy, les applications de LSCF dans les méthodes QM/MM ont par exemple concerné

l'hydrolyse de la thermolysine [52].

Peu de comparaisons ont été faites entre les différentes approches de la frontière QM/MM. Monard et Merz, Jr. les ont résumées en 1999 [53]. Reuter *et al.* [54] ainsi que Hall et Hillier [55] ont comparé plus directement les différentes approches, mais il semble qu'aucune n'ait l'avantage sur les autres. Cependant, leurs résultats se limitent uniquement au cas le plus favorable de coupure sur une liaison purement covalente entre atomes de carbone et ne peuvent être étendus à toutes les situations envisageables (notamment dans les cas de liaisons polaires, ioniques, datives . . . ).

A travers le développement de la méthode LSCF/MM *ab initio* et son application à divers problèmes chimiques, nous tenterons de proposer une solution sophistiquée au problème de la liaison frontière dans les simulations QM/MM.

## 3.2 La méthode LSCF/MM *ab initio*

Les liaisons qui relient les sous-systèmes quantique et classique sont décrites par des orbitales de liaison strictement localisées (SLBO). Ces orbitales sont gelées grâce au formalisme LSCF présenté dans la partie précédente. Dans cette partie, nous verrons comment sont déterminées ces SLBO et nous analyserons les erreurs causées par leur introduction dans un calcul quantique et les moyens d'y remédier. Nous verrons également que les SLBO sont facilement dérivables et permettent de calculer les forces agissant sur tous les atomes du système.

Afin d'optimiser complètement la géométrie du macrosystème, un potentiel de frontière a été mis au point. Celui-ci ne dépend que de la nature des atomes qui portent une SLBO. Nous verrons enfin que l'ajustement des charges ponctuelles à la frontière permet d'obtenir une bonne description énergétique du système réactionnel.

### 3.2.1 Les orbitales de liaison strictement localisées

En accord avec l'hypothèse de transférabilité des propriétés des liaisons, les SLBO sont déterminées sur des molécules modèles représentatives de ces liaisons. Nous choisirons toujours la molécule modèle la plus simple possible (par exemple, le formamide pour une liaison peptidique). La détermination de ces orbitales de liaison localisées est simple : une procédure de localisation est appliquée sur les orbitales moléculaires canoniques de la molécule modèle. Le choix du critère de localisation est discuté dans la suite. Puis les coefficients des orbitales atomiques (OA) des atomes ne portant pas la liaison sont annulés et la SLBO obtenue est renormée.

Une SLBO ne possède donc que des coefficients non-nuls pour les OA centrées sur les atomes qui définissent la liaison. On peut considérer cette SLBO comme un vecteur défini par ces deux atomes : cette caractéristique nous permettra de pivoter les SLBO dans l'espace et de calculer leurs dérivées premières par rapport aux coordonnées nucléaires.

**Procédures de localisation.** Il existe de nombreuses méthodes de localisation des OM. Les plus répandues dépendent de la minimisation ou de la maximisation d'un critère fixé a priori. Ces critères peuvent être de deux types : global ou local, c'est-à-dire qui localise l'ensemble des orbitales moléculaires ou qui localise uniquement une orbitale sans s'occuper des autres.

Nous commencerons par citer trois méthodes globales. La première, la méthode d'Edminston et de Ruedenberg [56] repose sur la maximisation du critère  $D = \sum_l J_{ll}$ , c'est-à-dire sur la maximisation des interactions électroniques à l'intérieur d'une orbitale ( $J$  représente l'interaction coulombienne). Cependant, cette méthode est très coûteuse en temps de calcul puisqu'elle nécessite de calculer de nombreuses intégrales biélectroniques.

La méthode de Boys et Foster [57] cherche à maximiser la distance entre les barycentres des orbitales. Cela revient alors à minimiser le critère  $\sum_l \langle \varphi_l | \frac{1}{r^2} | \varphi_l \rangle$ . Ces intégrales sont monoélectroniques et sont donc plus rapides à calculer. Ceci explique que cette méthode soit plus utilisée que la précédente.

Enfin, la méthode de Pipek et Mezey [58] repose sur la minimisation de la délocalisation des orbitales moléculaires. Cette méthode a l'avantage de conserver la séparation des orbitales de type  $\sigma$  et celles de type  $\pi$ .

Cependant, ces critères globaux ne sont pas les mieux adaptés à notre problème : nous ne sommes intéressés que par la localisation d'une seule orbitale de liaison, sans considération pour les autres orbitales. L'utilisation de critères locaux à une liaison (on l'appellera dans la suite liaison X-Y) semble plus appropriée. Particulièrement, nous avons utilisé les méthodes de Weinstein-Pauncz et de Magnasco-Perico [59]. La méthode de Weinstein-Pauncz (CWP pour critère Weinstein-Pauncz) maximise le critère

$$\sum_{\mu \in X, \nu \in Y} c_{\mu l} S_{\mu\nu} c_{\nu l} \quad (3.4)$$

et celle de Magnasco-Perico (CMP pour critère Magnasco-Perico) est très semblable :

$$\sum_{\mu, \nu \in X, Y} c_{\mu l} S_{\mu\nu} c_{\nu l} \quad (3.5)$$

Dans la première méthode, la somme porte sur les orbitales atomiques de X et Y séparément alors que

la deuxième fait la somme sans distinction de l'appartenance des OA à X ou Y (elle correspond à la norme d'une orbitale de liaison strictement localisée). En conséquence, la méthode CMP peut éventuellement localiser un doublet libre alors que cela s'avère impossible avec la méthode CWP. Cette différence se traduit par le caractère plutôt covalent des SLBO obtenues avec CWP et par le caractère plus ionique des liaisons obtenues à partir du critère CMP.

En conséquence, nous utiliserons le critère CWP pour être certain d'obtenir une orbitale de liaison. Si le besoin s'en fait sentir, ces orbitales peuvent être raffinées avec le critère CMP. Par la suite, nous noterons CWP et CWP+CMP les deux procédures de localisation retenues.

Une fois le critère de localisation choisi, celui-ci est calculé pour chaque orbitale moléculaire et on classe les orbitales moléculaires par valeur de critère décroissante. La première est donc celle que l'on veut localiser. Elle est ensuite mélangée à une autre orbitale à l'aide d'une matrice unitaire de rotation  $2 \times 2$

$$\begin{aligned} |\varphi'_1\rangle &= \cos \alpha |\varphi_1\rangle + \sin \alpha |\varphi_i\rangle \\ |\varphi'_i\rangle &= -\sin \alpha |\varphi_1\rangle + \cos \alpha |\varphi_i\rangle \end{aligned} \quad (3.6)$$

où  $\alpha$  est optimisé afin de maximiser la valeur du critère. Cette procédure est répétée jusqu'à ce que la valeur du critère n'augmente plus de façon significative.

Tous ces critères de localisation sont utilisés a posteriori sur la fonction d'onde SCF, c'est-à-dire sur les orbitales canoniques. D'autres méthodes existent qui permettent d'obtenir directement les SLBO. Citons par exemple les travaux de Couty [60] et ceux de Sironi [61, 62].

**Rotation des SLBO.** Lors de l'optimisation de la géométrie moléculaire, les atomes qui portent une SLBO sont libres de bouger et la SLBO doit s'adapter à ce déplacement. Il est donc nécessaire de définir une procédure qui permette cette adaptation.

L'énergie LSCF dépend fortement de la définition des SLBO, c'est-à-dire de leur forme, de leur polarité, etc... Le choix initial des SLBO que l'on gèle définit une hypersurface de potentiel que l'on ne doit plus quitter durant l'optimisation de géométrie. Il est donc très important que la polarité de la liaison ainsi que le rapport des interactions  $\sigma/\pi$  ne varient pas lorsque la SLBO est adaptée à la nouvelle géométrie. Dans le cas contraire, le système pourrait facilement passer d'une hypersurface à une autre et ne jamais atteindre la convergence.

Comme remarqué précédemment, une SLBO étant définie par deux atomes uniquement (c'est donc un vecteur), la réorientation de la SLBO s'opère facilement grâce à l'utilisation de matrices de rotation



adaptées. Les axes de rotation doivent être définis de manière univoque tout au long de l'optimisation : trois atomes sont alors nécessaires pour définir un repère cartésien local qui respecte l'unicité des axes. Les deux premiers sont naturellement les deux atomes définissant la liaison et le troisième est choisi arbitrairement (à condition qu'il ne soit pas colinéaire aux deux premiers). Nous avons repris la méthodologie que propose l'équipe de Friesner [47]. Notant X et Y les deux atomes portant la liaison et R le troisième atome, le repère local est défini par  $(\vec{A}, \vec{B}, \vec{C})$  tel que  $\vec{A} = \overrightarrow{XY}$ ,  $\vec{B} = \overrightarrow{XY} \wedge \overrightarrow{XR}$  et  $\vec{C} = \vec{B} \wedge \overrightarrow{XY}$  (cf. figure 3.1).

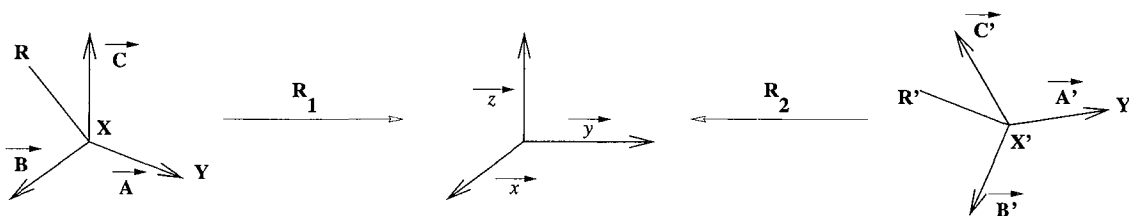


FIG. 3.1 – *Référentiels local et absolu*

Dans la nouvelle géométrie, ce repère est défini de la même façon par  $(\vec{A}', \vec{B}', \vec{C}')$ . Les deux matrices de rotation  $R_1$  et  $R_2$  qui permettent de passer de ces repères locaux au repère cartésien absolu  $(\vec{x}, \vec{y}, \vec{z})$  sont données par

$$R_1 = \begin{pmatrix} \frac{c_x}{c} & \frac{c_y}{c} & \frac{c_z}{c} \\ \frac{b_x}{b} & \frac{b_y}{b} & \frac{b_z}{b} \\ \frac{a_x}{a} & \frac{a_y}{a} & \frac{a_z}{a} \end{pmatrix} \quad (3.7)$$

et

$$R_2 = \begin{pmatrix} \frac{c'_x}{c'} & \frac{c'_y}{c'} & \frac{c'_z}{c'} \\ \frac{b'_x}{b'} & \frac{b'_y}{b'} & \frac{b'_z}{b'} \\ \frac{a'_x}{a'} & \frac{a'_y}{a'} & \frac{a'_z}{a'} \end{pmatrix} \quad (3.8)$$

telles que  $\vec{A} = a_x \vec{x} + a_y \vec{y} + a_z \vec{z}$  et  $a = \|\vec{A}\|$ . La rotation qui permet de passer du repère local initial au repère local final est donc  $R = R_1 \cdot R_2^{-1}$ . On peut maintenant procéder à la rotation des fonctions de base impliquées dans la SLBO. Il est évident que les fonctions de type  $s$  ne sont pas modifiées, elles sont invariantes par rotation grâce à leur symétrie sphérique. Un jeu de fonctions de type  $p$  se transforme comme  $(\vec{x}, \vec{y}, \vec{z})$ , on peut donc lui appliquer directement la matrice de rotation  $R$  définie ci-dessus.

Nous la noterons

$$R = \begin{pmatrix} R_{xx} & R_{xy} & R_{xz} \\ R_{yx} & R_{yy} & R_{yz} \\ R_{zx} & R_{zy} & R_{zz} \end{pmatrix} \quad (3.9)$$

Les matrices de rotation pour les fonctions de moment angulaire supérieur à 1 peuvent être déduites de  $R$ . Elles sont construites pour des fonctions de type  $d$  et  $f$ . Par exemple, l'expression du coefficient de la fonction de base  $d_{x'x'}$  dans la SLBO est la suivante

$$d_{x'x'} = p_{x'} \cdot p_{x'} = (R_{xx}p_x + R_{xy}p_y + R_{xz}p_z) \cdot (R_{xx}p_x + R_{xy}p_y + R_{xz}p_z) \quad (3.10)$$

Développant cette expression, on obtient

$$d_{x'x'} = R_{xx}^2 d_{xx} + R_{xy}^2 d_{yy} + R_{xz}^2 d_{zz} + 2R_{xx}R_{xy}d_{xy} + 2R_{xx}R_{xz}d_{xz} + 2R_{xy}R_{xz}d_{yz} \quad (3.11)$$

Opérant de même avec les autres fonctions de base, on forme une matrice  $6 \times 6$  pour les fonctions de type  $d$  et une matrice  $10 \times 10$  pour les fonctions de type  $f$ . On doit noter que ces transformations ne sont valides que si la base de fonctions est purement cartésienne. Si elle est de type sphérique (cinq fonctions de type  $d$  et sept fonctions de type  $f$ ), il faut préalablement transformer la base en base cartésienne, opérer les rotations puis retourner dans la base originale [63].

Le dernier atome du trièdre de référence ne peut être choisi au hasard : il doit permettre une bonne représentation de la dissymétrie de révolution autour de la liaison. Prenons comme exemple la rotation autour d'une liaison carbone-carbone dans l'éthane. Puisque la SLBO possède la symétrie de révolution autour de la liaison C-C, tous les atomes d'hydrogène sont équivalents, en choisir un arbitrairement n'aura aucune incidence. En revanche, dans le cas de la rotation autour de la liaison carbone-oxygène du méthanol, la symétrie axiale est perdue à cause des doublets non-liants de l'oxygène : la SLBO est fortement asymétrique. Le choix d'un hydrogène du groupement méthyle comme troisième atome ne permet pas de prendre en compte cette asymétrie, ce qui aboutira à un mauvais profil énergétique de rotation autour de la liaison C-O (figure 3.2).

En revanche, le choix de l'hydrogène alcoolique permet de garder cette asymétrie et le profil énergétique est alors en très bon accord avec celui obtenu avec un calcul SCF.

Graphiquement (figure 3.3), on remarque que la SLBO n'a pas suivi le mouvement de rotation du proton alcoolique dans le premier cas (noté H(C)) alors qu'elle s'est bien « retournée » dans le deuxième

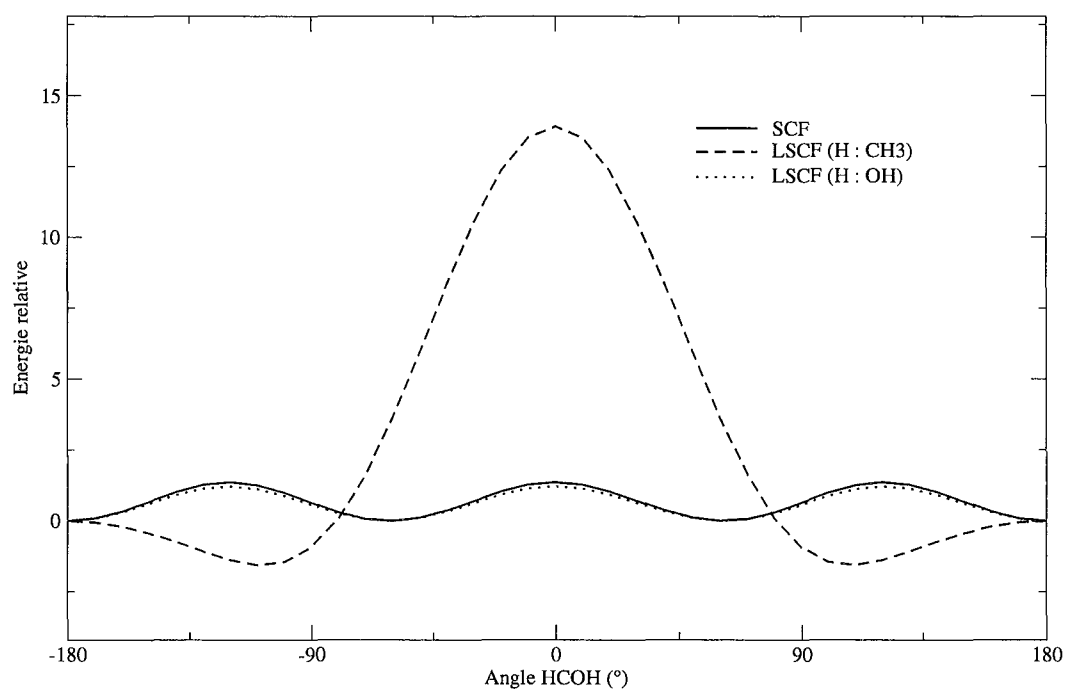


FIG. 3.2 – Energie HF/6-31G\* du méthanol en fonction de l'angle diédral H-C-O-H. La courbe en tirets correspond au profil énergétique dans le cas où un hydrogène du groupement méthyle est choisi comme troisième atome, la courbe en pointillés correspond au profil obtenu pour l'hydrogène alcoolique pris comme troisième atome (les deux premiers atomes du trièdre sont le carbone et l'oxygène).

cas (noté H(O)) pour une rotation de  $180^\circ$  de l'angle dièdre.

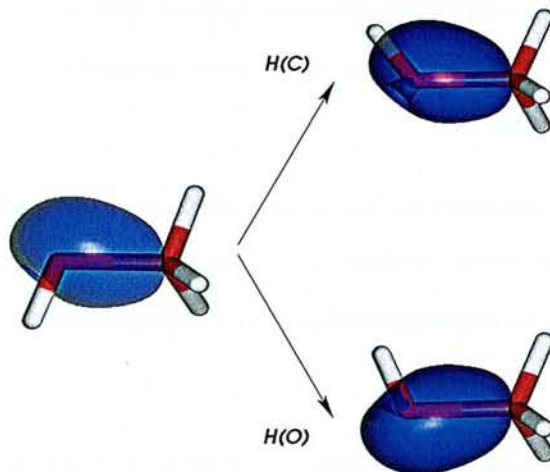


FIG. 3.3 – Evolution de la SLBO définie sur la liaison C-O du méthanol.

La procédure exposée ci-dessus permet de prendre en compte la réorientation dans l'espace d'une SLBO. Concernant le changement de longueur de la liaison qui porte la SLBO, une simple renormation de la SLBO est suffisante.

**Dérivées premières des SLBO.** Pour que les atomes frontières puissent bouger librement, il faut calculer les dérivées premières des SLBO par rapport aux coordonnées nucléaires. Dans leur version de la méthode LSCF/MM, Friesner et son groupe proposent la méthode suivante [47]. Pour une  $n^{\text{ième}}$  géométrie donnée, une SLBO est déduite de la même SLBO dans la géométrie précédente ( $n - 1$ ) par une rotation  $R_{n-1}$  :

$$SLBO_n = R_{n-1} \cdot SLBO_{n-1} \quad (3.12)$$

Les dérivées de  $SLBO_n$  s'obtiennent à l'aide des dérivées de  $R_{n-1}$  et de celles de  $SLBO_{n-1}$ . Ces dernières étaient obtenues à partir de  $R_{n-2}$  et de celles de  $SLBO_{n-2}$ . Et ainsi de suite jusqu'à la situation de départ dans la géométrie de la molécule modèle (notée  $SLBO_0$ ). A ce point, Friesner fait une hypothèse surprenante : les dérivées de  $SLBO_0$  sont nulles ! Les dérivées de  $SLBO_n$  s'obtiennent donc à partir de  $n$  dérivées des matrices de rotation  $R_i$  pour  $i = 0, n - 1$ .

Il nous semble évident que l'hypothèse de la nullité des dérivées des SLBO dans la situation modèle est fautive. *Les coefficients des SLBO sont des constantes durant un calcul SCF d'énergie pour une géométrie donnée. En revanche, il n'y a aucune raison que les dérivées par rapport aux coordonnées nucléaires soient nulles, même dans la situation de la molécule modèle.*

A notre connaissance, il n'y a pas de moyens de calculer ces dérivées autrement que numériquement. Deux possibilités s'offrent à nous :

- nous calculons numériquement les dérivées de  $SLBO_0$  (dans la situation modèle) une fois pour toutes et nous les conservons. A chaque pas d'optimisation de la géométrie, nous calculons les dérivées des matrices de rotation qui font passer de  $SLBO_0$  à  $SLBO_n$

ou

- nous calculons numériquement les dérivées de  $SLBO_n$  à chaque pas d'optimisation sans nous servir des matrices de rotation.

Comme le calcul des dérivées numériques de  $SLBO_n$  est très rapide, nous avons choisi la deuxième possibilité.

Evidemment, ces dérivées sont non nulles uniquement pour les trois atomes définissant le repère local attribué à chaque SLBO. Chacun de ces trois atomes est déplacé selon un des axes du repère cartésien absolu d'une quantité  $\delta\alpha$ , puis la SLBO est tournée en conséquence et renormée. Ainsi, les dérivées sont obtenues par différenciation numérique

$$\partial a_{\mu l}^{\alpha} = \frac{a_{\mu l}^{\alpha+\delta\alpha} - a_{\mu l}^{\alpha-\delta\alpha}}{2\delta\alpha} \quad (3.13)$$

où  $a_{\mu l}^{\alpha}$  est le coefficient de l'OA  $\mu$  dans la SLBO  $l$  à la position  $\alpha$  ( $\alpha = x, y, z$ ). Une fois ces dérivées numériques calculées, on peut les utiliser dans la forme analytique des dérivées de l'énergie électronique LSCF (chapitre précédent).

Les SLBO ne sont donc jamais considérées comme figées, elles s'adaptent à n'importe quelle modification géométrique. Les seules propriétés qui restent constantes sont la polarité de la liaison et le rapport des interactions  $\sigma/\pi$ .

**Les erreurs dues à l'utilisation des SLBO.** Avant d'étudier les détails du développement de notre méthode QM/MM, il faut d'abord se pencher sur les conséquences de l'utilisation des SLBO dans un calcul purement quantique. Nous pourrions ainsi séparer les erreurs provenant du modèle QM/MM de celles introduites par l'utilisation d'une SLBO. Pour cela, quelques petites molécules modèles ont été

sélectionnées comme représentatives d'un certain nombre de liaisons courantes. Nous nous sommes limités à l'étude de petites molécules pour que l'effet des SLBO soit maximal. La géométrie de chacune de ces molécules a été optimisée dans un premier temps avec un calcul SCF puis, dans un deuxième temps un calcul d'énergie LSCF sur la géométrie précédente est réalisé en gelant une SLBO obtenue selon les critères de localisation CWP et CWP+CMP.

Nous avons choisi le méthane (liaison C-H), l'ammoniac (liaison N-H), l'eau (liaison O-H), l'éthane (liaison C-C) et le méthanol (liaison C-O). Dans le tableau I sont reportés les résultats sur l'énergie, les charges de Mulliken et le moment dipolaire.

Nous constatons que l'introduction d'une SLBO dans un calcul quantique entraîne une erreur sur l'énergie entre 2 et 8 kcal.mol<sup>-1</sup>. On peut remarquer que cette différence est toujours positive (l'énergie LSCF est supérieure à l'énergie SCF), ce qui respecte le principe variationnel. Ces erreurs sont plus grandes lorsque le raffinement CMP est ajouté à CWP. Par contre CMP a peu d'influence sur les charges de Mulliken et sur les moments dipolaires obtenus. Il est évident qu'il est inutile d'introduire un caractère ionique dans une SLBO devant décrire une liaison purement covalente (C-H dans le méthane et C-C dans l'éthane). Mais il semble aussi que ce raffinement ait un effet trop marqué sur des liaisons polaires de type O-H ou C-O, puisque le moment dipolaire de l'eau et du méthanol devient alors trop grand. En conséquence, le critère CWP+CMP devra être utilisé avec discernement et uniquement dans le cas de liaisons au caractère ionique très marqué.

L'erreur sur l'énergie absolue peut paraître inacceptable au premier abord. Cependant, lors de réactions chimiques, on est intéressé uniquement par des énergies relatives de réaction. Plutôt que les énergies absolues LSCF soient en bon accord avec les énergies absolues SCF, il importe davantage que l'erreur énergétique introduite par l'utilisation d'une SLBO reste constante tout au long du chemin réactionnel. Dans la suite de nos travaux, nous verrons de nombreuses fois que cette condition est vérifiée, tant que la SLBO n'est pas immédiatement impliquée dans la réaction chimique étudiée. Cependant, nous pouvons déjà justifier cette condition avec l'exemple suivant. Considérons l'ionisation du méthane : au niveau de théorie HF/6-311++G\*\*, le résultat SCF est 13,317 eV. Le résultat LSCF (une SLBO sur une liaison C-H) est 13,314 eV. La différence de 3.10<sup>-3</sup> eV justifie pleinement l'hypothèse énoncée ci-dessus.

