



## 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>

U.F.R. Sciences et Techniques de la Matière et des Procédés  
Ecole Doctorale Lorraine de Chimie et Physique Moléculaires (SESAMES)  
Département de Formation Doctorale en Chimie Informatique et Théorique

# Thèse

présentée pour l'obtention du titre de

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

en Chimie Informatique et Théorique

par **Yohann MOREAU**

## **Développement d'une méthode théorique pour l'étude de la réactivité chimique des systèmes moléculaires complexes en solution**

Soutenance publique le 13 octobre 2005, salle Jean Barriol

Membres du Jury :

Président :	M. Manuel F. RUIZ-LOPEZ	Directeur de recherche, C.N.R.S., Nancy
Rapporteurs :	Mme. Isabelle DEMACHY-VACUS	Maître de Conférence, Université Paris-Sud
	M. Stéphane HUMBEL	Professeur, Université Aix-Marseille
Examineurs :	M. Xavier ASSFELD	Professeur, Université Henri Poincaré, Nancy-I
	M. Maurizio SIRONI	Professeur, Università degli Studi di Milano



## *Remerciements*

*Je tiens d'abord à remercier Monsieur Ruiz-López, Directeur de Recherche, pour m'avoir accueilli au sein de l'Equipe de Chimie et Biochimie théoriques. Je tiens aussi à remercier le Professeur Jean-Louis Rivail qui m'a permis de d'entrer dans le monde de la chimie théorique.*

*Je voudrais aussi exprimer ma profonde reconnaissance au Professeur Xavier Assfeld qui a accepté de m'encadrer durant mon stage de DEA puis durant les trois années de cette thèse. Je le remercie plus particulièrement pour la patience dont il a su faire preuve.*

*Madame Isabelle Demachy-Vacus, Maître de Conférence et le Professeur Stéphane Humbel ont accepté de rapporter ce travail. Je les remercie pour le temps qu'ils ont consacré à la lecture de ce mémoire et suis très honoré de les compter parmi les membres du jury. Je remercie aussi le Professeur Maurizio Sironi de l'Université de Milan ainsi que Monsieur Manuel F. Ruiz-López, Directeur de recherche au CNRS, d'avoir accepté de faire partie du jury.*

*Au Laboratoire, je ne manquerai pas de remercier Gérald Monard et François Ogliaro pour leur aide et leurs conseils. Que Monsieur Rinaldi soit aussi remercié : son aide en matière d'informatique (matérielle et logicielle) m'a été indispensable plus d'une fois. Je voudrais enfin remercier l'ensemble des personnes du laboratoire pour leur sympathie, avec une mention particulière à l'occupant de l'"autre bureau du fond", Monsieur Cartier, pour sa constante bonne humeur.*

*Un grand merci aux personnes rencontrées au Laboratoire mais avec qui nous n'avons pas fait que de la science. Je pense en particulier à Jean-Marc, Jean-Paul, Laurent, Muhannad, Hassan, Jérôme et à bien d'autres encore.*

*Stéphanie, sache seulement que durant les trois ans qu'ont duré cette thèse, j'ai pu trouver auprès de toi une présence qui me reste plus que précieuse et qui compte aussi pour l'aboutissement de ce travail.*

*Pour terminer, je voudrais dire un grand merci à mes Parents. En acceptant et en soutenant mes choix tout au long de ma scolarité puis de mes études, vous avez su trouver la meilleure façon de m'aider.*



# Table des matières

<b>Table des figures</b>	<b>9</b>
<b>Liste des tableaux</b>	<b>11</b>
<b>Introduction</b>	<b>15</b>
<b>1 Outils de modélisation</b>	<b>17</b>
1.1 Méthodes usuelles de la chimie quantique . . . . .	18
1.1.1 Principes de base . . . . .	18
1.1.2 La méthode Hartree-Fock . . . . .	20
1.1.3 Les équations de Roothaan . . . . .	21
1.1.4 Méthodes post Hartree-Fock . . . . .	23
1.1.5 Les méthodes de la fonctionnelle de la densité . . . . .	25
1.2 La mécanique moléculaire . . . . .	26
1.3 Les méthodes hybrides . . . . .	28
1.3.1 Approches à interaction directe . . . . .	29
1.3.2 L'approche ONIOM . . . . .	32
1.4 La méthode LSCF/MM <i>ab initio</i> . . . . .	33
1.4.1 LSCF . . . . .	34
1.4.2 Le champ de forces LSCF/MM . . . . .	37
1.5 Détermination du Hessien par blocs . . . . .	41
1.5.1 Principe général . . . . .	42
1.5.2 Procédure de calcul . . . . .	43
1.5.3 Caractérisation d'un état de transition . . . . .	44
1.5.4 Applications au dimère QM/MM d'eau . . . . .	47

<b>2</b>	<b>LSCF/MM/SCRF : définition et application</b>	<b>51</b>
2.1	Le modèle implicite de solvant . . . . .	52
2.1.1	Généralités . . . . .	52
2.1.2	Le modèle à cavité de Nancy . . . . .	53
2.2	Définition de la méthode LSCF/MM/SCRF . . . . .	55
2.3	Effets du solvant dans la réaction de Diels-Alder . . . . .	56
2.3.1	Le système réactif . . . . .	57
2.3.2	Rappels des résultats existants . . . . .	58
2.3.3	Détails calculatoires . . . . .	60
2.3.4	Effets du solvant sur la sélectivité . . . . .	63
2.3.5	Influence de l'encombrement stérique . . . . .	66
2.3.6	Modification du mécanisme réactionnel . . . . .	70
2.4	Conclusion . . . . .	75
<b>3</b>	<b>Traitement explicite du solvant</b>	<b>77</b>
3.1	Les méthodes hybrides de type soluté-solvant . . . . .	79
3.2	Description de la liaison hydrogène en QM/MM . . . . .	80
3.2.1	Protocole . . . . .	80
3.2.2	Résultats, discussion . . . . .	84
3.2.3	Conclusions de l'étude . . . . .	94
3.3	Définition d'une approche soluté-solvant adaptée . . . . .	95
3.3.1	Schéma général . . . . .	96
3.3.2	Détermination des paramètres adaptés . . . . .	97
3.3.3	Test des paramètres . . . . .	101
3.3.4	Application au dimère de l'eau . . . . .	105
3.4	Effets du solvant sur le réarrangement de Claisen . . . . .	108
3.4.1	Détails calculatoires . . . . .	110
3.4.2	Résultats, discussion . . . . .	110
3.4.3	Conclusion sur l'étude du réarrangement de Claisen . . . . .	116
3.5	Conclusion . . . . .	117
	<b>Conclusion générale</b>	<b>119</b>
	<b>Bibliographie</b>	<b>121</b>

<b>A</b>	<b>Modifications des logiciels</b>	<b>129</b>
A.1	Le programme Gaussian . . . . .	130
A.2	Méthode LSCF/MM/SCRF . . . . .	131
A.3	Implantation de LSCF/MM dans Gaussian03 . . . . .	134
A.4	Insertion de molécules d'eau discrètes . . . . .	136
A.5	Calcul de la matrice hessienne par blocs . . . . .	139
<b>B</b>	<b>Données associées au premier chapitre</b>	<b>141</b>
<b>C</b>	<b>Données associées au deuxième chapitre</b>	<b>147</b>
<b>D</b>	<b>Données associées au troisième chapitre</b>	<b>155</b>



# Table des figures

1.1	Notation des atomes au niveau de la frontière . . . . .	39
1.2	répartition des différents blocs de la matrice hessienne en fonction de la méthode utilisée pour traiter les atomes correspondants . . . . .	43
1.3	les deux procédures proposées pour le calcul par bloc du hessien . . . . .	45
1.4	Dimère, partition et numérotation des atomes utilisés . . . . .	47
1.5	différences de valeurs pour les éléments du hessien du dimère d'eau pour les deux approches, la numérotation des éléments correspond à celle de la figure 1.4. Les zones grisées sont calculées numériquement . . . . .	47
2.1	Modèle de continuum . . . . .	52
2.2	réaction de l'acrylate de (-)-menthyle avec le cyclopentadiène et produits formés . . . . .	57
2.3	Partition de l'acrylate de menthyle en sous-systèmes classique et quantique	61
2.4	Conformations du groupement isopropyle . . . . .	67
2.5	Liaisons prises en compte pour la détermination de l'asynchronicité : CC2 est la liaison la plus proche de l'auxiliaire chiral . . . . .	71
3.1	Structures utilisées dans l'étude. Les traits bleus signalent les liaisons hydrogène prises en compte. Les structures sont désignées par la nature du groupement en interaction avec l'eau. . . . .	81
3.2	Valeurs géométriques et énergétiques des calculs menés avec Amber comparées aux références quantiques . . . . .	85
3.3	Différence de disposition de la molécule d'eau par rapport au soluté entre calculs quantiques et classiques . . . . .	86
3.4	Valeurs géométriques et énergétiques des calculs menés avec une SLBO, comparés aux références quantiques . . . . .	87

3.5	Valeurs géométriques et énergétiques des calculs menés avec LSCF/MM comparées aux références quantiques. La frontière se situe au sein du soluté . . . . .	90
3.6	Structure pontée de l'éthanolate dans l'approche QM/MM : la partie où seules les liaisons sont représentées est traitée au niveau classique . . .	91
3.7	Valeurs géométriques et énergétiques des calculs menés avec LSCF/MM comparées aux références quantiques. Le soluté est décrit par la mécanique quantique et l'eau avec le modèle TIP3P . . . . .	93
3.8	Complexe de l'éthanol avec deux molécules d'eau . . . . .	103
3.9	Energies d'interaction pour l'éthanol quantique en interaction avec une (3.9(a)) ou deux (3.9(b)) molécules d'eau classiques. Le référence est MP2/6-311++G(2d,2p), NP se rapporte aux "Nouveaux Paramètres" .	104
3.10	Les deux formes du dimère de l'eau en QM/MM : la molécule quantique est celle qui est représentée en "boules et bâtons" . . . . .	106
3.11	réarrangement de Claisen pour l'allyle vinyle éther . . . . .	108
C.1	Numérotation des atomes du diénophile . . . . .	148

# Liste des tableaux

1.1	liste des interactions QM/MM non électrostatiques. <sup>(a)</sup> : terme de liaison spécifique au champ de forces LSCF/MM. . . . .	40
1.2	Fréquences des différents modes de vibration suivant la méthode utilisée	48
2.1	Les différents conformations conduisant aux seize structures de transition	59
2.2	Energies des orbitales frontières du diénophile en fonction du jeu de charges classiques utilisé (tableau C.1, annexe C) . . . . .	62
2.3	énergies relatives en kcal.mol <sup>-1</sup> des états de transition pour la phase gazeuse, pour le solvant avec uniquement les interactions électrostatique et comprenant les termes de type CDS. la valeur $\Delta$ est la différence entre cette dernière valeur et celle correspondante dans le vide. . . . .	65
2.4	différence d'énergie (kcal.mol <sup>-1</sup> ) pour les structures de transition se différenciant par l'approche de la face prochirale (encombrée - libre) .	67
2.5	différence d'énergie (kcal.mol <sup>-1</sup> ) pour les structures de transition se différenciant par l'attaque (endo-exo) . . . . .	68
2.6	différence d'énergie (kcal.mol <sup>-1</sup> ) pour les structures de transition se différenciant par la conformation du diénophile (anti-syn) . . . . .	69
2.7	Asynchronicités calculées à partir des longueurs et ordres des liaisons CC1 et CC2 et effet du solvant sur ces grandeurs . . . . .	72
2.8	Différence des longueurs des liaisons (Å) CC1 et CC2 avec les résultats obtenus en considérant l'acrylate de méthyle . . . . .	73
2.9	moments dipolaires en Debye dans le vide et le solvant des différents isomères de l'acrylate de (-)-menthyle et différences associées . . . . .	74

3.1	Somme des charges de Mulliken des atomes dans le groupement méthyle des différentes structures et charge totale de la molécule d'eau. <sup>a</sup> : les solutés seuls sont les mêmes que les précédents (OH (d) et SH(d)), seul l'arrangement avec la molécule d'eau change. . . . .	88
3.2	Somme des charges de Mulliken des atomes dans le groupement méthyle des différentes structures dans les approches QM et QM/MM. . . . .	94
3.3	Table de vérité des interactions différenciées . . . . .	97
3.4	Paramètres des différents modèles d'eau. $\sigma$ en donné en Å et $\varepsilon$ en kcal.mol <sup>-1</sup> . . . . .	101
3.5	Energies d'interactions et paramètres géométriques du dimère de l'eau traité en QM/MM. Laaksonen correspond aux valeurs obtenues par cet auteur avec le jeu de paramètres déjà donné (tableau 3.4) et N.P. est pour nos paramètres. . . . .	107
3.6	Energies d'interaction en kcal.mol <sup>-1</sup> dans le vide entre le réactif et une molécule d'eau suivant l'approche quantique ou hybride. La différence d'énergie calculée entre les deux approches est aussi donnée. . . . .	111
3.7	Longueurs des liaisons hydrogène en Å formées suivant l'approche (quantique ou hybride), dans le vide et avec le continuum. . . . .	111
3.8	Energies d'activation et de réaction pour les différentes approches : dans le vide, eau traitée au niveau quantique ou au niveau classique. les valeurs analogues sont aussi reportées pour les calculs conduits dans le continuum . . . . .	112
3.9	longueurs des liaisons (Å) modifiées au cours de la réaction de Claisen pour les différentes approches (sans molécule d'eau, avec une molécule quantique ou avec une classique) sans le continuum . . . . .	114
3.10	longueurs des liaisons (Å) modifiées au cours de la réaction de Claisen pour les différentes approches (sans molécule d'eau, avec une molécule quantique ou avec une classique) avec influence du continuum . . . . .	114
3.11	Charges du fragment ne contenant pas l'atome d'oxygène dans les trois état de transition pris en compte et calculées à partir des charges atomiques de Mulliken. . . . .	116
A.1	numéro et rôles des overlays . . . . .	130

B.1	Fréquences de vibration en $\text{cm}^{-1}$ calculées avec une procédure numérique complète pour l'état de transition dans la réaction de Diels-Alder . . .	142
B.2	Fréquences de vibration en $\text{cm}^{-1}$ calculées sans prise en compte des interactions électrostatiques entre les sous parties quantiques et classiques pour l'état de transition dans la réaction de Diels-Alder . . . . .	143
B.3	Fréquences de vibration en $\text{cm}^{-1}$ calculées en prenant en compte les interactions électrostatiques entre les sous parties quantiques et classiques pour l'état de transition dans la réaction de Diels-Alder . . . . .	144
B.4	Expressions du vecteur de transition suivant la méthode de calcul utilisée dans l'espace des degrés de liberté cartésiens des atomes. Seuls les déplacements liés aux atomes quantiques sont reportés, les valeurs liés aux atomes classiques étant systématiquement nulles. . . . .	145
C.1	charges de Mulliken des atomes de la partie quantique du diènophile, suivant le jeu de charges classiques portées par la partie classique de la même molécule, la numérotation des atomes se fait comme indiqué par la figure C.1 . . . . .	148
C.2	Energies absolues (u.a.) des réactifs dans le vide et dans le solvant . . .	149
C.3	Energies absolues des TS et produits et énergies d'activation et de réaction associées dans le vide . . . . .	149
C.4	Energies absolues des TS et produits et énergies d'activation et de réaction associées dans le solvant . . . . .	150
C.5	Longueurs des liaisons CC1 et CC2 dans les TS et produits : vide et le solvant . . . . .	151
C.6	Ordres des liaisons de Mayer CC1 et CC2 dans les TS et produits : vide et le solvant . . . . .	152
C.7	Charges de Mulliken sur les différents isomères du diènophile dans le vide et en solvant et variation des charges du au passage du vide au solvant. La numérotation est celle de la figure C.1 . . . . .	153
D.1	Jeux de charges calculés pour les différents solutés envisagés à l'aide de la méthode RESP au niveau MP2/6-31G* . . . . .	156
D.2	Energies d'interaction soluté/eau suivant les approches considérées en $\text{kcal.mol}^{-1}$ . . . . .	157
D.3	Longueurs des liaisons hydrogène considérées en Ångstroms . . . . .	157

- D.4 Energies d'interaction soluté/eau ( $\text{kcal.mol}^{-1}$ ) en utilisant les paramètres adaptés, et références quantiques auxquelles elles sont comparées . . . 158
- D.5 Longueurs des liaisons hydrogène considérées en Ångstroms en utilisant les paramètres adaptés, et références quantiques auxquelles elles sont comparées . . . . . 158
- D.6 Energies d'interaction entre une molécule quantique d'éthanol et une molécule d'eau traitée au niveau classique suivant le jeu de paramètres adopté, la base d'orbitales et la méthode. suivant le jeu de paramètres adopté. La référence quantique MP2/6-311++G(2d,2p) est à  $-5,319 \text{ kcal.mol}^{-1}$  159
- D.7 Energies d'interaction ( $\text{kcal.mol}^{-1}$ ) entre une molécule quantique d'éthanol et deux molécules classiques d'eau suivant le jeu de paramètres adopté, la base d'orbitales et la méthode. La référence quantique MP2/6-311++G(2d,2p) est à  $-14,420 \text{ kcal.mol}^{-1}$  . . . . . 159

# Introduction

La compréhension des processus chimiques ou biochimiques passe aujourd'hui souvent par l'étude théorique des différentes réactions chimiques qu'ils mettent en jeu. L'environnement du centre réactif peut avoir une influence déterminante sur le chemin réactionnel suivi. Nous citerons ainsi l'accélération de certaines réactions chimiques lorsqu'on augmente la polarité du solvant, comme dans le cas de la cycloaddition de Diels-Alder. On peut aussi mentionner le contrôle que peut avoir le chimiste sur la stéréosélectivité d'une réaction en ajoutant un auxiliaire chiral à l'un des réactifs.

Par ces deux exemples, on voit que l'environnement chimique peut être séparé en deux groupes. D'une part on trouve le solvant et d'autre part les parties des réactifs qui n'entrent pas directement dans le processus réactionnel. La prise compte des effets de ces deux facteurs revêt donc une importance cruciale dans la modélisation de la réactivité chimique.

Depuis plus de trente ans maintenant, de nombreux efforts ont été faits pour permettre la prise en compte de l'environnement, que ce soit le solvant ou les parties d'une des espèces réactives qui n'interviennent pas directement dans la réaction mais dont l'influence est importante. Ainsi, au laboratoire de Nancy, a été développée la méthode nommée SCRF (*Self Consistent Reaction Field*) qui traite le solvant comme un milieu continu polarisable. L'approche LSCF/MM (*Local Self Consistent Field/Molecular Mechanics*), elle aussi développée à Nancy, permet de traiter les systèmes moléculaires complexes (c'est à dire constitué d'un grand nombre d'atomes) en les séparant en deux parties : l'une est traitée par la mécanique quantique et l'autre par la mécanique moléculaire (QM/MM). Ces deux approches ont été développées séparément dans le but de prendre en compte à chaque fois un seul type d'environnement.

L'objectif principal de ce travail a été le développement d'outils permettant la prise en compte simultanée de l'influence du solvant et de l'environnement moléculaire autour du centre réactif. Cela a donc d'abord nécessité le développement d'une méthode

couplant les approches QM/MM et implicite citées plus haut. Nous nous sommes ensuite intéressés à l'introduction de quelques molécules de solvant explicites à proximité du site de réaction afin de modéliser les interactions soluté/solvant spécifiques, en particulier, la liaison hydrogène, mal décrite par le modèle de continuum.

Dans le premier chapitre de ce mémoire, nous avons choisi de rappeler les aspects méthodologiques et théoriques auxquels nos travaux font appel. Nous y avons aussi inclus une méthode de calcul de la matrice hessienne lorsque le système est étudié à l'aide de la méthode LSCF/MM afin de pouvoir optimiser les géométries des états de transition dans un temps de calcul raisonnable.