Pour le méthane, l'introduction d'une SLBO sur une liaison C-H brise la symétrie T<sub>d</sub> de la molécule. Ceci provoque l'apparition d'un moment dipolaire, inacceptable pour décrire correctement le méthane. Pour résoudre ce problème, il est possible de modifier la SLBO en multipliant les coefficients des OA du carbone par un facteur, et donc de modifier la polarité de la liaison représentée par la SLBO. Des optimisations de géométrie du méthane ont été réalisées au niveau de théorie HF/6-31G\* en prenant

TAB. I – Comparaison de quelques propriétés électroniques pour des calculs SCF et LSCF à la géométrie SCF ( $E$ , énergie absolue (en hartrees), est obtenue au niveau de théorie RHF/6-31G\*,  $\Delta E$  est la différence entre les énergies LSCF and SCF (en kcal.mol<sup>-1</sup>),  $q(X)$  est la charge de Mulliken de l'atome  $X$  (en électrons),  $\mu$  est le moment dipolaire de la molécule (en Debye))

	SCF	LSCF (CWP)	LSCF (CWP+CMP)
CH <sub>4</sub>	E=-40,195171  q(C)=-0,661 q(H)=0,165  $\mu=0$	E=-40,191399 $\Delta E=2,367$ q(C)=-0,646 q(H)=0,160 q(H <sub>SLBO</sub> )=0,165 $\mu=0,125$	E=-40,191199 $\Delta E=2,492$ q(C)=-0,647 q(H)=0,161 q(H <sub>SLBO</sub> )=0,165 $\mu=0,129$
NH <sub>3</sub>	E=-56,184356  q(N)=-0,996 q(H)=0,332  $\mu=1,915$	E=-56,178469 $\Delta E=3,694$ q(N)=-0,984 q(H)=0,329 q(H <sub>SLBO</sub> )=0,326 $\mu=1,908$	E=-56,177760 $\Delta E=4,139$ q(N)=-0,981 q(H)=0,324 q(H <sub>SLBO</sub> )=0,333 $\mu=2,096$
H <sub>2</sub> O	E=-76,010747  q(O)=-0,868 q(H)=0,434  $\mu=2,199$	E=-76,006970 $\Delta E=2,370$ q(O)=-0,864 q(H)=0,435 q(H <sub>SLBO</sub> )=0,429 $\mu=2,175$	E=-76,001688 $\Delta E=5,685$ q(O)=-0,868 q(H)=0,433 q(H <sub>SLBO</sub> )=0,435 $\mu=2,416$
C <sub>2</sub> H <sub>6</sub>	E=-79,228755  q(C)=-0,476 q(H)=0,159  $\mu=0$	E=-79,220787 $\Delta E=5,000$ q(C)=-0,464 q(H)=0,155  $\mu=0$	E=-79,219849 $\Delta E=5,589$ q(C)=-0,466 q(H)=0,155  $\mu=0$
CH <sub>3</sub> OH	E=-115,035418  q(C)=-0,166 q(H)=0,174 q(H)=0,141 q(O)=-0,726 q(H)=0,437 $\mu=1,868$	E=-115,028455 $\Delta E=4,369$ q(C)=-0,169 q(H)=0,172 q(H)=0,135 q(O)=-0,714 q(H)=0,440 $\mu=1,848$	E=-115,023755 $\Delta E=7,319$ q(C)=-0,163 q(H)=0,172 q(H)=0,137 q(O)=-0,722 q(H)=0,437 $\mu=1,978$

pour le facteur multiplicatif des valeurs comprises entre 1,5 et 0,5. Les longueurs de liaison, les charges de Mulliken, les moments dipolaires et les énergies sont reportés dans le tableau II.

TAB. II – Influence de la polarité d'une SLBO C-H sur des propriétés du méthane ( $d(C-H)_1$  est la longueur de la liaison décrite par la SLBO et  $d(C-H)_2$  est la longueur des autres liaisons (en Å),  $q_1$  est la charge de Mulliken de H (SLBO) et  $q_2$  est la charge de Mulliken des autres hydrogènes (en électron),  $\mu$  est le moment dipolaire du méthane,  $E$ , énergie absolue (en hartrees), est obtenue au niveau RHF/6-31G\*,  $\Delta E$  est la différence entre les énergies LSCF and SCF (en kcal.mol<sup>-1</sup>))

Facteur	$d(C-H)_1$	$d(C-H)_2$	$q_1$	$q_2$	$\mu$	$E$	$\Delta E$
SCF	1,083	1,083	0,165	0,165	0	-40,195171	0
1,5	1,070	1,084	0,384	0,150	0,614	-40,177094	11,343
1,25	1,081	1,083	0,289	0,154	0,305	-40,186839	5,228
1,15	1,086	1,083	0,244	0,156	0,151	-40,19606	3,492
1,1	1,088	1,083	0,219	0,157	0,067	-40,190590	2,875
1,07	1,089	1,082	0,203	0,158	0,013	-40,191020	2,605
1,06	1,090	1,082	0,198	0,158	0,006	-40,191134	2,533
1,05	1,090	1,082	0,193	0,159	0,024	-40,191232	2,472
1,02	1,091	1,082	0,176	0,159	0,081	-40,191426	2,350
1,0	1,092	1,082	0,165	0,160	0,121	-40,191467	2,324
0,9	1,094	1,081	0,104	0,163	0,335	-40,190370	3,013
0,75	1,095	1,080	-0,003	0,169	0,713	-40,183126	7,558
0,5	1,081	1,078	-0,230	0,181	1,502	-40,144226	31,968

Tout d'abord, remarquons qu'il n'est pas possible de décrire correctement toutes les propriétés avec le même facteur. Pour obtenir de bonnes longueurs de liaisons, il faut un facteur proche de 1,25, mais les charges de Mulliken et le moment dipolaire ne sont pas satisfaisants. Pour annuler le moment dipolaire, le facteur doit être égal à 1,06. Les longueurs de liaisons et l'énergie sont correctes, mais les charges de Mulliken sont un peu trop grandes. Avec un facteur proche de 1, on obtient une bonne énergie et des charges de Mulliken en bon accord avec les résultats SCF. Mais à cause de la compression de la liaison portant la SLBO, le moment dipolaire n'est pas nul.

Il semble donc nécessaire de déterminer le meilleur facteur au cas par cas, en fonction des propriétés que l'on veut reproduire. Cependant, le cas du méthane est particulier et il semble que le facteur unité soit suffisamment adapté à la majorité des cas que l'on rencontrera par la suite. Il était néanmoins utile de signaler cette possibilité offerte à l'utilisateur de LSCF.

Pour conclure sur les effets induits par une SLBO dans un calcul quantique, il faut donc être très prudent quant à l'utilisation du critère CWP+CMP. Ainsi, pour les liaisons organiques habituelles, il est



recommandé d'utiliser le critère CWP uniquement. Les SLBO obtenues avec ce critère semblent bien reproduire les propriétés de la liaison, sauf cas particulier où il est alors possible de modifier la polarité de la liaison.

### 3.2.2 Le champ de forces hybride LSCF/MM

L'Hamiltonien d'un système QM/MM est la somme de trois termes (équation 3.1), en conséquence de quoi l'énergie du système est donnée par :

$$E = E_{QM} + E_{MM} + E_{QM/MM} \quad (3.14)$$

Le premier terme est l'énergie du sous-système quantique (incluant la répulsion nucléaire), le deuxième terme est l'énergie du sous-système classique (décrit par un champ de forces comme exposé précédemment). Nous allons nous intéresser plus particulièrement au troisième terme qui représente l'énergie d'interaction entre les deux sous-systèmes. Avant d'aborder les différentes contributions qui entrent dans ce terme, nous allons définir une notation qui permettra d'identifier sans ambiguïté les atomes de la macromolécule. D'après la figure 3.4, les atomes classiques sont notés C, les atomes quantiques Q. Pour une liaison frontière, on notera X l'atome du côté quantique et Y l'atome du côté classique.

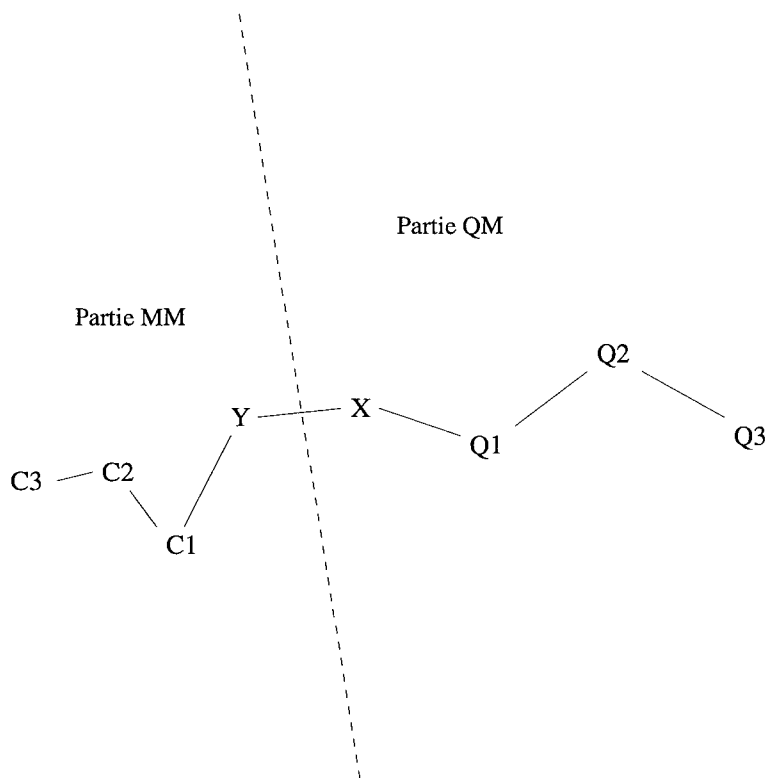


FIG. 3.4 – Notations des atomes d'un macrosystème QM/MM

Nous pouvons classer les interactions QM/MM en deux types : les interactions non-électrostatiques qui sont résumées dans le tableau III et les interactions électrostatiques entre la distribution de charges de la partie QM et les charges ponctuelles de la partie MM.

TAB. III – *Interactions non-électrostatiques QM/MM*

Potentiels de liaison	X-Y
Potentiels angulaires	X-Y-C1
Potentiels de torsion	Q1-X-Y-C1, X-Y-C1-C2
Potentiels vdW	Y-Qi (i>1), X-Ci (i>1), Ci-Qj

En effet, nous avons choisi de polariser la fonction d'onde électronique par toutes les charges MM. Les éléments de la matrice de Fock doivent donc être modifiés :

$$F_{\mu\nu} = H_{\mu\nu} + \sum_i^{MM} \left\langle \mu \left| \frac{Q_i}{r_i} \right| \nu \right\rangle + \sum_{\lambda\sigma} P_{\lambda\sigma} \left[ (\mu\nu|\lambda\sigma) - \frac{1}{2} (\mu\sigma|\lambda\nu) \right] \quad (3.15)$$

et un terme d'interactions noyaux de la partie quantique – charges classiques doit être ajouté à l'énergie du macrosystème. Si la description QM/MM du macrosystème nécessite l'utilisation de plusieurs liaisons frontières, d'autres termes peuvent être inclus : potentiels de liaison Y-Y, potentiels angulaires Y-Y-Y, Y-C-Y ...

Pour conserver l'électroneutralité du système, une charge ponctuelle de valeur +1 est placée sur l'atome frontière Y afin de compenser la charge de l'électron qu'il apporte à la liaison (dans le cas d'une liaison frontière covalente). Ces charges sont visibles de toutes les charges classiques, sans aucune restriction due à la condition « 1-4 » utilisée habituellement dans les champs de forces. Par contre, la charge MM portée par l'atome Y sera soumise à cette condition, comme n'importe quelle autre charge classique. Dans tous les cas, les formes et paramètres des potentiels classiques sont gardés tels quels, sans modification majeure (nous verrons par la suite que nous avons défini un potentiel de liaison frontière et ajusté les charges à la frontière). A l'inverse de Friesner [47] qui propose de reparamétrer en grande partie le champ de forces utilisé pour la partie MM, nous préférons nous focaliser sur la partie quantique et considérer la partie classique comme un terme perturbatif, de nature qualitative, agissant sur la partie quantique.

### 3.2.3 Potentiel de liaison frontière

Nous allons maintenant nous intéresser de plus près à la frontière (liaison X-Y) entre les deux sous-systèmes. Pour résumer, X est un atome quantique normal et Y est un atome hybride quantique-classique :

il possède les fonctions de base de l'atome quantique qu'il doit représenter, mais à la place de sa charge nucléaire, il possède une charge +1 à laquelle on ajoute une charge classique  $q_{MM}$  (cf. figure 3.5).

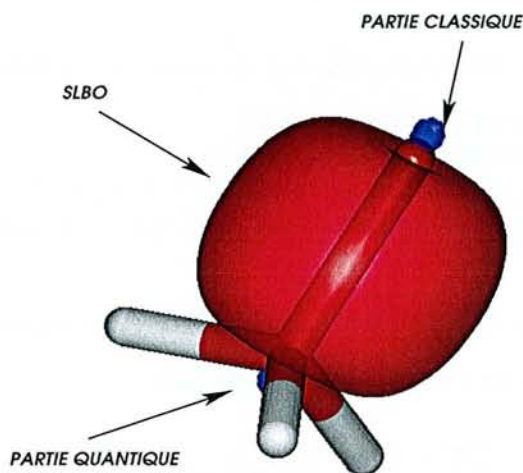


FIG. 3.5 – Frontière QM/MM : en haut, l'atome frontière Y ; en bas, l'atome frontière X.

Formellement, cet atome Y ressemble beaucoup à un atome d'hydrogène. Lors d'une optimisation de géométrie, si on laisse bouger librement les atomes X et Y qui portent une liaison frontière, la longueur de celle-ci va fortement diminuer et tendre vers une valeur proche de la longueur d'une liaison X-H. Il est donc nécessaire de corriger ce comportement en définissant un nouveau potentiel de liaison frontière.

Pour y parvenir, nous utilisons des molécules modèles très simples : l'éthane pour une liaison C-C, le méthanol pour une liaison C-O, la méthylamine pour une liaison C-N, ainsi que l'acétaldéhyde pour une liaison C-C avec un carbone hybridé  $sp^2$  et le formamide pour une liaison N-C, l'azote (formellement) et le carbone étant également hybridé  $sp^2$ . Les géométries de ces cinq molécules sont optimisées au niveau de théorie HF/6-31G\* et les SLBO correspondantes sont obtenues en utilisant le critère de Weinstein-Pauncz. Puis les profils énergétiques de dissociation des liaisons X-Y sont calculés :

- de façon entièrement quantique (noté SCF),
- de façon hybride LSCF/MM avec du côté MM uniquement la charge +1 sur l'atome Y (nous noterons ce calcul LSCF/+1).

Il n'y a donc aucune autre contribution QM/MM que l'électrostatique entre la charge +1 et la distribution de charges quantique. Typiquement, nous obtenons les profils de la figure 3.6.

De toute évidence, comme remarqué précédemment, les courbes sont différentes : les puits de poten-

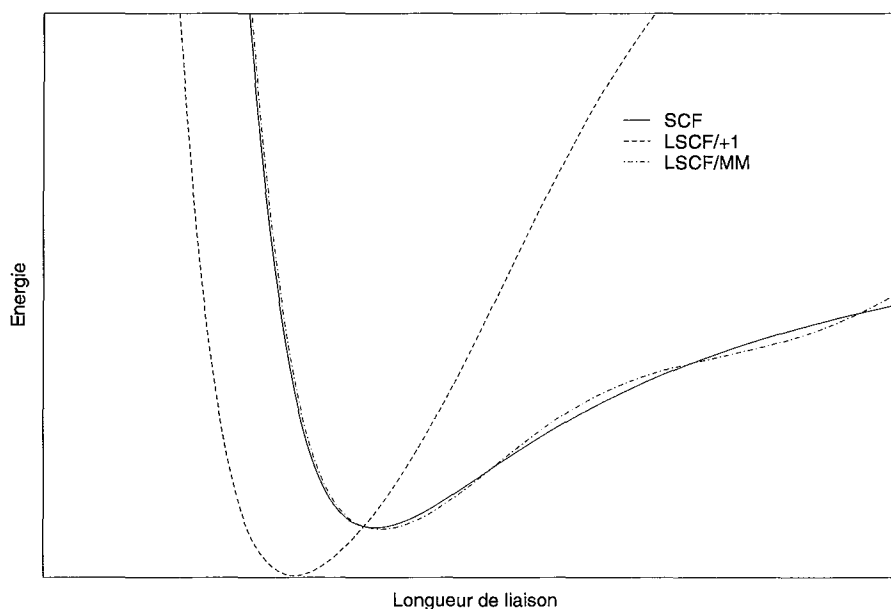


FIG. 3.6 – Profils énergétiques de dissociation d'une liaison frontière.

tiel n'ont pas la même forme et les distances d'équilibre sont différentes. Les erreurs causées par notre modèle de frontière QM/MM ont été analysées : deux contributions apparaissent principalement.

1. Lorsque la distance X-Y varie, l'évolution du recouvrement entre fonctions de base est négligée. La SLBO est définie pour une certaine longueur de liaison frontière et la distribution électronique qu'elle représente est gelée, sans pouvoir se réadapter.
2. Les interactions électrostatiques coeur-noyau, coeur-coeur et noyau-noyau (que l'on peut regrouper sous le terme ion-noyau) sont largement sous-estimées lorsqu'on remplace le noyau et le coeur de Y par une charge +1.

Un potentiel de liaison frontière à cinq composantes nous sert à corriger les deux erreurs énoncées ci-dessus :

$$E_{X-Y} = \underbrace{(A + Br + Cr^2)}_{\text{recouvrement}} e^{Dr} + \underbrace{\frac{E}{r}}_{\text{ion-noyau}} \quad (3.16)$$

où  $r$  est la distance de la liaison X-Y et  $A \dots E$  sont les paramètres à déterminer. La partie corrective du recouvrement provient du recouvrement entre deux fonctions  $1s$  de Slater dans la molécule  $H_2^+$ . Le paramètre  $E$  peut être perçu comme une charge nucléaire effective<sup>1</sup>. Les cinq composantes du potentiel

1. Plus exactement,  $E \propto Z_X (Z_Y - 1)$ ,  $Z_X$  étant la charge nucléaire de l'atome X.

sont ajustées par la méthode des moindres carrés afin de représenter au mieux la différence entre les deux courbes de dissociation SCF et LSCF/+1. Les valeurs des composantes sont reportées dans le tableau IV pour différents types de liaisons.

TAB. IV – Paramètres des potentiels de liaison frontière (en unités atomiques).

Type de liaison	A	B	C	D	E
C(sp <sup>3</sup> )-C(sp <sup>3</sup> )	-26,3145	5,2319	-1,0393	-0,3589	20,1330
C(sp <sup>3</sup> )-O	-25,0856	5,0907	-1,2836	-0,3848	18,8380
O-C(sp <sup>3</sup> )	-26,8871	4,8309	-0,9887	-0,3693	20,5643
C(sp <sup>3</sup> )-N	-28,3881	6,0076	-1,2983	-0,3736	20,9211
N-C(sp <sup>3</sup> )	-27,1145	5,0382	-1,0480	-0,3705	20,7263
C(sp <sup>2</sup> )-N(sp <sup>2</sup> )	-37,0482	8,9376	-2,0228	-0,4089	24,3843
N(sp <sup>2</sup> )-C(sp <sup>2</sup> )	-30,9690	6,2066	-1,3907	-0,3892	22,0465
C(sp <sup>2</sup> )-C(sp <sup>3</sup> )	-28,9594	6,1121	-1,2288	-0,3687	21,2616
C(sp <sup>3</sup> )-C(sp <sup>2</sup> )	-26,7094	5,3146	-1,0772	-0,3634	20,2739

Une fois les cinq paramètres déterminés pour les cinq types de liaison envisagés, les profils énergétiques sont recalculés à l'aide d'un calcul LSCF/MM (avec le potentiel de liaison frontière). Sur la figure 3.6, l'accord est très bon autour de la distance d'équilibre entre le profil SCF et le profil LSCF/MM. En revanche, aux grandes distances internucléaires, l'accord est nettement moins bon ; mais comme les liaisons frontières ne sont jamais directement impliquées dans une réaction chimique, il n'est pas nécessaire de bien décrire le comportement à grande longueur de liaison. Nous avons donc défini un potentiel frontière qui dépend uniquement du type et de l'hybridation des atomes qui la portent. De plus, lors de sa définition, aucun terme MM n'est intervenu : le potentiel de liaison frontière est donc indépendant du champ de forces utilisé dans le calcul QM/MM.

**Etude de la transférabilité du potentiel de liaison frontière.** Pour quantifier la validité de ce potentiel, les longueurs de liaison X-Y optimisées en LSCF/MM sont reportées dans le tableau V.

La partie quantique est traitée au niveau HF/6-31G\* et le champ de forces est l'un des trois évoqués précédemment. Les résultats sont en bon accord avec ceux des calculs purement quantiques. La plus grande différence est de l'ordre de 0,05 Å. Il semblerait que plus la liaison est polaire, plus l'erreur est grande. Par définition, les liaisons frontières ne seront jamais cassées, et la précision obtenue avec le potentiel de liaison frontière nous paraît satisfaisante. La précision du potentiel pourrait bien évidemment être améliorée, mais aux dépens de la transférabilité. De plus, nous vérifions que le potentiel est bien adapté à tous les champs de forces utilisés.

TAB. V – Distances d'équilibre en Å des différentes liaisons X-Y envisagées. Les indications « Non défini » correspondent à des potentiels angulaires non paramétrés dans Amber.

Liaison (X-Y)	SCF	LSCF/UFF	LSCF/Dreiding	LSCF/Amber
C(sp <sup>3</sup> )-C(sp <sup>3</sup> )	1,527	1,525	1,527	1,523
C(sp <sup>3</sup> )-O	1,400	1,355	1,356	1,354
O-C(sp <sup>3</sup> )	1,400	1,363	1,365	1,359
C(sp <sup>3</sup> )-N	1,454	1,463	1,464	1,450
N-C(sp <sup>3</sup> )	1,454	1,434	1,436	1,430
C(sp <sup>2</sup> )-C(sp <sup>3</sup> )	1,504	1,504	1,505	Non défini
C(sp <sup>3</sup> )-C(sp <sup>2</sup> )	1,504	1,504	1,506	Non défini
C(sp <sup>2</sup> )-N(sp <sup>2</sup> )	1,348	1,360	1,362	Non défini
N(sp <sup>2</sup> )-C(sp <sup>2</sup> )	1,348	1,359	1,361	Non défini

Nous allons maintenant tester la transférabilité du potentiel de liaison frontière envers les bases d'orbitales atomiques. En effet, les potentiels de liaison frontière ont été déterminés à partir des courbes de dissociation calculées dans la base 6-31G\*. Il est donc intéressant de tester leurs performances lorsqu'on passe dans une base plus ou moins étendue. Les géométries des petites molécules modèles sont donc à nouveau optimisées dans différentes bases. Les résultats sont reportés dans le tableau VI.

Dans tous les cas, les erreurs sont inférieures à 0,05 Å, ce qui est du même ordre que les erreurs acceptées précédemment. Le potentiel de liaison frontière semble donc suffisamment transférable d'une base à l'autre, sans la nécessité de le déterminer pour chaque base utilisée. Cependant, il est vrai que pour obtenir des accords sur la longueur de liaison frontière de l'ordre du pm, il faut alors recalculer ce potentiel pour chaque base en se focalisant autour de la distance d'équilibre. Par exemple, dans le tableau VII sont reportées les valeurs des cinq paramètres du potentiel de la liaison C-O optimisés pour différentes bases.

Nous avons choisi de privilégier la transférabilité du potentiel de liaison frontière plutôt que la précision. Remarquons que dans la méthode des pseudo-potentiels de groupe [37, 38], ceux-ci sont définis pour une base atomique donnée, ce qui leur permet d'atteindre des longueurs de liaison frontière de très bonne qualité.

Pour compléter cette étude sur la transférabilité des potentiels de liaison frontière, il reste à considérer différents niveaux de théorie pour la partie quantique. Nous avons choisi d'optimiser la géométrie du méthanol (groupement méthyle quantique, groupement alcool classique) en tenant compte de la corrélation électronique de deux façons : en utilisant la DFT ou en utilisant MP2 en base 6-31G\* avec le champ de forces UFF (le choix du méthanol est guidé par le tableau V : cette molécule est celle qui donne les

TAB. VI – Longueurs de liaison X-Y optimisées pour différentes bases, en Å.

Base	C(sp <sup>3</sup> )C(sp <sup>3</sup> )		C(sp <sup>3</sup> )N		C(sp <sup>3</sup> )O	
	SCF	LSCF/MM	SCF	LSCF/MM	SCF	LSCF/MM
STO-3G	1,538	1,520	1,485	1,472	1,433	1,432
3-21G	1,542	1,497	1,471	1,427	1,441	1,386
6-31G	1,530	1,511	1,450	1,434	1,431	1,393
6-31G*	1,527	1,525	1,454	1,463	1,400	1,355
6-31G**	1,527	1,526	1,452	1,465	1,399	1,407
6-311++G**	1,527	1,530	1,454	1,474	1,400	1,402

Base	C(sp <sup>2</sup> )N(sp <sup>2</sup> )		C(sp <sup>3</sup> )C(sp <sup>2</sup> )	
	SCF	LSCF/MM	SCF	LSCF/MM
STO-3G	1,403	1,397	1,537	1,521
3-21G	1,353	1,302	1,507	1,481
6-31G	1,349	1,320	1,495	1,482
6-31G*	1,348	1,360	1,504	1,504
6-31G**	1,348	1,375	1,503	1,504
6-311++G**	1,349	1,379	1,502	1,512

TAB. VII – Valeurs des paramètres de liaison frontière C-O, optimisés pour différentes bases, comparés aux paramètres transférables (dernière ligne).

Base	A	B	C	D	E
3-21G	-49,1494	12,9570	-3,4307	-0,4690	29,8439
6-31G*	-35,9888	10,0201	-2,1954	-0,3846	23,7594
6-311++G**	-31,1797	7,6785	-1,5070	-0,3398	23,1464
Potentiel transférable	-25,0856	5,0907	-1,2836	-0,3848	18,8380

résultats les moins bons). Les distances C-O sont données dans le tableau VIII.

TAB. VIII – Distances C-O (en Å) du méthanol pour différentes méthodes quantiques.

Méthode	SCF	LSCF/MM
HF	1,400	1,355
MP2	1,423	1,364
BLYP	1,435	1,420
B3PW91	1,412	1,416

Les résultats indiquent la même tendance à diminuer la distance C-O pour la méthode MP2 que pour Hartree-Fock avec cette fois-ci une erreur de 6 pm. En revanche, les résultats sont très bons pour la DFT, que ce soit avec une fonctionnelle « pure » telle que BLYP ou avec une fonctionnelle « hybride » telle que B3PW91.

Pour éprouver l'hypothèse de transférabilité des SLBO et du potentiel frontière entre la molécule modèle sur laquelle ils sont définis et une macromolécule, nous décidons d'étudier la structure d'une petite protéine, la crambine (cf figure 3.7). La géométrie de cette macromolécule (46 résidus, soit 642 atomes) est optimisée entièrement à divers niveaux, à partir de la structure cristalline PDB :

1. au niveau mécanique moléculaire uniquement (Amber),
2. au niveau LSCF/MM (HF/6-31G\* + Amber) avec le premier acide aminé (thyrosine) traité quantiquement (1 frontière de type C(sp<sup>2</sup>)-N(sp<sup>2</sup>)) noté LSCF/MM 1,
3. au niveau LSCF/MM (HF/6-31G\* + Amber) avec le vingt-troisième acide aminé (acide glutamique) traité quantiquement (2 frontières de type C(sp<sup>2</sup>)-N(sp<sup>2</sup>) et N(sp<sup>2</sup>)-C(sp<sup>2</sup>)) noté LSCF/MM 2.