Dans un deuxième temps, nous développons les caractéristiques de l'approche à trois niveaux LSCF/MM/SCRF. Cette dernière a pour particularité de coupler les caractéristiques de la méthode hybride LSCF/MM avec les possibilités du modèle implicite de solvant développé à Nancy. Cette nouvelle approche a été mise à profit pour permettre l'étude des effets couplés du solvant et de l'auxiliaire chiral dans la réaction de cyclo-addition de Diels-Alder entre le cyclopentadiène et l'acrylate de (-)-menthyle.

Nous nous sommes ensuite intéressés, dans le troisième chapitre, à la possibilité d'ajouter une ou quelques molécules explicites de solvant (ici l'eau) au système réactif. Cette possibilité nous a conduit à la détermination de paramètres spécifiques que nous testons et utilisons dans diverses applications.

Enfin, après avoir tiré les conclusions de nos travaux, les modifications apportées dans les programmes Gaussian98 [1] et Gaussian03 [2] permettant la mise en œuvre des méthodes sont récapitulées. Ces propos seront en outre complétés par quelques exemples de fichiers d'entrée spécifiques.

# Chapitre 1

## Outils de modélisation

Ce chapitre, divisé en deux parties distinctes est destiné à présenter les méthodes sur lesquelles s'appuient nos travaux. Nous avons d'abord voulu rappeler les principales méthodes de calcul de l'énergie électronique d'une molécule. Pour cela nous insistons plus particulièrement sur l'approche Hartree-Fock et la résolution des équations de Roothaan. Nous décrivons ensuite plus succinctement les méthodes permettant de prendre en compte la corrélation électronique, en particulier les méthodes multi-déterminantales ainsi que perturbatives. Nous rappelons enfin les principes généraux de la théorie de la fonctionnelle de la densité. Dans un deuxième temps, nous avons choisi d'évoquer les méthodes permettant l'étude des systèmes dont la taille rend prohibitive l'utilisation de calculs complètement quantiques. Pour cela, outre un rappel sur les méthodes de mécanique moléculaire, nous parlerons des méthodes hybrides, alliant mécanique classique et quantique (*QM/MM*) avec une description plus particulière de l'approche LSCF/MM *ab initio* sur laquelle nos travaux ont porté. Nous présentons enfin une approche permettant la détermination de la matrice hessienne d'un système décrit par une approche QM/MM. Cette détermination s'effectue par blocs en combinant des calculs analytiques et numériques.

## 1.1 Méthodes usuelles de la chimie quantique

Nous évoquons ici les méthodes quantiques utilisées lors de notre travail. Outre les références citées, les développements complets des équations présentées ici pourront être trouvés aux références [3] et [4].

### 1.1.1 Principes de base

La mécanique quantique stipule qu'un système peut être complètement décrit par sa fonction d'onde multiparticulaire  $\Psi_0$ , solution de l'équation de Schrödinger. Si on se place dans le cas non relativiste et qu'on considère que les propriétés étudiées ne dépendent pas du temps, celle-ci peut s'écrire :

$$\hat{H}_0 \Psi_0 = E_0 \Psi_0 \quad (1.1)$$

Où  $\hat{H}_0$  représente l'opérateur Hamiltonien du système, décrivant à la fois les termes cinétiques et potentiels du système étudié.  $\Psi_0(r_1 \dots r_n)$  (avec  $n$  le nombre de particules du système et  $r_i$  la coordonnée de la particule  $i$ ) est une fonction d'onde décrivant le système d'énergie  $E_0$ .  $\Psi_0(r_1 \dots r_n)$  est un vecteur propre de  $H_0$  associé à la valeur propre  $E_0$ . L'opérateur Hamiltonien  $\hat{H}_0$  peut être explicité de la façon suivante (en unités atomiques) :

$$\hat{H}_0 = \sum_{a=1}^{N_a} -\frac{1}{2M_a} \Delta_a + \sum_{i=1}^n -\frac{1}{2} \Delta_i - \sum_{a=1}^{N_a-1} \sum_{b>a}^{N_a} \frac{Z_a Z_b}{r_{ab}} - \sum_{a=1}^{N_a} \sum_{i=1}^n \frac{Z_a}{r_{ai}} + \sum_{i=1}^{n-1} \sum_{j>i}^n \frac{1}{r_{ij}} \quad (1.2)$$

Où les deux premiers termes correspondent aux opérateurs énergies cinétiques associées respectivement aux noyaux et aux électrons. Le terme suivant est associé à l'interaction entre noyaux et les deux derniers termes sont, dans l'ordre, les interactions noyaux-électrons et entre électrons.

La masse d'un électron étant près de deux mille fois inférieure à celle d'un proton, les mouvements respectifs des noyaux et des électrons peuvent être découplés. On fait alors l'approximation de Born et Oppenheimer [5]. On peut donc se ramener à l'étude du comportement des électrons dans le potentiel créé par les noyaux ayant une géométrie donnée. Ce dernier sera décrit par une fonction d'onde électronique, état propre de l'Hamiltonien électronique dans lequel n'intervient plus le premier terme de l'équation 1.2 relatif à l'énergie cinétique des noyaux. Il est à remarquer que le terme de répulsion

nucléaire, indépendant des coordonnées des électrons, n'agit que comme correction à l'énergie électronique du système.

La mécanique quantique n'admettant pas de solution analytique exacte pour l'équation de Schrödinger pour un nombre de particules supérieur à deux (appelé problème à  $n$  corps), on passe alors par une description monoélectronique du système. On décrit ainsi le comportement de chaque électron dans le champ créé par le reste du système. Dans l'approximation orbitale, chaque électron est décrit par une fonction monoélectronique, la spinorbitale, dépendant de la position et du spin de l'électron considéré. Celle-ci peut alors s'écrire pour l'électron  $i$  :  $\phi_i(r(i), s(i))$ .

De même, l'expression des spinorbitales peut être approchée si on considère que le mouvement des électrons est indépendant de la partie de spin de celui-ci, en d'autres termes, si on néglige le couplage entre le moment cinétique orbitaire et moment de spin.

La spinorbitale peut alors s'écrire comme le produit d'une fonction d'espace (orbitale, notée  $\varphi(r(i))$ , ne dépendant que de la position de l'électron) par une fonction de spin  $\sigma(i)$ , égale à  $\alpha(i)$  pour un spin de  $\frac{1}{2}$  et  $\beta(i)$  si celui ci vaut  $-\frac{1}{2}$ .

On a donc :  $\phi_i(r(i), s(i)) = \varphi(r(i)).\sigma(s(i))$

Les électrons sont des particules de spin demi-entier qu'on nomme fermions, qui sont donc soumis au principe d'exclusion de Pauli [6]. Ce dernier stipule que l'échange de deux électrons fait changer le signe de la fonction d'onde totale. Afin de respecter ce principe et en tenant compte des approximations énoncées précédemment, on exprime habituellement la fonction d'onde électronique comme un produit antisymétrisé, sous forme d'un déterminant de Slater [7] :

$$\Psi_0 = \frac{1}{\sqrt{n!}} \begin{vmatrix} \phi_1(1) & \phi_2(1) & \cdots & \phi_n(1) \\ \phi_1(2) & \phi_2(2) & \cdots & \phi_n(2) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_1(n) & \phi_2(n) & \cdots & \phi_n(n) \end{vmatrix} \quad (1.3)$$

Qu'on note de façon simplifiée :  $\Psi_0 = |\phi_1\phi_2\cdots\phi_n|$

### 1.1.2 La méthode Hartree-Fock

On veut maintenant déterminer un jeu de spinorbitales afin de construire le déterminant de Slater décrivant au mieux le système étudié. L'Hamiltonien électronique total s'écrit sous la forme :

$$\hat{H}_0 = \sum_{i=1}^n \left( -\frac{1}{2} \Delta_i + V_{Ni} \right) + \sum_{i=1}^{n-1} \sum_{j>i}^n \frac{1}{r_{ij}} \quad (1.4)$$

Où le premier terme, rassemblant les termes d'énergie cinétique des électrons et du potentiel dû au noyaux (noté  $V_{Ni}$ ), est un opérateur monoélectronique appelé Hamiltonien de cœur et noté  $\hat{H}^c(i)$ , tandis que le reste, correspondant à la répulsion électronique est un opérateur bi-électronique.

Dans la suite, on se place dans le cas où les spinorbitales sont mutuellement orthogonales, condition qu'on note :

$$\langle \phi_i | \phi_j \rangle = \delta_{ij} \quad (1.5)$$

Où  $\delta_{ij}$  est l'indice de Kronecker tel que  $\delta_{ij} = 1$  si  $i = j$  et zéro s'ils sont différents.

L'énergie associée au déterminant de Slater 1.3 peut donc s'écrire en notation de Dirac :

$$\begin{aligned} E_0 &= \sum_i^{occ} \langle \phi_i(1) | H^c(1) | \phi_i(1) \rangle \\ &+ \sum_{i=1}^{occ} \sum_{j>i} \left[ \langle \phi_i(1) \phi_j(2) | \frac{1}{r_{12}} | \phi_i(1) \phi_j(2) \rangle - \langle \phi_i(1) \phi_j(2) | \frac{1}{r_{12}} | \phi_j(1) \phi_i(2) \rangle \right] \end{aligned} \quad (1.6)$$

Le principe variationnel stipule que la meilleure fonction d'onde doit correspondre à un extremum (minimum) de l'énergie en fonction des spinorbitales. C'est à dire qu'une variation infinitésimale d'une spinorbitale ne doit pas changer l'énergie du système. Suivant cette considération, on écrit l'équation de Hartree-Fock qui permet la détermination des spinorbitales [8, 9] :

$$\hat{F} \phi_i = \varepsilon_i \phi_i \quad (1.7)$$

$\hat{F}$  étant un opérateur monoélectronique, appelé opérateur de Fock qui se décompose comme suit :

$$\hat{F}(1) = H^c(1) + \sum_i [\hat{J}_i(1) - \hat{K}_i(1)] \quad (1.8)$$

Avec  $H^c(1)$  l'Hamiltonien de cœur,

$\hat{J}_i(1)$  l'opérateur coulombien défini par :

$$\hat{J}_i(1) = \int_{(2)} \varphi_i^*(2) \frac{1}{r_{12}} \varphi_i(2) d\tau_2 \quad (1.9)$$

et  $\hat{K}_i(1)$  celui d'échange qui se définit par son action sur une fonction  $\varphi_j(1)$  :

$$\hat{K}_i(1)\varphi_j(1) = \varphi_i(1) \int_{(2)} \varphi_i^*(2) \frac{1}{r_{12}} \varphi_j(2) d\tau_2 \quad (1.10)$$

L'intégration se faisant sur les coordonnées de l'électron 2.

La méthode de Hartree-Fock présente l'avantage de ne faire intervenir que des opérateurs monoélectroniques. Cependant, les opérateurs  $\hat{J}_i(1)$  et  $\hat{K}_i(1)$  requièrent la connaissance préalable des spinorbitales que cette méthode est sensée fournir. La résolution d'un tel système d'équations doit donc passer par une méthode itérative dite du champ auto-cohérent désignée par l'acronyme anglais SCF (*Self Consistent Field*).

### 1.1.3 Les équations de Roothaan

Dans le cas d'une molécule, l'approximation LCAO (*Linear Combination of Atomic Orbitals*) permet aux spinorbitales (on parle alors de spinorbitales moléculaires) déterminées par la méthode de Hartree-Fock d'être développées sur une base d'orbitales atomiques sous forme de combinaison linéaire. Lorsque le système étudié est à couche fermée, c'est à dire que toute fonction d'espace l'est par deux électrons du fait de l'orthogonalité de leur fonction de spin (ils ont des spins opposés), on utilise l'expression due à Roothaan [10]. On se positionne alors dans le formalisme restreint ou RHF (*Restricted Hartree-Fock*).

Un jeu de  $N$  orbitales atomiques forme alors la base  $\{\chi_1 \cdots \chi_\mu \cdots \chi_\nu \cdots \chi_N\}$  sur laquelle on peut construire les orbitales moléculaires telles que :

$$\varphi_i = \sum_{\mu=1}^N C_{\mu i} \chi_\mu \quad (1.11)$$

On peut alors exprimer l'énergie du système étudié et ainsi définir l'expression d'un élément de l'opérateur de Fock dans la base des orbitales atomiques (OA) :

$$F_{\mu\nu} = H_{\mu\nu}^c + \sum_{\lambda} \sum_{\sigma} P_{\lambda\sigma} [\langle \mu\nu | \lambda\sigma \rangle - \frac{1}{2} \langle \mu\nu | \sigma\lambda \rangle] \quad (1.12)$$

avec  $H_{\mu\nu}^c = \langle \chi_\mu | \hat{H}^c | \chi_\nu \rangle$ , élément de l'Hamiltonien de cœur exprimé dans la base des orbitales atomiques,

l'intégrale bi-électronique :  $\langle \mu\nu | \lambda\sigma \rangle = \langle \chi_\mu(1)\chi_\nu(2) | \frac{1}{r_{12}} | \chi_\lambda(1)\chi_\sigma(2) \rangle$

et  $P_{\lambda\sigma}$ , l'élément de la matrice densité tel que :  $P_{\lambda\sigma} = 2 \sum_i^{occ} C_{\lambda i} C_{\sigma i}$

Dans sa forme matricielle, cette expression s'écrit :

$$\mathbf{F} = \mathbf{H} + \mathbf{G} \quad (1.13)$$

Avec  $\mathbf{F}$  la matrice de Fock,  $\mathbf{H}$  la matrice Hamiltonien de cœur et  $\mathbf{G}$  la matrice obtenue par contraction de la matrice densité avec les intégrales bi-électroniques.

L'introduction de multiplicateurs de Lagrange  $\varepsilon_i$  conduit à la dérivation des équations de Roothaan dont le détail peut être trouvé dans les ouvrages cités aux références [4] et [3]. Ces dernières s'écrivent :

$$\sum_{\mu} F_{\nu\mu} C_{\mu i} = \sum_{\mu} S_{\nu\mu} C_{\mu i} \varepsilon_i \quad (1.14)$$

avec l'introduction de l'élément  $S_{\nu\mu}$  tel que  $S_{\nu\mu} = \langle \chi_\nu | \chi_\mu \rangle$

Et dont la notation sous forme matricielle est (avec  $\mathbf{S}$  la matrice de recouvrement) :

$$\mathbf{FC} = \mathbf{SCE} \quad (1.15)$$

Où  $\mathbf{E}$  est la matrice diagonale des valeurs propres  $\varepsilon_i$ . La non orthogonalité de la base d'orbitales atomiques dans laquelle on travaille rend difficile la résolution de ce système d'équations. Afin de se ramener à un problème aux valeurs propres, on utilise une transformation unitaire par le biais de la matrice  $\mathbf{X}$  telle que  $\mathbf{X}^\dagger \mathbf{S} \mathbf{X} = \mathbf{1}$ . On utilise fréquemment l'orthogonalisation de Löwdin avec  $\mathbf{X} = \mathbf{S}^{-\frac{1}{2}}$ . L'équation 1.15 s'écrit alors :

$$\mathbf{F}'\mathbf{C}' = \mathbf{C}'\mathbf{E} \quad (1.16)$$

avec  $\mathbf{F}' = \mathbf{X}^\dagger \mathbf{F} \mathbf{X}$

et  $\mathbf{C}'$  telle que  $\mathbf{C} = \mathbf{X} \mathbf{C}'$

Comme nous l'avons dit précédemment, la connaissance préalable des orbitales moléculaires rend le problème autocohérent et nécessite donc une procédure itérative. Pour une géométrie donnée, on part avec une matrice  $\mathbf{P}$  de départ, puis on construit la matrice de Fock  $\mathbf{F} = \mathbf{H} + \mathbf{G}$ . Après passage dans la base orthogonale, on résout l'équation aux valeurs propres 1.16 dont on tire une nouvelle matrice  $\mathbf{C}'$  qui, après

retour dans la base des orbitales atomiques, donnera une nouvelle matrice densité  $\mathbf{P}$ . On calcule aussi l'énergie totale du système telle que  $E = \frac{1}{2}tr(\mathbf{P}(\mathbf{H} + \mathbf{F}))$ .

Habituellement, on considère que la convergence du processus est atteinte lorsque la variation de l'énergie calculée ainsi que la différence entre les matrices densité de deux itérations consécutives sont négligeables. La fonction d'onde ainsi déterminée répond au principe variationnel et on dit que le système a atteint sa cohérence interne.

Concernant les systèmes à couche ouverte, le formalisme appelé UHF (*Unrestricted Hartree-Fock*) développé à la fois par Pople-Nesbet [11] et Berthier [12] sépare en deux sous-ensembles les électrons de spin différent, permettant ainsi de d'écrire les états dans lesquels le nombre des électrons de spin  $\alpha$  et  $\beta$  est différent. La résolution des équations dans ce cas fait intervenir deux systèmes d'équations distincts (pour les électrons  $\alpha$  et  $\beta$ ) non indépendants et qui doivent être résolus simultanément. Le détail pourra de nouveau en être trouvé aux références [4] et [3].

En conclusion, les méthodes autocohérentes de détermination de la fonction d'onde peuvent être considérées comme la pierre angulaire du calcul quantique en chimie. En effet, elles servent souvent de point de départ à des méthodes plus poussées qu'on appelle post Hartree-Fock. Ces méthodes sont en effet développées afin de palier le défaut principal de l'approche Hartree-Fock qui est l'absence de description de la corrélation entre électrons. Ce défaut est dû au fait que l'interaction entre électrons n'est pas traitée explicitement mais plutôt de façon moyenne (on parle d'"Hamiltonien de champ moyen"). Ces méthodes, pouvant faire intervenir un ou plusieurs déterminants de Slater, permettent de fournir une partie de l'énergie de corrélation du système.

#### 1.1.4 Méthodes post Hartree-Fock

Comme dit précédemment, ces méthodes ont pour but de calculer une partie de la corrélation électronique qu'une approche Hartree-Fock ne donne pas. Nous choisissons ici de détailler deux approches pour leurs caractéristiques distinctes : d'abord une méthode de type perturbative qui s'appuie sur une approche monodéterminante et l'interaction de configurations qui prend en compte plusieurs déterminants de Slater. C'est cette dernière que nous développons en premier.

##### Interaction de configurations

Dans cette approche, la fonction d'onde multiélectronique est décrite par la combinaison linéaire de plusieurs déterminants orthogonaux. Outre le déterminant  $\Psi_0$  obtenu

par la méthode Hartree-Fock, on prend aussi en compte ceux construits en remplaçant dans ce dernier une ou plusieurs spinorbitales occupées  $(\phi_a, \phi_b, \dots)$  par autant de spinorbitales virtuelles  $(\phi_r, \phi_s, \dots)$ . La fonction d'onde ainsi calculée s'écrit :

$$\Phi = C_0 \Psi_0 + \sum_a \sum_r^{\text{occ. virt.}} C_a^r \Psi_a^r + \sum_{a,b} \sum_{r,s}^{\text{occ. virt.}} C_{a,b}^{r,s} \Psi_{a,b}^{r,s} + \dots \quad (1.17)$$

Où les indices des déterminants et des coefficients associés correspondent aux orbitales virtuelles qui remplacent les orbitales occupées dans le déterminant décrivant l'état fondamental. Les coefficients de l'équation 1.17 sont alors optimisés en diagonalisant l'Hamiltonien électronique 1.4 dans la base des configurations. Cette méthode permet d'approcher de façon variationnelle les fonctions d'ondes exactes pour les différents états, excités ou fondamental. Dans le cas d'une base complète (de taille infinie) et de la prise en compte de toutes les excitations (on parle de *Full CI*), la solution obtenue est exacte.