Dans le premier calcul, la géométrie est optimisée en coordonnées cartésiennes en utilisant la technique des microitérations. Pour les calculs QM/MM, l'optimisation se fait en coordonnées internes, grâce à l'algorithme de Németh et collaborateurs [64] développé au laboratoire. Il permet de passer très rapidement de la représentation en coordonnées cartésiennes à celle en coordonnées internes et vice-versa. Initialement, cet algorithme utilisait la méthode de minimisation « steepest descent ». Cette dernière étant peu efficace près du minimum, une partie de notre travail a été de coupler l'algorithme de Németh avec le minimiseur BFGS pour mémoire limitée de Nocedal et collaborateurs [65, 66]. Cette nouvelle méthode d'optimisation de géométrie croît linéairement avec la taille du système, ce qui est très intéressant dans des calculs QM/MM (par exemple, il faut environ six cents pas d'optimisation pour faire converger la structure de la crambine, définie à l'aide de plus de six mille coordonnées internes).



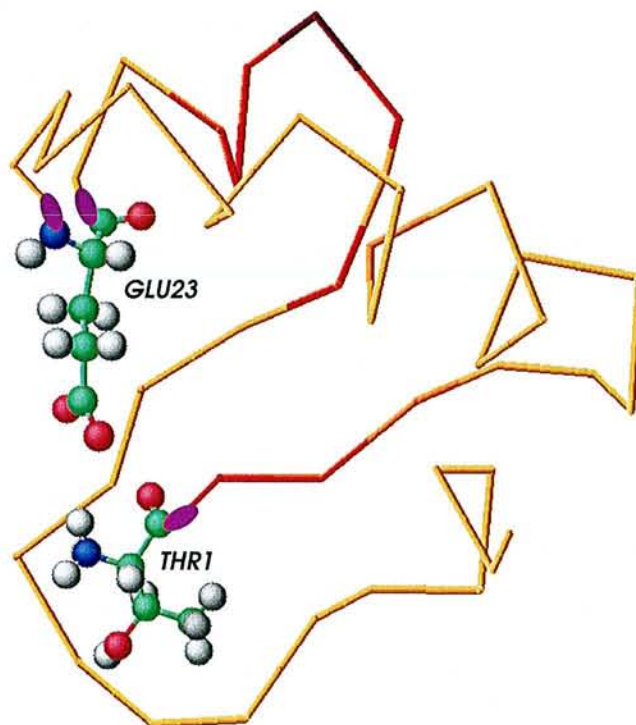


FIG. 3.7 – Structure de la crambine : le squelette est en bâton orangé, les fragments quantiques sont en boules et bâtons de couleur et les frontières sont en cylindres violets.

Les structures résultantes sont très satisfaisantes : les variations sur les longueurs de liaison et les angles de valence de la partie quantique sont toujours inférieures à 0,05 Å et 5 degrés. Concernant plus particulièrement les frontières, les longueurs de liaison optimisées sont reportées dans le tableau IX.

TAB. IX – Longueurs optimisées des liaisons frontières en Å. Les valeurs cristallographiques sont dans la colonne PDB, les résultats provenant du calcul de mécanique moléculaire sont dans la colonne MM, les valeurs résultant des calculs hybrides sont dans les colonnes LSCF/MM 1 et 2 (respectivement 1 et 2 frontières).

Frontière	PDB	MM	LSCF/MM 1	LSCF/MM 2
THR1-THR2	1,324	1,330	1,374	1,335
PRO22-GLU23	1,351	1,343		1,374
GLU23-ALA24	1,347	1,339		1,341

Au vu des résultats, il semble donc bien que l'hypothèse de transférabilité des SLBO et du potentiel de liaison frontière soit justifiée. Le bon comportement du potentiel de liaison frontière est donc bien vérifié, et il ne semble pas que les erreurs de description géométrique augmentent avec le nombre de SLBO. Bien évidemment, ces résultats n'ont que valeur de test pour notre potentiel de liaison frontière. Une seule structure optimisée n'est pas significative pour de si grandes molécules. Cependant, nous pouvons considérer que cette étude est le premier pas vers l'adaptation de notre champ de forces hybride aux dynamiques moléculaires.

Pour conclure, nous avons mis au point un potentiel de liaison frontière indépendant du champ de forces utilisé dans la partie classique, indépendant de la méthode et de la base utilisées dans la partie QM et transférable de la molécule modèle sur laquelle il a été déterminé à la macromolécule que l'on veut étudier. Ce potentiel de liaison frontière est un avantage certain de notre méthode hybride sur les autres puisqu'elle permet a priori de traiter n'importe quel type de frontière QM/MM.

### 3.2.4 Modification des charges à la frontière

A ce point de la « construction » de la méthode LSCF/MM, nous avons montré que le formalisme LSCF est bien adapté aux méthodes QM/MM à travers l'utilisation de SLBO. Grâce à la définition d'un potentiel de liaison frontière, très simple à mettre en oeuvre, nous avons obtenu une bonne description géométrique de la frontière QM/MM. Dans cette partie, nous allons maintenant analyser les erreurs induites par l'utilisation d'un champ de forces MM dans un calcul QM/MM et la manière dont nous pouvons corriger celles-ci en modifiant les charges MM à la frontière.

Dans ce but, nous choisissons d'étudier quelques réactions de substitution nucléophile  $SN_2$  très

simples entre le chlorométhane  $\text{CH}_3\text{Cl}$  et des amines primaires qui sont notées  $\text{CH}_3(\text{CH}_2)_i\text{NH}_2$ , avec  $i = 0, 1, 2, 3$ . Un exemple de complexe réactif est donné figure 3.8. Nous comparons les études LSCF/MM (LSCF/3-21G + UFF) aux études complètement quantiques (avec ou sans SLBO). Dans tous les cas, les SLBO sont définies sur la liaison entre atomes lourds, située la plus loin possible du centre réactionnel (sur la liaison C-N pour  $i = 0$  et sur la dernière liaison C-C dans les autres cas).

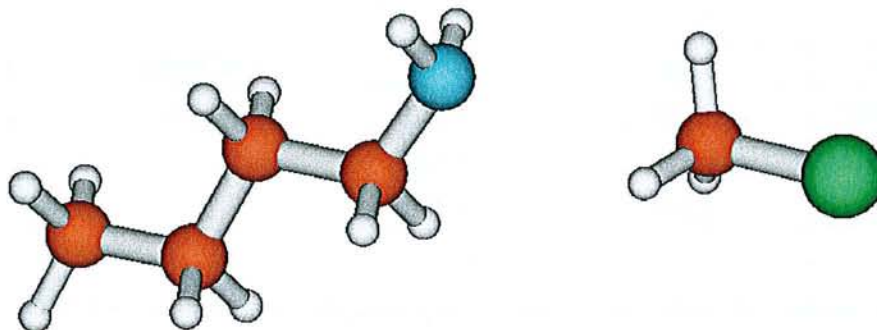
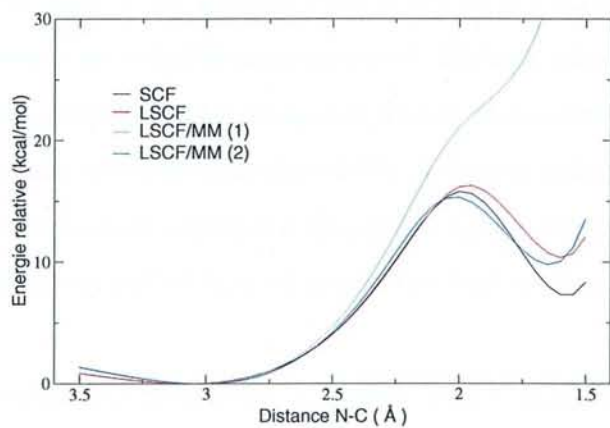


FIG. 3.8 – Réaction de Menshutkin :  $\text{CH}_3(\text{CH}_2)_3\text{NH}_2 + \text{CH}_3\text{Cl}$

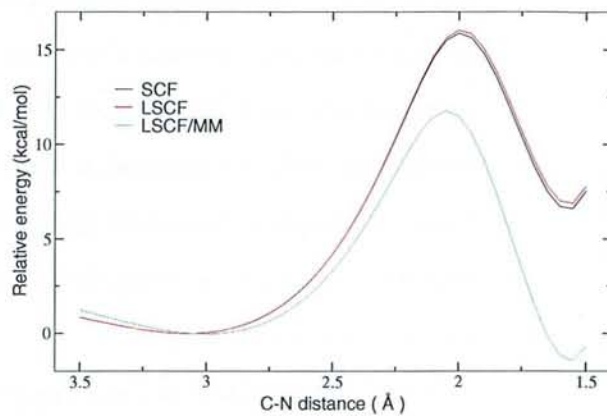
Les dérivées secondes analytiques ne sont pas disponibles pour la méthode LSCF. Pour obtenir les structures de transition de la réaction, nous avons adopté la démarche suivante : les atomes N, C et Cl sont contraints à rester alignés durant toute la réaction ; nous faisons varier la distance N-C de 3,5 Å à 1,5 Å (avec un pas de 0,05 Å) et la géométrie de la super-molécule est entièrement optimisée, en accord avec cette contrainte. Comme charges ponctuelles MM, nous choisissons d'utiliser les valeurs conventionnelles des champs de forces pour un groupement méthyle terminal (on aurait très bien pu utiliser des valeurs issues d'un calcul quantique comme les charges de Mulliken), c'est-à-dire -0,15 pour le carbone et +0,05 pour chacun des trois atomes d'hydrogène.

Nous obtenons les profils énergétiques de réaction de la figure 3.9. Comme les énergies absolues ne peuvent pas être comparées entre un calcul quantique et un calcul QM/MM, nous avons tracé ces profils par rapport au minimum énergétique du calcul quantique (3,05 Å).

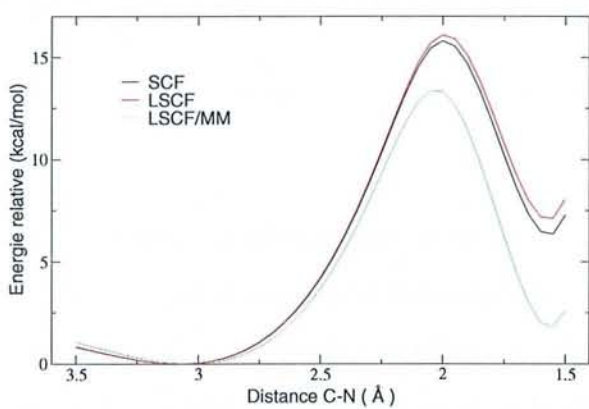
D'abord, il est intéressant de comparer les profils obtenus dans les études purement quantiques sans et avec une SLBO : l'accord est particulièrement bon lorsque la SLBO n'est pas définie sur un atome directement impliqué dans la réaction (cas  $i = 0$ ), la différence énergétique ne dépassant pas  $1 \text{ kcal.mol}^{-1}$ . Cette différence est due à la modification de la polarité de la liaison sur laquelle la SLBO est définie durant le processus chimique et qui ne peut être prise en compte dans le calcul LSCF avec la SLBO (celle-ci



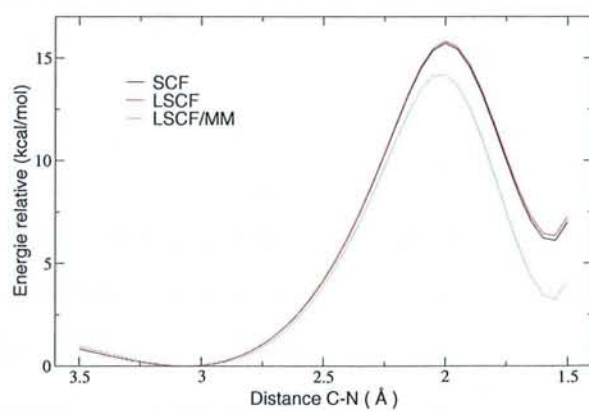
(a) i=0



(b) i=1



(c) i=2



(d) i=3

FIG. 3.9 – Profils énergétiques de la réaction  $\text{CH}_3(\text{CH}_2)_i\text{NH}_2 + \text{CH}_3\text{Cl}$

est gelée : sa polarité est fixée pour toute la réaction). Au vu des différences énergétiques, il n'est pas nécessaire de modifier une SLBO durant une réaction chimique lorsque celle-ci n'est pas connectée à un atome engagé dans la réaction.

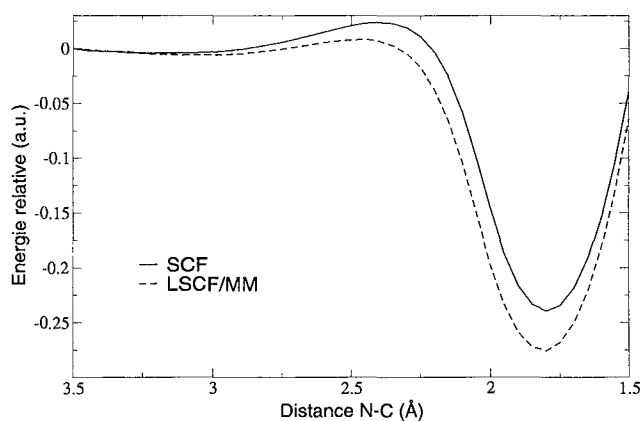
Les résultats LSCF/MM sont beaucoup moins satisfaisants. Pour obtenir des barrières de réaction en bon accord avec les résultats quantiques, la frontière QM/MM doit se situer à trois ou quatre liaisons du centre réactionnel. Il est évident que ce résultat n'est pas dû au formalisme LSCF lui-même, mais plutôt à la représentation hybride QM/MM du système étudié. Cette remarque est d'ailleurs très générale puisqu'elle s'applique à toutes les méthodes QM/MM : le sous-système MM doit agir comme une perturbation sur le sous-système QM. Pour cette raison, la taille de la partie QM doit toujours être suffisamment étendue. Cependant, il serait intéressant de pouvoir réduire cette taille en minimisant les erreurs causées par le formalisme QM/MM, notamment dans l'optique de simulations de dynamiques moléculaires *ab initio* pour lesquelles le pas limitant est bien évidemment le calcul de l'énergie et des forces de la partie quantique.

Pour illustrer un peu plus le rôle trop perturbateur de la partie MM sur les électrons de la partie QM, nous avons séparé les différentes contributions énergétiques tout au long du chemin réactionnel : énergie cinétique des électrons, interactions des électrons avec les noyaux quantiques et les charges classiques, répulsion électronique et répulsion nucléaire au cours de la réaction  $\text{CH}_3\text{CH}_2\text{NH}_2 + \text{CH}_3\text{Cl}$  (cf. figure 3.10).

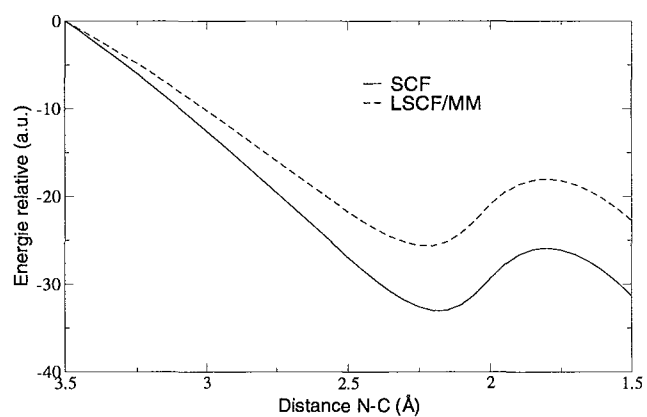
Ces courbes expriment les différentes énergies, relatives au premier point de la courbe. Le remplacement du groupement méthyle terminal quantique par un groupement classique n'est pas anodin. Les écarts entre les valeurs obtenues avec la méthode SCF et la méthode LSCF/MM sont loin d'être négligeables, à part pour la contribution cinétique des électrons. Les écarts maximaux pour les contributions potentielles électrons-électrons et électrons-charges (nucléaires ou classiques) sont de l'ordre de 5 unités atomiques (u.a.), alors qu'ils sont de l'ordre de 10 u.a. pour la contribution due aux interactions électrostatiques entre les noyaux et les charges (nucléaires ou classiques).

Les variations énergétiques montrent à quel point la séparation d'un macrosystème en deux parties QM et MM est critique. Comme l'amplitude des erreurs est énorme, les compensations d'erreurs seraient purement fortuites. En conséquence, une reparamétrisation du champ de forces classique est indispensable.

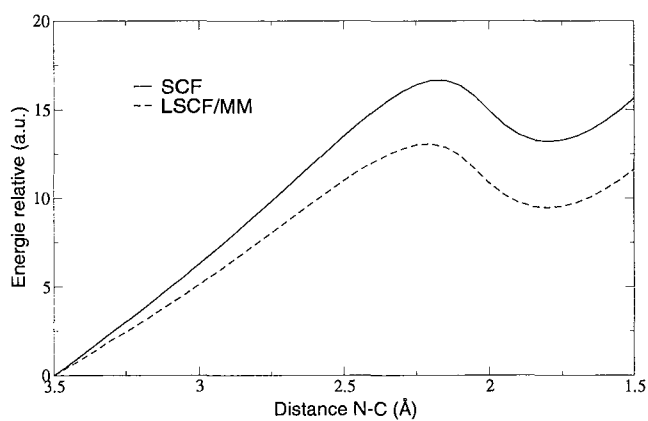
L'analyse des erreurs induites par l'utilisation d' « atomes MM » à la place d' « atomes QM »



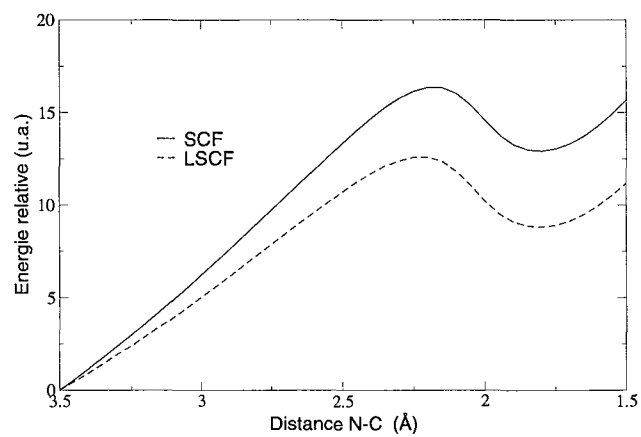
(a) Energie cinétique électronique



(b) Interactions électrons - (noyaux QM + charges MM)



(c) Répulsion électronique



(d) Interactions noyaux QM - (noyaux QM + charges MM)

FIG. 3.10 – Contributions quantiques à l'énergie QM/MM lors du chemin réactionnel

montre plusieurs contributions. La première que nous allons examiner est la définition de la connectivité, nécessaire au champ de forces. Les interactions QM/MM de van der Waals sont soumises à la « condition 1-4 », la connectivité doit donc être établie dans la partie quantique. Mais si les atomes de la partie MM sont trop proches du centre réactionnel, deux possibilités s'offrent à nous : soit une connectivité est définie entre les deux atomes qui réagissent, soit non, et ce, sans qu'on puisse la modifier durant la réaction. C'est le cas  $i = 0$  de la  $SN_2$  et nous l'avons illustré dans la figure 3.9 (courbes LSCF/MM (1) et (2), respectivement sans et avec connectivité N-C). Aux courtes distances N-C, les courbes sont très différentes. Pour garder une définition cohérente et appropriée de la connectivité avant et après la réaction chimique, il est donc impossible de placer le sous-système MM trop près du centre réactionnel. On note que dès le cas  $i = 1$ , cet effet disparaît puisque les termes liés QM/MM ne concernent plus que les atomes qui restent connectés à la partie classique tout au long de la réaction.

La plus forte interaction QM/MM est due au terme électrostatique. Il est donc nécessaire de se pencher sur l'influence des charges MM sur le profil énergétique de réaction, et plus particulièrement de celles qui sont proches de la partie quantique et que nous qualifierons de charges frontières. Les charges MM sont le plus souvent déterminées de façon à ce qu'elles reproduisent les interactions électrostatiques entre atomes non-liés (séparés par trois liaisons au moins). Cependant, dans notre formalisme QM/MM, chaque charge classique interagit avec la distribution de charges quantique sans la restriction « 1-4 ». De plus, les atomes frontières Y possèdent, en plus de leur charge MM, une charge +1 centrée sur le noyau et un électron décrit par une SLBO. Il est donc évident que les charges situées à la frontière QM/MM et plus particulièrement les charges portées par les atomes Y doivent être adaptées à cette situation hybride, afin d'obtenir une bonne polarisation de la frontière QM/MM.

Dans notre étude, nous avons fait varier les charges portées par le groupement méthyle terminal dans le cas  $i = 1$ . Ces charges sont comprises entre -0,15 et +0,45 pour le carbone et +0,05 et -0,15 pour les hydrogènes. Quelques profils réactionnels sont reportés dans la figure 3.11. Le meilleur choix s'avère être le jeu  $\{+0,39;-0,13\}$  pour C et H respectivement. Les profils énergétiques sont alors pratiquement identiques (voir figure 3.12). Quant aux géométries, elles ne diffèrent jamais plus de 0,02 Å. Nous pouvons également analyser les charges de Mulliken portées par les atomes quantiques. Nous avons choisi le dernier pas de la réaction, lorsque la substitution nucléophile a eu lieu. Les charges sont données sur la figure 3.13 : elles sont obtenues par un calcul quantique (noté SCF) et par deux calculs hybrides ; le premier n'utilise pas les charges adaptées à la frontière (noté LSCF/MM), le deuxième utilise ce jeu de charges  $\{+0,39;-0,13\}$  (noté LSCF/MM\*).

Les charges de Mulliken des atomes du sous-système quantique sont en très bon accord avec celles

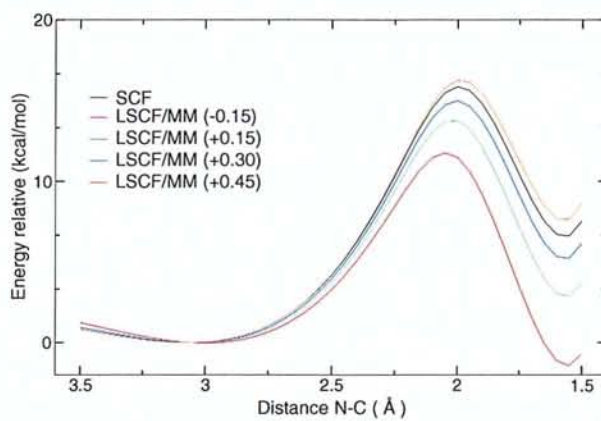


FIG. 3.11 – Profils énergétiques de la réaction  $\text{CH}_3\text{CH}_2\text{NH}_2 + \text{CH}_3\text{Cl}$  pour différentes valeurs de charges MM (celles du carbone sont notées entre parenthèses).

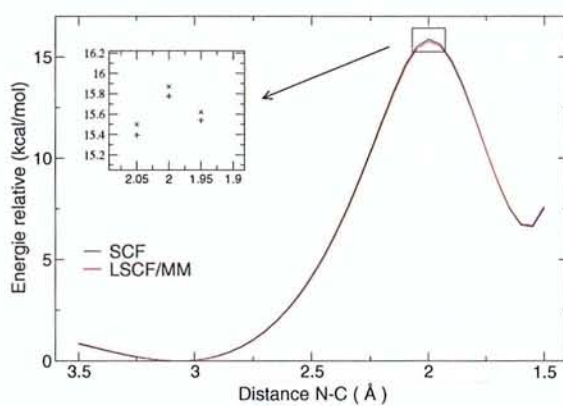


FIG. 3.12 – Profil énergétique de la réaction  $\text{CH}_3\text{CH}_2\text{NH}_2 + \text{CH}_3\text{Cl}$  en utilisant les charges adaptées.



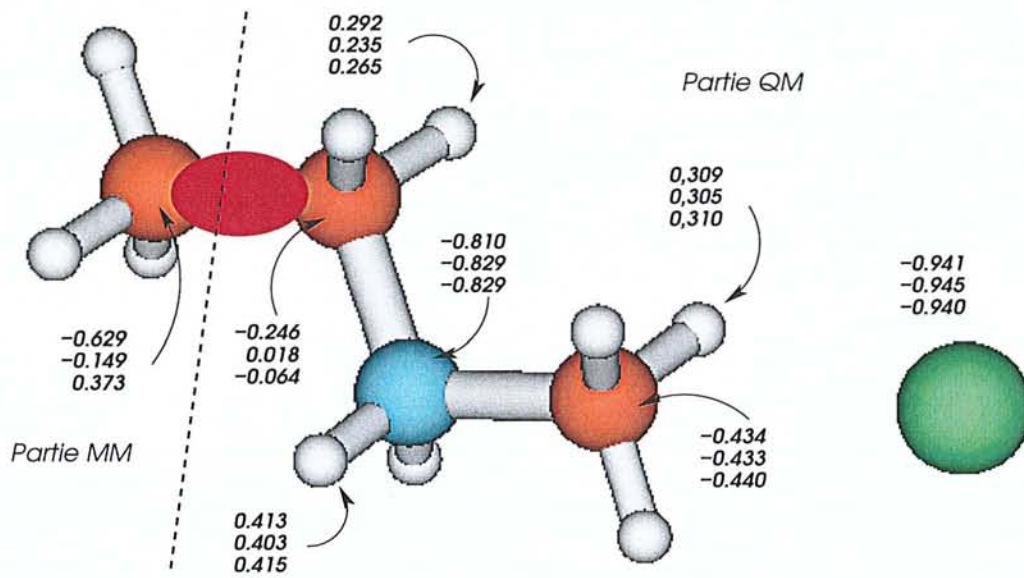


FIG. 3.13 – Charges de Mulliken (dans l'ordre : SCF, LSCF/MM et LSCF/MM\*) en électrons.

obtenues par le calcul quantique sauf pour le carbone qui porte la SLBO, et ce avec ou sans charges modifiées à la frontière. L'influence des charges à la frontière semble donc agir plus subtilement sur le système quantique. Le moment dipolaire du système semble être plus sensible que les charges de Mulliken à ces charges : 19,15 pour SCF, 21,03 pour LSCF/MM et 19,33 pour LSCF/MM\*. L'adaptation des charges à la frontière permet de corriger efficacement le moment dipolaire.