Cette méthode souffre principalement de son coût calculatoire qui ne la rend utilisable que pour des systèmes de petite taille. On se contente d'ailleurs souvent de tronquer l'espace des configurations en ne faisant intervenir que les doubles (CID) ou simples et doubles excitations (CISD). De même, on peut envisager l'optimisation simultanée des coefficients de l'équation 1.17 et des coefficients des orbitales atomiques dans les orbitales moléculaires. Ces méthodes multiconfigurationnelles (on parle de MCSCF) ont un coût calculatoire lui aussi élevé et nécessitent aussi de se limiter à un nombre restreint de configurations. La méthode CASSCF [13] (*Complete Active Space SCF*) est une méthode multiconfigurationnelle dans laquelle on définit un "espace actif" et pour lequel les différentes configurations retenues sont celles permises par l'espace considéré. Ce type de méthode est particulièrement avantageux dans le cas où le système étudié comporte des orbitales virtuelles de basse énergie ayant un poids important, comme par exemple, dans les systèmes quasi-dégénérés.

### Les méthodes perturbatives

Plusieurs approches de type perturbatif existent pour déterminer l'énergie de corrélation à partir de la fonction d'onde calculée par la méthode Hartree-Fock. Dans l'approche de Møller et Plesset [14], l'Hamiltonien  $H$  considéré est écrit comme la somme de l'Hamiltonien d'ordre zéro  $H_0$  plus un opérateur de perturbation  $V$ .  $H_0$  étant ici l'Hamiltonien de Fock :

$$\hat{H} = \sum_i^n \hat{F}_i + \hat{V} \quad (1.18)$$

La méthode MP2, pour laquelle le développement perturbatif est fait jusqu'à l'ordre deux et qui ne fait intervenir que les diexcitations est la plus couramment utilisée. On peut d'ailleurs considérer que la plus grande partie de l'énergie de corrélation est déjà contenue dans ce terme correctif. L'énergie totale s'écrit alors :

$$E_{tot.} = E_0 + E_{corr.}^{(2)} \quad (1.19)$$

avec

$$E_{corr.}^{(2)} = \sum_{a<b,r<s} \frac{|\langle \Psi_0 | \sum_{i<j} r_{ij}^{-1} | \Psi_{ab}^{rs} \rangle|^2}{\varepsilon_a + \varepsilon_b - \varepsilon_r - \varepsilon_s} \quad (1.20)$$

Où la notation correspondant aux diexcitations est la même que précédemment,  $i$  et  $j$  étant les indices des électrons et  $\varepsilon_n$  correspondant à l'énergie de la spinorbitale  $n$ , valeur propre de l'opérateur de Fock.

### 1.1.5 Les méthodes de la fonctionnelle de la densité

La théorie de la fonctionnelle de la densité permet aussi de calculer l'énergie électronique des systèmes chimiques. Cette approche, introduite par Hohenberg et Kohn [15], stipule que toute propriété de l'état fondamental d'un système est complètement déterminé par la densité électronique  $\rho(r)$ . L'énergie apparaît alors comme une fonctionnelle de cette densité. D'une façon analogue au principe variationnel, on peut en outre montrer que la densité est stationnaire pour l'état fondamental. L'introduction d'une approche de type orbitalaire telle que  $\rho(r) = \sum_i \varphi_i(r)^* \varphi_i(r)$ , conduit aux équations de Kohn-Sham dont la résolution peut se faire de façon analogue à celle de Hartree-Fock. Dans l'approche mise au point par Kohn et Sham [16], la fonctionnelle d'énergie  $E[\rho(r)]$  s'écrit comme la somme de trois termes :

$$E[\rho(r)] = V[\rho(r)] + T[\rho(r)] + E_{xc}[\rho(r)] \quad (1.21)$$

Où  $V[\rho(r)]$  regroupe les interactions entre les électrons et les noyaux et des électrons entre eux.  $T[\rho(r)]$  est un terme d'énergie cinétique d'un système d'électrons sans interaction.  $E_{xc}[\rho(r)]$  regroupe enfin l'énergie d'échange-corrélation et la différence entre le vrai terme cinétique et  $T[\rho(r)]$ .

Le potentiel conduisant à l'énergie d'échange-corrélation n'étant pas connu dans sa forme exacte, plusieurs formes approchées existent. On distingue notamment les

fonctionnelles de type LDA (*Local Density Approximation*) dans laquelle la densité est considérée localement comme un gaz homogène d'électrons. On trouve également des fonctionnelles qui font aussi intervenir le gradient de cette densité et nommées GGA (*Generalized Gradient Approximation*). Enfin, une troisième sorte de fonctionnelle, incorpore une partie de l'échange Hartree-Fock; on parle alors de fonctionnelles hybrides. Ces fonctionnelles sont aussi paramétrées pour reproduire certains résultats de référence et semblent donc particulièrement indiquées dans le traitement des problèmes chimiques. Nous citerons ici la fonctionnelle B3LYP [17], mise en œuvre lors de ce travail et qui utilise les fonctionnelles d'échange de Becke [18] et de corrélation de Lee, Yang et Parr [19].

Ces méthodes ont pour principal avantage de décrire une partie de la corrélation électronique pour un coût informatique qui reste du même ordre de grandeur que celui d'un calcul Hartree-Fock à système constant et base équivalente. Elles permettent donc d'étudier des systèmes de taille relativement importante.

## 1.2 La mécanique moléculaire

Les méthodes quantiques d'étude des espèces chimiques évoquées dans la partie précédente se caractérisent par le haut degré de précision qu'elles peuvent atteindre. Cependant, la contrepartie, comme déjà évoqué, en est leur coût calculatoire en terme de temps et de ressources informatiques. Il est en effet difficile d'envisager, même de nos jours et avec un niveau raisonnable de précision, l'étude de systèmes chimiques comportant plus de quelques centaines d'atomes, ce qui interdit leur usage pour l'étude de systèmes tels que les protéines, par exemple. On peut alors choisir de se tourner vers l'utilisation de la mécanique moléculaire.

Dans cette thèse, nous avons aussi fait appel à l'utilisation de la mécanique moléculaire mais par le biais des méthodes hybrides que nous détaillerons dans la prochaine partie de cette section.

Dans l'approche qui est celle de la mécanique moléculaire, l'énergie électronique est considérée comme une fonction paramétrique des coordonnées nucléaires, ce qui permet de ne plus passer par la résolution de l'équation de Schrödinger de l'Hamiltonien électronique. Ces paramètres sont déterminés à partir de données expérimentales ou par des calculs *ab initio* de haut niveau. Les molécules sont donc décrites comme des ensembles d'atomes définis par leurs paramètres. Ces derniers interagissent entre eux

par des potentiels prédéfinis suivant les lois de la mécanique newtonienne. On définit alors pour une molécule un champ de forces, constitué par les termes d'interactions entre atomes ainsi que les paramètres des atomes dont on connaît les positions. L'autre point-clé de ces méthodes classiques (par opposition aux méthodes quantiques) est qu'elles reposent sur l'hypothèse de transférabilité des propriétés chimiques. Ainsi, la liaison carbone-hydrogène, par exemple, a une longueur similaire dans toutes les molécules organiques. On peut donc définir un nombre restreint de paramètres sur des molécules ou des systèmes modèles qu'on utilisera ensuite pour construire le champ de forces en question.

Ces méthodes étant par définition paramétriques, il est évident que de nombreux modèles de champs de forces ont été définis en fonction des types de systèmes étudiés. Durant cette étude, nous avons utilisés les trois champs de forces présents dans le logiciel Gaussian [2], à savoir Amber [20], UFF [21] et Dreiding[22]. Nous choisissons de ne détailler ici que les termes de Amber décrit comme un potentiel minimaliste.

Dans ces champ de forces, on distingue deux types de termes définissant les potentiels d'interactions entre les atomes : les termes non liés, regroupant les termes électrostatiques et ceux de van der Waals et les termes liés, décrivant les interactions entre atomes d'une même molécule distants de une (liaison), deux (angle) ou trois (dièdre) liaisons.

On écrit l'énergie totale de la façon suivante :

$$\begin{aligned}
 E_{tot} = & \sum_{liaisons} k_r (r - r_{eq})^2 + \sum_{angles} k_\theta (\theta - \theta_{eq})^2 + \sum_{dièdres} \frac{V_n}{2} [1 + \cos(n\phi - \gamma)] \\
 & + \sum_{i>j} \left[ \frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^6} + \frac{q_i q_j}{\epsilon R_{ij}} \right]
 \end{aligned} \tag{1.22}$$

Les termes liés (première ligne de l'équation 1.22) de liaison et d'angles sont des potentiels harmoniques, prenant en paramètre une constante de force ( $k$ ) et en variable la différence entre la distance ou l'angle considéré et la valeur d'équilibre correspondante, suivant les atomes pris en compte. Pour les termes de torsion, on a un potentiel comprenant une barrière de rotation  $V_n$  d'ordre  $n$ ,  $\phi$  étant l'angle considéré et  $\gamma$  son déphasage par rapport à l'origine.

Pour les termes non liés (seconde ligne dans l'équation 1.22), la partie de van der Waals est décrite par un modèle de Lennard-Jones où  $r_{ij}$  est la distance entre les atomes  $i$  et  $j$ ,  $A_{ij}$  et  $B_{ij}$  étant des grandeurs calculées à partir des diamètres de collision  $\sigma_{ij}$  et

profondeurs de puits  $\varepsilon_{ij}$  propres à chaque atome  $i$  et  $j$  tels que :

$$\begin{aligned} A_{ij} &= 4\varepsilon_{ij}\sigma_{ij}^{12} \\ B_{ij} &= 4\varepsilon_{ij}\sigma_{ij}^6 \end{aligned} \quad (1.23)$$

Avec  $\varepsilon_{ij} = (\varepsilon_{ii}\varepsilon_{jj})^{\frac{1}{2}}$  et  $\sigma_{ij} = \frac{1}{2}(\sigma_{ii} + \sigma_{jj})$

Ces termes ont la particularité d'être soumis à la condition dite "1-4" qui annule les interactions non liées pour les atomes distants de moins de trois liaisons et applique un facteur d'échelle pour les atomes exactement distants de trois liaisons.

Du fait de leur faible coût calculatoire, ces méthodes permettent d'aborder des systèmes pouvant atteindre plusieurs dizaines de milliers d'atomes. On peut alors modéliser leur structure mais aussi leur évolution au cours du temps, via l'intégration des équations du mouvement. Un détail complet des méthodes de modélisation par la mécanique classique pourra être trouvé dans l'ouvrage cité à la référence [23].

Malgré cela, l'approche classique souffre d'une importante limitation. Les termes d'interactions, pour être définis, nécessitent la connaissance préalable de la connectivité du système qui recense toutes les liaisons chimiques. Le potentiel total d'un système dépendant explicitement de cette connectivité, il est obligatoire de la conserver pour rester sur la même surface définie par ce potentiel. L'impossibilité de créer ou de supprimer une liaison rend donc impossible, avec les champs de forces standards, la modélisation de la réactivité chimique des systèmes étudiés. Afin de pallier ce défaut, il est possible de faire appel à des méthodes mettant en œuvre la mécanique quantique pour décrire la partie réactive d'un système et la mécanique moléculaire pour le reste.

### 1.3 Les méthodes hybrides

Dans la suite, nous appellerons méthodes "hybrides" ou "QM/MM" (*Quantum Mechanics/Molecular Mechanics*) toutes les approches combinant mécanique quantique et moléculaire, combinaison qui pourra se faire de différentes manières.

Le principe de cette méthode tire parti du fait que les processus chimiques au sein d'une molécule ne concernent généralement qu'un petit nombre d'atomes. Cette assertion permet de mener l'étude de la réactivité d'un système en modélisant son centre actif par une molécule modèle mais le reste du système bien que peu ou pas touché a un effet sur le centre réactif non nul. L'idée de ces méthodes est donc de traiter le

centre réactif par la mécanique quantique et le reste de la molécule par le biais de la mécanique classique. L'Hamiltonien  $\hat{H}_{tot}$  du système peut alors se décomposer en trois termes :

$$\hat{H}_{tot} = \hat{H}_{QM} + \hat{H}_{MM} + \hat{H}_{QM/MM} \quad (1.24)$$

Où  $\hat{H}_{QM}$  et  $\hat{H}_{MM}$  sont les Hamiltoniens des sous-systèmes quantiques et classiques et  $\hat{H}_{QM/MM}$  l'Hamiltonien d'interactions entre les deux parties. C'est l'expression de ce dernier terme qui définit les spécificités des différentes approches. L'autre point clé apparaît quand les deux sous-systèmes appartiennent à une même molécule, rendant alors obligatoire la coupure d'une ou plusieurs liaisons chimiques. Enfin, il est important de préciser qu'on peut utiliser ces méthodes hybrides dans le cas où la partie classique sert à décrire un solvant, dans lequel est placé un soluté quantique, approche qui ne nécessite pas de couper de liaison chimique. Nous nous intéresserons à ce domaine spécifique dans le troisième chapitre de ce mémoire. Nous nous proposons donc de traiter ici les principales approches et leur spécificités.

### 1.3.1 Approches à interaction directe

Dans cette catégorie, on peut regrouper toutes les méthodes dans lesquelles l'interaction entre parties quantiques et classiques se fait directement, c'est à dire où les charges classiques sont ressenties par la partie quantique en étant intégrées dans l'Hamiltonien de cœur. La principale différence entre ces méthodes est la façon de traiter la coupure de certaines liaisons lors de la partition QM/MM. En effet, la coupure d'une ou plusieurs liaisons pour créer le sous-système quantique en ferait une espèce radicalaire dont les propriétés ne sont pas celles de ce sous-système au sein de la molécule. Les solutions proposées pour résoudre ce problème de "liaison pendante" sont assez diverses, et on peut grossièrement former deux groupes : dans le premier, la liaison est simplement complétée en ajoutant un atome monovalent tandis que dans le second, les modifications portent sur la liaison elle-même.

Lorsqu'on choisit d'ajouter un atome, la première approche, assez simple dans son concept, consiste à saturer la liaison coupée par un atome d'hydrogène. Cette méthode, appelée "link atom" [24], pose cependant plusieurs problèmes conceptuels, le premier étant l'augmentation du nombre de degrés de libertés du système. En effet, au système total sont ajoutés autant d'atomes d'hydrogène que de liaisons coupées, ce qui change la dimension de l'hypersurface de potentiel. Le problème du positionnement des atomes de liaison se pose, à savoir s'il doit ou non rester dans l'alignement de la liaison coupée

mais aussi si la liaison formée entre l'atome à la frontière et l'hydrogène doit être relaxée, fixe ou encore proportionnelle à la longueur de celle qu'elle remplace. Se pose ensuite la question des interactions de l'atome de liaison avec son environnement, en effet, si celles-ci sont comptées, elles n'ont pas de sens physique mais si elles ne le sont pas, il est possible de perdre certaines informations, comme la polarisation induite par l'environnement.

Enfin, l'inconvénient majeur est que du point de vue électronique puisque la liaison créée ne représente pas forcément bien celle qu'elle est sensée décrire. Le remplacement d'un atome lourd (carbone, oxygène....) par un hydrogène changera les caractéristiques de la liaison comme sa polarité par exemple.

Afin de faire face à ces défauts, on peut citer, entre autres, le remplacement de l'atome frontière par un pseudo-halogène [25] ou par un groupement atomique fantôme [26], qui permet, en particulier, d'ajuster la polarité de la liaison.

De même, Antes et Thiel proposent la méthode "Adjusted Connection atoms" [27] où l'atome de liaison est un atome standard pour le niveau classique et reçoit des paramètres spécifiques d'interaction avec les autres atomes de la partie quantique. Cette approche a été développée dans le cadre des méthodes semi-empiriques et a pour avantage, par rapport au link atom d'apporter une certaine flexibilité grâce aux paramètres mais aussi de ne pas ajouter de degrés de liberté supplémentaires puisque l'atome de connection est l'atome frontière classique.

Dans le concept de "pseudo-liaison" de Zhang *et al.* [28],[29], pour une liaison frontière entre deux atomes X et Y (X appartenant au sous-système quantique et Y à la partie classique), l'atome Y est remplacé par un atome  $Y_{ps}$  avec une couche de valence complète dont on a supprimé un électron. On ajoute alors à cet atome un pseudo-potentiel effectif de cœur pour que la liaison X- $Y_{ps}$  reproduise la liaison X-Y initiale.

Dilabo, Hurley et Christiansen proposent quant à eux de remplacer l'atome de liaison par un pseudo-potentiel qu'ils appellent QCP (*Quantum Capping Potential*)[30], qui ne demande que très peu de modifications du programme et permet d'améliorer la représentation de la liaison coupée au niveau quantique.

Swart propose de son côté la méthode "AddRemove" [31], dans laquelle la position de l'atome de liaison dépend de celle de l'atome classique lié à la frontière coupée, permettant ainsi de supprimer les degrés de liberté supplémentaires introduits. De plus l'introduction de cet atome est corrigée en retirant l'interaction que celui-ci aurait avec

le sous-système quantique s'il était classique.

On peut enfin citer l'approche de Das *et al.* [32] dite "double link atom" et pour laquelle un atome de liaison est aussi ajouté à la sous-partie classique, pour supprimer la polarité non physique induite par la coupure.

La seconde solution pour résoudre le problème des liaisons pendantes consiste à traiter directement la partie électronique de la liaison coupée. Warshell et Levitt [33] ont proposé d'utiliser des orbitales hybrides de liaison pour décrire la liaison coupée. C'est à partir de cette idée qu'a été développée à Nancy l'approche LSCF (*Local Self Consistent Field*) au niveau semi-empirique [34],[35] puis au niveau *ab initio* [36]. Cette dernière méthode, sur laquelle ont porté nos travaux fera l'objet d'une description détaillée dans la suite de ce chapitre.

L'approche GHO (*Generalized Hybrid Orbital*) [37],[38],[39] due à Gao *et al.* qui a été portée au niveau *ab initio* s'inspire de cette méthode. Friesner en a aussi repris les principes pour son propre champ de forces [40],[41].

Dans l'approche récente de Lin et Truhlar [42], nommée RC (*redistributed charge*), des charges ponctuelles sont positionnées entre l'atome de liaison et ses premiers voisins classiques. Selon les auteurs, cette approche peut être vue comme une adaptation au niveau classique de la méthode GHO.

L'autre aspect important dans l'approche QM/MM est le calcul des interactions entre les deux sous systèmes. Les interactions de type van der Waals étant décrites la plupart du temps au niveau classique en considérant les potentiels standard du champ de force de la partie MM pour les atomes QM, l'attention s'est donc portée principalement sur les termes électrostatiques. Dans la majorité des cas, on se contente de faire interagir des charges fixes positionnées sur les atomes avec la partie quantique. Ces charges doivent donc faire l'objet d'une attention spéciale, et en particulier, celle qui se trouve sur l'atome classique lié au sous système quantique. On citera en particulier les travaux de Reuter *et al.* [43] qui, outre une comparaison entre les différentes descriptions des frontières, traitent de la charge sur l'atome classique à la frontière et proposent de lui attribuer une charge faible mais non nulle.

Das *et al.* [32] proposent de délocaliser les charges classiques sous formes de gaussiennes, dans le but de décrire le caractère délocalisé du nuage électronique. Selon Thompson et Schenter, l'approche faisant intervenir des charges ponctuelles reste limitée puisque elle ne décrit que la polarisation de la partie quantique par l'environnement classique et pas l'inverse. Ils proposent de donner un caractère polarisable aux

atomes classiques [44].

On rappellera enfin que toutes ces méthodes ont pour but de décrire des systèmes de grande taille, ce qui induit un grand nombre de degrés de liberté que seule une approche statistique peut traiter. La plupart des méthodes énoncées ci-dessus est donc généralement utilisée dans le cadre de simulations dépendant du temps et font donc appel à un Hamiltonien semi-empirique pour la partie quantique pour des raisons de coût calculatoire. On citera alors aussi pour mémoire les travaux de Laio *et al.* dans lesquels la partie quantique est traitée par la méthode Car-Parrinello [45].