Grâce à notre formalisme LSCF/MM, il est donc possible de réduire la taille du sous-système quantique tout en gardant une bonne description géométrique et énergétique de la réaction. Les charges négatives portées par les hydrogènes peuvent sembler dénuées de sens, mais dans notre exemple, la partie MM est réduite à sa plus simple expression. Dans un système macromoléculaire, la partie MM est très étendue et l'excès de charges à la frontière peut être réparti sur les atomes proches, sans trop perturber les interactions MM. Enfin, notons que nous avons choisi de modifier uniquement les charges MM à la frontière QM/MM (ce sont elles qui perturbent le plus la partie QM du système) et qu'en conséquence les autres charges MM d'un système macromoléculaire sont choisies de la manière habituelle.

Afin de vérifier la validité de ces charges modifiées, la réaction correspondant au cas  $i = 3$  est de nouveau étudiée. La partie MM correspond à  $\text{CH}_3\text{-CH}_2\text{-CH}_2$  et la partie QM à  $\text{CH}_2\text{-NH}_2 + \text{CH}_3\text{-Cl}$ . Dans un premier calcul hybride (appelé LSCF/MM\*), les charges du groupement  $\text{CH}_2$  le plus proche de la partie QM sont celles déterminées ci-dessus et les valeurs des autres charges sont données par le champ de forces. Dans un second calcul (appelé LSCF/MM), nous n'utilisons pas nos charges adaptées. L'électroneutralité du système entier est respectée à chaque fois. Les profils énergétiques sont reportés dans la figure 3.14. Comme attendu, la courbe LSCF/MM dévie fortement de la courbe quantique : les minima ne sont pas les mêmes et les barrières énergétiques n'ont pas la même hauteur. En revanche, les résultats LSCF/MM\* sont très satisfaisants. L'erreur énergétique est de l'ordre de  $1 \text{ kcal.mol}^{-1}$ , ce qui est suffisamment précis pour un calcul QM/MM. De ce dernier exemple, il ressort donc que l'ajustement des charges à la frontière permet de garder une bonne description énergétique de la réaction, tout en gardant un sous-système MM proche du centre réactionnel. Il semble également que ces charges soient transférables du système dans lequel elles ont été calculées aux macrosystèmes correspondants.

### 3.2.5 Protocole de simulation

Pour résumer, la mise au point du champ de forces hybride LSCF/MM *ab initio* a nécessité deux innovations : la définition d'un potentiel de liaison frontière qui permet de décrire correctement la géométrie de cette liaison, représentée par une orbitale de liaison strictement localisée ; l'ajustement des charges ponctuelles portées par les atomes classiques frontières, notamment celle de l'atome qui porte la

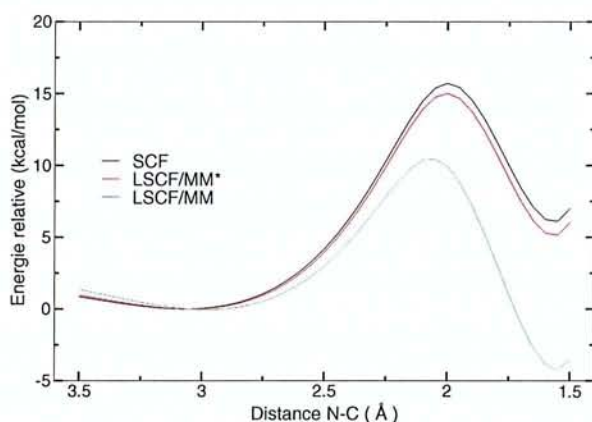


FIG. 3.14 – Profils énergétiques de la réaction  $\text{CH}_3(\text{CH}_2)_3\text{NH}_2 + \text{CH}_3\text{Cl}$ .

SLBO, afin d'assurer une bonne polarisation de la frontière. Il faut cependant remarquer que le premier est ajusté sur une molécule modèle la plus simple possible, et que le deuxième point s'est également avéré réalisable sur un système modèle du système réactif envisagé. Autrement dit, le potentiel de liaison frontière et les charges adaptées à la frontière sont des paramètres transférables.

En conséquence, nous avons défini un protocole de simulation LSCF/MM *ab initio* :

1. Pour une liaison X–Y coupée, la molécule modèle la plus simple est considérée (du type  $\text{H}_n\text{X}-\text{YH}_m$ ).
2. Localisation d'une SLBO selon un critère donné (il nous semble que le critère de Weinstein-Pauncz soit le plus adapté), sur la molécule modèle.
3. Calcul du profil de dissociation de la liaison (LSCF/+1) et détermination des cinq paramètres du potentiel de liaison frontière, en comparant avec la molécule modèle.
4. Calculs LSCF/MM afin d'obtenir le meilleur jeu de charges à la frontière (critères énergétiques ou géométriques), toujours en comparant avec le système modèle.
5. Simulation LSCF/MM\* du véritable macrosystème.

Cette procédure, qui doit permettre d'obtenir des résultats LSCF/MM suffisamment précis, est très simple et rapide à mettre en oeuvre. En effet, il faut remarquer que les points 1 à 4 sont réalisés sur une molécule modèle de la liaison (par exemple l'éthane pour une liaison C – C hybridés  $\text{sp}^3$ ). De plus, le potentiel frontière est déterminé une fois pour toute pour un type de liaison, il n'est donc pas nécessaire de le déterminer pour chaque étude. A terme, il devrait être possible de mettre au point tous ces potentiels et charges frontières pour tous les types de liaison envisageables, ce qui fera disparaître les points 3 et 4.

D'ores et déjà, nous disposons du tableau IV comme premier pas vers cette paramétrisation complète.

Dans la suite de ce mémoire, nous appliquerons ce protocole à l'étude LSCF/MM *ab initio* à la réactivité de petites molécules organiques catalysées par des zéolithes. D'autres équipes utilisent actuellement la méthode LSCF/MM dans divers domaines : adsorption d'un feuillet de sulfure de molybdène sur une surface d'alumine [67], catalyse asymétrique par un complexe du cuivre (en collaboration avec W. Harb et M.F. Ruiz-López au laboratoire de Nancy).

Mais avant de s'intéresser aux zéolithes, nous allons combiner l'approche LSCF/MM avec la capacité de la méthode LSCF quantique à modéliser les états ionisés et proposer une approche originale de caractérisation des effets d'environnement dans les protéines.

### 3.3 Application de LSCF/MM à l'étude d'états ionisés dans la crambine

Comme mentionné dans le chapitre précédent, la détermination des énergies d'ionisation des électrons de coeur par spectroscopie électronique est largement utilisée comme outil de caractérisation structurale.

Nous avons choisi d'étudier les états ionisés de coeur de l'acide glutamique dans la crambine grâce à l'approche LSCF/MM présentée ci-dessus. Plus particulièrement, nous nous sommes intéressés aux électrons 1s des différents atomes de carbone et d'oxygène de ce résidu (à l'exclusion des atomes impliqués dans les liaisons frontières). Pour mémoire, l'acide glutamique est traité au niveau HF/6-31G\*\* tandis que les quarante-cinq autres résidus sont modélisés avec le champ de forces Amber. Deux liaisons frontières (liaisons peptidiques) sont décrites par deux orbitales de liaison strictement localisées. Nous nous plaçons dans la géométrie optimisée obtenue auparavant.

Durant le calcul LSCF d'un état ionisé, cinq spinorbitales sont gelées (nous sommes dans le formalisme non-restreint) : deux SLBO de spin  $\alpha$ , deux SLBO de spin  $\beta$  et une spinorbitale de coeur identifiée à un électron 1s (arbitrairement choisie de spin  $\beta$ ).

Sept énergies d'ionisation (correspondants aux notations de la figure 3.15) sont déterminées avec ou sans partie MM et les valeurs sont données dans les deux dernières colonnes du tableau X.

La validation de nos résultats passe par la comparaison à l'expérience, mais il faudrait alors étudier tous les acides aminés de la protéine consécutivement. Cependant, comme il n'y a pas de données expérimentales concernant les énergies d'ionisation de coeur des protéines, nous ne pouvons que comparer les valeurs obtenues pour de petites molécules. Nous avons donc choisi HCOOH au même niveau de théorie que précédemment. Comparées à l'expérience [68], les différences sont comprises entre 1 et 2 eV. Nos résultats peuvent être considérés comme qualitativement bons et pourront donc être comparés.



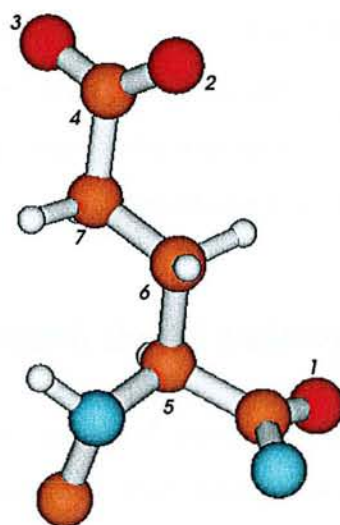


FIG. 3.15 – Acide glutamique dans la crambine. Les chiffres correspondent aux atomes ionisés.

TAB. X – Potentiels d'ionisation d'électrons  $1s$  de l'acide et de l'anion carboxylique, de l'acide glutamique (GLU) dans la crambine (en eV).

Orbitale	HCOOH (Exp.)	HCOOH (LSCF)	HCOO <sup>-</sup>	GLU	GLU+SLBO	Crambine
1				534,30	534,48	537,19
2	540,70	542,34	531,39	531,51	531,24	536,01
3	539,00	540,20	531,74	531,93	531,65	535,79
4	295,80	297,75	289,50	289,65	289,31	293,52
5				289,34	288,44	290,61
6				287,55	287,01	289,97
7				286,90	286,47	289,83

En comparant les deux dernières colonnes du tableau X, il semble que l'effet apparent de la protéine sur les énergies d'ionisation est d'attacher plus fortement ces électrons aux noyaux : on peut comparer ce comportement à la solvation qui diminue les énergies par rapport à la phase gazeuse. Cependant, cet effet est plus ou moins intense selon les cas. L'anion carboxylique semble être le plus affecté et l'effet environnemental diminue lorsqu'on se rapproche du squelette de la protéine. Cette variation des effets de la protéine sur les énergies est sûrement due à la plus forte interaction entre la charge -1 localisée sur l'anion et l'environnement.

Pour valider ces quelques remarques, nous avons cherché à séparer les différentes contributions dans le calcul LSCF/MM qui pourraient modifier les résultats. Si on remplace les SLBO définies sur les deux liaisons peptidiques qui terminent l'acide aminé par des atomes d'hydrogène (colonne GLU), les énergies d'ionisation varient peu : les différences sont inférieures à 1 eV, cependant une certaine évolution des effets se fait sentir du carbone du squelette, le plus proches des SLBO jusqu'aux atomes de carbone les plus éloignés. S'intéressant de plus près à l'anion carboxylique, les énergies d'ionisation des électrons de coeur des deux atomes d'oxygène et de l'atome de carbone (colonne HCOO<sup>-</sup>) sont très proches des valeurs correspondantes dans l'acide glutamique. Notons que ces calculs ont été effectués dans la géométrie de la protéine : l'optimisation des géométrie modifient très peu ces valeurs.

Le « déplacement » des valeurs d'énergie de l'acide glutamique de la situation isolée à la protéine est dû essentiellement aux effets environnementaux de celle-ci, c'est-à-dire à la polarisation de la partie QM par la partie MM.

La méthode LSCF permet la détermination d'énergie d'ionisations dans les systèmes étendus comme les protéines. Les résultats, qualitatifs, sont analysables en terme d'effet de l'environnement sur la partie quantique et peuvent aider à la compréhension des spectres électroniques de ces composés.

### 3.4 Conclusion

Dans ce chapitre, nous avons décrit la mise au point du champ de forces hybride LSCF/MM dont la principale caractéristique provient de la description des liaisons qui constituent la frontière entre la partie quantique et la partie classique.

L'utilisation d'orbitales de liaisons strictement localisées permet de s'affranchir des effets néfastes des « link atoms » : elles n'introduisent pas de degrés de liberté supplémentaires, elles sont facilement « modulables » (polarité, rapport des interactions  $\sigma/\pi$ ) ce qui permet a priori de décrire n'importe quel type de liaison. De plus, leur stricte localisation sur la liaison d'intérêt permet de calculer les dérivées

premières de l'énergie : l'optimisation de la géométrie complète du macrosystème est possible. Enfin, les erreurs dues à leur introduction dans un calcul quantique restent faibles et constantes : les énergies relatives peuvent être comparées et il n'est pas besoin de recalculer les SLBO pour chaque géométrie, ni de considérer la polarisation auto-cohérente des SLBO par le reste des orbitales moléculaires.

La reparamétrisation du champ de forces a été limitée à deux interventions. Elle a nécessité la définition de potentiels de liaisons frontières, de forme analytique simple et qui sont transférables entre champs de forces classiques, entre méthodes et bases quantiques. De plus, les charges ponctuelles à la frontière ont été adaptées afin de permettre une bonne polarisation de celle-ci. Un protocole de simulation a été mis au point afin de déterminer, si nécessaire, ces potentiels et ces charges frontières : il est très simple et rapide à mettre en oeuvre.

Cependant, comparée à l'utilisation des « link atoms », notre approche est plus compliquée à gérer et ne peut donc pas encore considérée comme une « boîte noire ». Les dérivées secondes de l'énergie ne sont pas encore disponibles, ce qui ne permet pas de caractériser un chemin réactionnel.

Ces désavantages seront dépassés dans un futur proche : une paramétrisation complète du champ de forces hybride est possible bien que ce soit un travail de longue haleine (mais qui bénéficiera des futures études utilisant notre méthode), les dérivées secondes sont envisageables après une étude théorique des équations correspondantes.

L'avenir de la méthode LSCF/MM passera par des études sur des sujets originaux qui devront se démarquer des autres méthodes se contentant pour la majorité d'entre elles à considérer les frontières QM/MM sur des liaisons carbone – carbone dans des systèmes covalents.

Il est également envisagé de la coupler avec les méthodes de type continuum afin d'incorporer les effets de solvant. La description de l'environnement par un champ de forces classique est parfois limitée. Plutôt que d'utiliser un autre champ de forces qui sera polarisable, il devrait être possible de le remplacer par une approche semi-empirique : les interactions « *ab initio* / semi-empirique » devront bien évidemment être paramétrées pour s'adapter à cette situation. Enfin, afin de traiter des systèmes très étendus, il sera alors possible de formuler une approche LSCF *ab initio* / LSCF semi-empirique / MM.

## Bibliographie

- [1] W. Yang. *J. Mol. Struct. (Theochem)* **255**, 461 (1992).

- [2] J.C. Burant, G.E. Scuseria et M.J. Frisch. *J. Chem. Phys.* **105**, 8969 (1996).
- [3] A. Van der Vaart, V. Gogonea, S.L. Dixon et K.M. Merz Jr. *J. Comp. Chem.* **21**, 1494 (2000).
- [4] 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. Cladwell et P.A. Kollman. *J. Am. Chem. Soc.* **117**, 5179 (1995).
- [5] A.K. Rappé, C.J. Casewit, K.S. Colwell, W.A. Goddard III et W.M. Skiff. *J. Am. Chem. Soc.* **114**, 10024 (1992).
- [6] S.L. Mayo, B.D. Olafson et W.A. Goddard III. *J. Phys. Chem.* **94**, 8897 (1990).
- [7] M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, V.G. Zakrzewski, J.A. Montgomery, Jr., R.E. Stratmann, J.C. Burant, S. Dapprich, J.M. Millam, A.D. Daniels, K.N. Kudin, M.C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G.A. Petersson, P.Y. Ayala, Q. Cui, K. Morokuma, D.K. Malick, A.D. Rabuck, K. Raghavachari, J.B. Foresman, J. Cioslowski, J.V. Ortiz, A.G. Baboul, B.B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R.L. Martin, D.J. Fox, T. Keith, M.A. Al-Laham, C.Y. Peng, A. Nanayakkara, C. Gonzales, M. Challacombe, P.M.W. Gill, B. Johnson, W. Chen, M.W. Wong, J.L. Andres, C. Gonzales, M. Head-Gordon, E.S. Replogle et J.A. Pople. *Gaussian 98, Revision A.9* (1998). Gaussian, Inc., Pittsburgh PA.
- [8] F. Maseras et K. Morokuma. *J. Comp. Chem.* **16**, 1170 (1995).
- [9] S. Dapprich, I. Komáromi, K. Suzie Byun, K. Morokuma et M.J. Frisch. *J. Mol. Struct. (Theochem)* **461-462**, 1 (1999).
- [10] M. Svensson, S. Humbel, R.D.J. Froese, T. Matsubara, S. Sieber et K. Morokuma. *J. Phys. Chem.* **100**, 19357 (1996).
- [11] T. Matsubara, F. Maseras, N. Koga et K. Morokuma. *J. Phys. Chem.* **100**, 2573 (1996).
- [12] J.R. Shoemaker, L.W. Burggraf et M.S. Gordon. *J. Phys. Chem. A* **103**, 3245 (1999).
- [13] T.K. Woo, L. Cavallo et T. Ziegler. *Theor. Chem. Acc.* **100**, 307 (1998).
- [14] J. Gao. *Reviews in Computational Chemistry*, tome 7, chapitre Methods and applications of combined quantum mechanical and molecular mechanical potentials, pages 119–185. VCH, New York (1995).
- [15] T. Wesolowski et A. Warshel. *J. Phys. Chem.* **97**, 8050 (1993).
- [16] T. Wesolowski et A. Warshel. *J. Phys. Chem.* **98**, 5183 (1994).
- [17] I. Tuñón, M.T.C. Martins-Costa, C. Millot et M.F. Ruiz-López. *J. Mol. Model.* **1**, 196 (1995).



- [18] I. Tuñón, M.T.C. Martins-Costa, C. Millot, M.F. Ruiz-López et J.L. Rivail. *J. Comp. Chem.* **17**, 19 (1996).
- [19] S. Chalmet et M.F. Ruiz-López. *Chem. Phys. Lett.* **329**, 154 (2000).
- [20] R.A. Bryce, M.A. Vincent et I.H. Hillier. *J. Phys. Chem. A* **103**, 4094 (1999).
- [21] Stéphanie Chalmet. *Simulation de processus chimiques en phase liquide par des méthodes hybrides QM/MM*. Thèse de doctorat, Université Henri Poincaré Nancy 1 (2000).
- [22] M.J. Field, P.A. Bash et M. Karplus. *J. Comp. Chem.* **11**, 700 (1990).
- [23] M. Eichinger, P. Tavan, J. Hutter et M. Parrinello. *J. Chem. Phys.* **110**, 10452 (1999).
- [24] J. Bentzien, R.P. Muller, J. Florián et A. Warshel. *J. Phys. Chem. A* **102**, 2293 (1998).
- [25] P. Amara, A. Volbeda, J.C. Fontecilla-Camps et M.J. Field. *J. Am. Chem. Soc.* **121**, 4468 (1999).
- [26] P.D. Lyne, M. Hodoscek et M. Karplus. *J. Phys. Chem. A* **103**, 3462 (1999).
- [27] V. Moliner, J. Andrés, M. Oliva, V.S. Safont et O. Tapia. *Theor. Chem. Acc.* **101**, 228 (1999).
- [28] T.K. Woo, P.M. Margl, L. Deng, L. Cavallo et T. Ziegler. *Catal. Today* **50**, 479 (1999).
- [29] S. Hayashi et I. Ohmine. *J. Phys. Chem. B* **104**, 10678 (2000).
- [30] G. Li, S.B.M. Bosio et W.L. Hase. *J. Mol. Struct. (Theochem)* **556**, 43 (2000).
- [31] M. Garcia-Viloca, À. González-Lafont et J.M. Lluch. *J. Am. Chem. Soc.* **123**, 709 (2001).
- [32] P.L. Cummins et J.E. Gready. *J. Phys. Chem. B* **104**, 4503 (2000).
- [33] S. Ranganathan et J.E. Gready. *J. Phys. Chem. B* **101**, 5614 (1997).
- [34] I. Antes et W. Thiel. *J. Phys. Chem. A* **103**, 9290 (1999).
- [35] Y. Zhang, T.-S. Lee et W. Yang. *J. Chem. Phys.* **110**, 46 (1999).
- [36] Y. Zhang, H. Liu et W. Yang. *J. Chem. Phys.* **112**, 3483 (2000).
- [37] R. Poteau, I. Ortega, F. Alary, A. Ramirez Solis, J.-C. Barthelat et J.-P. Daudey. *J. Phys. Chem. A* **105**, 198 (2001).
- [38] R. Poteau, F. Alary, H. Abou El Makarim, J.-L. Heully, J.-C. Barthelat et J.-P. Daudey. *J. Chem. Phys. A* **105**, 206 (2001).
- [39] A. Warshel et M. Levitt. *J. Mol. Biol.* **103**, 227 (1976).
- [40] V. Théry, D. Rinaldi, J.L. Rivail, B. Maigret et G.G. Ferenczy. *J. Comp. Chem.* **15**, 269 (1994).
- [41] G. Monard, M. Loos, V. Théry, K. Baka et J.L. Rivail. *Int. J. Quant. Chem.* **58**, 153 (1996).
- [42] X. Assfeld et J.-L. Rivail. *Chem. Phys. Lett.* **263**, 100 (1996).
- [43] X. Assfeld, N. Ferré et J.-L. Rivail. *ACS Symposium: Combined quantum mechanical and molecular mechanical methods*, chapitre The Local Self Consistent Field principles and applications to

- combined quantum mechanical - molecular mechanical computations on biomacromolecular systems, pages 234–249. 712. A.C.S., Washington, D.C. (1998).
- [44] J. Gao, P. Amara, C. Alhambra et M.J. Field. *J. Phys. Chem A* **102**, 4714 (1998).
- [45] C. Alhambra, J.C. Corchado, M.L. Sánchez, J. Gao et D.G. Truhlar. *J. Am. Chem. Soc.* **122**, 8197 (2000).
- [46] P. Amara, M.J. Field, C. Alhambra et J. Gao. *Theor. Chem. Acc.* **104**, 336 (2000).
- [47] D.M. Philipp et R. A. Friesner. *J. Comp. Chem.* **20**, 1468 (1999).
- [48] R.B. Murphy, D.M. Philipp et R.A. Friesner. *Chem. Phys. Lett.* **321**, 113 (2000).
- [49] R.B. Murphy, D.M. Philipp et R.A. Friesner. *J. Comp. Chem.* **21**, 1442 (2000).
- [50] V. Kairys et J. H. Jensen. *J. Phys. Chem. A* **104**, 6656 (2000).
- [51] M.S. Gordon, M.A. Freitag, P. Bandyopadhyay, J.H. Jensen, V. Kairys et W.J. Stevens. *J. Phys. Chem. A* **105**, 293 (2001).
- [52] S. Antonczak, G. Monard, M.F. Ruiz-López et J.L. Rivail. *J. Am. Chem. Soc.* **120**, 8825 (1998).
- [53] G. Monard et K.M. Merz Jr. *Acc. Chem. Res.* **32**, 904 (1999).
- [54] N. Reuter, A. Dejaegere, B. Maigret et M. Karplus. *J. Chem. Phys.* **104**, 1720 (2000).
- [55] R.J. Hall, S.A. Hindle, N.A. Burton et I.H. Hillier. *J. Comp. Chem.* **21**, 1433 (2000).
- [56] C. Edminston et K. Ruedenberg. *Rev. Mod. Phys.* **35**, 457 (1963).
- [57] J.M. Foster et S.F. Boys. *Rev. Mod. Phys.* **32**, 300 (1960).
- [58] J. Pipek et P.G. Mezey. *J. Chem. Phys.* **90**, 4916 (1989).
- [59] H. Weinstein, R. Pauncz et M. Cohen. *Advances in atomic and molecular physics*, chapitre Localized molecular orbitals, pages 97–140 (1971).
- [60] M. Couty, C.A. Bayse et M.B. Hall. *Theor. Chem. Acc.* **97**, 96 (1997).
- [61] M. Sironi et A. Famulari. *Theor. Chem. Acc.* **103**, 417 (2000).
- [62] M. Sironi, A. Famulari, M. Raimondi et S. Chiesa. *J. Mol. Struct. (Theochem)* **529**, 47 (2000).
- [63] H.B. Schlegel et M.J. Frisch. *Int. J. Quant. Chem.* **54**, 83 (1995).
- [64] K. Németh, O. Coulaud, G. Monard et J. Ángyán. *J. Chem. Phys.* **113**, 5598 (2000).
- [65] C. Zhu, R.H. Byrd, P. Lu et J. Nocedal. *L-BFGS-B: a limited memory FORTRAN code for solving bound constrained optimization problems*. Rapport technique, NAM-11, EECS Department, Northwestern University (1994).
- [66] R.H. Byrd, P. Lu, J. Nocedal et C. Zhu. *SIAM J. Scientific Computing* **5**, 1190 (1995).

- [67] O. Maresca. *Etude théorique d'un catalyseur d'hydrotraitement : étude du support alumine  $\gamma$  et adsorption du sulfure de molybdène*. Thèse de doctorat, Université des Sciences et Technologies de Lille (2000).
- [68] W.L. Jolly, K.D. Bomben et C.J. Eyermann. *Atomic Data and Nuclear Data Tables* **31**, 433 (1984).

## Chapitre 4

# La méthode LSCF/MM/Madelung : définition et paramétrisation

Les principes de notre approche LSCF/MM étant maintenant bien définis, nous pouvons entreprendre l'étude de systèmes étendus. Nous avons choisi de nous intéresser à la réactivité dans les zéolithes pour les raisons exposées dans l'introduction de ce mémoire. Il est important de rappeler que le pouvoir catalytique des zéolithes dépend fortement de la force du site acide de Brønsted sur lequel se déroule la réaction. La relation structure – acidité sera au centre de notre étude.

Le modèle LSCF/MM présenté au chapitre précédent pourrait être appliqué tel quel, en suivant le protocole de simulation indiqué. Cependant, pour bien représenter les interactions électrostatiques du fragment quantique avec le reste du solide, le cluster MM devrait contenir plusieurs milliers ou dizaines de milliers d'atomes, puisque les interactions électrostatiques opèrent encore à longue distance.

Plutôt que de définir un système QM/MM très étendu, nous avons préféré nous limiter à quelques centaines d'atomes que nous plongeons dans le potentiel électrostatique du reste du cristal, appelé potentiel de Madelung [1]. Le couplage de notre méthode QM/MM avec un potentiel externe a été mis au point et est exposé dans une première partie. Nous le validerons sur des exemples simples et comparerons notre approche aux méthodes existantes. Cette méthode a été développée avec János Ángyán au laboratoire.