Afin de compléter cette partie, nous choisissons de citer quelques références qui donnent un point de vue assez global sur ce sujet. Ainsi, nous commencerons par citer l'article de Monard et Merz [46]. Les articles de Reuter *et al.* [43] ainsi que de Hall *et al.* [47] comparent eux aussi les différentes approches.

Dans un article plus récent, Monard *et al.* traitent ces méthodes dans le cadre de l'étude de la réactivité enzymatique [48]. L'article d'Amara et Field [49] s'intéresse quant à lui à l'évaluation de l'approche "link atom" en comparaison avec d'autres modèles. Les auteurs stipulent notamment que le facteur déterminant dans de telles approches reste l'interaction entre les deux sous-parties. Nous finirons enfin par citer le numéro spécial de la revue Journal of Molecular Structure : THEOCHEM consacré aux méthodes QM/MM [50] qui offre un panel récent des méthodes et applications de l'approche hybride en général.

L'approche QM/MM peut aussi se faire d'une façon différente que nous avons choisi de décrire séparément. Il s'agit de l'approche de type ONIOM.

### 1.3.2 L'approche ONIOM

L'acronyme ONIOM [51] (pour *Our own N-layered Integrated molecular Orbital and molecular Mechanics*) désigne en fait une approche générale des calculs faisant intervenir plusieurs niveaux de théories, cette approche, comme nous allons voir diffère des précédentes par sa philosophie, c'est pourquoi nous avons choisi de la traiter à part. En effet, pour calculer l'énergie dans ce cas, plusieurs calculs sont faits. Dans le cas de la prise en compte de deux niveaux de calculs, on procède comme suit. On considère d'abord le système entier (désigné comme réel), dont on calcule l'énergie à un bas niveau ( $E_1$ ). On calcule ensuite l'énergie de la partie réactive (système modèle), dont les éventuelles liaisons coupées sont complétées par des atomes de liaison (généralement l'hydrogène), au niveaux bas ( $E_2$ ) et haut ( $E_3$ ).

Ces trois valeurs sont ensuite combinées pour obtenir l'énergie totale du système suivant l'équation :

$$E_{tot} = E_1 - E_2 + E_3 \quad (1.25)$$

On remarque que, dans cette approche, les énergies sont calculées indépendamment les unes des autres, ce qui permet d'utiliser n'importe quelle méthode pour déterminer ces dernières. On a donc principalement deux approches. La première consiste à coupler deux méthodes quantiques [52] (IMOMO : *Integrated Molecular Orbitals Molecular Orbitals*) avec, typiquement, un haut niveau *ab initio* et un bas niveau semi-empirique. Dans la seconde, le bas niveau est traité par la mécanique moléculaire (IMOMM : *Integrated Molecular Orbitals Molecular Mechanics*).

De la même façon, l'approche peut être étendue au nombre de niveaux décidé. dans la pratique, trois niveaux [53] sont suffisants (un niveau classique, un semi-empirique et le dernier traité par une méthode quantique de haut niveau).

Malgré cette flexibilité, l'approche ONIOM souffre d'un défaut majeur lié à sa nature même. En effet, puisque les différents niveaux sont traités indépendamment, les interactions entre les différents parties ne sont prises en compte qu'au niveau de calcul le plus bas. Dans le cas de IMOMM, la fonction d'onde de la partie quantique n'est donc pas polarisée par les charges présentes sur les atomes du système classique, seules les interactions stériques et électrostatiques évaluées au bas niveau de théorie sont prises en compte. Ce défaut n'est vraiment gênant que dans le cas où l'effet prépondérant de l'environnement est électrostatique. On note par ailleurs que, dans la version de cette méthode implantée dans Gaussian 03 [2], cette limitation peut être corrigée en incorporant les charges de la partie classique dans le calcul du système modèle.

En raison de sa facilité de mise en œuvre, cette méthode connaît un grand succès dans son utilisation. On citera, parmi les développements récents, le couplage de l'approche ONIOM avec la méthode de dynamique quantique ADMP (*Atom-centered Density Matrix Propagation*) [54].

## 1.4 La méthode LSCF/MM *ab initio*

Comme nous l'avons écrit plus haut, nos travaux ont eu pour base principale la méthode QM/MM LSCF/MM et c'est pourquoi nous choisissons de la développer plus précisément dans cette partie.

La particularité de la méthode LSCF/MM vient de la façon de traiter la liaison frontière entre les sous-parties quantique et classique. En effet, la liaison pendante créée par sa position frontalière est décrite par une orbitale localisée qui n'évolue pas au cours de la procédure auto-cohérente d'optimisation de la fonction d'onde du sous-système quantique. Nous commencerons donc par décrire la méthode quantique LSCF qui permet le "gel" d'une ou plusieurs orbitales moléculaires puis nous ferons de même avec le champ de force hybride LSCF/MM qui en est issu.

### 1.4.1 LSCF

Comme nous venons de l'évoquer, l'objectif de la méthode LSCF [55] est de pouvoir optimiser une fonction d'onde électronique sous la contrainte de garder inchangées une ou plusieurs spinorbitales durant le calcul auto-cohérent. Ces spinorbitales sont fixées au début du calcul par l'utilisateur et elles représentent en outre une densité électronique constante qui peut interagir avec le reste des électrons du système quantique.

#### Le formalisme LSCF

Si on se place dans le formalisme LCAO, seuls le jeu de coefficients des orbitales atomiques qui définit ces spinorbitales dans la base considérée ainsi que leurs nombres d'occupation doivent être connus au préalable. Aucune contrainte n'est donc imposée concernant la forme et le degré d'occupation d'une spinorbitale gelée. Dans la suite du développement, nous nous plaçons dans le formalisme restreint, où le nombre d'occupation d'une orbitale gelée (OG) est zéro (orbitale vide) ou deux (occupée).

Cette méthode est applicable aux approches basées sur un formalisme orbitalaire, telles que celles prenant en compte l'hamiltonien de Hartree-Fock comme point de départ (comme celles destinées à prendre en compte la corrélation électronique). Cette remarque est aussi valable pour les méthodes de la DFT basées sur une approche de type Kohn-Sham.

L'orbitale gelée numéro  $i$ , notée  $|l_i\rangle$  est exprimée sur une base  $\{\chi_1 \cdots \chi_\mu \cdots \chi_\nu \cdots \chi_N\}$  de  $N$  orbitales atomiques :

$$|l_i\rangle = \sum_{\mu}^N \alpha_{\mu i} |\chi_{\mu}\rangle \quad (1.26)$$

avec  $\alpha_{\mu i}$  les coefficients satisfaisant les conditions de normation telles que :

$$\langle l_i | l_i \rangle = \sum_{\mu}^N \sum_{\nu}^N \alpha_{\mu i}^* S_{\mu\nu} \alpha_{\nu i} = 1 \quad (1.27)$$

Notre but est de trouver le jeu des  $N - L$  orbitales restantes (avec  $L$  le nombre d'orbitales gelées) qui minimisent l'énergie électronique du système que nous noterons "OM" (Orbitales Moléculaires). Afin d'éviter la difficulté calculatoire due à la non orthogonalité entre orbitales gelées et relaxées, nous choisissons de construire nos OMs sur une base de  $N$  orbitales atomiques de telle façon à ce qu'elles soient orthogonales aux OGs. Pour plus de clarté, les orbitales sont désignées par leurs indices. On prend pour convention d'attribuer des lettres latines aux orbitales de type moléculaire (dont font partie les OGs) et grecques aux orbitales atomiques.

Le premier point est d'abord l'orthogonalisation mutuelle des OGs par la méthode de Löwdin [56] pour former  $L$   $|l'_i\rangle$  orbitales gelées orthogonales (OGOs) :

$$|l'_i\rangle = \sum_{\mu}^N a_{\mu i} |\chi_{\mu}\rangle \quad (1.28)$$

Les  $N$  orbitales atomiques de la base sont ensuite rendues orthogonales aux ces OGs en soustrayant leur projection sur ces OGOs. On note  $|\tilde{\mu}\rangle$  une de ces fonctions telle que :

$$|\tilde{\mu}\rangle = \sum_{\nu} \left[ 1 - \sum_i^L S_{\mu i}^2 \right]^{-\frac{1}{2}} \cdot \left[ \delta_{\mu\nu} - \sum_i^L \sum_{\eta} a_{\nu i} a_{\eta i} S_{\eta\mu} \right] |\nu\rangle \quad (1.29)$$

Avec  $S_{\mu i}$  l'intégrale de recouvrement entre les fonctions  $\mu$  et  $i$  ( $\langle \mu | i \rangle$ ),  $S_{\nu\eta}$  exprime  $\langle \nu | \eta \rangle$  et  $\delta_{\mu\nu}$  est le symbole de Kronecker. Cette équation peut être exprimée simplement par :

$$|\tilde{\mu}\rangle = \sum_{\nu} m_{\nu\tilde{\mu}} |\nu\rangle \quad (1.30)$$

où  $m_{\nu\tilde{\mu}}$  est l'élément  $(\nu, \tilde{\mu},)$  de la matrice  $\mathbf{M}$  d'orthogonalisation des OA aux OG de dimension  $N \times N$ .

Ces fonctions ne sont cependant pas linéairement indépendantes du fait de leur implication dans les  $L$  combinaisons linéaires que sont les orbitales gelées. La matrice  $\tilde{\mathbf{S}}$  de recouvrement des fonctions ainsi construite aura  $L$  valeurs propres nulles. De même la matrice unitaire  $\mathbf{U}$  des vecteurs propres correspondante aura  $L$  vecteurs propres indéfinis. La suppression de ces  $L$  vecteurs et valeurs propres nuls permet de construire la matrice d'orthogonalisation canonique[4]  $\mathbf{X}$  de dimension  $(N \times (N - L))$  telle que :

$$\mathbf{X} = \mathbf{U}\tilde{\mathbf{S}}^{-\frac{1}{2}} \quad (1.31)$$

Les  $N-L$  fonctions ainsi déterminées sont linéairement indépendantes, orthogonales mutuellement ainsi qu'aux orbitales gelées, ce qui nous permet alors de déterminer les OMs dans les conditions désirées.

On peut par ailleurs remarquer que les deux transformations 1.29 et 1.31 peuvent être faites en une seule opération grâce à une matrice  $\mathbf{B}$  telle que :

$$\mathbf{B} = \mathbf{M}\mathbf{X} \quad (1.32)$$

où  $\mathbf{B}$  est de dimension  $(N \times (N-L))$  et peut directement remplacer la matrice  $\mathbf{S}^{-\frac{1}{2}}$  dans la procédure standard SCF telle que décrite au début de ce chapitre. La matrice associée à l'opérateur de Fock dans la base orthogonalisée ainsi définie s'écrira donc :

$$\mathbf{F}' = \mathbf{B}^\dagger \mathbf{F} \mathbf{B} \quad (1.33)$$

et sera de dimension  $((N-L) \times (N-L))$ .

La matrice densité totale  $\mathbf{P}^T$  aura quant à elle la particularité de s'écrire :

$$\mathbf{P}^T = \mathbf{P}^Q + \mathbf{P}^G \quad (1.34)$$

Où  $\mathbf{P}^Q$  est associée au système des OM calculées par la procédure SCF et  $\mathbf{P}^G$  est associée aux OGs.

La méthode LSCF peut aussi être utilisée dans le cadre non restreint, il suffit pour cela de construire deux matrices densités associées aux OG pour les électrons  $\alpha$  et  $\beta$ . De la même façon on déterminera deux matrices  $\mathbf{B}^\alpha$  et  $\mathbf{B}^\beta$ . Son utilisation dans le cadre des méthodes post Hartee-Fock telles qu'évoquées dans ce chapitre est possible. Dans le cas d'une approche perturbative, il ne faut cependant pas prendre en compte les configurations faisant intervenir des orbitales gelées, celles-ci n'étant pas fonctions propres de l'opérateur de Fock.

Il est important de remarquer que, puisque les OG ne sont pas modifiées lors du calcul auto cohérent, celles-ci n'interviennent que comme paramètres du problème et servent ainsi à définir l'hypersurface de potentiel ainsi que les positions des noyaux.

### Domaines d'application de la méthode LSCF

La méthode purement quantique LSCF permet d'optimiser la fonction d'onde d'un système tout en conservant certaines orbitales moléculaires inchangées et ce, quels que soient la forme de ces dernières et leur degré d'occupation. Cette caractéristique permet plusieurs applications de la méthode. Comme nous l'avons déjà évoqué dans ce chapitre, en localisant une orbitale et en la gelant, on peut décrire une liaison coupée par une frontière (que nous nommerons liaison frontière) dans le cadre des méthodes QM/MM. Ce domaine sera repris en détails dans la section suivante. Une autre application de cette méthode est la détermination des énergies d'excitation électronique et d'ionisation des molécules.

Le principe est de partir du déterminant de Slater de l'état fondamental  $\Psi$ , d'en geler une orbitale et de la vider tout en gardant le même nombre d'électrons. On ré-optimise ensuite la fonction d'onde pour obtenir un nouveau déterminant  $\Psi'$  décrivant un état dans lequel une orbitale précédemment occupée est vidée au profit d'un niveau qui était vide, état qu'on peut qualifier d'excité. Les deux déterminants étant orthogonaux entre eux [57], l'énergie d'excitation s'écrit alors comme la différence entre les énergies totales de  $\Psi$  et  $\Psi'$ . Le calcul d'énergies d'ionisation se fait de la même manière mais en considérant un déterminant  $\Psi'$  auquel on soustrait un électron.

Ces approches ont été utilisées pour calculer les énergies d'excitation et d'ionisation de cœur sur de petites molécules et, couplées à une description hybride, pour l'étude des mêmes phénomènes au sein de polymères[58].

#### 1.4.2 Le champ de forces LSCF/MM

En ce qui concerne nos travaux, nous avons utilisé l'approche LSCF couplée avec un champ de forces classique dans le cadre des méthodes QM/MM. La frontière entre les deux sous-systèmes est décrite par une (ou plusieurs, si besoin est) orbitale de liaison gelée et strictement localisée que nous désignerons SLBO (pour *Strictly Localized Bond Orbital*).

#### Définition et utilisation des orbitales localisées

La première étape d'un calcul LSCF/MM est donc, pour pouvoir définir le sous-système quantique, la définition des SLBO qui connecteront les deux parties. En considérant l'hypothèse de transférabilité des propriétés chimiques, les orbitales loca-

lisées sont déterminées sur de petites molécules modèles des liaisons frontières. Ainsi, si on coupe une liaison entre deux atomes de carbones hybridés  $sp^3$ , on peut procéder à la détermination de la SLBO sur une molécule d'éthane, plus petite molécule contenant ce type de liaison.

Pour ce faire, on dispose de plusieurs méthodes qui se différencient par le critère de localisation. Les méthodes globales (toutes les orbitales moléculaires occupées sont localisées) de Boys et Fosters [59], Edminston-Ruedenberg [60] ou Pipek-Mezey [61] sont disponibles. Cependant, puisque le but est d'obtenir une seule orbitale localisée, les méthodes à critère local (localisation d'une seule orbitale) sont plus intéressantes. La méthode de Weinstein-Pauncz apparaît alors comme méthode de choix [58], le critère de Magnasco-Persico pouvant ensuite venir affiner le résultat obtenu [62].

On remarquera que dans le cas où la liaison se situe entre deux atomes de natures différentes, cette procédure permet de reproduire, du moins en partie, la polarité de la liaison ainsi décrite.

Toutes ces méthodes étant appliquées après le calcul SCF, les critères cités ne localisent pas complètement les orbitales sur la liaison souhaitée. Certaines orbitales atomiques non centrées sur les atomes définissant la liaison à localiser possèdent donc encore des coefficients non nuls dans l'orbitale localisée. Afin de définir une orbitale strictement localisée, on supprime ces coefficients de l'orbitale considérée, ce qui nécessite ensuite de renormer la SLBO ainsi créée.

Il est par ailleurs envisageable de produire des orbitales strictement localisées directement avec le calcul SCF, comme dans l'approche ELMO de Sironi *et al.* [63] par exemple, une étude comparative entre cette approche et les précédentes a par ailleurs été menée récemment [64].

Ces SLBOs qui sont le moyen de représenter la liaison coupée par une approche QM/MM ne sont donc pas modifiées par le calcul LSCF. Cependant, lors d'une optimisation de géométrie, les atomes entre lesquels elles sont définies doivent pouvoir bouger et les SLBOs doivent alors pouvoir s'adapter.

Comme nous l'avons précisé précédemment, les orbitales gelées (dont font partie les SLBOs) sont des paramètres qui servent à définir la surface d'énergie potentielle du système étudié. Pour ne pas quitter cette surface durant une modification de géométrie, les SLBOs doivent conserver leurs caractéristiques de polarité et de rapport entre les interactions  $\sigma$  et  $\pi$ . L'adaptation d'une SLBO durant une optimisation de géométrie se fera donc uniquement par rotation et adaptation en longueur (une simple renormation

suffit) de celle-ci. Afin de pouvoir effectuer une telle rotation, on définit un trièdre unique de référence avec trois atomes du système formé des deux atomes de la liaison considérée ainsi que d'un troisième qui doit rendre compte de la dissymétrie de la rotation autour de la liaison. On applique ensuite la transformation aux différentes fonctions de base de la SLBO par le biais de matrices de rotations, telles que décrites dans la référence [57].

On doit enfin pouvoir connaître les dérivées premières des SLBO selon les coordonnées nucléaires pour permettre aux atomes frontières de bouger lors d'une optimisation. Une procédure numérique rapide permet alors de calculer les coefficients des orbitales atomiques dans les SLBOs et ainsi de les utiliser dans le calcul analytique des dérivées de l'énergie électronique LSCF dont on pourra de nouveau trouver le détail du calcul dans les références [57] et [58].

Nous avons ainsi détaillé la façon dont sera utilisée l'approche LSCF pour définir les frontières au niveau quantique dans un calcul QM/MM. Nous pouvons donc à présent décrire le champ de forces LSCF/MM lui-même.

### Définition du champ de forces LSCF/MM

Dans la suite, nous adopterons la notation illustrée par la figure 1.1 où les atomes classiques sont notés C et ceux de la partie quantique Q. Ils sont indicés dans l'ordre de la connectivité depuis l'atome frontière X pour la partie quantique et Y pour la partie classique.

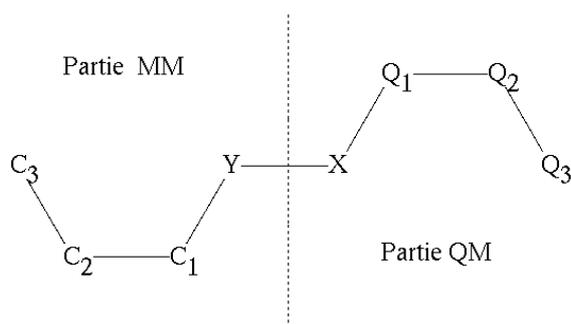


FIG. 1.1 – Notation des atomes au niveau de la frontière

L'atome frontière Y a pour particularité d'avoir à la fois des caractéristiques classiques et quantiques. En tant qu'atome quantique, il possède les fonctions de base de l'atome qu'il représente et apporte généralement un électron au système. Schématiquement,

il sert comme "point d'accroche" à la SLBO du côté classique. Afin de respecter l'équilibre des charges du système étudié, il possède aussi une charge nucléaire égale au nombre d'électrons apportés, généralement un. Ce nombre peut aussi être égal à zéro ou deux dans le cas d'une liaison dative, en fonction de la nature d'accepteur ou de donneur de Y. En tant qu'atome classique, il possède tous les paramètres de celui qu'il représente.

Les interactions entre les différents atomes des deux sous-systèmes sont calculées, à l'exception de la partie électrostatique, au niveau du champ de forces. Ces interactions sont résumées dans le tableau 1.1.

liaisons	$X - Y^{(a)}$
angles	$X - Y - C_1$
torsions	$Q_1 - X - Y - C_1, X - Y - C_1 - C_2$
van der Waals	$X \cdots C_i (i > 1), Q_i \cdots Y (i > 1)$ et $Q_i \cdots C_i, \forall i$

TAB. 1.1 – liste des interactions QM/MM non électrostatiques. <sup>(a)</sup> : terme de liaison spécifique au champ de forces LSCF/MM.

Si on ne considère que les interactions QM/MM, on remarque que seuls les atomes frontières possèdent un potentiel de liaison, ce dernier étant spécifique. Si on considère les atomes X et Y du point de vue quantique, X est un atome habituel tandis que Y possède les fonctions de base de l'atome qu'il représente mais seulement une charge nucléaire +1 et une éventuelle charge classique. De ce fait, cet atome ressemble à un atome d'hydrogène et lors d'une optimisation de géométrie, la liaison X-Y voit sa longueur tendre vers celle d'une liaison X-H [58]. Afin de corriger ce comportement, il a donc été proposé de rajouter un potentiel de liaison frontière entre X et Y [58]. Celui-ci est déterminé pour reproduire le profil énergétique de liaison considérée au voisinage de la distance d'équilibre. Ce potentiel, fait donc partie intégrante du champ de forces LSCF/MM.

Les interactions électrostatiques, sont déterminées au niveau quantique en incluant les charges portées par les atomes classiques dans le calcul de la fonction d'onde. Ces charges interagissent donc directement avec les électrons et les noyaux de la partie quantique. La polarisation de ce sous-système se fait en ajoutant un terme à l'expression 1.12 de l'opérateur de Fock. L'élément  $(\mu, \nu)$  de cet opérateur s'écrit alors dans la base des orbitales atomiques  $(\phi_\mu, \phi_\nu)$  :

$$F_{\mu\nu} = F_{\mu\nu}^0 + \sum_i^{MM} \langle \phi_\mu | \frac{q_i}{r_i} | \phi_\nu \rangle \quad (1.35)$$

avec  $F_{\mu\nu}^0$  l'élément de la matrice de Fock sans prise en compte des charges et  $q_i$  la charge de l'atome classique  $i$ . On remarque par ailleurs que l'interaction électrostatique n'est pas soumise à la condition "1-4" évoquée précédemment : toutes les charges classiques sont ressenties par la partie quantique.

Enfin, il faut noter que, pour tous les paramètres intervenant dans l'interaction QM/MM, le choix a été fait d'utiliser ceux fournis par le champ de forces de la partie classique. Avec les SLBOs, les potentiels frontières (lorsqu'ils ne sont pas définis dans le code) et les charges classiques sont donc les seuls paramètres à déterminer avant de pouvoir effectuer un calcul LSCF/MM.

Nous venons de décrire les fonctionnalités de la méthode LSCF/MM lors du début de nos travaux. Au cours de ces derniers, nous avons été amenés à lui adjoindre quelques fonctionnalités, dont les plus importantes sont décrites dans les chapitres suivants. D'autre part, nous avons aussi programmé une méthode permettant de calculer la matrice hessienne plus rapidement et avec un niveau d'approximation raisonnable. C'est cet aspect qui est développé dans le paragraphe suivant.

## 1.5 Détermination du Hessien par blocs

Dans l'étude des systèmes moléculaires complexes, la détermination de points critiques de la surface d'énergie potentielle construite à partir des degrés de liberté du système chimique est essentielle pour pouvoir décrire la réactivité de ce dernier.

Le terme de point critique recouvre à la fois les minima de la surface ainsi que les points selle, caractéristiques des structures de transition.

La méthode LSCF/MM ne permet pas encore le calcul analytique des dérivées secondes de l'énergie en fonction des coordonnées nucléaires. La caractérisation de points critiques nécessitant le calcul du hessien, tels que les états de transition, passe donc obligatoirement par une détermination numérique des dérivées secondes où chaque élément  $i, j$  de la matrice hessienne (noté aussi  $\mathcal{H}_{i,j}$ ) se calcule tel que :

$$\frac{\partial^2 E}{\partial \chi_i \partial \chi_j} = \frac{1}{2\delta} \left\{ \frac{\partial E(r, \chi_i + \delta)}{\partial \chi_j} - \frac{\partial E(r, \chi_i - \delta)}{\partial \chi_j} \right\} \quad (1.36)$$

Avec  $\chi_i$  la  $i^{ieme}$  coordonnée du système,  $r$  désignant l'ensemble des coordonnées différentes de  $\chi_i$  et  $\delta$  un petit déplacement de la coordonnée  $\chi_i$ .

Dans le cas de coordonnées cartésiennes (problème à  $3N$  dimensions, avec  $N$  le nombre d'atomes du système), cette procédure nécessite donc au moins  $2 \times 3N + 1$  déterminations de gradients de l'énergie.

Cette procédure est donc particulièrement longue si on s'intéresse à des systèmes de grandes tailles et n'est plus envisageable pour ceux correspondant au domaine d'étude des approches hybrides QM/MM.

Lorsqu'on étudie la réactivité de tels systèmes chimiques, on s'intéresse particulièrement à la région, de taille restreinte, décrite par la mécanique quantique. Les coordonnées de réaction d'intérêt faisant alors essentiellement intervenir des degrés de libertés associés aux atomes de cette partie.

### 1.5.1 Principe général

Afin de réduire significativement le temps de calcul de la matrice hessienne, nous adoptons une stratégie dans laquelle les constantes de force associées aux coordonnées des atomes quantiques sont déterminées numériquement tandis que celles associées aux atomes classiques le sont analytiquement, en utilisant uniquement les termes du champs de force. Cette stratégie implique un certain nombre d'approximations dont nous testerons la validité.

Ce principe de calcul différencié a été utilisé dans différents travaux. Notons en particulier ceux de Prat-Resina *et al.* [65] qui proposent d'utiliser une matrice hessienne dans laquelle les éléments du sous-bloc quantique sont calculés exactement tandis que, parmi les autres éléments, seuls ceux de la diagonale sont non nuls. Associée à des algorithmes de "mise à jour" du hessien, cette approche permet la localisation d'états de transition dans le cas de systèmes de grandes tailles.

Notre stratégie est plus proche des travaux de Truhlar *et al.* [66] puisque, de la même façon, le système chimique est découpé en deux sous-parties. La partie d'intérêt particulier est traitée à un haut niveau de théorie pour lequel le calcul analytique des dérivées secondes est impossible alors que le reste l'est avec une théorie permettant ce calcul.

Lorsqu'on traite un système de  $N$  atomes avec une approche QM/MM, la matrice hessienne, exprimée sur la base des coordonnées cartésiennes, est de dimension  $3N \times 3N$ . Dans la suite, nous nous plaçons dans le cas où  $l$  atomes sont traités au niveau QM

et  $m$  le sont avec la mécanique moléculaire. Pour plus de clarté dans les schémas, on suppose par ailleurs que les premiers atomes (et les  $3l$  coordonnées correspondantes) sont ceux traités quantiquement, ces derniers étant regroupés. Ceci peut être représenté par la structure en blocs de la matrice hessienne (figure 1.2)

$$\mathcal{H} = \left( \begin{array}{c|c} QM & QM/MM \\ \hline QM & MM \\ /MM & \end{array} \right)$$

FIG. 1.2 – répartition des différents blocs de la matrice hessienne en fonction de la méthode utilisée pour traiter les atomes correspondants

La partie QM, de taille  $l \times l$ , regroupe tous les termes  $\mathcal{H}_{i,j}$  pour lesquels  $i$  et  $j$  sont des coordonnées d'atomes quantiques. Pour les parties QM/MM, quand  $i$  se rapporte à un atome quantique,  $j$  correspond à un atome classique et inversement. On notera par ailleurs que la matrice Hessienne est diagonale symétrique et donc que  $\mathcal{H}_{i,j} = \mathcal{H}_{j,i}$ . Enfin dans la partie MM,  $i$  et  $j$  désignent des coordonnées d'atomes traités par la mécanique moléculaire.

### 1.5.2 Procédure de calcul

La procédure de détermination se fait en deux temps : la détermination numérique des constantes de force liées aux atomes quantiques et le calcul analytique au niveau classique pour les coordonnées associées aux atomes de la partie traitée par la mécanique moléculaire.

Dans le traitement numérique, le système entier est pris en compte, y compris les termes du champ de forces et tous les termes du gradient sont donc calculés à chaque pas. Lorsqu'on calcule la constante de force suivant la coordonnée  $\chi_i$ , tous les éléments  $\frac{\partial^2 E}{\partial \chi_i \partial \chi_j}$ , pour tout  $j$ , sont déterminés, y compris lorsque  $j$  ne se réfère pas à une coordonnée d'un atome quantique. Cette propriété permet donc, outre le sous-bloc QM, de décrire les sous-blocs QM/MM sans approximation.

En ce qui concerne le dernier sous-bloc (MM), si le calcul analytique ne se fait qu'avec le champ de forces, on perd alors contributions dues aux termes d'interaction entre la sous-partie quantique et la sous-partie classique, modifiant ainsi la surface de

potentiel prise en compte. Le calcul des constantes de force dans ce cas ne donnera donc pas les mêmes termes que dans le cas d'une détermination complète exacte.

On peut cependant considérer que négliger les interactions électrostatiques entre les sous-parties quantiques et classiques dans le calcul des constantes de force de la partie MM est une approximation acceptable. En effet, si on prend l'exemple de la recherche d'un état de transition pour une réaction chimique, le vecteur de transition est principalement construit sur les coordonnées des atomes quantiques, région où se produit la réaction, ce qui ne nécessite pas forcément un calcul précis pour les composantes appartenant aux autres sous-blocs.

La prise en compte de cette interaction pour la détermination des constantes de force associées aux atomes classiques peut quand même présenter un intérêt comme dans le cas où la réaction subit l'influence d'une liaison hydrogène.

Nous proposons donc, dans le calcul lié au sous-bloc MM, d'incorporer un jeu de charges ponctuelles ajusté sur le potentiel électrostatique [67] de la fonction d'onde de la partie quantique. Ces charges occupent les positions des atomes quantiques. Il est donc possible de définir deux procédures : soit l'interaction électrostatique entre les sous parties classiques et quantiques de la molécules est négligée, soit elle est considérée par le biais du jeu de charges.

Les deux procédures possibles sont décrites par la figure 1.3.

Afin d'illustrer l'utilisation de ces procédures, nous avons choisi de les appliquer à des systèmes étudiés au cours de nos travaux.

### 1.5.3 Caractérisation d'un état de transition

Le premier usage de notre méthode approchée du calcul du hessien est bien entendu de fournir une matrice hessienne de départ aux algorithmes de recherche d'états de transition. La prise en compte d'une matrice de bonne qualité (se rapprochant de la matrice exacte) permet de diminuer le nombre de pas pour atteindre le point désiré. Cependant, notre méthode peut aussi être utilisée pour caractériser un point selle. Dans l'exemple qui suit, nous nous intéressons à une des structures de transition obtenues dans l'étude présentée au chapitre deux (état de transition nommé : endo s-cis RE syn). Les constantes de force, fréquences et modes de vibrations, obtenus à l'aide des deux approches présentées, sont comparés aux valeurs correspondantes déterminés par une procédure numérique standard (complète), prise comme référence.

Sur l'ensemble des éléments du Hessien, la différence maximale entre la référence

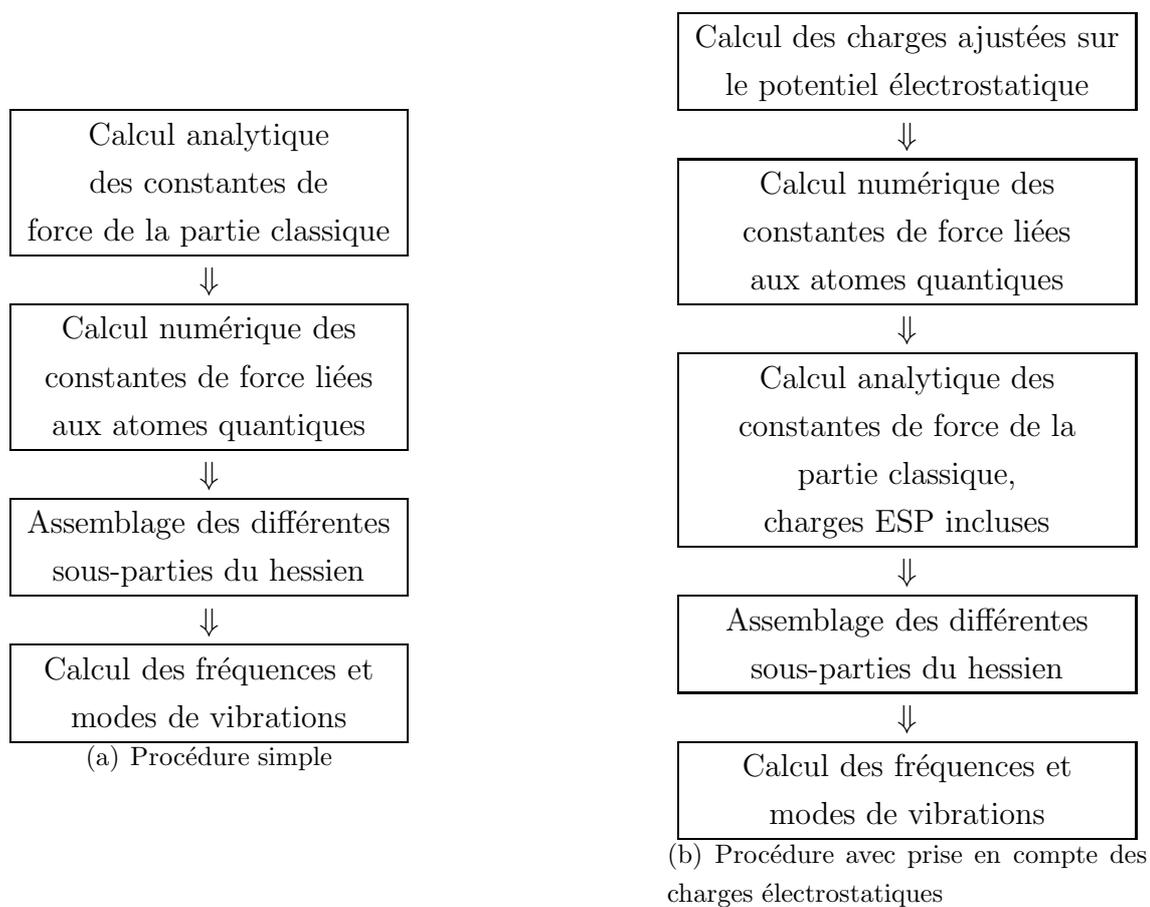


FIG. 1.3 – les deux procédures proposées pour le calcul par bloc du hessien

et le calcul sans prise en compte des charges est de  $2,18 \times 10^{-2}$  Ha.bohr<sup>-2</sup>. Cette différence est la même pour la procédure incluant les charges dans le calcul analytique des éléments du hessien associés aux atomes classiques.

Lorsqu'on compare les différentes fréquences de vibrations (reportées pour les trois approches en annexe : tableaux B.1, B.2 et B.3), la moyenne de la déviation absolue est de  $1,29 \text{ cm}^{-1}$  dans les deux cas, ce qui indique une bonne reproduction de la matrice hessienne. Si on s'intéresse plus particulièrement au mode dont la fréquence a une valeur imaginaire et portant le vecteur de transition, on remarque que la fréquence associée est bien reproduite puisque la référence donne  $747,1344 \text{ i cm}^{-1}$ . La procédure sans prise en compte des charges donne  $747,1315 \text{ i cm}^{-1}$  tandis que lorsque celles-ci sont incluses, on obtient  $747,2526 \text{ i cm}^{-1}$ .

La représentation dans l'espace des degrés de liberté cartésiens du vecteur propre associé au mode en question est identique dans les trois calculs (tableau B.4). Il faut en outre remarquer que les éléments de ce dernier sont principalement des degrés de liberté de la partie quantique. Le bloc du hessien correspondant à cette dernière étant calculé sans approximation par rapport à la référence, il est donc logique que le mode imaginaire soit bien reproduit.

Enfin, en ce qui concerne le temps de calcul, le facteur limitant étant le calcul numérique, on devrait s'approcher d'un rapport, entre un calcul numérique complet et un calcul approché, proche de celui entre le nombre d'atomes quantiques et le nombre total d'atomes. Ici, le système en compte quarante-huit dont vingt sont quantiques, le rapport idéal devrait donc être de  $\frac{20}{48} = 0,417$ . Or on constate que ce rapport est de 0,510 dans les deux cas.

L'exemple que nous avons pris ici montre que notre méthode approchée permet de caractériser avec un niveau d'approximation très raisonnable un état de transition à un moindre coût calculatoire. Cependant, les deux approches (avec ou sans prise en compte de charges dans le calcul analytique du sous bloc associé à la partie classique) conduisent à un résultat équivalent. L'incorporation des charges ne semble pas améliorer le résultat dans le cas présent.

Afin d'étudier l'effet de cette prise en compte, nous proposons de nous intéresser dans la partie suivante à un système pour lequel les interactions électrostatiques sont primordiales : le dimère de l'eau.

### 1.5.4 Applications au dimère QM/MM d'eau

Le dimère de l'eau que nous étudions (figure 1.4) est décrit au niveau MP2/6-31G\* pour la partie quantique et TIP3P [68] pour la partie classique. La molécule classique est donneuse de proton.

Pour chacune des méthodes, on calcule la différence entre chaque élément  $\mathcal{H}_{ij}$  avec son homologue dans le hessien obtenu par calcul numérique complet. Cette différence est représentée dans la figure 1.5 à l'aide d'une échelle de couleurs.

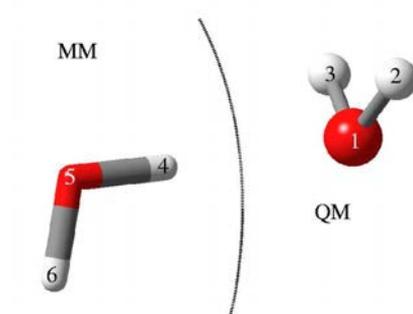


FIG. 1.4 – Dimère, partition et numérotation des atomes utilisés

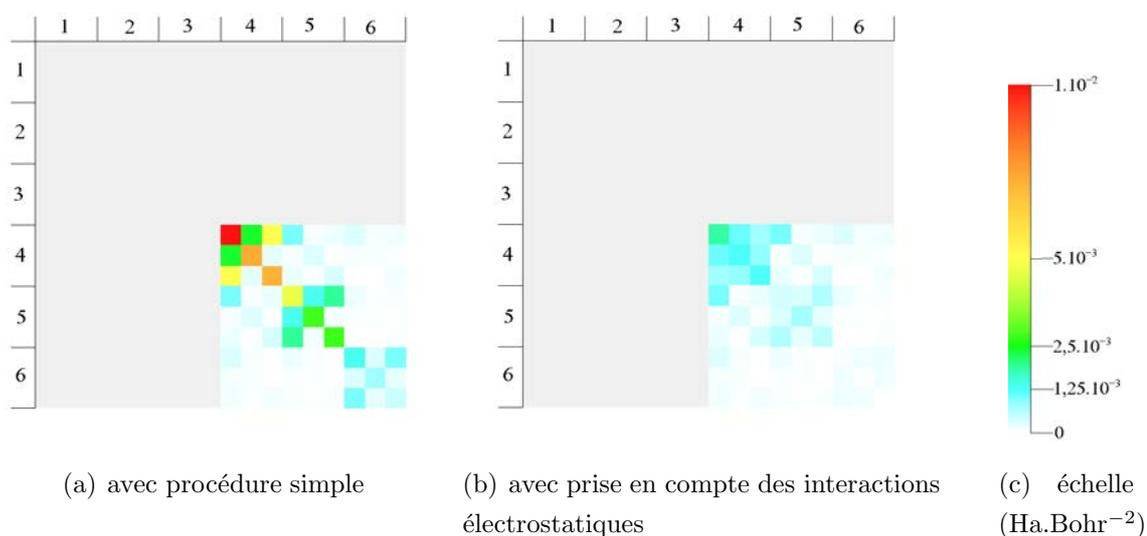


FIG. 1.5 – différences de valeurs pour les éléments du hessien du dimère d'eau pour les deux approches, la numérotation des éléments correspond à celle de la figure 1.4. Les zones grisées sont calculées numériquement

D'après cette représentation, on peut remarquer que l'inclusion de charges classiques ajustées sur le potentiel électrostatique permet de diminuer la différence avec la référence. En effet, la différence maximale constatée entre l'approche numérique complète et l'approche simple est de  $0,99 \times 10^{-1}$  Ha.bohr<sup>-2</sup> alors qu'avec les charges, cette même erreur passe à moins de  $2,00 \times 10^{-3}$  Ha.bohr<sup>-2</sup>.