Enfin nous appliquerons le protocole de simulation tel qu'il est défini au chapitre précédent afin de construire le champ de forces hybride adapté à cette situation. Les liaisons que nous devons couper sont partiellement ioniques, ce qui entraînera des problèmes liés à la polarité des SLBO qui les représenteront. Nous insisterons sur les difficultés que nous avons rencontrées lors de cette paramétrisation, afin d'illustrer les avantages et les inconvénients de notre méthode.

Nous donnerons également quelques résultats préliminaires et quelques pistes pour le futur.

## 4.1 Interactions entre un potentiel externe et une distribution de charges

Les interactions électrostatiques entre une distribution de charges quantiques et un potentiel externe  $V_{ext}(\vec{r})$  (de n'importe quel type) sont souvent modélisées en substituant au potentiel externe un jeu de charges ponctuelles  $\{q_i\}$  (ou de multipôles) sensées représenter au mieux ce potentiel. L'énergie électronique est alors modifiée en incluant dans la matrice de l'Hamiltonien de cœur les intégrales d'interaction électrostatique entre la distribution électronique (exprimée dans une base de  $N$  orbitales atomiques  $\{\chi_\mu\}$ ) et les  $N_q$  charges  $q_i$  :

$$\sum_i^{N_q} \left\langle \chi_\mu \left| \frac{q_i}{r_i} \right| \chi_\nu \right\rangle \quad (4.1)$$

où  $r_i$  représente la distance électron – charge et en ajoutant un terme d'interaction entre les  $N_a$  noyaux et les charges :

$$\sum_a^{N_a} \sum_i^{N_q} \frac{Z_a q_i}{R_{ai}} \quad (4.2)$$

avec  $Z_a$  la charge nucléaire de l'atome  $a$  et  $R_{ai}$  la distance noyau – charge.

Cependant, cette représentation présente certains désavantages. Le premier concerne l'interchangeabilité des descriptions. Considérons deux molécules A et B en interaction. Si, dans un premier temps, A est traitée quantiquement et B est représentée par un jeu de multipôles, et si dans un deuxième temps les descriptions sont inversées, alors les énergies d'interaction électrostatique calculées dans les deux situations sont généralement différentes.

Deuxièmement, la représentation discrète du potentiel externe n'est pas toujours évidente. Dans le cas des zéolithes et des solides en général, le potentiel de Madelung [1] est souvent déterminé par la méthode des sommes d'Ewald [2] et il est difficile de le représenter par un jeu de charges ou de multipôles (ces approches sont habituellement appelées «embedding» et présentées dans l'introduction de ce mémoire). Il est alors souvent nécessaire d'utiliser un jeu très étendu de charges et le nombre d'intégrales (équation 4.1) devient très élevé. D'autres solutions existent, comme le calcul explicite des intégrales d'interaction des électrons et d'une expression analytique du potentiel de Madelung [3].

La méthode que nous proposons permet de conserver l'interchangeabilité des descriptions et de considérer l'influence du potentiel externe sur la distribution de charges quantiques sans aucune autre approximation que la façon de générer ce potentiel. Notre approche fait l'objet d'une publication et nous en rappelons ici les grandes lignes.

### 4.1.1 Théorie

Une distribution de charges quantiques (représentée dans la base  $\{\chi_\mu\}$  par la matrice densité  $\mathbf{P}$ ) qui s'écrit sous la forme :

$$\rho(\vec{r}) = \sum_a^{N_a} Z_a \delta(\vec{r} - \vec{r}_a) - \sum_{\mu\nu}^N P_{\mu\nu} \chi_\mu^*(\vec{r}) \chi_\nu(\vec{r}) \quad (4.3)$$

est en interaction avec un potentiel externe  $V_{ext}(\vec{r})$ . Nous cherchons à modéliser cette interaction par celle entre ce même potentiel et une approximation de la distribution de charges quantiques  $\tilde{\rho}(\vec{r})$  construite à partir d'un jeu de multipôles  $Q_{lm}^a$ , d'ordre  $l$  et de rang  $m$  et centrés sur les atomes  $a$ .

**Détermination des multipôles adaptés au potentiel électrostatique.** Pour conserver la généralité du problème, les multipôles sont déterminés sans référence au potentiel externe. Nous nous plaçons dans le cas d'une molécule isolée dont nous voulons représenter au mieux la distribution électronique. Autrement dit, nous voulons remplacer cette « propriété continue » par une approximation de celle-ci qui est « discrète ».

Les façons de générer ces multipôles sont nombreuses. Par exemple, la théorie des atomes dans les molécules de Bader [4] permet d'obtenir des multipôles transférables, avec cependant l'inconvénient que la longueur de l'expansion multipolaire peut parfois être très grande. La méthode de Stone [5], basée sur les propriétés multiplicatives des fonctions gaussiennes, permet de garder cette expansion à un ordre assez bas. Nous avons préféré utiliser un modèle de multipôles ajustés au potentiel électrostatique, car il donne une description compacte et précise de la distribution électronique.

La détermination de ces multipôles se fait en minimisant la différence (au sens des moindres carrés)

$$\sum_k \sum_{lm} q_{lm}^k \left( \tilde{V}_{lm}^k - V_{lm}^k \right)^2 \quad (4.4)$$

où  $V_{lm}^k$  est la composante  $lm$  du potentiel électrostatique exact dû à la distribution de charges quantiques au point  $k$  de l'espace sur lequel est centré un jeu de multipôles arbitraires  $q_{lm}^k$  :

$$V_{lm}^k = \int d\vec{r}' T_{lm}(\vec{r}'_k, \vec{r}') \rho(\vec{r}') \quad (4.5)$$

et  $\tilde{V}_{lm}^k$  est cette même composante obtenue à partir de la représentation approchée de la distribution de

charges quantiques :

$$\tilde{V}_{lm}^k = \sum_a^{N_a} \sum_{l'm'} Q_{l'm'}^a T_{lm,l'm'}(\vec{r}_k, \vec{r}_a) \quad (4.6)$$

Notre approximation remplace bien une somme continue par une somme discrète.

Sous forme matricielle, cela revient à résoudre les équations suivantes :

$$\mathbf{T}^\dagger \mathbf{q} \mathbf{T} \mathbf{Q} = \mathbf{T}^\dagger \mathbf{q} \mathbf{V} \quad (4.7)$$

La matrice  $\mathbf{T}$  représente l'opérateur d'interaction électrostatique cartésien entre les multipôles centrés sur les atomes et les multipôles situés arbitrairement dans l'espace. Un élément de  $\mathbf{T}$ , la matrice  $T^{KA}$ , est de la forme

$$T^{KA} = \begin{pmatrix} \frac{1}{r_{ka}} & -\frac{x_k-x_a}{r_{ka}^3} & -\frac{y_k-y_a}{r_{ka}^3} & -\frac{z_k-z_a}{r_{ka}^3} & \dots \\ \frac{x_k-x_a}{r_{ka}^3} & \frac{3(x_k-x_a)^2-r_{ka}^3}{r_{ka}^5} & \frac{3(x_k-x_a)(y_k-y_a)}{r_{ka}^5} & \frac{3(x_k-x_a)(z_k-z_a)}{r_{ka}^5} & \dots \\ \frac{y_k-y_a}{r_{ka}^3} & \frac{3(x_k-x_a)(y_k-y_a)}{r_{ka}^5} & \frac{3(y_k-y_a)^2-r_{ka}^3}{r_{ka}^5} & \frac{3(y_k-y_a)(z_k-z_a)}{r_{ka}^5} & \dots \\ \frac{z_k-z_a}{r_{ka}^3} & \frac{3(x_k-x_a)(z_k-z_a)}{r_{ka}^5} & \frac{3(y_k-y_a)(z_k-z_a)}{r_{ka}^5} & \frac{3(z_k-z_a)^2-r_{ka}^3}{r_{ka}^5} & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \quad (4.8)$$

où  $K$  représente l'ensemble des multipôles d'ordre  $l$  et de rang  $m$  centrés au point  $k$  et  $A$  symbolise l'ensemble des multipôles  $l'm'$  centrés sur l'atome  $a$ . Les dimensions de la matrice  $\mathbf{T}$  sont  $N_K \times N_A$  (respectivement le nombre de multipôles centrés sur les points de la grille et sur les centres atomiques).

Le vecteur  $\mathbf{Q}$  réunit l'ensemble des vecteurs  $Q^A$  dont les composantes sont les multipôles atomiques  $Q_{lm}^a$  à déterminer. De la même façon, le vecteur  $\mathbf{q}$  est composé de vecteurs  $q^K$  des multipôles  $q_{lm}^k$  centrés au point  $k$ ; le vecteur  $\mathbf{V}$  est composé des vecteurs  $V^K$  qui collectent les composantes  $lm$  du potentiel électrostatique exact de la distribution de charges quantiques aux différents points  $k$ .

Les solutions aux équations 4.7 sont les suivantes :

$$Q^A = Z^A - \sum_B \sum_K \left( (\mathbf{T}^\dagger \mathbf{q} \mathbf{T})^{-1} \right)^{AB} T^{KB} q^K \sum_{\mu\nu} P_{\mu\nu} V_{\mu\nu}^K \quad (4.9)$$

Dans cette expression, le potentiel électrostatique au point  $K$  est égal à la contraction de la matrice densité  $\mathbf{P}$  avec la matrice des intégrales monoélectroniques  $\mathbf{V}^K$  (dont un élément est  $V_{\mu\nu}^K =$

$\langle \chi_\mu \left| \frac{q_{lm}^k}{r_k} \right| \chi_\nu \rangle$ ). Cette expression peut être écrite plus simplement en

$$Q^A = Z^A - \sum_{\mu\nu} P_{\mu\nu} Q_{\mu\nu}^A \quad (4.10)$$

où les éléments  $A$  du vecteur  $\mathbf{Q}_{\mu\nu}$  s'identifient à

$$Q_{\mu\nu}^A = \sum_B \sum_K \left( (T^\dagger q T)^{-1} \right)^{AB} T^{KB} q^K V_{\mu\nu}^K \quad (4.11)$$

Dans la suite, nous avons choisi de négliger les multipôles  $\mathbf{q}$  (c'est-à-dire d'égaliser à 1) pour garder la généralité de notre approche (dans le cas contraire, ils pourraient jouer le rôle de facteurs de pondération). En utilisant une notation matricielle, l'expression finale des multipôles adaptés au potentiel électrostatique est :

$$\mathbf{Q}_{\mu\nu} = (\mathbf{T}^\dagger \mathbf{T})^{-1} \mathbf{T}^\dagger \mathbf{V}_{\mu\nu} \quad (4.12)$$

telle que  $\mathbf{Q}_{\mu\nu}$  est un vecteur de dimension  $N_A$ ,  $\mathbf{T}$  est une matrice de dimensions  $N_K \times N_A$  et  $\mathbf{V}_{\mu\nu}$  est un vecteur de longueur  $N_K$ .

**Energie et forces d'interaction entre les multipôles adaptés et le potentiel externe.** Après avoir obtenu la meilleure approximation de la distribution de charges par des multipôles, nous pouvons la faire interagir avec le potentiel externe  $V_{ext}(\vec{r})$ . Cette interaction se traduit par un terme supplémentaire dans l'Hamiltonien, dont un élément de matrice doit être corrigé de la quantité :

$$\Delta H_{\mu\nu} = \sum_A Q_{\mu\nu}^A V_{ext}^A \quad (4.13)$$

qui contracte les éléments du vecteur  $\mathbf{Q}_{\mu\nu}$  avec les composantes  $V_{ext}^A$  correspondantes du potentiel externe  $V_{ext}(\vec{r})$  en chaque point  $A$ . La correction à l'énergie correspondante est  $\Delta E =$

$$- \sum_{\mu\nu} P_{\nu\mu} \Delta H_{\mu\nu} + \sum_a Z_a V_{ext}(\vec{r}_a) = \sum_a \sum_k \sum_{lm} \sum_{l'm'} Q_{lm}^a T_{lm,l'm'} q_{l'm'}^k \quad (4.14)$$

La continuité des descriptions est vérifiée dans cette égalité : le membre de gauche, qualifié de « quantique » est égal au membre de droite, dit « classique ». De plus, notons que la fonction d'onde est bien polarisée par le potentiel externe.

La forme analytique des forces résultant de ces interactions est simple à obtenir. Les dérivées pre-



mières de l'expression 4.14 par rapport à une coordonnée nucléaire  $x$  vont corriger les dérivées de l'énergie totale :

$$\frac{\partial E}{\partial x} = \frac{\partial E_0}{\partial x} - \sum_{\mu\nu} \frac{\partial P_{\nu\mu}}{\partial x} \Delta H_{\mu\nu} - \sum_{\mu\nu} P_{\nu\mu} \frac{\partial \Delta H_{\mu\nu}}{\partial x} + \sum_a Z_a \frac{\partial V(\vec{r}_a)}{\partial x} \quad (4.15)$$

Les dérivées des éléments de la matrice densité se traitent de la manière habituelle [6]. Le dernier terme, classique, se calcule facilement. Les dérivées des éléments du terme correctif à l'Hamiltonien (nous les noterons  $\partial \Delta H_{\mu\nu}$  pour plus de commodité) sont assez simples à obtenir : en dérivant l'expression 4.13, il vient

$$\partial \Delta H_{\mu\nu} = \sum_a \sum_{lm} (\partial Q_{lm}^a V_{lm}^a + Q_{lm}^a \partial V_{lm}^a)_{\mu\nu} \quad (4.16)$$

qui s'évalue très simplement en dérivant l'expression 4.12 :

$$\partial \mathbf{Q}_{\mu\nu} = \partial (\mathbf{T}^\dagger \mathbf{T})^{-1} \mathbf{T}^\dagger \mathbf{V}_{\mu\nu} + (\mathbf{T}^\dagger \mathbf{T})^{-1} \partial \mathbf{T}^\dagger \mathbf{V}_{\mu\nu} + (\mathbf{T}^\dagger \mathbf{T})^{-1} \mathbf{T}^\dagger \partial \mathbf{V}_{\mu\nu} \quad (4.17)$$

et en remarquant que

$$\partial (\mathbf{T}^\dagger \mathbf{T})^{-1} = - (\mathbf{T}^\dagger \mathbf{T})^{-1} (\partial \mathbf{T}^\dagger \mathbf{T} + \mathbf{T}^\dagger \partial \mathbf{T}) (\mathbf{T}^\dagger \mathbf{T})^{-1} \quad (4.18)$$

### 4.1.2 Mise en oeuvre et validation

Les expressions développées ci-dessus ont été incorporées dans le logiciel Gaussian98 [7] grâce à la création d'un nouveau «link» (1397) et à la modification du link 701. Elles n'ont pas nécessité le calcul d'intégrales nouvelles : les intégrales du potentiel électrostatique et leurs dérivées premières sur une grille de points quelconques sont suffisantes. En effet, si on développe l'équation 4.13 :

$$\Delta H_{\mu\nu} = \sum_a \sum_k \sum_{lm} \left( (\mathbf{T}^\dagger \mathbf{T})^{-1} \mathbf{T}^\dagger \right)_{lm}^{ak} V_{\mu\nu}^k V_{lm}^a \quad (4.19)$$

et que la somme sur les sites multipolaires  $A$  ( $\sum_a \sum_{lm} \left( (\mathbf{T}^\dagger \mathbf{T})^{-1} \mathbf{T}^\dagger \right)_{lm}^{ak} V_{lm}^a$ ) est effectuée, il reste

$$\Delta H_{\mu\nu} = \sum_k Z_k V_{\mu\nu}^k \quad (4.20)$$

c'est-à-dire que chaque point  $k$  de la grille sur laquelle est calculé le potentiel électrostatique de la distribution de charges quantiques est pondéré par un élément du vecteur  $\mathbf{Z}$ . Le même raisonnement s'applique au calcul des forces.

Dans le cas de solides comme les zéolithes, le potentiel externe est calculé par les sommes d'Ewald, en utilisant le programme MadEwa [8]. Les détails du couplage Gaussian98/MadEwa sont donnés en annexe.

Nous allons maintenant valider notre méthode en vérifiant l'hypothèse d'interchangeabilité des descriptions, puis en l'appliquant à quelques cas simples.

**Interchangeabilité.** Comme expliqué précédemment, l'interchangeabilité caractérise la continuité des descriptions entre une molécule quantique et le potentiel externe qui lui est appliqué. Dans notre cas, elle dépend des ordres des développements multipolaires qui :

1. génèrent le potentiel externe ;
2. représentent la distribution électronique.

Ces ordres doivent être égaux pour que l'interchangeabilité puisse être vérifiée. Deux systèmes moléculaires ont été considérés : le complexe formamide – eau (noté FW) et le dimère de l'eau (noté WW). Dans chacun des cas, quatre calculs sont réalisés. La première molécule est quantique et la deuxième est classique (décrite par des charges ajustées au potentiel électrostatique de la molécule isolée) ; on utilise alors soit le formalisme développé ici (en se limitant à une description de la distribution électronique par des charges), soit l'opérateur d'interaction « direct » (équation 4.1). Ensuite les rôles sont inversés. Dans chacun des calculs, les grilles qui servent à déterminer les charges classiques servent également à la détermination des charges représentant la distribution électronique. De plus, la première estimation de la matrice densité est la fonction d'onde optimisée pour la molécule isolée.

Ainsi, plusieurs informations peuvent être obtenues : l'énergie d'interaction électrostatique pure, cette même énergie après polarisation de la fonction d'onde et l'énergie totale.

Cette méthodologie a été appliquée pour différentes valeurs de distance intermoléculaire (figure 4.1), sans optimisation de géométrie, au niveau de théorie HF/6-31G\*\*. Les résultats sont reportés sur la courbe 4.2(a) pour le complexe formamide – eau et sur la courbe 4.2(b) pour le dimère de l'eau. Elles expriment un taux d'interchangeabilité obtenu selon la formule :

$$\text{Interchangeabilité} = 100 - 100 \times \left| \frac{E(A_Q B_C) - E(A_C B_Q)}{E(A_C B_Q)} \right| \quad (4.21)$$

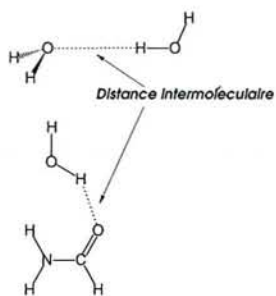


FIG. 4.1 – Définition des distances intermoléculaires dans le dimère de l'eau et dans le complexe formamide-eau.

où  $E$  est une énergie d'interaction (avant ou après polarisation) des molécules A et B, décrites soit quantiquement ( $Q$ ) soit classiquement ( $C$ ).

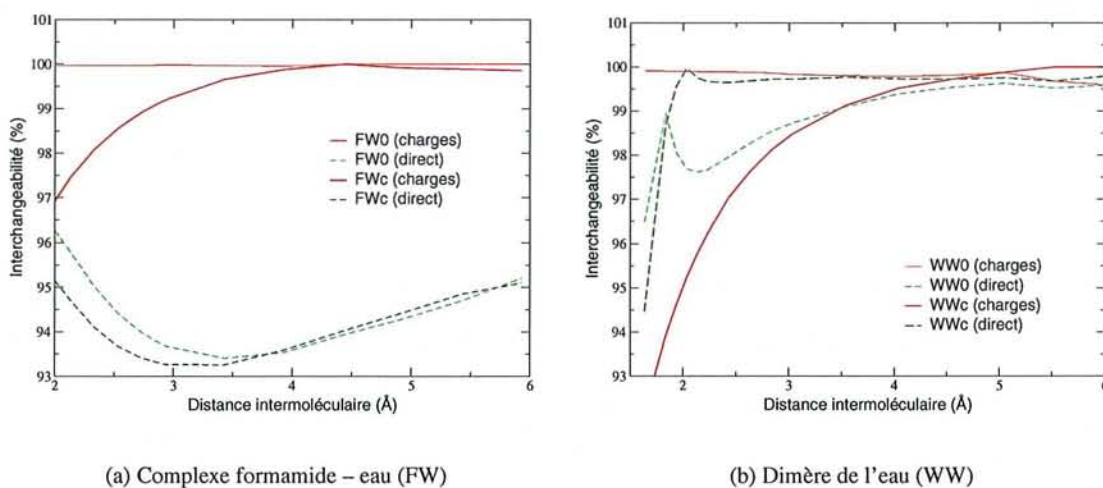


FIG. 4.2 – Interchangeabilité des descriptions pour deux systèmes moléculaires avant polarisation de la fonction d'onde (noté 0) et après polarisation (noté c).

Au vu des courbes 4.2, il est évident que notre formalisme décrit les systèmes quantique et classique de manière cohérente puisque, pour la contribution électrostatique pure (notée 0 sur les courbes), l'interchangeabilité est proche de 100 % quel que soit le système et quelle que soit la distance intermoléculaire. Cette propriété n'est vérifiée qu'aux grandes distances pour la méthode directe.

Nous notons aussi que les comportements de l'énergie de polarisation (notée c) sont différents d'un système à l'autre. Notre méthode autorise l'échange des descriptions des deux sous-unités à partir de 3 Å (interchangeabilité supérieure à 99 %) pour le complexe formamide-eau. Quant à la méthode directe, elle ne permet d'atteindre cette cohérence qu'aux très grandes distances. Cette tendance s'inverse dans le

cas du dimère de l'eau. Dès la distance intermoléculaire d'équilibre, la méthode directe fait preuve d'une très bonne interchangeabilité (supérieure à 99 %). Notre formalisme ne permet d'atteindre ce niveau qu'à partir de 4,5 Å. Ceci est sûrement dû au fait que la description de la molécule d'eau avec des charges est obligatoirement limitée : lorsque le plan contenant la molécule classique contient aussi le moment dipolaire de la molécule quantique, l'utilisation de charges suffit ; dans le cas contraire, la description doit être renforcée à l'aide de dipôles atomiques.

Il est aussi très important de noter qu'aux courtes distances, les charges classiques pénètrent le nuage électronique et en modifie la « forme », ce qui est bien pris en compte par l'approche « directe ». En revanche, si la distribution électronique est représenté par des charges, ce phénomène n'intervient plus.

D'une manière générale, l'interchangeabilité de notre approche est vérifiée, surtout à partir du double de la distance d'équilibre intermoléculaire. Cependant, le niveau de développement des multipôles adaptés au potentiel électrostatique doit être choisi judicieusement selon le système chimique étudié.

Cette dernière remarque sera également illustrée dans les applications suivantes qui nous permettront de valider notre approche tant du point de vue énergétique que géométrique.

**Choix de la grille de points.** A part le choix de l'ordre du développement multipolaire, l'autre paramètre ajustable dans notre approche est le type de la grille de points sur laquelle le potentiel électrostatique électronique est calculé, c'est-à-dire sa forme, sa densité de points, etc . . . Dans ce paragraphe, nous allons discuter de ce choix en nous basant à nouveau sur les deux systèmes de la partie précédente : le dimère de l'eau et le complexe formamide-eau.

La géométrie d'une molécule quantique (Hartree-Fock en base 6-31G\*\*) est optimisée en présence d'une molécule d'eau classique, représentée par le modèle TIP3P (géométrie expérimentale de l'eau liquide, trois charges ponctuelles : -0,834 électron centré sur l'atome d'oxygène et +0,417 centré sur chacun des hydrogènes). Le premier calcul utilise l'opérateur d'interaction habituel (équation 4.1), il sera considéré comme la référence<sup>1</sup>. Les autres utilisent la méthodologie développée ci-dessus, les multipôles ajustés au potentiel électrostatique étant soit des charges (expérience notée 1), soit des charges et des dipôles (expérience notée 2) qui interagissent avec le potentiel crée par la molécule d'eau TIP3P. Pour ces calculs, quatre différents types de grille sont utilisés : Merz-Kollman-Singh (MKS) [10], Francl (CHELP) [11], Breneman (CHELPG) [12] et PNT. Les trois premiers étaient déjà présents dans Gaussian98, le dernier provient du programme Grid3.1 [13]. Tous les paramètres qui permettent de définir une grille sont bien sûr contrôlables, mais il semble que le type de grille importe plus que les caractéristiques

---

1. Nous notons que la véritable énergie d'interaction électrostatique entre deux molécules quantiques pourraient être obtenue à l'aide de calculs de type SAPT [9], mais ces calculs sont très lourds et n'ont pas été entrepris.

propres d'une grille comme l'espacement entre deux points ou l'extension. Pour cette raison, les résultats concernant le dimère de l'eau reportés dans le tableau I comparent uniquement les types de grille entre eux.

TAB. I – Dimère de l'eau.

Différences d'énergie ( $\Delta E$  en kcal.mol<sup>-1</sup>) par rapport à la référence (dans la même géométrie). Géométries optimisées (en Å et en degrés).

	Ref	MKS		CHELP		CHELPG		PNT	
		1	2	1	2	1	2	1	2
$\Delta E$	0,000	0,609	0,183	0,714	0,339	0,623	0,206	0,685	0,172
Distance OH	0,9446	0,9446	0,9444	0,9444	0,9442	0,9446	0,9444	0,9445	0,9444
Angle HOH	105,46	105,53	105,60	105,65	105,74	105,52	105,61	105,43	105,59
Distance intermoléculaire	2,032	2,032	2,032	2,033	2,033	2,032	2,032	2,032	2,032

L'utilisation de charges et de dipôles ajustés au potentiel électrostatique (expérience 2) permet de se rapprocher de façon systématique des résultats de référence, par rapport aux calculs ne faisant intervenir que des charges (expérience 1). La molécule d'eau étant très polaire, l'utilisation de charges pour représenter la distribution électronique ne peut être que limitée (nous avons fait la même remarque concernant l'interchangeabilité). Mais dans tous les cas, les différences d'énergie sont inférieures à 1 kcal.mol<sup>-1</sup>. Les géométries d'équilibre varient très peu d'une grille à l'autre et sont très proches des valeurs de référence.

Les mêmes résultats concernant le complexe formamide-eau sont donnés dans le tableau II. Les géométries d'équilibre de la molécule quantique variant très peu d'une méthode à l'autre, nous ne reportons pas leurs valeurs.

TAB. II – Complexe formamide-eau.

Différences d'énergie ( $\Delta E$  en kcal.mol<sup>-1</sup>) par rapport à la référence (dans la même géométrie). Géométries optimisées (en Å, l'indice *f* indique le formamide, l'indice *w* indique l'eau).