Dans le même ordre d'idée, nous présentons dans le tableau 1.2 les fréquences de vibrations obtenues avec les deux approches et comparées à celles du calcul numérique complet.

modes	ref. (cm <sup>-1</sup> )	sans charges (cm <sup>-1</sup> )	avec charges (cm <sup>-1</sup> )
1	147,55	104,36 <i>i</i>	134,70
2	223,16	154,69	205,82
3	243,87	162,49	227,78
4	281,38	272,87	280,33
5	510,17	471,78	499,09
6	761,83	635,71	733,06
7	1739,30	1739,30	1739,29
8	2361,75	2346,64	2360,80
9	3660,44	3686,12	3656,09
10	3723,03	3734,87	3723,39
11	3781,45	3781,80	3781,44
12	3913,49	3913,49	3913,49

TAB. 1.2 – Fréquences des différents modes de vibration suivant la méthode utilisée

Dans ce tableau les six premiers modes correspondent à des vibrations impliquant les deux molécules sans déformation interne tandis que les six suivants sont des modes internes de chaque molécule.

Dans les deux cas, les fréquences de vibrations les plus éloignées de la référence sont celle attachées aux modes "inter-moléculaires". Les modes internes sont reproduits dans les deux cas avec une très bonne précision. Ceci est cohérent avec une approximation faite sur les interactions non liées. On remarque par ailleurs que les fréquences sont reproduites avec une erreur maximale (sans compter la fréquence imaginaire) de près de 80 cm<sup>-1</sup> (mode 3) sans prise en compte des charges tandis que dans le cas contraire,

on descend à moins de  $30 \text{ cm}^{-1}$  (mode 6). Cet exemple montre donc que la prise en compte des charges ponctuelles dans le calcul analytique des constantes de force appartenant au sous-bloc MM (figure 1.2) permet de se rapprocher sensiblement des valeurs obtenues par un calcul numérique complet.

Nous disposons donc ici d'un outil efficace permettant de calculer plus rapidement les constantes de force d'un système. Cette approche nous permet de disposer d'un hessien de départ de bonne qualité dans la recherche d'états de transition, ce qui en facilitera l'aboutissement. En outre, la caractérisation de ces derniers peut se faire avec une bonne précision. En effet, dans le cas d'une structure de transition, la coordonnée de réaction est construite essentiellement avec les degrés de liberté des atomes proches du centre réactif, décrits par la mécanique quantique. Or nous avons pu voir que la partie du hessien ainsi que les modes de vibration associés à cette partie étaient de très bonne qualité dans les deux cas. Cependant, lorsque les interactions électrostatiques sont importantes entre les sous-systèmes quantique et classique, l'ajout, dans le calcul analytique des constantes de force du sous-bloc MM, de charges ajustées sur le potentiel électrostatique de la fonction d'onde améliore sensiblement le hessien obtenu.



## Chapitre 2

# LSCF/MM/SCRF : définition et application

Dans ce chapitre, nous détaillons le couplage de la méthode hybride LSCF/MM avec le modèle de solvant implicite autocohérent SCRF (*Self Consistent Reaction Field*). Ayant précédemment défini la méthode hybride, nous commencerons donc par donner quelques repères méthodologiques concernant les modèles de solvant implicite. Nous détaillerons plus particulièrement celui développé à Nancy et avec lequel nous avons travaillé. Nous présenterons ensuite les résultats obtenus avec notre nouvelle approche dans l'étude des effets de solvant sur la réaction de Diels-Alder entre l'acrylate de (-)-menthyle et le cyclopentadiène. Les travaux présentés dans ce chapitre ont par ailleurs fait l'objet d'une publication[69].

## 2.1 Le modèle implicite de solvant

### 2.1.1 Généralités

Comme nous l'avons évoqué dans le chapitre précédent, les méthodes quantiques de la chimie théorique ne permettent encore aujourd'hui que de modéliser des systèmes de taille réduite. L'étude de la réactivité d'un système se fait donc souvent en ne considérant que les espèces directement impliquées et pas l'environnement (dans la majorité des cas un solvant) qui peut, dans certains cas, avoir une influence déterminante sur le chemin réactionnel suivi. La prise en compte du solvant est donc un domaine de la modélisation qui a fait l'objet d'importantes recherches depuis les travaux de Rivaill *et al.* en 1973 [70]. Les différents modèles développés sont basés sur les travaux de Born [71], Kirkwood [72] et Onsager [73] et ont pour caractéristique commune de traiter le solvant comme un milieu continu polarisable, supposé homogène et isotrope (qu'on désignera par continuum) et caractérisé par sa constante diélectrique relative statique notée  $\epsilon_0$ . On crée ensuite au sein du continuum une cavité dans laquelle on insère le soluté quantique, comme l'illustre la figure 2.1. La distribution des charges électrostatiques de la molécule de soluté polarise le continuum qui, en réponse, crée un champ électrique autour du système (champ de réaction) et en modifie alors la fonction d'onde, qui elle-même modifiera ensuite la polarisation du solvant. Ce processus auto-cohérent donne son nom à ce type de méthode, qu'on désigne par les initiales SCRF pour *Self Consistent Reaction Field*.

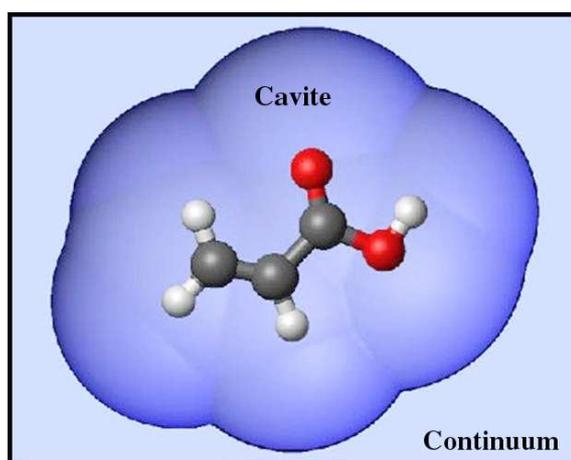


FIG. 2.1 – Modèle de continuum

Les différences entre les approches disponibles viennent principalement de trois points : d'abord la forme de la cavité, ensuite, l'algorithme utilisé pour le calcul du champ de réaction et enfin la détermination des contributions non électrostatiques, grandeur nécessaire au calcul de l'énergie libre de solvatation.

Parmi les différents modèles nous citons ceux qui, actuellement, sont les plus utilisés. Nous commençons donc par le modèle à cavité de Pise, PCM (*Pisa Cavity Model*) développé par Tomasi *et al.* [74, 75, 76] qui représente la réponse du continuum (le champ de réaction) à l'aide d'un ensemble de densités de charge distribuées sur la surface de la cavité.

L'approche développée par Cramer et Truhlar [77],[78] appelée de façon générale SMx et dont le dernier développement est désigné par SM5.43R[79] utilise l'approximation de Born généralisée (GB) avec l'algorithme de Still *et al.* [80]. Dans cette approche, le terme électrostatique est calculé à partir des charges partielles atomiques du soluté et les effets non électrostatiques, grâce à des tensions superficielles atomiques.

Nous citerons enfin aussi pour mémoire le modèle Cosmo [81] développé par Klamt *et al.* dont la principale caractéristique est d'utiliser un milieu continu dont la constante diélectrique est infinie.

Avant de passer à la description du modèle développé à Nancy, nous souhaitons de nouveau mentionner l'article de revue écrit par Cramer et Truhlar [78] et celui, plus récent, écrit par Tomasi [82] qui pourront fournir une vue générale de ces méthodes.

### 2.1.2 Le modèle à cavité de Nancy

L'approche développée à Nancy [70, 83, 84, 85, 86, 87, 88], qui nous intéresse plus particulièrement, repose sur les travaux de Kirkwood [89]. Il a montré que l'énergie libre d'interaction soluté-solvant peut être exprimée en un point comme un développement de Taylor du potentiel électrostatique créé par la distribution de charge du soluté au sein du continuum.

Dans notre cas, la distribution des charges du soluté est décrite comme un développement multipolaire qui utilise des moments sphériques, notés  $M_l^m$ , désignant ainsi un multipôle d'ordre  $l$  avec  $m$  composantes telles que  $-l \leq m \leq +l$ . L'interaction entre la molécule de soluté et le champ de réaction est décrite par l'opérateur que l'on notera  $H'$  tel que :

$$H' = - \sum_{l=0}^{\infty} \sum_{m=-l}^l R_l^m M_l^m \quad (2.1)$$

avec

$$R_l^m = \sum_{l'=0}^{\infty} \sum_{m'=-l'}^{l'} f_{ll'}^{mm'} M_{l'}^{m'} \quad (2.2)$$

Où  $f_{ll'}^{mm'}$  est le facteur du champ de réaction qui ne dépend que de la forme de la cavité et de la constante diélectrique  $\varepsilon$ , si on considère que la réponse de l'environnement (le continuum) est linéaire.

Le champ de réaction étant considéré comme une perturbation de l'Hamiltonien  $H_0$  de la molécule isolée, le nouvel Hamiltonien s'écrit :

$$H_{tot.} = H_0 + H' \quad (2.3)$$

La variation d'énergie libre électrostatique s'écrit :

$$\Delta G_{el} = -\frac{1}{2} \sum_{l=0}^{\infty} \sum_{m=-l}^l \langle R_l^m \rangle \langle M_l^m \rangle \quad (2.4)$$

Le facteur  $\frac{1}{2}$  est dû au fait que l'énergie nécessaire à la polarisation du milieu vaut la moitié de l'énergie d'interaction.

Le terme d'énergie électrostatique est alors inclus dans l'opérateur de Fock, dont on écrit ici l'élément de la matrice associée,  $\phi_\mu$  et  $\phi_\nu$  étant deux fonctions de base :

$$F_{\mu\nu} = F_{\mu\nu}^0 + \sum_{l=0}^{\infty} \sum_{m=-l}^l R_l^m \langle \phi_\mu | M_l^m | \phi_\nu \rangle \quad (2.5)$$

avec  $F_{\mu\nu}^0$  l'élément de la matrice de Fock associée à l'opérateur non perturbé.

L'énergie électrostatique est associée à deux autres termes pour constituer l'énergie libre totale d'interaction soluté-solvant :

$$\Delta G_{tot.} = \Delta G_{el} + \Delta G_{cav} + \Delta G_{disp} \quad (2.6)$$

Avec  $\Delta G_{cav}$  l'énergie libre de cavitation qui peut se concevoir comme la variation d'énergie libre nécessaire à la création de la cavité au sein du continuum et  $\Delta G_{disp}$  celle de dispersion.

Il est enfin important de remarquer que la définition du volume et de la forme de la cavité joue aussi un rôle important dans ce type d'approche. Plusieurs choix peuvent être considérés. Outre l'utilisation d'une cavité sphérique, on peut aussi envisager des formes ellipsoïdales[84].

Le volume de la cavité est déterminé à partir des sphères de van der Waals dont les rayons atomiques sont multipliés par un facteur d'échelle ajusté pour une série de molécules afin de se rapprocher du volume molaire expérimental.

Actuellement, on utilise généralement une cavité dont la forme est adaptée à la surface de la molécule [85, 87, 88], comme la surface de Richards[90]. Cette dernière est déterminée en faisant rouler une sphère dont le rayon est lié à la nature du solvant (rayon de sonde) sur la surface de van der Waals de la molécule. Ceci est fait dans le but d'éliminer les singularités liées aux intersections des sphères associées aux atomes.

Nous finirons enfin cette section en évoquant quelques unes des possibilités offertes par ce modèle. Outre le fait de pouvoir choisir entre différentes formes de cavités, on peut en contrôler le volume en modifiant le facteur d'échelle sur les rayons de van der Waals des différents atomes. Ceci permet, entre autres, de traiter des systèmes dont une partie de la densité électronique peut se situer à l'extérieur de la cavité, comme dans le cas d'espèces anioniques, par exemple. L'effet principal d'une augmentation de la taille de la cavité est la diminution de l'interaction soluté/continuum.

L'utilisateur peut encore préciser l'ordre auquel sera limité le développement multipolaire, permettant ainsi d'arrêter le développement à l'ordre auquel il converge. Un développement jusqu'à l'ordre six est généralement suffisant. Le développement multipolaire peut par ailleurs être fait au centre de la cavité ou en plusieurs centres (généralement les positions atomiques), permettant ainsi une meilleure description du champ électrostatique créé par le soluté. Ceci permet généralement une meilleure convergence du développement multipolaire.

Enfin, cette méthode a pour avantage de permettre le calcul analytique des dérivées de l'énergie jusqu'à l'ordre deux [91, 92], ce qui rend l'optimisation de géométrie et la localisation de structures de transition aisées.

C'est donc cette méthode que nous avons choisi de coupler avec l'approche hybride LSCF/MM afin de pouvoir modéliser les effets du solvant sur un système de grande taille. Il est par ailleurs important de noter qu'une approche couplant les méthodes ONIOM [51] et PCM [74] a été proposée par Vreven *et al.* [93] dans le même but.

## 2.2 Définition de la méthode LSCF/MM/SCRF

Notre premier travail a été de coupler la représentation du solvant avec le formalisme QM/MM pour conduire à la méthode à trois niveaux LSCF/MM/SCRF. Cette dernière

a été implantée dans le logiciel Gaussian98 [1]. Quelques précisions concernant ce travail pourront être trouvés dans l'annexe A. Le couplage entre les deux méthodes est assez évident mais nous devons cependant spécifier quelques points.

Tout d'abord, nous choisissons d'inclure dans la cavité l'ensemble du système QM/MM, de telle façon que le continuum soit polarisé à la fois par le sous-système quantique mais aussi par les charges classiques. La géométrie du système complet subira alors l'effet du champ de réaction. Bien entendu, seule la partie quantique sera polarisée, le potentiel du champ de force n'étant représenté que par des charges ponctuelles.

Ensuite, du fait de la présence simultanée de la partie classique et du continuum, l'opérateur est modifié comme précisé par les équations 1.35 et 2.5. L'élément  $(\mu, \nu)$  de l'opérateur s'exprime alors dans la base d'orbitales atomiques  $(\mu, \nu)$  :

$$F_{\mu\nu} = F_{\mu\nu}^0 + \sum_i^{MM} \langle \mu | \frac{q_i}{r_i} | \nu \rangle + \sum_{l=0} \sum_m R_m^l \langle \mu | \hat{M}_{l,e}^m | \nu \rangle \quad (2.7)$$

avec  $F_{\mu\nu}^0$  est l'opérateur associé à la fonction d'onde non perturbée et  $M_{l,e}^m$  l'opérateur du champ de réaction associé à la partie électronique.

Enfin, du fait de la polarisation du continuum par les charges classiques, l'opérateur multipolaire, entrant dans le calcul de  $R_l^m$  est modifié tel que :

$$M_{l,tot}^m = \sum_A^{noyaux} Z_A \hat{M}_{l,A}^m + \sum_{\mu,\nu} P_{\mu\nu}^T \langle \phi_\mu | \hat{M}_{l,e}^m | \phi_\nu \rangle + \sum_C^{MM} Q_c \hat{M}_{l,C}^m \quad (2.8)$$

Le premier terme étant lié aux noyaux indicés  $A$ , le deuxième concernant la partie électronique qui prend en compte la matrice densité totale  $P_{\mu\nu}^T$ , incluant aussi les effets dûs aux SLBOs. Quant au troisième terme, il concerne les charges classiques indicées  $C$ .

Cette méthode offre l'ensemble des possibilités permises par les deux méthodes ainsi couplées. De même, les dérivées premières analytiques de l'énergie en fonction des coordonnées nucléaires étant utilisables, il sera possible d'optimiser la géométrie du système LSCF/MM en solution.

## 2.3 Effets du solvant dans la réaction de Diels-Alder

Afin d'illustrer les possibilités de la méthode ainsi décrite, nous avons choisi de l'appliquer à l'étude des effets du solvant sur la sélectivité et le mécanisme dans la réaction de Diels-Alder de l'acrylate de (-)-menthyle avec le cyclopentadiène (figure

2.2). Notre choix s'est porté sur cette réaction particulière du fait de la présence des deux facteurs que permet de prendre en compte notre méthode à trois niveaux. Il s'agit de l'influence du solvant mais aussi de la présence d'un auxiliaire chiral dont l'effet principalement stérique autorise à le traiter au niveau classique. De plus, cette réaction a déjà fait l'objet de différentes études à la fois expérimentales et théoriques [94],[95],[96]. Cela nous permettra donc de valider nos résultats en les comparant à ceux de la littérature mais aussi de compléter ces derniers.

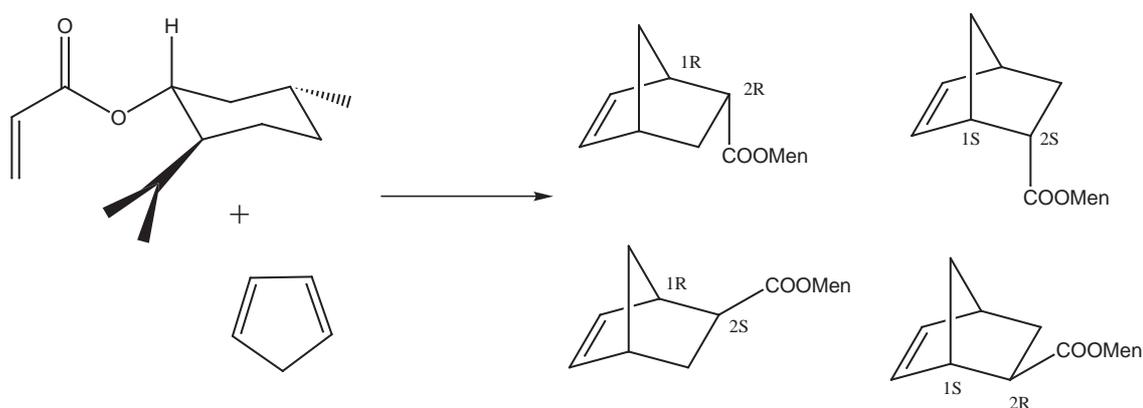


FIG. 2.2 – réaction de l'acrylate de (-)-menthyle avec le cyclopentadiène et produits formés

### 2.3.1 Le système réactif

La réaction que nous avons étudiée conduit à quatre produits différents qu'on note (de gauche à droite et de bas en haut sur la figure 2.2), en fonction de la face d'attaque et de la configuration des carbones chiraux : endo-1R-2R, endo-1S-2S, exo-1R-2S et exo-1S-2R. Il faut noter que chaque produit possède plusieurs conformations. Le nombre de structures de transitions (et donc de chemins réactionnels) possibles est assez important (tableau 2.1). En effet, outre les approches ENDO ou EXO entre les réactifs et le choix entre faces SI et RE de l'acrylate pour l'attaque, on a le choix entre les configurations S-CIS et S-TRANS pour la partie réactive. De plus, le groupement menthyle peut adopter deux conformations stables par rapport au groupement acrylate. Pour les définir, on prend en compte l'angle dièdre H-C-O-O (sur la figure 2.2, il implique dans l'ordre le seul atome d'hydrogène représenté, le carbone qui le porte, l'atome d'oxygène qui y est lié et l'oxygène carbonyle). Dans la première conformation, notée Syn, l'angle dièdre

est d'une trentaine de degrés tandis que dans la forme Anti, celui ci est d'environ cent quatre-vingts degrés.