	Ref	MKS		CHELP		CHELPG		PNT	
		1	2	1	2	1	2	1	2
$\Delta E$	0,000	0,548	0,686	1,116	0,528	0,517	0,764	0,607	0,628
Distance O <sub>f</sub> H <sub>w</sub>	1,907	1,909	1,912	1,901	1,909	1,908	1,912	1,903	1,911
Distance O <sub>w</sub> H <sub>f</sub>	2,056	2,057	2,058	2,054	2,057	2,057	2,058	2,051	2,059

Sauf le cas de la grille CHELP, l'ajout de dipôles aux charges ajustées n'améliore pas l'accord avec le calcul de référence, mais ici encore, les variations d'énergie restent petites. Cette tendance s'observe

également dans les distances intermoléculaires (sauf, légèrement, pour la grille CHELP et la grille PNT). La grille PNT semble posséder le comportement le plus stable.

Cette première étude montre l'efficacité de notre méthodologie : même dans le cas d'une molécule polaire, une description de la distribution électronique par des charges (et éventuellement des dipôles) ajustées au potentiel est satisfaisante.

Concernant le choix d'un type de grille, il est difficile à décider, tant les résultats sont similaires. Cependant, nous avons choisi de ne plus utiliser dans la suite les grilles de type CHELP (ses résultats semblent les moins précis) et de type CHELPG (les grilles générées comportent souvent un grand nombre de points, ce qui demande beaucoup de temps de calcul et de ressources informatiques).

Dans cette première application, les énergies d'interaction électrostatique sont relativement faibles, comparées à l'énergie totale du système. L'augmentation de leurs proportions devrait nous permettre de mieux quantifier les résultats de notre méthode.

**Solvation du formamide.** Dans cette nouvelle application, la molécule de formamide est placée dans une boîte de deux cent seize molécules d'eau toujours représentées par le modèle TIP3P (l'arête de la boîte est 18,8 Å pour une densité de 1). Comme précédemment, la géométrie du formamide est optimisée au niveau HF/6-31G\*\*. Dans cette étude, nous nous basons encore sur des résultats de référence obtenus avec l'opérateur usuel. Dans les calculs utilisant notre formalisme, trois grilles sont utilisées : MKS, PNT et une grille « moléculaire » dont les points sont centrés sur les atomes des molécules de solvant. Les multipôles adaptés au potentiel sont limités aux charges. Ce choix a été guidé par les résultats de l'application précédente.

Les géométries optimisées avec les trois différentes grilles sont toutes en excellent accord (les variations de longueur de liaison sont toutes inférieures au pm). Pour une géométrie donnée, les différences énergétiques par rapport à la référence restent faibles : 0,007 kcal.mol<sup>-1</sup> pour la grille PNT, 0,038 kcal.mol<sup>-1</sup> pour la grille MKS et 0,155 kcal.mol<sup>-1</sup> pour la grille « moléculaire ». Comme cette dernière n'est pas assez régulière pour permettre de prendre en compte toutes les fluctuations du potentiel électrostatique, l'expansion limitée aux charges n'est alors pas suffisante. En revanche, si l'on utilise des charges et des dipôles, la différence d'énergie diminue à 0,006 kcal.mol<sup>-1</sup>, confirmant ainsi notre analyse.

La grille PNT semblant donner les meilleurs résultats (et d'autres études que nous ne reportons pas ici indiquent la même tendance), nous la choisissons comme grille par défaut dans la suite.

**Premier test QM/MM/Madelung : NaCl.** Pour cette dernière application, nous avons choisi de tester le champ de forces hybride QM/MM/Madelung ainsi défini sur un système solide typique. Le choix s'est porté naturellement sur le chlorure de sodium NaCl. Dans sa géométrie expérimentale, le solide est cubique à faces centrées et la distance entre deux atomes voisins est 2,82 Å. Nous allons montrer l'intérêt qu'il y a à utiliser une description hybride afin d'obtenir une valeur théorique de cette distance suffisamment proche.

Bien entendu, le modèle le plus simple consiste en un atome de sodium et un atome de chlore. Quelle que soit la méthode quantique et la base envisagée, la distance d'équilibre entre ces deux atomes est toujours largement inférieure à la valeur expérimentale. La valeur obtenue au niveau de théorie BPW91/6-311+G\* sera notre point de départ : 2,396 Å.

La première amélioration apportée à notre modèle consiste à simuler son environnement par un continuum diélectrique de constante relative  $\epsilon = 6,12$  (constante diélectrique expérimentale du chlorure de sodium solide). La molécule de NaCl est placée au centre d'une cavité créée au sein du continuum. L'interaction est traitée de façon auto-cohérente. Au niveau de théorie BPW91/6-311+G\*, la distance d'équilibre Na-Cl s'établit à 2,476 Å, ce qui indique clairement que la prise en compte, même sommaire, des interactions électrostatiques est essentielle.

Cette description sommaire du solide considère l'environnement isotrope et continu. Il est possible d'affiner cette représentation en traitant de manière explicite les plus proches voisins : une boîte de cent huit atomes (cube de trois motifs NaCl de côté) est construite. Un traitement purement classique (champ de forces UFF [14]) permet d'obtenir une distance Na-Cl de 2,499 Å.

Maintenant, au centre de la boîte, un atome de sodium et un atome de chlore sont décrits quantiquement (BPW91/6-311+G\*); le reste des atomes est traité par la mécanique moléculaire. Les interactions QM/MM sont uniquement de types non-liées (électrostatique et van der Waals) telles que définies dans UFF. Le reste du cristal est d'abord négligé puis il est pris en compte par l'interaction de son potentiel de Madelung (calculé par les sommes d'Ewald) avec le cluster QM/MM. Nous utilisons alors le formalisme décrit plus haut : la distribution électronique est ajustée par des charges uniquement.

Les distances Na-Cl quantiques résultantes sont 2,485 Å et 2,620 Å, respectivement pour le calcul QM/MM et le calcul QM/MM/Madelung. La première valeur est du même ordre que celles obtenues avec le diélectrique ou le champ de forces. Elle confirme ainsi la nécessité de considérer l'environnement proche de la molécule quantique. Les effets électrostatiques à longue distance sont spectaculaires : la distance Na-Cl augmente d'environ 15 pm par rapport au calcul QM/MM. Cet exemple justifie pleinement l'incorporation du potentiel de Madelung dans notre modèle.

Il est vrai que le résultat QM/MM/Madelung est encore assez éloigné de la valeur expérimentale. Plusieurs facteurs expliquent cet accord qualitatif. Le premier met en cause l'utilisation de UFF comme champ de forces classique, qui n'est pas le plus adapté à la description des solides ioniques comme NaCl (et qui justifie le fait que la position des atomes classiques soit gelée). Le deuxième provient de la dissymétrie des descriptions : le fragment quantique n'est pas juste au centre du cluster QM/MM, les effets ressentis à « droite » ne sont pas équivalents à ceux ressentis à « gauche », ce qui provoque une distorsion de la géométrie du cluster.

Néanmoins, ces résultats montrent l'importance des effets environnementaux dans les solides, surtout s'ils sont à caractère ionique. A ce point du développement méthodologique, nous pouvons maintenant définir le modèle hybride QM/MM/Madelung dédié à l'étude de la réactivité dans les zéolithes (ou les solides cristallisés en général).

## 4.2 Le champ de forces hybride LSCF/MM/Madelung

Avant de débiter le protocole de simulation, les systèmes qui nous serviront à effectuer la reparamétrisation du champ de forces doivent être définis. Nous choisissons la chabazite qui est une zéolithe de petite maille unitaire, et pour cette raison, très étudiée théoriquement et expérimentalement. Elle est essentiellement constituée de silicium et d'oxygène. Cependant la substitution d'un atome de silicium par un atome d'aluminium permet d'envisager quatre sites acides potentiels. Nous étudierons ces différents sites acides (notés O1, O2, O3 et O4 dans la figure 4.3).

Ces sites se différencient par le type d'anneaux dans lesquels ils sont engagés (anneaux à 4, 6 ou 8 atomes d'oxygène). Dans le tableau III sont récapitulées ces structures.

TAB. III – *Engagements des différents sites acides dans les structures en anneaux.*

Site	Anneaux à 4	Anneaux à 6	Anneaux à 8
O <sub>1</sub>	2	0	1
O <sub>2</sub>	1	1	1
O <sub>3</sub>	2	1	0
O <sub>4</sub>	1	0	2



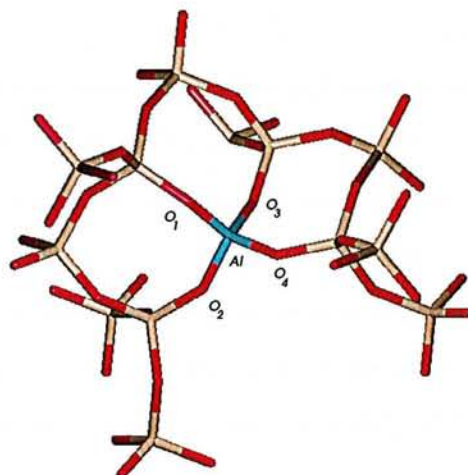


FIG. 4.3 – Site de substitution dans la chabazite : quatre sites possibles pour placer un proton.

#### 4.2.1 Etude préliminaire

Avant de chercher à comparer ces quatre sites, nous avons voulu valider notre approche sur un seul d'entre eux ( $O_1$ ) en réduisant la partie quantique à sa plus simple expression : Si–OH–Al et les six SLBO qui terminent ce fragment, au niveau BPW91/6-31G\*. La partie classique (décrite par le champ de forces UFF) complète les trois anneaux (deux anneaux à quatre oxygènes et un anneau à huit).

Dans un premier temps, nous avons comparé notre modèle à différentes autres approches :

- deux clusters quantiques  $SiAlOH_7$  et  $SiAlO_7H_7$  au niveau BPW91/6-31G\* ;
- notre système modèle du site acide défini ci-dessus et traité par UFF puis ONIOM (BPW91/6-31G\*:UFF) ;
- un modèle périodique quantique de la chabazite avec une substitution et le site acide, par VASP [15].

Nos calculs se contenteront d'optimiser les positions des atomes O et H du site acide uniquement. Deux calculs LSCF/MM sont réalisés : le premier sans couplage avec le potentiel de Madelung (dans ce cas, les atomes de silicium de la partie MM composant les anneaux sont saturés avec des atomes d'hydrogène) et le deuxième en incorporant le potentiel généré par le reste du solide. Le système adopté ici permet de nous affranchir momentanément des problèmes liés à la description des frontières QM/MM puisque celles-ci sont fixes.

Dans le tableau IV sont comparés les différents termes géométriques qui caractérisent le site acide.

TAB. IV – Comparaison des géométries du site acide  $O_1$  pour différents méthodes (distances en Å et angles en degrés). La dernière colonne donne des valeurs moyennes telles que postulées par Sauer [16] à partir de calculs quantiques sur différents clusters.

Paramètres	SiAlOH <sub>7</sub>	SiAlO <sub>7</sub> H <sub>7</sub>	UFF	ONIOM	LSCF / MM	LSCF / MM / Mad	VASP	Sauer
Si–O	1,720	1,742	1,641	1,678	1,748	1,743	1,702	1,69
Al–O	1,979	1,926	1,919	1,885	1,983	1,881	1,904	1,94
O–H	0,976	0,978	0,855	0,967	0,969	0,971	0,974	0,96
Si–O–Al	126,6	128,7	136,3	136,2	124,7	120,5	132,9	130
Si–O–H	116,6	117,1	118,0	116,6	115,9	115,5	116,4	
Al–O–H	116,8	112,8	105,7	107,2	105,0	109,8	110,6	
Si–O–Al–H	1,0	14,7	0,04	0,3	46,8	44,1	0,6	

La géométrie du site acide décrit par notre approche est en accord qualitatif avec les calculs les plus sophistiqués de type VASP. Le plus gros désaccord concerne l'angle diédral caractérisant la position du proton acide par rapport au plan formé par les atomes de silicium, d'oxygène et d'aluminium. L'atome d'oxygène étant coordonné trois fois, il nous semble logique que le site acide soit pyramidal, ce que montre bien notre modèle et qui est confirmé par l'expérience [17]. Cependant, d'autres paramètres comme les longueurs de liaison Si–O et Al–O sont mal décrits : notre modèle n'est bien sûr pas encore assez développé à ce point. On note que l'incorporation des effets du potentiel de Madelung ne sont pas négligeables, ils corrigent notamment la longueur de liaison Al–O de 0,1 Å. Cette exemple permet de mettre en lumière la nécessité de mettre au point un champ de forces hybride dédié à l'étude des zéolithes.

Malgré les problèmes de paramétrisation, nous pouvons espérer que les erreurs se conservent lorsqu'on change de système moléculaire étudié. Nous avons comparé l'évolution des paramètres géométriques du site acide entre la chabazite substitué et son équivalent silicoaluminophosphate HSAPO–34 (par rapport à la chabazite, les atomes de silicium sont remplacés alternativement par des atomes d'aluminium et de phosphore et les aluminium de substitution par des silicium). Les calculs sont effectués sur le même cluster que précédemment. Les variations chabazite → HSAPO–34 sont reportées dans le tableau V.

D'après les tendances exprimées, il semble que le couplage d'une méthode QM/MM avec le champ de Madelung du cristal soit complètement justifié. Le champ de forces UFF n'est pas correctement paramétré, surtout pour les systèmes de type silicoaluminophosphates. ONIOM modifie peu la structure du

TAB. V – Variation des longueurs de liaison caractéristiques entre la chabazite et HSAPO-34 en pm pour différentes méthodes.

	UFF	ONIOM	LSCF/MM	LSCF/MM/Mad	VASP
$\Delta\text{Si-O}$	+4	+3	+3	+5	+6
$\Delta\text{Al-O}$	+10	0	-5	-5	-6

site acide, ce qui est sûrement dû à l'absence de polarisation du fragment quantique par l'environnement classique. La méthode LSCF/MM différencie bien les deux structures et l'ajout du potentiel de Madelung permet de se rapprocher davantage des tendances observées dans un calcul périodique.

#### 4.2.2 Molécule modèle

La comparaison des quatre sites acides dans la chabazite ne sera possible du point de vue QM/MM que si les quatre études utilisent la même partition du système. De plus, comme la prise en compte de la corrélation électronique s'avère importante dans ces systèmes, nous utiliserons la DFT avec la fonctionnelle BPW91 et la base polarisée 6-31G\* (ce niveau se rapproche le plus de celui utilisé avec le logiciel VASP). Ces deux considérations nous ont conduit à envisager deux types de coupure QM/MM qui permettent de traiter quantiquement les quatre sites acides possibles tout en gardant dans la partie quantique un nombre de fonctions de base raisonnable (figure 4.4). Le premier niveau de coupures concerne les douze liaisons Si-O les plus proches du site de substitution, tandis que le deuxième coupe les douze liaisons O-Si suivantes. Pour des raisons de capacités informatiques (en mémoire et en temps), la première partition est bien sûr la plus attrayante.

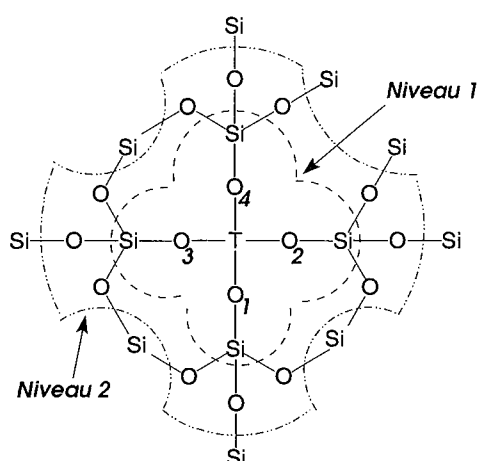


FIG. 4.4 – Coupures QM/MM envisagées dans la chabazite : douze liaisons frontières sont nécessaires. Le futur site de substitution est au centre du fragment QM.

Le champ de forces utilisé pour décrire le sous-système classique est UFF [14] car c'est le seul dans Gaussian98 [7] à traiter le silicium, l'aluminium et l'oxygène (les paramètres de ce dernier sont spécialement adaptés aux zéolithes).

Le cluster QM/MM retenu comprend plus de quatre cents atomes, le fragment quantique étant placé aux environs de son centre. Lors d'une optimisation de géométrie, il n'est pas nécessaire de relaxer la position de tous ces atomes à cause de la rigidité de l'édifice cristallin. En conséquence, seuls les atomes situés jusqu'à cinq liaisons du centre de substitution pourront bouger librement. Ce nombre est le même que celui utilisé dans les études QM/MM effectuées par de Vries [18] : les résultats obtenus par notre méthode pourraient alors être comparés aux siens.

L'application du protocole de simulation LSCF/MM, défini au chapitre précédent, aux zéolithes n'est pas très aisée. En effet, la frontière entre les sous-systèmes quantique et classique met en jeu des liaisons au caractère à la fois covalent et ionique. Pour représenter au mieux ces liaisons, nous avons retenu le critère de localisation de Weinstein-Pauncz, l'utilisation du critère de Magnasco-Perico ne semblant pas apporter une meilleure description de la liaison.

La première étape du protocole concerne la définition de la molécule modèle des liaisons frontières. D'après les coupures QM/MM envisagées, trois modèles ont été étudiés. Les deux premiers modèles se rapportent à la première coupure (silicium quantique et oxygène classique) :

- $\text{H}_3\text{Si-OH}$  car c'est le modèle envisageable le plus simple ;
- $\text{Si(OH)}_4$ , car trois liaisons frontières Si-O sont portées par un seul atome de silicium.

Pour la deuxième coupure envisagée (l'oxygène est quantique et le silicium est classique), nous avons uniquement étudié le modèle  $\text{HO-SiH}_3$ .

La validation du modèle sera basée sur deux critères qui seront comparés à des calculs quantiques :

- charges de Mulliken des atomes frontières ;
- géométries de la partie quantique (la longueur de la liaison frontière étant fixée).

**Premier modèle de frontière.** Les résultats obtenus ne sont pas satisfaisants. Les charges de Mulliken portées par les deux atomes frontières ne sont pas du tout correctes (tableau VI). L'atome frontière côté classique (noté Y dans le chapitre précédent) n'apporte qu'un électron au système quantique. En revanche, l'ensemble des fonctions de base associées à cet atome est utilisé pour former la SLBO. Bien entendu, cette dernière utilise majoritairement les fonctions de valence, et les fonctions de cœur restent vides. Celles-ci sont d'autant plus basses en énergie que l'atome Y est lourd et elles vont donc « attirer »

les électrons du système quantique, ce qui perturbe beaucoup le système et conduit à une mauvaise polarisation de la frontière.

TAB. VI – Charges de Mulliken de  $H_3SiOH$  en électrons. Les valeurs entre parenthèses sont les charges « unifiées » qui ajoutent à la charge de l'atome lourd celles des hydrogènes liés. La charge donnée pour les atomes d'hydrogène liés au silicium est la moyenne des trois.

Atome	SCF	$Si_q-O_{+1}$	$Si_q-O_{+3}$
Si	0,897 (0,402)	1,163	0,235
O	-0,865(-0,402)	-0,787	0,125
H(O)	0,464	classique	classique
H(Si)	$\simeq -0,165$	$\simeq -0,130$	$\simeq -0,120$

Les solutions possibles sont de deux types : soit on incorpore explicitement les électrons de cœur de l'atome frontière Y dans le calcul LSCF/MM, soit ceux-ci sont représentés par un potentiel de cœur effectif. Dans le cas de l'oxygène, la deuxième solution n'est pas réalisable avec les pseudo-potentels usuels. Il faudrait alors mener toute une étude en vue de déterminer un pseudo-potentel spécifique à l'oxygène. Nous avons donc préféré la première solution : deux électrons sont ajoutés au système. Au total, l'oxygène frontière apportant trois électrons, une charge +3 est alors nécessaire pour conserver l'électronneutralité du système. Les charges de Mulliken qui résultent de cette stratégie sont également reportées dans le tableau VI.

Malheureusement, il semble que l'ajout explicite des deux électrons de cœur n'améliore pas la description de la frontière Si–O, les charges de Mulliken sont encore plus erronées que dans le cas précédent.

A ces problèmes de polarisation de la frontière s'ajoutent également des problèmes géométriques qui se sont révélés lorsque nous avons utilisé ces SLBO Si–O dans le système zéolithique que nous désirions étudier (pour rappel, quatre frontières mettant en jeu trois SLBO chacune, de type Si–O). Les structures optimisées présentent alors des anomalies sévères : les liaisons purement quantiques issues des atomes de silicium frontières sont beaucoup trop longues (de 0,1 à 0,2 Å). Le fait que trois SLBO soient portées par le même atome quantique entraîne une très mauvaise description de la quatrième liaison. L'ajout des électrons de cœur des atomes d'oxygène Y ne suffit pas non plus à ramener cette liaison à une longueur convenable.

Tous ces problèmes nous ont amenés à rejeter cette molécule modèle de la frontière.

**Deuxième modèle de frontière.** Dans le modèle précédent, les SLBO étaient déterminées sans prendre en compte la particularité de la première coupure : trois liaisons frontières par atome quantique X. Le choix de la nouvelle molécule modèle s'est alors naturellement porté sur  $SiO_4H_4$ . La procédure de loca-

lisation portera ainsi sur une densité électronique qui reflète bien la situation à la frontière.

Sur cette molécule, les trois SLBO sont d'abord déterminées selon le critère de Weinstein-Pauncz (CWP), puis par le critère de Pipek-Mezey (CPM). Le premier critère conduit à une localisation indépendante des deux autres liaisons à localiser, tandis qu'avec le deuxième critère, les trois orbitales de liaison sont localisées simultanément.

De la même façon que pour le premier modèle, nous nous sommes intéressés aux charges de Mulliken portées par les atomes frontières ainsi qu'aux géométries optimisées, afin de trouver la meilleure représentation des liaisons frontières. Ainsi les calculs LSCF quantique et les calculs LSCF/MM (sans ou avec les électrons de cœur des oxygènes frontières, respectivement notés LSCF/MM<sub>+1</sub> et LSCF/MM<sub>+3</sub>) sont comparés avec un calcul SCF dans le tableau VII (niveau de théorie BPW91/6-31G\* et UFF). Comme nous l'avons montré dans le chapitre précédent, il est possible de modifier la polarité de la SLBO en multipliant les coefficients des orbitales atomiques centrées sur un des deux atomes par un facteur à ajuster. D'autres calculs LSCF/MM utilisant ces SLBO modifiées ont été effectués et sont notés LSCF/MM<sub>p</sub>. Nous ne reportons que l'exemple donnant les meilleurs résultats dans le tableau VII.

TAB. VII – Charges de Mulliken (en électrons) et distance d'équilibre (en Å) de la liaison Si–O (ne portant pas de SLBO) dans la molécule modèle SiO<sub>4</sub>H<sub>4</sub>. Les oxygènes frontières sont notés O<sub>f</sub> et les valeurs indicées c sont les charges classiques portées par les hydrogènes classiques afin de conserver l'électroneutralité du système.

	SCF	CWP				CPM	
		LSCF	LSCF/MM <sub>+1</sub>	LSCF/MM <sub>p</sub>	LSCF/MM <sub>+3</sub>	LSCF/MM <sub>+1</sub>	LSCF/MM <sub>+3</sub>
q <sub>Si</sub>	0,94	0,93	1,76	1,36	1,51	1,77	2,33
q <sub>O<sub>f</sub></sub>	-0,64	-0,63	-0,69	-0,55	-0,64	-0,70	-0,80
q <sub>O</sub>	-0,64	-0,65	-0,80	-0,81	-0,73	-0,80	-0,72
q <sub>H<sub>c</sub></sub>	0,40	0,40	0,25 <sub>c</sub>	0,25 <sub>c</sub>	0,25 <sub>c</sub>	0,25 <sub>c</sub>	0,25 <sub>c</sub>
q <sub>H</sub>	0,40	0,42	0,36	0,35	0,39	0,37	0,41
d(Si–O)	1,65	1,67	1,86	1,73	1,86	1,74	1,73

De nouveau, les résultats ne sont pas satisfaisants. Les valeurs obtenues par LSCF sont en bon accord avec les résultats SCF, ce qui montre encore une fois que ce ne sont pas les SLBO qui perturbent le système. En revanche, dès qu'on utilise un modèle LSCF/MM, la charge de Mulliken de l'atome de silicium augmente considérablement, c'est-à-dire qu'il perd des électrons au profit des oxygènes voisins. La seule liaison Si–O non représentée par une SLBO tend à s'allonger, ce qu'illustre la dernière ligne du tableau VII. Ni l'ajout des électrons de cœur, ni la modification de la polarité des SLBO ne permettent d'obtenir une description correcte.

De toute évidence, la proximité des trois liaisons frontières perturbe énormément la distribution électronique et il semble vain de chercher à corriger celle-ci.

**Dernier modèle de frontière.** La décision qui s'impose alors est de diminuer les interactions entre liaisons frontières. Pour cela, les frontières QM/MM doivent être déplacées une liaison plus loin du site de substitution dans la chabazite.

Aucun atome quantique frontière ne porte plusieurs SLBO, il n'est donc pas nécessaire d'avoir recours à une molécule modèle compliquée. En conséquence, nous utilisons à nouveau le modèle minimal  $H_3Si-OH$ , mais cette fois-ci, l'atome d'oxygène est quantique tandis que l'atome de silicium est classique. Comme pour les deux modèles précédents, une étude portant sur les charges de Mulliken est réalisée (tableau VIII).

TAB. VIII – Charges de Mulliken de  $H_3SiOH$  en électrons. Les valeurs entre parenthèses sont les charges « unifiées » qui ajoutent à la charge de l'atome lourd celles des hydrogènes liés.

Atome	SCF	$Si_{+1}-O_q$	$Si_{pseudo}-O_q$
Si	0,897 (0,401)	0,173	0,451
O	-0,865(-0,401)	-0,612	-0,856
H(O)	0,464	0,439	0,405
H(Si)	$\simeq$ -0,165	classiques	classiques

Le silicium est un atome qui possède dix électrons de cœur. Comme remarqué précédemment, les orbitales atomiques de cœur du silicium frontière ne sont que très peu engagées dans la SLBO définie sur la liaison frontière et peuvent donc se remplir avec les électrons du système. Avec la base 6-31G\*, la polarité de la liaison frontière est alors erronée, comme en témoigne la colonne  $Si_{+1}-O_q$  du tableau VIII. La description du cœur du silicium par un pseudo-potentiel est alors possible afin de simuler ces électrons. Nous avons choisi la base de qualité double- $\zeta$  LANL2DZ. La polarité de la liaison frontière est alors mieux représentée : la charge de Mulliken portée par l'oxygène diffère de 0,01 électron de celle résultant d'un calcul quantique, tandis que la charge du silicium est en bon accord avec la charge « unifiée » du groupement  $SiH_3$  (0,451 au lieu de 0,401). Nous notons cependant que la charge de Mulliken de l'atome d'hydrogène alcoolique est inférieure à la valeur de référence et que le pseudo-potentiel tend à diminuer cette valeur. L'influence de celui-ci peut se résumer très schématiquement par «  $Si \Rightarrow O \rightarrow H$  » où la multiplicité des flèches symbolise l'intensité des effets du pseudo-potentiel de cœur.