Ces différentes possibilités sont illustrées dans le tableau 2.1. On considérera donc dans la suite seize ( $2^4$ ) structures de transitions.

### 2.3.2 Rappels des résultats existants

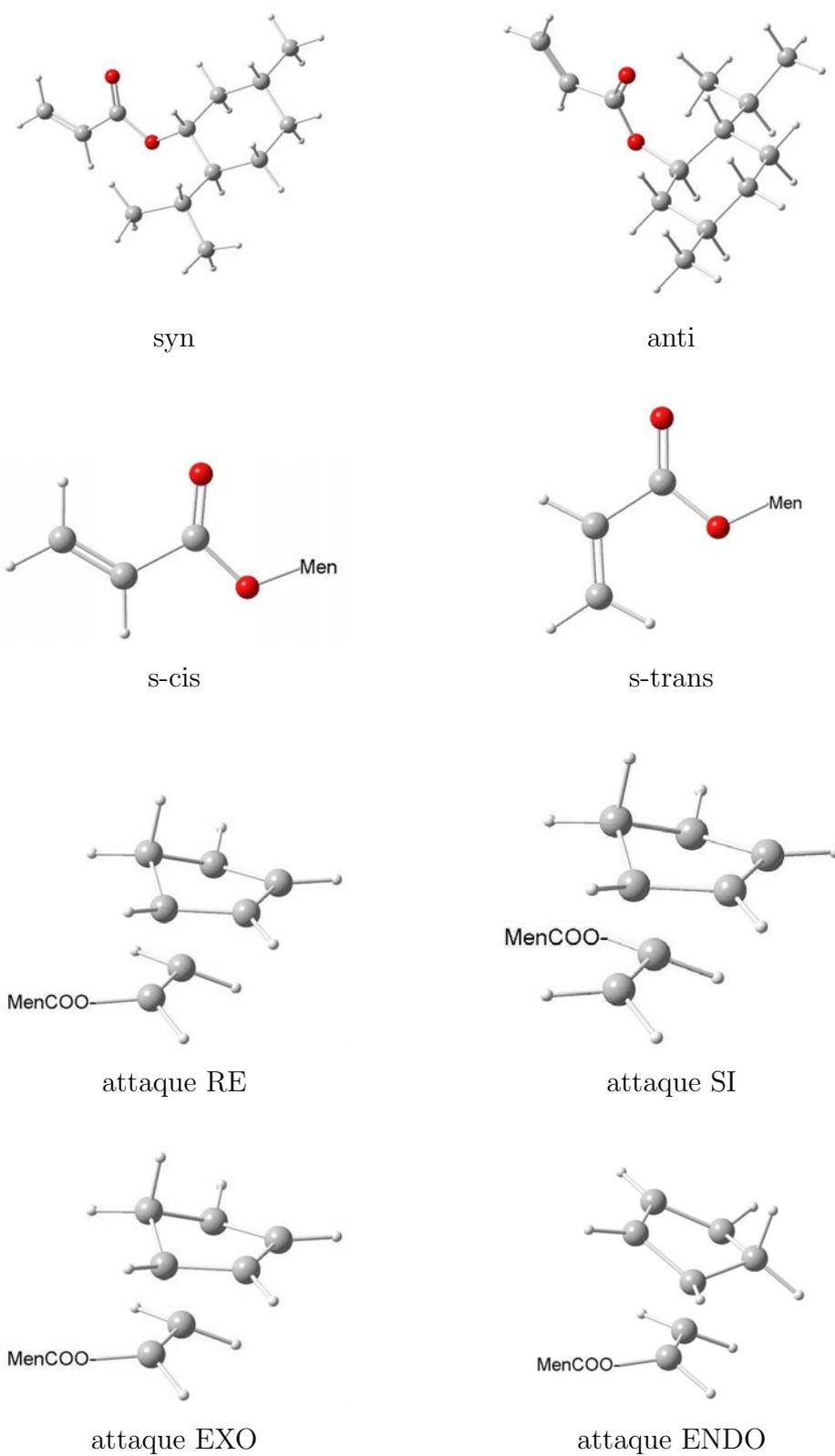
Dans cette partie, nous rappelons brièvement les conclusions tirées d'études précédentes. Pour cela nous nous appuyons sur trois points. D'abord, nous considérons les résultats expérimentaux obtenus par Cativiela *et al.* [94] mais aussi des études théoriques menées dans le même temps sur un système modèle d'acrylate de méthyle, prenant en compte les effets du solvant par le biais du modèle de continuum de Nancy [95]. Dans un troisième temps, nous considérons aussi l'étude théorique de notre réaction avec un modèle hybride semi-empirique/classique dont le but est la prise en compte de l'auxiliaire chirale. Cependant, cette approche ne rend pas compte des effets du solvant sur la réactivité [96].

Les deux études théoriques évoquées ont permis l'étude soit des effets du solvant, soit de l'influence du groupement (-)-menthyle sur la sélectivité et le mécanisme réactionnel. Notre méthode permet par contre l'étude simultanée des deux facteurs.

Du point de vue expérimental [94], la réaction a été étudiée dans plusieurs solvants de polarités différentes. Dans tous les cas, ce sont les produits issus d'une attaque ENDO qui sont majoritaires et parmi les deux produits possibles, c'est celui désigné 1R-2R (figure 2.2) qui est formé préférentiellement, l'excès diastereofacial étant corrélé avec la polarité du solvant.

Dans l'étude théorique des effets du solvant (modélisé comme étant l'eau) et prenant en compte l'acrylate de méthyle, quatre états de transition ont été étudiés. Trois conclusions principales se dégagent de cette étude. D'une part, le solvant augmente l'asynchronicité de la réaction. On peut quantifier cette dernière comme la différence entre les longueurs de liaisons carbone-carbone en formation dans la structure de transition. Ce phénomène est dû à une polarisation du centre réactif du diénophile. D'autre part, l'excès diastereomérique est systématiquement en faveur des structures ENDO par rapport aux formes EXO. Enfin, les états de transition s-trans sont plus stabilisés que les s-cis par la présence du solvant. Ceci corrobore le fait que le produit ENDO est favorisé par le solvant.

L'approche hybride [96] qui prenait aussi en compte la présence d'un catalyseur



TAB. 2.1 – Les différents conformations conduisant aux seize structures de transition

Al-Cl<sub>3</sub> a quant à elle mis en évidence la possibilité qu'a le groupement isopropyle de tourner suivant l'axe de la liaison carbone-carbone le reliant au cycle dans le menthyle. Cette propriété expliquant le faible effet stérique du groupement menthyle, même pour des chemins réactionnels pour lesquels l'approche du cyclopentadiène se fait par la face encombrée.

### 2.3.3 Détails calculatoires

#### Partition quantique/classique

Dans notre étude, nous avons traité le cyclopentadiène à l'aide de la mécanique quantique uniquement tandis que l'acrylate de menthyle est partitionné comme indiqué par la figure 2.3, où Y est l'atome frontière comme nous l'avons défini plus haut et qui remplace un atome de carbone. Le choix de cette partition est assez évident, en effet, une coupure plus proche du centre réactif risquerait d'influencer les propriétés de ce dernier. Ceci est d'autant plus vrai que dans notre cas, la partie quantique inclut les atomes d'oxygène dont l'effet sur la double liaison  $\pi$  réactive ne doit pas être négligé. D'un autre côté, éloigner la frontière du centre réactif conduirait à couper le cycle du substituant menthyle, ce qui n'est pas souhaitable non plus puisque cela nécessiterait l'introduction d'une seconde SLBO. La coupure entre l'acrylate et le menthyle semble donc la plus logique, tant du point de vue chimique que technique.

#### Détermination de l'orbitale frontière

Comme il a été indiqué dans le chapitre précédent, la détermination de l'orbitale frontière a été menée sur la plus petite molécule comportant la liaison à couper, dans notre cas, le méthanol. Pour cela nous avons choisi de prendre une molécule dont la géométrie reprend celle de la liaison et de son environnement dans le réactif concerné dans sa forme la plus stable (acrylate de (-)-menthyle s-trans syn). Les atomes entourant la liaison carbone-oxygène sont remplacés par des hydrogènes et seules les distances entre atomes lourds et hydrogènes sont optimisées, ceci afin de déterminer la liaison frontière sur une molécule modèle qui reproduit au mieux celle présente dans le réactif optimisé au niveau quantique. Le critère de localisation est celui de Weinstein-Pauncz [62] et le troisième atome, nécessaire à la définition du trièdre de rotation de la SLBO lors des optimisations de géométrie est défini comme l'atome de carbone du groupement carbonyle de l'acrylate.

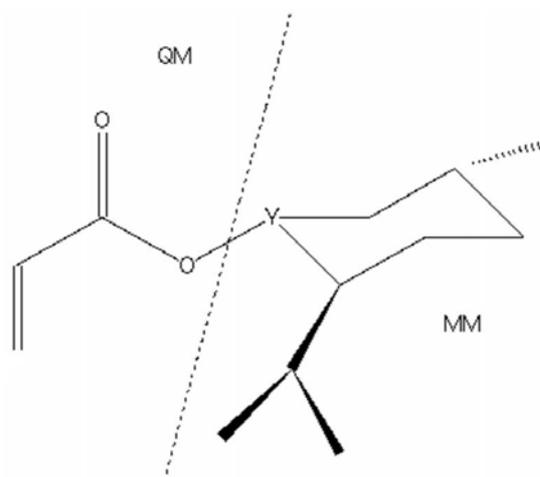


FIG. 2.3 – Partition de l'acrylate de menthyle en sous-systèmes classique et quantique

### Choix de la base et du champs de force

Le niveau de théorie que nous employons est RHF/6-31G\*, il a en effet été montré [95] que les énergies relatives auxquelles nous nous intéressons sont très peu sensibles à la corrélation électronique et à la taille de la base, à condition que cette dernière soit au moins de qualité double- $\zeta$  polarisée.

Concernant le champ de force classique nous nous sommes tournés vers Dreiding [22] car tous les paramètres liés et non électrostatiques nécessaires ici sont disponibles avec ce dernier.

Les charges classiques du menthyle ont été déterminées dans le but de reproduire au mieux la polarisation de la partie réactive. Pour cela, nous avons testé différents jeux de charges classiques déterminées à l'aide de plusieurs méthodes (Merz-Kollman[67], Chelp [97], ChelpG[98] et Opep[99]) permettant de reproduire le potentiel électrostatique de l'acrylate de menthyle s-trans syn au niveau RHF/6-31G\*.

Afin de choisir entre les différents résultats, nous comparons les charges de Mulliken [100] (valeurs reportées en annexe, tableau C.1) obtenues avec les différents jeux de charges ainsi que les énergies des orbitales frontières aux valeurs correspondantes dans la référence quantique (tableau 2.2)

D'après ces résultats, on peut voir que les énergies des orbitales frontières sont

	QM	MK	MK×1.1	Chelp	ChelpG	OPEP
$\varepsilon_{HOMO}$	-0.3913	-0.3830	-0.3829	-0.3938	-0.3879	-0.3729
$\varepsilon_{LUMO}$	0.1172	0.1137	0.1166	0.0265	0.0839	0.1405

TAB. 2.2 – Energies des orbitales frontières du diénophile en fonction du jeu de charges classiques utilisé (tableau C.1, annexe C)

assez bien reproduites bien que ces dernières soient assez sensibles au jeu de charges utilisé. On note alors que les charges de type Merz-Kollman sont celles qui donnent le meilleur résultat. Celui-ci étant légèrement amélioré si les charges sont multipliées par un facteur 1,1. Par contre, les charges de Mulliken sont moins sensibles au choix du jeu de charges classiques. Pour la suite de notre étude nous choisissons donc d'utiliser les charges de Merz-Kollman multipliées par un facteur 1,1. Par ailleurs, la charge de l'atome frontière est ajustée pour garder la neutralité du sous-système classique.

### Modélisation du solvant

Les paramètres utilisés pour le modèle de continuum sont assez standards. La constante diélectrique du continuum est celle de l'eau ( $\varepsilon = 78,4$ ), nous souhaitons en effet modéliser la réaction dans un solvant aqueux. On utilise une cavité dont la forme est basée sur celle de la molécule, le volume en étant déterminé à l'aide des rayons de van der Waals multipliés par 2,0 au lieu de 1,3084 habituellement, ceci pour assurer une convergence plus facile de la fonction d'onde. Les valeurs de l'énergie libre de solvation ainsi calculées seront d'amplitude moins importante qu'avec une cavité plus petite mais ne devraient pas modifier le raisonnement qui porte sur des différences d'énergies.

Les développements multipolaires multicentriques sont faits jusqu'à l'ordre six, valeur suffisante pour converger l'énergie SCRF. Si on prend l'exemple de l'état de transition correspondant à une attaque du cyclopentadiène sur la face RE du conformère le plus stable de l'acrylate de menthyle s-cis, les contributions, aux ordres allant de zéro à six sont respectivement : 0,000, -0,416, -0,429, -0,273, -0,061, 0,033 et 0,001 kcal.mol<sup>-1</sup>.

Il est enfin important de préciser que notre modèle de solvant ne rend compte des interactions spécifiques que sont les liaisons hydrogène que de façon moyenne. Ces dernières peuvent influencer la réactivité d'espèces analogues à celles que nous étudions [101, 102, 103, 104], particulièrement en ayant pour effet d'augmenter la constante de

réaction.

La liaison hydrogène fait intervenir une interaction électrostatique mais aussi un transfert de charge entre le soluté et une ou plusieurs molécules de solvant. Le modèle de continuum que nous utilisons ici décrit de façon efficace la première composante, mais pas complètement la seconde. Dans notre cas, la partie réactive peut être influencée par la formation d'une liaison H avec la fonction carbonyle de l'acrylate, que nous considérons comme accessible à des molécules de solvant suffisamment petites (l'eau), quelque soit le système. On peut donc considérer que la différence entre les énergies d'interaction de liaisons H pour les différents états de transition sera moins importante que l'énergie de polarisation. Jorgensen *et al.* [101] ont montré que les différences entre énergies de liaisons hydrogène étaient inférieures à la kcal.mol<sup>-1</sup>, ce qui correspond principalement à la différence de la partie électrostatique. Si nous considérons que les interactions spécifiques de liaison hydrogène sont similaires dans toutes les structures de transitions envisagées et comme nous sommes intéressés par des différences énergétiques entre ces structures, nous pouvons donc ignorer cette composante.

Le logiciel AMSOL [105] a été utilisé pour mener des calculs semi-empiriques AM1 [106] sur les structures de transition avec le modèle de solvant SM2 [107] afin de connaître les contributions non électrostatiques à l'énergie libre de solvation (termes de cavitation, dispersion et spécifiques : CDS), les contributions liées à la liaison hydrogène étant déjà incluses dans cette méthode, au moins de façon implicite.

### Détermination des structures de transition

Nous avons procédé à la détermination des structures de transition avec et sans l'influence du solvant. Les dérivées secondes de l'énergie LSCF/MM en fonction des positions nucléaires ne pouvant pas encore être calculées de façon analytique, la localisation des états de transition a été faite en prenant en compte les constantes de forces obtenues par un calcul quantique complet au niveau RHF/3-21G. Un calcul numérique des constantes de force a ensuite permis de confirmer que les structures obtenues étaient bien des états de transition.

#### 2.3.4 Effets du solvant sur la sélectivité

Les résultats que nous avons ainsi obtenus vont donc nous permettre de nous intéresser à plusieurs facteurs que sont la sélectivité, l'influence de l'auxiliaire chiral et du solvant sur celle-ci mais aussi sur le mécanisme réactionnel en lui même.

Afin de ne pas surcharger la discussion, nous avons choisi de reporter les grandeurs qui n'entrent pas directement dans le raisonnement dans l'annexe C où figurent donc, entre autres, les énergies absolues des espèces étudiées mais aussi certains de leurs paramètres géométriques.

Afin de déterminer l'influence des différents facteurs évoqués ci-dessus sur la sélectivité, nous nous intéressons plus particulièrement aux différences énergétiques entre structures. Le tableau 2.3 contient les énergies des seize états de transition, relativement au plus stable d'entre eux, pour les calculs sans et avec l'influence du solvant. Dans ce cas les valeurs comprenant uniquement l'interaction électrostatique mais aussi celles incluant les termes CDS sont reportées,

Enfin, le produit obtenu pour chaque structure de transition est reporté dans la dernière colonne avec la dénomination associée aux carbones chiraux telle qu'elle a été donnée en association avec la figure 2.2.

On rappelle par ailleurs que, expérimentalement, le produit 1R-2R est obtenu de façon majoritaire, cette sélectivité augmentant avec la polarité du solvant [94]. Les chemins réactionnels conduisant à ce produit devraient donc être les plus favorables en terme d'énergie. Malgré cela, comme cela a déjà été constaté dans le cas de l'acrylate de méthyle [95], les états de transitions les plus stables ne sont pas ceux qui conduisent au produit majoritairement formé. Cette remarque nous indique donc que le niveau de théorie auquel nous travaillons est insuffisant pour donner des résultats fiables à quelques dixièmes de kilocalories par mole près. Néanmoins, nous pouvons nous intéresser aux effets de la présence ou non du solvant en comparant les données dans le continuum à celle correspondantes dans le vide.

La première remarque concerne les interactions non-électrostatiques (notées CDS). En effet on peut voir d'une part que, comparé aux valeurs de l'interaction électrostatique, le terme considéré est généralement inférieure. Ceci est cohérent avec l'assertion selon laquelle l'énergie de stabilisation dans le solvant est principalement due à l'interaction électrostatique. Cependant, comme nous travaillons avec une cavité dont le volume est agrandi (par le facteur multiplicatif appliqué aux rayons atomique de van der Waals pour déterminer le volume) par rapport au volume idéal, l'interaction électrostatique est diminuée. C'est pourquoi, il ne faut pas perdre de vue que l'interaction électrostatique est sous estimée dans nos calculs. Cette remarque va de nouveau dans le sens d'une prépondérance de la partie électrostatique dans l'interaction soluté-solvant.

D'un autre côté, les valeurs de cette contribution ne modifie que très peu l'ordre

Face d'attaque	TS	Vide	Solvant		$\Delta$	Produit
			Electrostat.	Elec.+CDS		
libre	endo s-cis RE syn	0,00	0,00	0,00	0,00	1S-2S
encombrée	endo s-cis RE anti	7,86	7,97	7,89	0,03	1S-2S
encombrée	endo s-cis SI syn	0,81	0,96	1,07	0,26	1R-2R
libre	endo s-cis SI anti	6,30	6,27	6,30	0,00	1R-2R
encombrée	endo s-trans RE syn	2,65	2,22	2,12	-0,53	1S-2S
libre	endo s-trans RE anti	8,28	7,92	8,01	-0,27	1S-2S
libre	endo s-trans SI syn	1,35	1,00	1,04	-0,31	1R-2R
encombrée	endo s-trans SI anti	2,08	1,87	1,88	-0,20	1R-2R
libre	exo s-cis RE syn	0,54	0,60	0,64	0,10	1R-2S
encombrée	exo s-cis RE anti	8,27	8,37	8,47	0,20	1R-2S
encombrée	exo s-cis SI syn	1,68	1,95	1,89	0,21	1S-2R
libre	exo s-cis SI anti	6,94	6,98	7,03	0,09	1S-2R
encombrée	exo s-trans RE syn	3,33	3,28	3,09	-0,24	1R-2S
libre	exo s-trans RE anti	8,33	7,96	7,93	-0,40	1R-2S
libre	exo s-trans SI syn	2,02	1,69	1,76	-0,26	1S-2R
encombrée	exo s-trans SI anti	10,08	9,63	9,68	-0,40	1S-2R

TAB. 2.3 – énergies relatives en kcal.mol<sup>-1</sup> des états de transition pour la phase gazeuse, pour le solvant avec uniquement les interactions électrostatique et comprenant les termes de type CDS. la valeur  $\Delta$  est la différence entre cette dernière valeur et celle correspondante dans le vide.

relatif des états de transition en terme d'énergies. Dans la suite, les raisonnements concernant la sélectivité peuvent être menés en ne considérant que les facteurs stériques et électrostatiques qui sont donnés par notre méthode.