Ce dernier modèle de la frontière est le plus performant que nous ayons testé et nous avons décidé de le retenir pour la suite du protocole de simulation LSCF/MM/Madelung.

L'augmentation de la taille du fragment quantique nous oblige également à alléger les calculs, et la modélisation des électrons de cœur pour tous les atomes de silicium et d'aluminium par des pseudo-potentiels permet effectivement de résoudre ce problème de taille.

Le système quantique fixé à cette taille comporte deux cent dix électrons dont cinquante traités par les pseudo-potentiels de cœur, les autres sont représentés à partir de trois cent cinquante deux fonctions de base (nous avons choisi la base 6-31G\* pour les atomes d'oxygène et la base LANL2DZ pour les atomes de silicium et d'aluminium).

### 4.2.3 Mise au point du champ de forces hybride pour la liaison O–Si

L'étape suivante, après avoir choisi la molécule modèle de la frontière, concerne la paramétrisation du potentiel de liaison frontière. Celui-ci est déterminé sur le système  $\text{H}_3\text{SiOH}$  dans la base mixte LANL2DZ pour le silicium et 6-31G\* pour les autres atomes. Les paramètres ajustés sont donnés dans le tableau IX.

TAB. IX – Paramètres du potentiel de liaison frontière O–Si (en unités atomiques)

A	B	C	D	E
-10,30	4,43	-0,84	-0,94	0,75

La différence la plus remarquable avec les potentiels obtenus pour les liaisons organiques habituelles (chapitre précédent) concerne le dernier paramètre E qui est très petit. Cela signifierait que la prise en compte des électrons de cœur à travers un pseudo-potentiel est suffisante pour corriger en grande partie la répulsion ion-noyau. Cette remarque est justifiée lorsqu'on détermine ce même paramètre sans utiliser de pseudo-potentiel : la valeur de E est alors très élevée (environ 40).

L'étape suivante du protocole concerne l'ajustement des charges classiques à la frontière ; dans notre cas, il s'agira surtout de la charge portée par l'atome de silicium. Notons que les charges portées par les hydrogènes classiques ne servent qu'à neutraliser la charge portée par le silicium. La charge totale portée par chacun des sous-systèmes quantique et classique est nulle.

Alors que dans le chapitre précédent le critère était énergétique, nous avons préféré dans ce cas utiliser un critère géométrique, qui permette d'ajuster au mieux la distance Si–O. De plus, nous veillerons à ce que la charge de Mulliken portée par l'atome de silicium soit proche de celle obtenue dans un calcul quantique.

La valeur retenue est 0,5 électron (en plus de la charge +1 qui permet de conserver l'électroneutralité du système quantique) : elle permet d'avoir un bon compromis géométrie–charges. Les valeurs correspondantes sont reportées dans le tableau X.

Les résultats sont très satisfaisants, excepté pour l'angle de liaison Si–O–H, trop petit de  $10^\circ$  et les longueurs des liaisons classiques Si–H gérées par UFF. Mais ces divergences ne sont pas inquiétantes,



TAB. X – Comparaison des géométries optimisées de  $H_3Si-OH$  avec les méthodes SCF et LSCF/MM après paramétrisation du potentiel de liaison frontière et ajustement des charges à la frontière (distances en Å, angles en degrés et charges de Mulliken en électron). Les atomes d'hydrogène portés par le silicium sont classiques dans le calcul LSCF/MM et leurs charges sont adaptées aux charges de Mulliken obtenues dans le calcul SCF ; les valeurs entre parenthèses sont les charges de Mulliken « unifiées » du silicium.

	SCF	LSCF/MM
Si–O	1,647	1,638
O–H	0,946	0,947
Si–O–H	119,0	108,2
H–Si	1,478 ; 1,470	1,380 ; 1,391
$q_{Si}$	0,897 (0,402)	0,958 (0,458)
$q_O$	-0,865	-0,875
$q_{H(O)}$	0,464	0,417
$q_{H(Si)}$	-0,173 ; -0,150	-0,175 ; -0,150

elles sont dues au modèle minimaliste retenu et elles devraient disparaître dans les calculs portant sur le macro-système.

Dans le modèle classique retenu pour la chabazite, les atomes de silicium sont chargés +1 et les oxygènes -0,5 (nous considérons que pour chaque liaison Si–O, l'atome d'oxygène prend 0,25 électrons à chaque atome de silicium voisin). A la frontière QM/MM, un atome de silicium est engagé pour un quart dans la partie quantique : la charge classique qu'il porte devrait alors être 0,75. Or, la charge adaptée à la frontière est 0,5 : il faut donc compenser cette différence en répartissant les 0,25 électrons par atome de silicium frontière sur les atomes classiques voisins. Cette modification reste mineure et ne doit pas beaucoup perturber les interactions QM–MM et MM–MM.

Le champ de forces hybride LSCF/MM/Madelung est maintenant prêt. La dernière étape consiste donc à le valider.

### 4.3 Validation du champ de forces : application LSCF/MM/Madelung à la chabazite pure

Le champ de forces hybride LSCF/MM/Madelung est maintenant mis au point. Cependant, son applicabilité au système-modèle présenté auparavant reste à prouver. Pour cela, nous avons optimisé la géométrie des soixante-six atomes (c'est-à-dire, tous les atomes quantiques ainsi que les deux premières « couches » classiques) centrés sur le futur site de substitution (que nous noterons T) dans le modèle de la chabazite pure.

Pour résumer notre modèle, le fragment quantique est composé de vingt-et-un atomes et cent soixante

électrons pour trois cent cinquante deux fonctions de base (LANL2DZ pour les atomes de silicium et 6-31G\* pour les atomes d'oxygène) traités au niveau DFT avec la fonctionnelle BPW91. Le reste des atomes composant le cluster QM/MM (environ trois cent cinquante atomes) est décrit avec le champ de forces classique UFF, les charges portées par les atomes sont +1 pour Si et -0,5 pour O, sauf dans le cas des atomes frontières Si dont la charge vaut +0,5. Douze liaisons frontières de type  $O_q-Si_c$  sont définies pour connecter les sous-systèmes quantique et classique. Le potentiel de Madelung du cristal est déterminé par les sommes d'Ewald dans un cristal parfait composé de charges ponctuelles +1 et -0,5. L'interaction entre ce potentiel et le cluster QM/MM utilise l'approche exposée dans la première partie de ce chapitre : la distribution électronique est représentée par des charges ponctuelles ajustées au potentiel électrostatique calculé sur une grille de type PNT.

Les géométries optimisées sont reportées dans le tableau XI et comparées d'abord à un calcul LSCF/MM puis aux résultats obtenus par Jeanvoine [15] à partir d'un calcul quantique périodique et, enfin, à l'expérience.

TAB. XI – Géométries optimisées autour du futur site de substitution T (en Å et en degrés). Les données expérimentales sont obtenues par rayons X ( $a=9,421$  Å et  $\alpha=94,2$  ° dans le groupe d'espace R-3m) ou par diffraction des neutrons sur la chabazite substituée [17].

	LSCF/MM	LSCF/MM /Madelung	VASP	Rayons X	Neutrons
T-O <sub>1</sub>	1,701	1,672	1,616	1,643	1,617
T-O <sub>2</sub>	1,654	1,664	1,605	1,642	1,599
T-O <sub>3</sub>	1,665	1,665	1,612	1,650	1,615
T-O <sub>4</sub>	1,636	1,654	1,612	1,652	1,613
Si-O <sub>1</sub>	1,644	1,671	1,616	1,643	1,617
Si-O <sub>2</sub>	1,694	1,691	1,605	1,642	1,600
Si-O <sub>3</sub>	1,668	1,665	1,612	1,650	1,615
Si-O <sub>4</sub>	1,700	1,675	1,612	1,652	1,615
T-O <sub>1</sub> -Si	135,7	136,2	154,1	146,5	144,8
T-O <sub>2</sub> -Si	151,2	147,5	150,6	149,2	149,4
T-O <sub>3</sub> -Si	148,6	148,1	150,2	147,6	147,8
T-O <sub>4</sub> -Si	165,0	167,3	144,1	140,9	150,0
O <sub>1</sub> -T-O <sub>2</sub>	104,2	106,8	109,7	111,8	110,3
O <sub>1</sub> -T-O <sub>3</sub>	104,3	107,5	109,0	112,0	110,2
O <sub>1</sub> -T-O <sub>4</sub>	108,0	110,2	111,8	107,9	107,9
O <sub>2</sub> -T-O <sub>3</sub>	112,5	111,1	109,5	107,7	109,7
O <sub>2</sub> -T-O <sub>4</sub>	116,2	112,7	107,9	106,1	108,4
O <sub>3</sub> -T-O <sub>4</sub>	110,6	108,5	108,9	111,1	110,2

La nécessité d'incorporer les effets électrostatiques à longue portée est parfaitement illustrée en comparant les résultats LSCF/MM et LSCF/MM/Madelung. L'écart maximal entre les longueurs des liaisons autour du silicium central est beaucoup plus important dans le premier cas que dans le deuxième : 6,5

pm contre 3,7 pm. Le potentiel de Madelung du reste du cristal tend donc à symétriser ces liaisons. Les angles Si–O–Si et O–Si–O varient relativement peu.

La comparaison au calcul périodique et à l'expérience est plus difficile. Si l'expérience de diffraction des neutrons sert de référence (il s'agit donc d'une géométrie moyenne de la chabazite substituée), nos résultats surestiment les longueurs de liaison de 4 à 9 pm. Quant aux angles, les valeurs LSCF/MM/Madelung sont en assez bon accord, excepté les angles T–O<sub>1</sub>–Si et T–O<sub>4</sub>–Si qui diffèrent de 8,6° et 17,3° respectivement. En revanche, en se basant sur la structure de la chabazite pure générée à partir des données cristallographiques, les écarts sur les longueurs de liaison sont alors inférieurs à 5 pm (quant aux angles, ils varient peu selon l'expérience considérée). Plusieurs explications de ces divergences sont possibles.

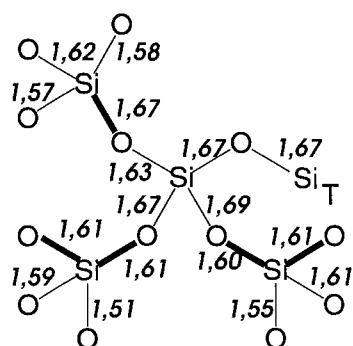
Une source possible d'erreur provient des coupures QM/MM. La multiplication de celles-ci provoque nécessairement des perturbations importantes qui diminuent lorsqu'on s'éloigne des frontières. Dans notre modèle, les atomes quantiques sont au plus à quatre liaisons de la frontière la plus proche. Cet effet est illustré dans la figure 4.5.

Deux cas de figure peuvent se présenter pour les frontières : l'atome de silicium de type Y est engagé dans une ou deux liaisons frontières (on parlera alors de frontières « simple » ou « double »). Dans le premier cas, la longueur de la liaison frontière varie entre 1,658 et 1,701 Å ; dans le deuxième cas, cette longueur varie entre 1,591 et 1,652 Å. Il est évident que le potentiel de liaison frontière, tel que défini précédemment, n'est pas adapté à la deuxième situation, ce qui provoque ce rétrécissement des deux liaisons frontières. En conséquence, les liaisons immédiatement voisines s'allongent parfois exagérément (de 1,674 à 1,727 Å au lieu de 1,626 à 1,652 Å pour les frontières « simples »).

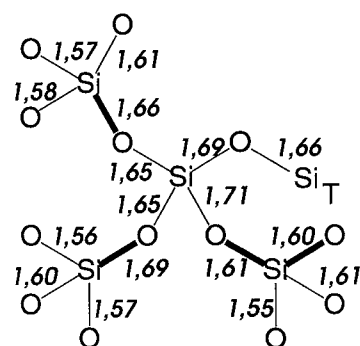
D'après ces résultats, il est clair que la présence de frontières multiples est source d'erreurs supplémentaires. Une solution consisterait alors à définir un potentiel de liaison frontière qui dépend explicitement du nombre de frontières portées par l'atome Y (et donc de sa charge). Le travail de détermination des paramètres du potentiel s'en trouverait alors compliqué, et surtout, la transférabilité des potentiels serait perdue. Une deuxième solution nécessiterait de situer la frontière QM/MM encore plus loin du site de substitution. Mais à l'heure actuelle, les calculs qui en découleraient seraient difficilement envisageables.

Concernant les liaisons définies par des atomes classiques uniquement, celles qui sont issues d'un silicium frontière sont très courtes (de 1,478 à 1,621 Å) ; les autres varient entre 1,60 et 1,70 Å approximativement. Ce dernier résultat montre que le champ de forces UFF n'est que peu adapté à notre problème, malgré la définition de paramètres spéciaux pour les atomes d'oxygène.

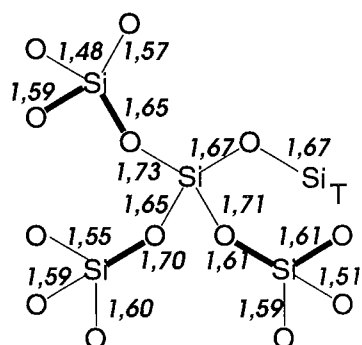
Au vu de ce test, le modèle LSCF/MM/Madelung peut donc être considéré comme partiellement



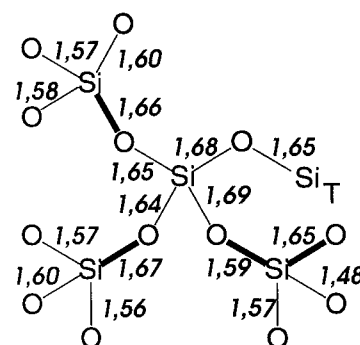
(a) Fragment O1



(b) Fragment O2



(c) Fragment O3



(d) Fragment O4

FIG. 4.5 – Longueurs de liaisons (en Å) pour différents fragments correspondants aux quatres sites acides possibles. Les liaisons frontière sont en gras.

achevé. Bien évidemment, il peut encore être amélioré, grâce à un raffinement des frontières. Mais ceci entraînera sûrement une perte de transférabilité.

En outre, cette application constitue un cas limite de la méthode LSCF/MM/Madelung en ce sens que nous cherchons à simuler une structure *très* symétrique à l'aide d'une méthodologie hybride, asymétrique par essence. Lors de la substitution d'un atome de silicium par un atome d'aluminium, cette symétrie disparaît et notre méthode devrait s'appliquer plus facilement.

## 4.4 Conclusion

Dans ce dernier chapitre dédié à la mise au point du champ de forces hybride QM/MM/Madelung, nous avons dans un premier temps établi une nouvelle méthode de traitement des interactions entre un cluster QM/MM et un potentiel externe. Ce développement, guidé par l'idée d'incorporer les effets du potentiel de Madelung d'un solide parfait dans un modèle théorique d'un centre acide zéolithique, a été ensuite validé sur des exemples de taille croissante. L'interchangeabilité des descriptions classiques et quantiques a également été vérifiée.

Dans un deuxième temps, le champ de forces hybride a été défini. La liaison  $O_q-Si_c$  s'est avérée être la seule à donner des résultats encourageants. Un potentiel de liaison frontière et des charges ajustées ont alors été déterminés.

Ainsi défini, le champ de forces hybride a été testé sur un modèle élaboré de la chabazite pure. A la fin de ces travaux, nous pensons que notre approche est partiellement validée, en ce sens qu'elle nécessiterait encore une certaine paramétrisation du champ de forces hybride. De plus, l'utilisation dans le logiciel d'un champ de forces mieux adapté aux zéolithes s'avère nécessaire.

Bien d'autres améliorations sont possibles pour perfectionner notre modèle. La qualité de la première estimation de la matrice densité est capitale et on ne peut se contenter de la façon standard de la générer à partir d'un calcul de type semi-empirique ou Hückel. Les interactions QM/MM dans ces systèmes sont loin d'être parfaitement décrites dans le champ de forces classique utilisé. L'utilisation d'un champ de forces spécialement dédié au matériau zéolithique (comme celui de Sauer [19]) ne peut qu'améliorer les résultats.

Les futures applications de notre approche sont nombreuses : étude comparée des différents sites acides après substitution du silicium par un aluminium ; applications à des systèmes de grande maille unitaire ; multiplication des sites de substitution . . . Tous ces projets sont en cours à l'heure actuelle.

## Bibliographie

- [1] E. Madelung. *Physik. Z.* **19**, 524 (1918).
- [2] P.P. Ewald. *Ann. Physik* **64**, 253 (1921).
- [3] G. Ferenczy, G.I. Czonka, G. Náray-szabó et J.G. Ángyán. *J. Comp. Chem.* **19**, 38 (1998).
- [4] R.F.W. Bader. *Atoms in Molecules – A Quantum Theory*. Oxford University Press (1990).
- [5] A.J. Stone. *Chem. Phys. Lett.* **83**, 233 (1981).
- [6] J.A. Pople, R. Krishnan, H.B. Schlegel et J.S. Binkley. *Int. J. Quant. Chem.: Quant. Chem. Symp.* **13**, 225 (1979).
- [7] M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, V.G. Zakrzewski, J.A. Montgomery, Jr., R.E. Stratmann, J.C. Burant, S. Dapprich, J.M. Millam, A.D. Daniels, K.N. Kudin, M.C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G.A. Petersson, P.Y. Ayala, Q. Cui, K. Morokuma, D.K. Malick, A.D. Rabuck, K. Raghavachari, J.B. Foresman, J. Cioslowski, J.V. Ortiz, A.G. Baboul, B.B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R.L. Martin, D.J. Fox, T. Keith, M.A. Al-Laham, C.Y. Peng, A. Nanayakkara, C. Gonzales, M. Challacombe, P.M.W. Gill, B. Johnson, W. Chen, M.W. Wong, J.L. Andres, C. Gonzales, M. Head-Gordon, E.S. Replogle et J.A. Pople. *Gaussian 98, Revision A.9* (1998). Gaussian, Inc., Pittsburgh PA.
- [8] J. Ángyán, G. Ferenczy et P. Nortier. *Madewa 1.0* (1996).
- [9] R. Moszynski, P.E.S. Wormer, B. Jeziorski et Ad van der Avoird. *J. Chem. Phys.* **103**, 8058 (1995).
- [10] B.H. Besler, K.M. Merz Jr. et P.A. Kollman. *J. Comp. Chem.* **11**, 431 (1990).
- [11] L.E. Chirlian et M.M. Francl. *J. Comp. Chem.* **8**, 894 (1987).
- [12] C.M. Breneman et K.B. Wiberg. *J. Comp. Chem.* **11**, 361 (1990).
- [13] C. Chipot et J. Ángyán. *Grid 3.1 point multipoles derived from molecular electrostatic properties* (1996).
- [14] A.K. Rappé, C.J. Casewit, K.S. Colwell, W.A. Goddard III et W.M. Skiff. *J. Am. Chem. Soc.* **114**, 10024 (1992).
- [15] Y. Jeanvoine. *Modélisation ab initio des silicates, aluminosilicates et silicoaluminophosphates. Etude structurale et dynamique des sites acides de Bronsted dans HSSZ-13 et HSAPO-34*. Thèse de doctorat, Université Henri Poincaré, Nancy I (1997).
- [16] J. Sauer, P. Ugliengo, E. Garrone et V.R. Saunders. *Chem. Rev.* **94**, 2095 (1994).

- [17] L.J. Smith, A. Davidson et A.K. Cheetham. *Catal. Lett.* **49**, 143 (1997).
- [18] A.H. de Vries, P. Sherwood, S.J. Collins, A.M. Rigby, M. Rigutto et G.J. Karner. *J. Phys. Chem. B* **103**, 6133 (1999).
- [19] J. Sauer et M. Sierka. *J. Comp. Chem.* **26**, 1470 (2000).

## Chapitre 5

# Conclusion générale

L'objectif principal de ce travail était le développement d'un nouvel outil dédié à l'étude de la réactivité catalysée par les zéolithes. Pour cela, nous avons effectué un travail en trois étapes.

- La première a permis de développer une méthode quantique qui détermine la fonction d'onde électronique, connaissant a priori certaines orbitales moléculaires qui restent gelées durant le calcul SCF. Les équations de Roothaan ont été modifiées et les dérivées premières de l'énergie ont été obtenues. Des perspectives concernant la modélisation d'états excités ou ionisés ont été données ; nous nous sommes particulièrement intéressés aux ionisations des électrons de cœur.
- La deuxième étape a consisté à coupler la méthode LSCF avec un champ de forces classique grâce à la définition et à la paramétrisation de potentiels de liaison frontière et également en adaptant les charges classiques à la frontière. Un protocole de simulation LSCF/MM a été proposé et il a été démontré que le champ de forces hybride obtenu peut être appliqué à l'étude de protéines ou de réactions organiques. Nous avons également appliqué avec succès le principe d'étude des ionisations de cœur au cas de la crambine.
- La troisième étape a permis d'incorporer dans la description du système les effets électrostatiques à longue distance dus au potentiel de Madelung du reste du solide supposé parfait. Une méthodologie nouvelle a été développée, basée sur une représentation approchée de la distribution de charges QM/MM, et a été validée. Puis le protocole de simulation LSCF/MM a été appliqué sur un modèle LSCF/MM/Madelung de la chabazite pure. Un champ de forces hybride a ainsi été défini et paramétré pour les liaisons frontières de type silicium-oxygène. Enfin, celui-ci a été validé sur le modèle de la chabazite.



Les potentialités du logiciel développé sont nombreuses. La méthode quantique LSCF accepte :

- un formalisme restreint (RHF) ou non-restreint (UHF) ;
- des spinorbitales prédéfinies de n'importe quelle forme ;
- un nombre d'occupation de ces spinorbitales variable (0 ou 1) ;
- le calcul des dérivées premières analytiques de l'énergie par rapport aux coordonnées nucléaires ;
- la prise en compte de la corrélation électronique aux niveaux DFT, MPn et CI.

La modification de Gaussian98 offre les possibilités suivantes pour les calculs hybrides QM/MM :

- calcul QM/MM avec ou sans SLBO ;
- optimisation complète de la géométrie du système QM/MM ;
- polarisation du fragment quantique par les charges ponctuelles de l'environnement ;
- couplage avec trois champs de forces (Amber, UFF, Dreiding) ;
- potentiels de liaison frontière pour de nombreuses liaisons chimiques typiques ;
- charges ponctuelles adaptées à la frontière ;
- algorithme d'optimisation de géométrie en coordonnées internes à croissance linéaire, couplé avec le minimiseur L-BFGS-B .

Enfin, la méthode hybride à trois niveaux QM/MM/Madelung permet :

- taille quelconque de la cellule unitaire ;
- défauts chimiques traités explicitement (QM ou MM) ;
- interaction entre le cluster QM/MM et le potentiel externe via des multipôles d'ordre 0 ou 1 ;
- quatre types de grille, de tailles et de formes modulables ;
- optimisation complète du cluster QM/MM ;
- potentiel externe généré par les sommes d'Ewald (programme MadEwa) ou les sommes directes de charges et/ou dipôles.

L'application du formalisme QM/MM à l'étude de la réactivité dans les zéolithes, comme dans tous les sites macromoléculaires complexes, ne peut être envisagée sans suivre une démarche stricte, qui passe forcément par la validation de la méthode. Ces trois années ont été l'occasion de tester les différentes composantes du champ de forces hybride sur de nombreux cas typiques ou atypiques. Ce n'est que lorsque cette étape a été menée à bien que la méthode peut être « automatisée ». Un premier pas dans ce sens a été accompli en définissant un protocole de simulation QM/MM.

De nombreux développements restent à accomplir. Chaque conclusion de chapitre a permis de mettre en lumière le futur de chaque composante. Nous énumérerons ici les plus importantes : LSCF «linear scaling », étude des ionisations de cœur dans les systèmes étendus, couplage LSCF *ab initio* et LSCF semi-empirique, etc ...

Concernant la simulation de la réactivité dans les zéolithes, nous avons maintenant à notre disposition un outil adapté et performant. Les études à mener sont nombreuses. Il sera particulièrement intéressant de modéliser les effets de plusieurs substitutions sur des sites en interaction. L'activité catalytique en est sûrement modifiée. De plus, l'étude des zéolithes de grande maille unitaire (et d'intérêt industriel) est maintenant accessible.

La dernière remarque de ce mémoire aura pour thème l'originalité. La validation d'une technique nouvelle en chimie théorique passe d'abord par la reproduction de résultats que d'autres ont obtenus soit par l'expérience, soit par des calculs utilisant une méthodologie plus ou moins proche. Les champs d'application qui permettent une telle validation sont souvent assez réduits. Il est malheureusement dommage que beaucoup d'entre elles se cantonnent dans ces domaines et n'essaient pas d'explorer de nouvelles thématiques. Autrement dit, ces méthodes, qui ne manquent pas d'originalité dans leurs conceptions, en manquent souvent beaucoup dans leurs applications et se contentent de calculer ce que d'autres ont déjà calculé. Le formalisme LSCF et les techniques qui l'utilisent sont suffisamment novateurs pour que leur avenir soit assuré si ce « piège » est évité ; les pistes données vont dans ce sens.



## Annexe A

# Modifications du logiciel Gaussian98

Dans cette dernière partie, nous passerons en revue les modifications qui ont été effectuées dans Gaussian98 [1] (en abrégé G98). Ce logiciel de chimie quantique est constitué d'un ensemble de programmes (appelés «links») exécutés successivement, qui communiquent et partagent des informations. Au cours de nos travaux, nous avons modifié les links 101, 301, 402, 502, 601, 701 et nous avons créé le link 397.

Les travaux qui allient calcul quantique et calcul classique n'étant pas prévus dans G98, nous devons utiliser une route non-standard dont un exemple suit (cas d'une optimisation de géométrie QM/MM) dans le tableau I. Nous reprenons ici les notations de G98 et nous nous limitons au premier pas d'optimisation. Bien évidemment, cette procédure est répétée jusqu'à convergence de la géométrie.

1/ ... /1,3	Informations moléculaires ; optimisation de géométrie
2/ ... /2	Vérification des données moléculaires
4/ ... /2	Calcul MM (énergie et forces)
3/ ... /1,2,3,97	Intégrales ; Madelung
4/ ... /1	Estimation de la matrice densité
5/ ... /2	(L)SCF
6/ ... /1	Analyse de population ; localisation des SLBO
7/ ... /1,2,3,16	Calcul des forces
1/ ... /3	Optimisation de géométrie

TAB. I – Route non-standard d'un calcul QM/MM/Madelung typique. Les ... figurent les options que l'on passe aux links symbolisés par x/y (x=3 et y=1 pour le link 301 par exemple).

Les options que l'on passe aux links seront notées dans la suite IOP(ref)=val avec ref et val des nombres entiers. Nous allons maintenant détailler chaque link modifié ou créé. Il faut noter que les modifications permettent également d'effectuer des calculs QM/MM sans liaisons frontières et donc n'utilisant pas LSCF.

## Link 101

Ce programme est dédié à la lecture et au traitement des spécifications moléculaires : charge, multiplicité, type d'atomes, géométrie, connectivité . . .

C'est donc dans ce link que nous allons distinguer les atomes classiques des atomes quantiques, avec une attention particulière portée aux atomes frontières de type Y (du côté MM de la frontière). Après les coordonnées géométriques de l'atome, une lettre symbolise l'appartenance de celui-ci à un sous-système : H («High») pour un atome quantique, L («Low») pour un atome classique et M («Middle») pour un atome Y. On peut également spécifier le nombre d'électrons qu'apporte cet atome en plaçant un nombre entier N strictement positif après la lettre M. Ces informations sont stockées dans un tableau LTYPE de dimension NATOMS (nombre d'atomes de la molécule) tel que

Atome H	LTYPE(Atome) = 0
Atome M	LTYPE(Atome) = N
Atome L	LTYPE (Atome) = -1

En cas d'absence d'une information, les défauts sont H pour le type d'atome et 1 pour le nombre d'électrons.

Dans ce link sont aussi analysés les types MM des atomes et codés sous forme numérique (ces codes sont stockés dans un tableau IATTYP). Par exemple, l'oxygène zéolithique est noté O\_3\_z dans le champ de forces UFF : il est alors codé 100108004 (quatrième type de l'atome de numéro atomique 8 dans le champ de forces codé 10). Le signe de ce code permet d'«activer» ou de «désactiver» un atome (si le code est positif, l'atome doit être considéré ; s'il est négatif, l'atome doit être ignoré), ce qui sera très utile pour, par exemple, ne pas attribuer de fonctions de base aux atomes classiques.

Une option a été ajoutée afin de permettre les calculs QM/MM (avec ou sans frontières) : IOP(98) différent de 0.

## Link 402

Ce link est dédié aux calculs semi-empiriques et, pour ce qui nous intéresse, aux calculs de mécanique moléculaire avec un des trois champs de forces proposés (sélectionné par l'IOP(20)={9 pour Dreiding [2], 10 pour UFF [3], 11 pour Amber [4]). L'activation du code QM/MM se fait automatiquement.

A l'entrée de ce link, tous les codes MM (tableau IATTYP) sont positifs afin de permettre la définition des interactions classiques, y compris entre les sous-systèmes quantique et classique.

La première partie des modifications apportées à ce link concerne l'attribution des charges ponctuelles MM aux atomes. Bien évidemment, les atomes QM n'ont pas de charges MM. Les atomes MM ont des charges, soit lues dans le fichier de données, soit déterminées par l'algorithme dit d'«équilibrage de charges» de Rappé et Goddard III [5]. Dans le cas d'un atome frontière Y, une charge égale à LTYPE(Y) est ajoutée à la charge MM. La charge totale MM est calculée et affichée pour vérification.

Ensuite, les paramètres du champ de forces utilisé doivent être déterminés. Ceux-ci sont déterminés pour chaque liaison, angle, dièdre et terme non-lié défini dans la connectivité sans prise en compte de la qualité (QM ou MM) des atomes mis en jeu. A ces paramètres sont ajoutés ceux qui définissent les potentiels de liaison frontière. Ils sont au nombre de cinq par liaison frontière.

Arrive ensuite le calcul de chaque contribution à l'énergie MM : une table de vérité est établie pour chaque type de contribution (par exemple, neuf cas possibles pour le potentiel de liaison A–B, résumés dans le tableau II).

A	B	Compté ?
Q	Q	NON
Q	Y	NON
Q	C	NON
Y	Q	NON
Y	Y	OUI
Y	C	OUI
C	Q	NON
C	Y	OUI
C	C	OUI

TAB. II – Table de vérité des termes de liaison dans le champ de forces hybride QM/MM. Un atome quantique est noté Q, un atome classique C et un atome frontière côté MM Y.

Evidemment, certains cas n'ont aucun sens : ils correspondent aux termes faisant intervenir une liaison Q–C (par exemple, un dièdre C–C–Q–Q). Cependant, leur absence nous permet de vérifier la validité de la connectivité utilisée.

Concernant les termes non-liés, nous distinguons les interactions de van der Waals (potentiel de Lennard-Jones dans tous les champs de forces utilisés) des interactions électrostatiques. Les premières sont soumises à la condition «1-4» même pour les termes QM/MM, elles sont donc toutes calculées dans ce link. Les secondes ne sont soumises à la condition «1-4» qu'entre charges MM. En conséquence, seules celles-ci seront calculées ici, les interactions électrostatiques QM/MM le seront dans le link 301 avec la répulsion nucléaire.

Mentionnons tout de suite le cas particulier d'une interaction électrostatique entre deux atomes Y. Le

premier, noté  $Y_a$ , possède une charge égale à  $q_{MM}(a) + LTYPE(a)$  et le deuxième,  $Y_b$  une charge  $q_{MM}(b) + LTYPE(b)$ . Seul le terme  $\frac{q_{MM}(a) \cdot q_{MM}(b)}{r_{ab}}$  soumis à la condition «1-4», est calculé dans ce link, les trois autres le sont dans le link 301.

Les contributions des liaisons frontières sont calculées dans la subroutine XYOL, commune à tous les champs de forces.

Bien entendu, si nécessaire, les forces correspondantes sont calculées.

Il faut noter qu'une option permet d'activer ou non chacun des types d'interaction : IOP(63)= 111111 (ancienne valeur par défaut) active toutes ces contributions, un 0 à la place d'un 1 les désactivent. Un autre 1 est considéré : IOP(63)=1111111 (nouveau défaut) active les potentiels de liaison frontière, 0111111 les désactive.

Enfin, les calculs classiques étant achevés, tous les atomes MM voient leurs codes IATTYP changés de signe.

## Link 301

Ce link initialise tout ce qui est nécessaire au calcul des intégrales (bases, pseudo-potentiels ... ). Il calcule également toutes les contributions non-électroniques à l'énergie. Les modifications QM/MM sont activées automatiquement.

Grâce au tableau IATTYP, aucune fonction de base n'est attribuée aux atomes classiques (sauf les atomes frontières Y). Le décompte des électrons du système est modifié : pas d'électrons pour les atomes MM et LTYPE(Y) électrons pour un atome frontière de type Y.

Comme mentionné dans la section précédente, les interactions électrostatiques non-soumises à la condition «1-4» sont calculées en même temps que la répulsion nucléaire.

## Link 397 / MadEwa

Ce link permet de prendre en compte un potentiel externe dans un calcul QM/MM (dans le cas d'un cristal, le potentiel de Madelung). Il a été presque entièrement créé : il comprend dix-huit sous-routines soit mille quatre cents lignes de code. A cela, il faut ajouter le programme MadEwa [6] qui est lié au link 397 sous forme de librairie et qui est entièrement dédié au calcul du potentiel de Madelung par la méthode des sommes d'Ewald. Il a également été modifié, et nous y reviendrons à la fin de cette section.

Le link est guidé par la routine principale : Cryst. Celle-ci effectue les opérations suivantes :

1. Calculer les valeurs du potentiel externe (et de ses dérivées premières et secondes par rapport aux coordonnées nucléaires) sur les sites atomiques QM et MM. Le potentiel externe peut être obtenu par le programme MadEwa, ou plus simplement, en spécifiant dans le fichier de données un certain nombre de charges ponctuelles (et, éventuellement, de dipôles ; la syntaxe est identique à celle utilisée avec le mot-clé CHARGE). Le choix est guidé par l'IOP(99) : positif pour utiliser MadEwa, négatif pour les charges.
2. Générer une grille de points sur lesquels le potentiel électrostatique dû aux électrons sera calculé. Cette grille peut être lue ou calculée à partir d'une des quatre procédures disponibles. La décision est prise à partir de la valeur que prend l'IOP(97) : Merz-Kollman (1), CHELP (2), CHELPG (3), PNT (4). La forme et la taille de cette grille peuvent être contrôlées à partir de cet IOP également.
3. Calculer le tenseur cartésien  $T$  qui permettra de déterminer les charges / charges + dipôles ajustées au potentiel électrostatique. La décision d'une représentation est donnée par l'IOP(99) : charges (1), charges + dipôles (2). Le vecteur  $Z$  peut ensuite être calculé à partir de  $T$  et du potentiel externe.
4. Calculer les intégrales du potentiel électrostatique électronique sur la grille. Chaque point de la grille est alors pondéré par la valeur de  $Z$  correspondante.
5. Mettre à jour l'Hamiltonien de coeur avec ces intégrales.
6. Calculer les contributions non-électroniques à l'énergie totale : noyaux QM et charges MM sont soumis au potentiel externe.
7. En cas d'optimisation de géométrie, calculer les dérivées du vecteur  $Z$  par rapport aux coordonnées nucléaires.

Il faut alors sauvegarder en mémoire le potentiel externe, la grille, le vecteur  $Z$  et ses dérivées.

**MadEwa** Comme mentionné précédemment, ce programme permet de calculer (entre autres) le potentiel de Madelung d'un cristal parfait représenté par des charges ponctuelles (ou des multipôles) en un certain nombre de points. Il a été écrit par J. Ángyán, G. Ferenczy et P. Nortier. La version qui a été utilisée et modifiée date de 1996.

La communication entre G98 et MadEwa se fait par l'intermédiaire de fichiers de données et de résultats utilisés par MadEwa. Le fichier de données doit être préparé à l'avance et nommé «fort.55». Il doit contenir toutes les spécifications nécessaires à la modélisation du cristal ainsi que les coordonnées



des atomes dont les charges sont retirées du potentiel. Quand G98 appelle MadEwa, il lui passe alors les coordonnées des atomes QM et MM. MadEwa peut alors être exécuté et G98 cherche dans le fichier de résultats «fort.56» les valeurs du potentiel de Madelung et de ses dérivées.

Un exemple de fichier de données est représenté ci-dessous :

```
$ Essai sur la chabazite
CALC EWAL
Si 1 .104400 .333800 .874900 1.
...
O 1 .672300 .975200 .975200 -0.5
CELL A 9.421 B 9.421 C 9.421 ALPHA 94.2 BETA 94.2 GAMMA 94.2
SYMM 1 x,y,z
COOR INPUT FRAC UNIT ANG
COOR QUANTUM CART UNIT ANG
COOR OUTPUT CART UNIT BOHR
ENER UNIT AU
DERI 2
QUAN X 0.977840 Y 3.058770 Z -1.480920
...
QUAN X 15.663290 Y 17.316700 Z 16.091669
SITE ...
```

Le mot-clé QUAN introduit les coordonnées des atomes dont la contribution au potentiel est à retirer.

Le mot-clé SITE est suivi des coordonnées des atomes QM et MM.

Les modifications apportées à MadEwa ont permis de retirer les contributions d'atomes ne se situant pas forcément dans la maille (0,0,0) et de calculer le pontentiel de Madelung en tout point du cristal. Ces extensions étaient nécessaires car le cluster QM/MM s'étend souvent sur plus d'une maille unitaire.

Nous avons également permis l'expression des coordonnées des atomes à supprimer dans le repère cartésien ou fractionnaire.

## Link 502

Ce link est dédié à la résolution des équations SCF. Il a d'abord été modifié par X. Assfeld pour inclure la méthode LSCF: seuls des calculs d'énergie en formalisme restreint étaient possibles. Nous l'avons donc modifié pour permettre :

- n'importe quel calcul LSCF : HF ou DFT, en formalisme restreint ou non-restreint ;
- l'utilisation de l'algorithme DIIS («Direct Inversion of Iterative Subspace») [7] ;

- la lecture des OG en fichier de données ou en mémoire ainsi que les spins associés et leur nombre d'occupation, dans le cas de SLBO le trièdre d'atomes de référence est également lu ;
- la modification des SLBO par rotation à chaque nouvelle géométrie ;
- le calcul de l'énergie totale QM/MM/Madelung ;
- le calcul de la matrice **W** nécessaire aux dérivées premières de l'énergie LSCF ;
- la préparation des OG à un calcul post-Hartree-Fock.

Une option gouverne LSCF : l'IOP(99) indique le nombre d'OG et de SLBO.

## Link 601

Ici est exécutée l'analyse de population de la fonction d'onde SCF (orbitales moléculaires, charges de Mulliken . . . ). C'est également dans ce link qu'a été placée par X. Assfeld la localisation des SLBO. Outre les critères de Weinstein-Pauncz et de Magnasco-Perico qu'il avait déjà introduits, nous avons ajouté les critères de Boys-Foster et de Pipek-Mezey. Le choix est gouverné par l'IOP(98) : Weinstein-Pauncz (1), Weinstein-Pauncz + Magnasco-Perico (2), Boys-Foster (3), Pipek-Mezey (4) et Boys-Foster + Pipek-Mezey (5). Pour ces trois nouveaux critères, ce sont uniquement des orbitales localisées qui sont obtenues.

De plus, la polarité des SLBO obtenues par les deux premiers critères peut être modifiée en multipliant les coefficients des orbitales atomiques centrées sur le deuxième atome de la liaison par un facteur fixé par l'IOP(99) \* 0,01.

Bien entendu, plusieurs SLBO peuvent être calculées en même temps sur la même molécule.

## Link 701

Ce link permet de calculer les dérivées premières des termes non-électroniques de l'énergie et des intégrales monoélectroniques. Nous l'avons modifié pour permettre de calculer les dérivées LSCF, particulièrement les contributions dues aux SLBO, et les dérivées des interactions avec le potentiel externe. Il n'y a pas d'option particulière, tout est géré automatiquement. Cependant, il est possible d'imposer la vérification de la nullité des dérivées du recouvrement entre SLBO orthonormées en mettant IOP(99)=1.

**Dérivées LSCF** Il est nécessaire de calculer les contributions des SLBO dans les dérivées premières. Pour cela, les dérivées numériques des SLBO orthonormées sont d'abord calculées. Puis, nous vérifions

que les dérivées de la matrice de recouvrement entre SLBO orthonormées est bien nulle (si ce n'est pas le cas, cela signifie qu'un trièdre de référence pour une SLBO est devenu linéaire). Enfin, nous calculons les contributions des SLBO aux dérivées en utilisant l'équation donnée dans le chapitre consacré à la méthode LSCF.

**Dérivées potentiel externe** Les informations nécessaires, calculées dans le link 397, sont récupérées. Plusieurs contributions sont calculées : les plus simples sont les dérivées dues aux interactions charges QM/MM avec le potentiel externe. Celles qui résultent des interactions électrons-potentiel sont calculées en deux temps : d'abord les dérivées du potentiel électrostatique électronique sur la grille, dont les points sont pondérés par le vecteur  $Z$  puis ce même potentiel électrostatique qu'on « multiplie » par les dérivées du vecteur  $Z$ .

## Autres modifications de G98

- La validation de notre méthode passait par des calculs d'énergie de polarisation et d'interaction électrostatique entre les électrons et le potentiel externe (cf. chapitre QM/MM/Madelung). Pour obtenir ces valeurs, nous avons également modifié le link 302, consacré aux intégrales monoélectroniques et encore le 502.
- L'optimisation de géométrie de la crambine dans le chapitre QM/MM a été effectuée en coordonnées internes grâce à l'algorithme de Németh [8] que nous avons modifié : l'algorithme de minimisation L-BFGS-B [9] a été introduit dans le link 103.

## Annexe B

# Quelques exemples de fichiers de données LSCF

### Ionisation de coeur : 2 calculs

1. Etat fondamental de la molécule CO au niveau MP2 (tous les électrons sont corrélés) en base quintuple- $\zeta$ . L'impression des orbitales moléculaires est demandée.

```
%MEM=20MW
```

```
#P RMP2=FULL/cc-pV5Z SCF=TIGHT IOP(6/7=3) DENSITY=MP2 NOSYM
```

```
CO fundamental
```

```
0 1
```

```
C
```

```
O 1 R
```

```
R 1.1282
```

2. Etat ionisé de coeur (1s du carbone), toujours au même niveau mais en formalisme non-restreint. Il y a une orbitale gelée qui ne doit pas être utilisée dans la partie MP du calcul. Dans le fichier de données, l'orbitale gelée est introduite par le mot-clé FO. Ensuite vient la notation « -10 ». Le chiffre -1 signifie orbitale de spin  $\beta$  et le 0 signifie qu'elle est vide. A la fin les coefficients de la FO terminés par \*\*\*\*.

```
%SUBST L502 /s23/ferre/NG98
%MEM=20MW
#P UMP2=FULL/cc-pV5Z SCF=TIGHT IOP(5/98=101,6/7=3) NOSYM
```

```
CO ionisé
```

```
1 2
```

```
C
```

```
O 1 R
```

```
R 1.1282
```

```
FO
```

```
-10
```

```
.94591
```

```
.00101
```

```
.07817
```

```
.00148
```

```
.00202
```

```
.00030
```

```
.00000
```

```
.00000
```

```
.00191
```

```
...
```

```
.00000
```

```
****
```

A la fin des deux calculs, l'énergie d'ionisation est égal à la différence entre les deux énergies obtenues.

## Calcul QM/MM/Madelung : cas de la chabazite

Le fichier qui suit nous a servi pour optimiser la structure de la chabazite telle que cela a été décrit dans le chapitre 4.

```
%SUBST L101 /s23/ferre/NG98
%SUBST L301 /s23/ferre/NG98
%SUBST L397 /s23/ferre/NG98
%SUBST L402 /s23/ferre/NG98
%SUBST L502 /s23/ferre/NG98
%SUBST L701 /s23/ferre/NG98

%CHK=ChaX

%MEM=250MB

#P NONSTD

1/18=20,30=1,26=3,38=1,57=2,98=1/1,3;
2/15=1,17=6,18=5/2;
4/20=10,22=1,24=1,64=40,67=2/2;
3/5=7,7=1,8=12,11=1,16=1,17=8,25=1,30=1,97=3000,99=1/1,2,3,97;
4/11=1/1;
5/5=2,7=256,32=2,38=4,42=408,99=12/2;
6/7=2,8=2,9=2,28=1/1;
7/30=1/1,2,3,16;
1/18=20/3(1);
99//99;
2/11=1,15=3,17=6,18=5/2;
4/16=2,20=10,22=1,24=1,64=40/2;
3/5=7,6=1,7=1,8=12,11=1,16=1,17=8,25=1,30=1,97=3000,99=1/1,2,3,97;
4/5=5,7=1,16=2/1;
5/5=2,32=2,38=4,42=408,99=12/2;
7/30=1/1,2,3,16;
1/18=20/3(-6);
2/11=1,15=3/2;
6/7=2,8=2,9=2,28=1/1;
99/9=1/99;

Chabazite pure BPW91/6-31G*:UFF:MAD

0 1
Si-Si3-1.      -1    0.77840    3.058770    -1.480920    L
...
```

O-O_3_z--.5	-1	0.232280	3.060560	-0.009580	L
O-O_3_z--.5	0	9.134030	8.438600	4.988010	H
O-O_3_z--.5	0	9.134030	5.592640	8.050620	H
O-O_3_z--.415	0	6.296970	8.663500	8.050620	L
...					
Si-Si3-0.5	0	8.88470	5.594430	9.521960	M
Si-Si3-1.	-1	1.71730	8.321900	14.993010	L
Si-Si3-1.	-1	6.39840	0.680750	17.312469	L
...					
Si-Si3-1.	-1	0.977840	12.454470	7.250110	L
Si-Si3-0.5	0	8.19590	9.726990	1.779060	M 2
Si-Si3-1.	-1	3.126480	17.368151	-0.540400	L
Si-Si3-1.	-1	6.239840	17.315830	-0.589020	L
Si-Si3-1.	0	8.88470	9.906120	4.882030	H

65 1.00 25 1.00 24 1.00 14 1.00

240 1.00 26 1.00 22 1.00 15 1.00

135 1.00 27 1.00 23 1.00 13 1.00

...

360 1.00 283 1.00

351 1.00 186 1.00

O 0

6-31G\*

\*\*\*\*

Si 0

LANL2DZ

\*\*\*\*

Si 0

LANL2DZ

SLBO

11 67 75 250

...

11 417 283 360

...

\*\*\*\*

La première partie de ce fichier établit une route non-standard pour Gaussian. Toutes les options utilisées soit sont documentées dans le manuel de G98, soit ont été décrites ci-dessus.

Après le titre, la charge et la multiplicité (ces deux derniers ne concernent bien évidemment que la partie quantique du calcul), la géométrie moléculaire est donnée (ici en coordonnées cartésiennes, mais elle peut également l'être en coordonnées internes). Comme précisé au début de l'annexe, deux types d'atome doivent être donnés : d'abord quantique (par exemple Si), puis classique (par exemple Si3), suivi éventuellement par la charge classique portée par cet atome. Notons que les charges portées par les atomes quantiques n'ont pas de signification et ne sont pas prises en compte. Les types d'atome sont séparés des coordonnées cartésiennes par 0 ou -1 : les atomes sont libres de bouger ou sont fixes respectivement. C'est une manière pratique de faire une optimisation de géométrie contrainte en coordonnées cartésiennes.

Puis la connectivité est donnée : chaque atome possède une ligne qui contient les numéros des atomes qui lui sont liés et le type de liaison. Un atome non-lié sera symbolisé par une ligne vide.

Puis les bases d'orbitales atomiques sont données, ainsi que les pseudo-potentiels.

Nous introduisons la section dédiée aux 12 orbitales gelées par le mot-clé SLBO. Pour chaque SLBO sont donnés dans l'ordre : le code d'occupation (11 = occupée de type  $\alpha$ ) et le triplet d'atomes de référence, les coefficients des fonctions de base centrées sur l'atome X, les coefficients des fonctions de base centrées sur l'atome Y. La fin de la section est signalée par \*\*\*\*.

A noter : à côté de ce fichier, il faut également préparer un fichier pour MadEwa, tel que décrit précédemment et qui portera obligatoirement le nom « fort.55 ».

## Bibliographie

- [1] M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, V.G. Zakrzewski, J.A. Montgomery, Jr., R.E. Stratmann, J.C. Burant, S. Dapprich, J.M. Millam, A.D. Daniels, K.N. Kudin, M.C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G.A. Petersson, P.Y. Ayala, Q. Cui, K. Morokuma, D.K. Malick, A.D. Rabuck, K. Raghavachari, J.B. Foresman, J. Cioslowski, J.V. Ortiz, A.G. Baboul, B.B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R.L. Martin, D.J. Fox, T. Keith, M.A. Al-Laham, C.Y. Peng, A. Nanayakkara, C. Gonzales, M. Challacombe, P.M.W. Gill,



- B. Johnson, W. Chen, M.W. Wong, J.L. Andres, C. Gonzales, M. Head-Gordon, E.S. Replogle et J.A. Pople. *Gaussian 98, Revision A.9* (1998). Gaussian, Inc., Pittsburgh PA.
- [2] S.L. Mayo, B.D. Olafson et W.A. Goddard III. *J. Phys. Chem.* **94**, 8897 (1990).
- [3] A.K. Rappé, C.J. Casewit, K.S. Colwell, W.A. Goddard III et W.M. Skiff. *J. Am. Chem. Soc.* **114**, 10024 (1992).
- [4] 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. Cladwell et P.A. Kollman. *J. Am. Chem. Soc.* **117**, 5179 (1995).
- [5] A.K. Rappé et W.A. Goddard III. *J. Phys. Chem.* **95**, 3358 (1991).
- [6] J. Ángyán, G. Ferenczy et P. Nortier. *Madewa 1.0* (1996).
- [7] P. Pulay. *J. Comp. Chem.* **3**, 556 (1982).
- [8] K. Németh, O. Coulaud, G. Monard et J. Ángyán. *J. Chem. Phys.* **113**, 5598 (2000).
- [9] C. Zhu, R.H. Byrd, P. Lu et J. Nocedal. *L-BFGS-B: a limited memory FORTRAN code for solving bound constrained optimization problems*. Rapport technique, NAM-11, EECS Department, Northwestern University (1994).



Le pouvoir catalytique des zéolithes est à la base de nombreuses applications chimiques industrielles. Cependant, la compréhension de la réactivité chimique catalysée par ces solides iono-covalents n'est pas encore achevée. Les modèles théoriques doivent prendre en compte certaines caractéristiques particulières de ces systèmes : taille, périodicité, défauts chimiques, sites acides.

Notre travail a consisté à développer un modèle hybride alliant mécanique quantique et mécanique moléculaire (QM/MM), dans lequel le centre réactionnel est traité par la mécanique quantique, l'environnement proche est représenté par un champ de forces classique, le tout ressentant le potentiel de Madelung du reste du solide supposé parfait.

La mise au point du modèle a nécessité plusieurs étapes :

- développement théorique de la méthode Local Self Consistent Field *ab initio* ;
- couplage avec un champ de forces classique, définition d'un potentiel de liaison frontière et ajustement des charges à la frontière QM/MM ;
- mise au point d'une méthode originale de prise en compte d'un potentiel externe dans un calcul QM/MM ;
- paramétrisation d'un champ de forces hybride dédié aux zéolithes et validation sur la chabazite.

Le programme écrit pour cette méthode est un outil qui permet d'aborder les problèmes de réactivité dans les zéolithes de grande maille unitaire et présentant des défauts et des sites de substitution répartis aléatoirement autour du centre réactif.

Cette méthode a été testée sur divers systèmes chimiques (petites molécules, protéines, ...) et nous avons également calculé des potentiels d'ionisation d'électrons de coeur avec une grande précision.

## Mots-clés

CHIMIE QUANTIQUE

QM/MM

SYSTEMES MACROMOLECULAIRES COMPLEXES

ZEOLITHES

AB INITIO

MADELUNG

PROTEINES

IONISATIONS DE COEUR