Nous nous intéressons plus précisément à l'effet du solvant par le biais des différences entre les énergies relatives calculées dans le vide et dans le solvant pour une même structure. Cette donnée correspond à la valeur notée  $\Delta$  dans le tableau 2.3 si on veut prendre en compte les contributions spécifiques de dispersion et de cavitation. Le raisonnement reste cependant identique si on ne tient compte que de la partie électrostatique aussi présentée.

On remarque que, de façon systématique, les effets du solvant favorisent les formes *s-trans* par rapport aux structures *s-cis*. Ce constat rejoint les observations expérimentales mais aussi les résultats de l'étude menée sur l'acrylate de méthyle.

Cette propriété est la seule d'ordre général qui puisse être dégagée de nos résultats. En effet, si on s'intéresse maintenant à la sélectivité entre formes ENDO et EXO, les résultats ne sont pas aussi nets que dans l'étude sur l'acrylate de méthyle. Dans cette approche, le solvant a pour effet de stabiliser systématiquement les structures de forme ENDO par rapport aux conformères EXO. Dans notre cas, cette propriété n'est pas systématique. Pour nous rapprocher de l'étude précédente, nous commençons par considérer les espèces les plus stables et pour lesquelles l'attaque se fait par la face non encombrée (voir tableau 2.3). Dans ce cas, les formes ENDO sont effectivement plus stabilisées par le solvant que les formes EXO correspondantes. Lorsqu'on s'intéresse aux conformères pour lesquels l'encombrement stérique est le plus faible, les résultats rejoignent ceux obtenus sur un modèle plus simple (acrylate de méthyle) mais aussi aux observations expérimentales. Cependant, quand l'encombrement stérique est plus important, aucune tendance générale ne peut être décrite. Ce point montre que les effets couplés du solvant et de l'auxiliaire chiral ne sont pas simplement additifs, ce qui implique donc la nécessité de mener les calculs sur toutes les structures possibles, justifiant par ailleurs notre approche couplant le champ de force hybride avec le modèle de solvant.

### 2.3.5 Influence de l'encombrement stérique

Pour analyser l'influence de l'encombrement stérique, nous nous focalisons sur les différences d'énergies entre conformères se distinguant par un seul facteur à la fois.

Considérons donc d'abord l'effet proprement dit du groupement isopropyle présent

encombrée - libre	vide	solvant
endo s-cis syn	0,81	1,07
endo s-cis anti	1,56	1,59
endo s-trans syn	1,30	1,08
endo s-trans anti	-6,20	-6,13
exo s-cis syn	1,14	1,25
exo s-cis anti	1,33	1,44
exo s-trans syn	1,31	1,33
exo s-trans anti	1,75	1,75

TAB. 2.4 – différence d'énergie ( $\text{kcal.mol}^{-1}$ ) pour les structures de transition se différenciant par l'approche de la face prochirale (encombrée - libre)

sur le (-)-menthyle. Cette influence peut être étudiée en prenant en compte les différences d'énergies entre TSs ne se distinguant que par la face d'addition prochirale (tableau 2.4). On remarque que les différences d'énergie, que ce soit dans le vide ou avec prise en compte du solvant, est de l'ordre de une à deux  $\text{kcal.mol}^{-1}$  favorisant la face libre, ce qui va dans le sens de l'étude au niveau semi-empirique/classique[96]. Ceci est cependant faux pour les isomères endo s-trans anti où la différence est de plus de 6  $\text{kcal.mol}^{-1}$ .

Cette importante différence d'énergie (comparée aux autres valeurs) en faveur du conformère pour lequel l'addition se fait par la face encombrée est due à une propriété du groupement isopropyle qui avait déjà été mise en évidence [96]. En effet, ce dernier a la possibilité d'adopter différentes conformations désignées en adoptant la nomenclature proposée par Maddaluno [108] et modifiée par Salvatella *et al.* [96].(figure 2.4). Dans notre cas, comme dis précédemment, nous avons considéré seize états de transitions pour lesquels la conformation était "axiale".

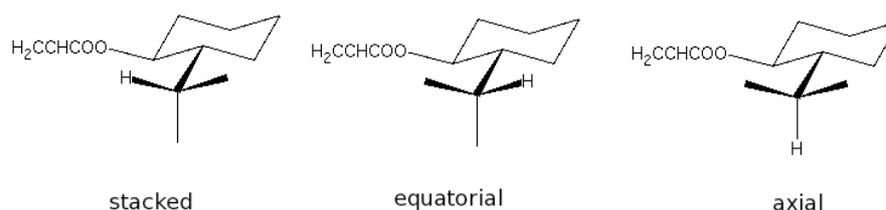


FIG. 2.4 – Conformations du groupement isopropyle

Ce choix a été fait en considérant la conformation la plus stable dans les réactifs. Cependant, dans le cas précis de la structure ENDO s-trans SI anti et uniquement dans ce cas, l'interaction stérique entre le groupement isopropyle et son environnement est telle qu'il adopte systématiquement une configuration "stacked" (figure 2.4). Toutes les tentatives d'optimisation de structures avec une conformation axiale ayant à chaque fois redonné la conformation "stacked".

On doit voir dans cette stabilisation une préférence pour la conformation "stacked" plutôt que pour l'approche par la face encombrée. N'ayant pas considéré explicitement ce phénomène de rotation dans notre étude, ce qui aura imposé l'étude d'un minimum de quarante-huit structures ( $16 \times 3$ ), nous choisissons donc de ne pas conclure quant à la stabilité de cette structure par rapport aux autres.

Lorsqu'on s'intéresse aux différences d'énergies entre conformères ne se différenciant que par l'attaque (ENDO ou EXO), on remarque que la différence d'énergie (inférieure à  $1 \text{ kcal.mol}^{-1}$ ) se fait toujours en faveur de la forme ENDO (tableau 2.5), ce qui est cohérent avec les résultats existants [95]. Cette préférence est à peu près équivalente que ce soit dans le vide ou dans le solvant. Dans notre cas, contrairement à ce qui a déjà été mentionné, ce dernier n'exacerbe pas cette sélectivité. Encore une fois, la rotation de l'isopropyle induit une forte stabilisation du TS ENDO s-trans SI anti par rapport à son homologue EXO.

ENDO-EXO	vide	solvant
s-cis RE syn	0,54	0,64
s-cis RE anti	0,41	0,58
s-cis SI syn	0,87	0,82
s-cis SI anti	0,64	0,73
s-trans RE syn	0,68	0,97
s-trans RE anti	0,05	-0,08
s-trans SI syn	0,67	0,72
s-trans SI anti	8,00	7,80

TAB. 2.5 – différence d'énergie ( $\text{kcal.mol}^{-1}$ ) pour les structures de transition se différenciant par l'attaque (endo-exo)

Nous choisissons enfin de nous intéresser aux différences induites par la passage d'une forme syn à une conformation anti de l'acrylate de (-)-menthyle. Dans tous

les cas, la forme syn est stabilisée de 5 à 8 kcal.mol<sup>-1</sup> par rapport à la forme anti (tableau 2.6) à l'exception des formes endo s-trans SI faisant intervenir la structure dont l'isopropyle est en conformation "stacked". Cette préférence est de nouveau en accord avec les résultats existant [96], qui prévoyaient une différence de l'ordre de la kcal.mol<sup>-1</sup>.

syn-anti	vide	solvant
endo s-cis RE	7,86	7,89
endo s-cis SI	5,49	5,23
endo s-trans RE	5,63	5,89
endo s-trans SI	0,73	0,84
exo s-cis RE	7,73	7,83
exo s-cis SI	5,26	5,14
exo s-trans RE	5,00	4,84
exo s-trans SI	8,06	7,92

TAB. 2.6 – différence d'énergie (kcal.mol<sup>-1</sup>) pour les structures de transition se différenciant par la conformation du diénophile (anti-syn)

Le fait que la différence énergétique entre les forme syn et anti est plus importante dans notre étude peut être expliquée de deux façons différentes. En premier lieu, il est important de noter que l'étude au niveau AM1 a été menée en présence d'un acide de Lewis (catalyseur) en interaction avec l'atome d'oxygène du groupement carbonyle. Cela peut expliquer que les résultats ne coïncident pas avec ceux de notre étude où seuls les réactifs sont pris en compte.

L'autre interprétation possible est liée au phénomène de rotation de l'isopropyle. En effet, dans l'étude au niveau semi-empirique, la différence d'énergie entre les formes syn (la plus stable) et anti (Table 1 de la ref. [96]) est de 1,2 kcal.mol<sup>-1</sup>, le groupement isopropyle étant systématiquement fixé dans la position stacked.

Dans notre approche, nous n'avons considéré que la conformation "axiale" de l'isopropyle, sauf dans le cas de la structure endo s-trans SI anti. Or, c'est pour le couple faisant intervenir cette dernière structure que la différence énergétique est proche des calculs semi-empiriques.

Il semble donc que la différence entre les calculs existants et nos résultats a pour origine un plus grand encombrement stérique lié à la conservation de la conformation

”axiale” de l’isopropyle.

### 2.3.6 Modification du mécanisme réactionnel

Un autre point sur lequel il est intéressant de se pencher est celui de la modification du mécanisme réactionnel induit par la présence à la fois du solvant et du groupement (-)-menthyle. Cet aspect a aussi été traité dans la référence [95], ce qui nous donne de nouveau un point de comparaison quant à l’apport de notre méthode.

Nous considérons ici la réaction de Diels-Alder uniquement comme un processus concerté, le mécanisme biradicalaire qui pourrait être envisagé ne pouvant pas être traité au niveau de théorie auquel nous travaillons. Nous nous focalisons donc sur l’asynchronicité du mécanisme, que la présence d’un diénophile asymétrique ainsi que la polarisation due au solvant sont susceptibles d’induire. L’asynchronicité peut être vue comme la différence dans l’avancement de la formation des deux liaisons carbone-carbone formées au cours de la réaction de Diels-Alder.

En toute rigueur, le degré d’avancement devrait être calculé tout au long de la coordonnée de réaction, nécessitant alors des calculs dépendant du temps. Dans notre cas, ne disposant pas de telles données, mais seulement des structures de transition, nous ferons l’hypothèse que l’asynchronicité peut être déterminée en se basant uniquement sur le degré d’avancement de formation des liaisons à l’état de transition. Deux façons de calculer ce degré d’avancement sont possibles : soit en considérant l’aspect géométrique, en prenant en compte les longueurs de liaisons en cours de formation, soit un aspect plus électronique, en s’intéressant aux ordres de liaison [109]. Dans les deux cas, on peut calculer le degré d’avancement de formation de la liaison  $i$  noté  $\Gamma_i$  de manière analogue. Pour un degré géométrique, on a :

$$\Gamma_i^D = \frac{D_i^P}{D_i} \quad (2.9)$$

Où  $D_i^P$  la longueur de la liaison  $i$  dans le produit et  $D_i$  sa longueur dans la structure de transition.

Lorsque l’aspect électronique est pris en compte, on a :

$$\Gamma_i^B = \frac{B_i^P}{B_i} \quad (2.10)$$

Où  $B_i^P$  et  $B_i$  sont respectivement les ordres de liaison de Mayer d’une même liaison dans le produit final et à l’état de transition.

Cette procédure a pour avantage de conduire à une valeur de  $\Gamma$  de zéro pour les réactifs (infiniment distants, sans interaction) et de un pour le produit final.

On propose de calculer l'asynchronicité notée  $A$  de la façon suivante, en considérant deux liaisons notées 1 et 2 :

$$A = \frac{\Gamma_1 - \Gamma_2}{\Gamma_1} \quad (2.11)$$

Avec  $\Gamma_1$  la valeur la plus grande. Ainsi, lorsque  $\Gamma_1$  et  $\Gamma_2$  sont égaux,  $A$  est nulle, ce qui correspond à une réaction totalement synchrone. Par contre dans le cas limite où l'une des liaisons n'est pas formée, ( $\Gamma$  correspondant nul),  $A$  vaut 1, indiquant une réaction totalement asynchrone.

Les longueurs des liaisons (telles que décrites par la figure 2.5) et ordres de liaisons de Mayer associés sont reportés dans l'annexe C à la fois pour les états de transition mais aussi pour les produits finaux correspondant (état final stable le plus proche du TS correspondant) dans le tableau C.5.

On donne par ailleurs les asynchronicités ainsi calculées dans le tableau 2.7, ceci pour toutes les structures de transition, dans le vide et dans le solvant. De plus, pour étudier l'effet du solvant, le rapport des asynchronicités dans le solvant et dans le vide est aussi calculé et reporté dans le tableau 2.7 tel que :

$$S = \frac{A_{solv}}{A_{vide}} \quad (2.12)$$

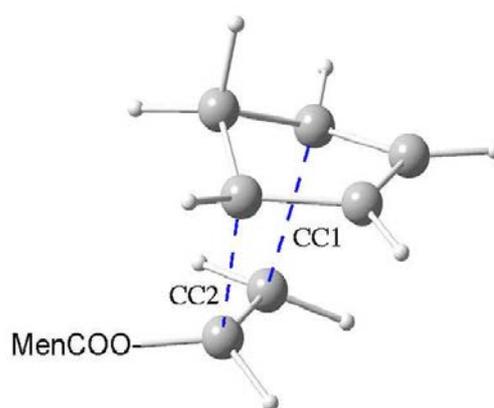


FIG. 2.5 – Liaisons prises en compte pour la détermination de l'asynchronicité : CC2 est la liaison la plus proche de l'auxiliaire chiral

molécule	Longueur de liaisons			Ordre de liaison		
	vide	solvant	S	vide	solvant	S
endo s-cis RE syn	0,09	0,09	1,02	0,25	0,26	1,03
endo s-cis RE anti	0,09	0,09	1,04	0,24	0,25	1,02
endo s-cis SI syn	0,10	0,10	1,02	0,26	0,26	1,03
endo s-cis SI anti	0,09	0,09	1,03	0,25	0,25	1,03
endo s-trans RE syn	0,07	0,08	1,12	0,22	0,24	1,09
endo s-trans RE anti	0,08	0,09	1,10	0,23	0,25	1,08
endo s-trans SI syn	0,08	0,09	1,11	0,23	0,25	1,09
endo s-trans SI anti	0,09	0,09	1,09	0,23	0,25	1,08
exo s-cis RE syn	0,08	0,08	1,05	0,23	0,24	1,05
exo s-cis RE anti	0,07	0,09	1,25	0,26	0,27	1,02
exo s-cis SI syn	0,09	0,09	1,04	0,23	0,24	1,03
exo s-cis SI anti	0,09	0,09	1,03	0,22	0,23	1,04
exo s-trans RE syn	0,08	0,09	1,08	0,22	0,24	1,07
exo s-trans RE anti	0,07	0,08	1,15	0,20	0,22	1,08
exo s-trans SI syn	0,08	0,08	1,11	0,22	0,24	1,09
exo s-trans SI anti	0,07	0,08	1,12	0,21	0,23	1,09

TAB. 2.7 – Asynchronicités calculées à partir des longueurs et ordres des liaisons CC1 et CC2 et effet du solvant sur ces grandeurs

Du point de vue technique, on remarquera simplement que les deux valeurs du rapport des asynchronicités, calculées différemment donnent des résultats similaires. Ce rapport permet en outre de constater que la présence du solvant (par le biais des interactions électrostatiques soluté-solvant) a systématiquement pour effet d'augmenter l'asynchronicité de la réaction, en effet, cette valeur est toujours supérieure à l'unité. Cet effet dû au solvant est cohérent avec les résultats obtenus dans l'étude sur l'acrylate de méthyle [95].

Si on s'intéresse maintenant plus particulièrement aux longueurs de liaisons (tableau C.5, annexe C), on remarque que l'augmentation de l'asynchronicité passe, au niveau du TS, par une diminution de la longueur de la distance CC1 (la moins proche de l'auxiliaire chiral) tandis que CC2 est allongée. On constate que les ordres de liaisons (tableau C.6, annexe C) varient en accord.

En ce qui concerne les effets de l'introduction de l'auxiliaire chiral complet sur le mécanisme, un point de comparaison est fait dans le tableau 2.8 avec les résultats obtenus en considérant l'acrylate de méthyle, à la fois dans le vide et dans le solvant. Bien que les niveaux de calculs soient différents dans les deux études, les différences sont susceptibles de supporter un raisonnement au moins qualitatif.

	endo s-cis		endo s-trans		exo s-cis		exo s-trans	
	vide	solvant	vide	solvant	vide	solvant	vide	solvant
référence [95]	0,181	0,206	0,166	0,199	0,169	0,194	0,164	0,199
RE syn	0,211	0,216	0,165	0,185	0,203	0,219	0,183	0,204
RE anti	0,225	0,229	0,183	0,202	0,189	0,224	0,175	0,201
SI syn	0,238	0,243	0,185	0,206	0,230	0,231	0,191	0,207
SI anti	0,212	0,218	0,201	0,219	0,201	0,219	0,185	0,204

TAB. 2.8 – Différence des longueurs des liaisons (Å) CC1 et CC2 avec les résultats obtenus en considérant l'acrylate de méthyle

Dans les travaux pris comme référence, l'asynchronicité était définie comme la différence de longueurs de liaisons. Nous utilisons donc ici cette grandeur. Les valeurs reportées dans ce tableau nous indiquent que l'introduction de l'auxiliaire chiral par rapport à un simple méthyle augmente aussi l'asynchronicité du processus étudié à la fois dans le vide et dans le solvant. Ceci est tout à fait logique en raison de l'augmentation de la gêne stérique supplémentaire.

Comme dit précédemment, l'ajout du solvant augmente aussi l'asynchronicité. On remarquera cependant que cette augmentation, liée à un facteur indépendant de l'auxiliaire chiral est moins importante dans notre cas que pour l'acrylate de méthyle. L'explication se trouve sans doute dans le facteur multiplicatif des rayons atomiques de van der Waals servant à déterminer la cavité. Dans notre cas, il est fixé à deux, contrairement à l'étude de référence pour lequel il vaut environ 1,3 (valeur par défaut). Ce facteur plus grand, conduit (pour une même géométrie) à une cavité de volume plus important et donc interagissant moins avec le soluté.

On considère donc que l'asynchronicité est plus influencée par la présence du solvant que par l'encombrement stérique. Si on s'intéresse aux valeurs propres des orbitales frontières des réactif, on voit que la réaction est principalement influencée par l'interaction entre la HO du cyclopentadiène et la BV de l'acrylate ( $\Delta_{HO-BV} = 0.42237$  u.a., à comparer aux 0.5275 u.a. pour une interaction entre la HO de l'acrylate et la BV du diène). Ces valeurs n'étant que peu modifiées par le solvant. Les effets du solvant se font par le biais d'une polarisation de l'acrylate principalement (le cyclopentadiène ayant un faible moment dipolaire et du fait de sa symétrie, les effets du solvant sur le diène ne peuvent modifier l'asynchronicité).

Le solvant tend donc à augmenter le moment dipolaire du diénophile (tableau 2.9). Cette polarisation se fait vers l'oxygène du carbonyle, comme le montrent les charges de Mulliken (tableau C.7 en annexe) des réactifs en question. En suivant un raisonnement analogue à celui fait dans la référence [95], on peut dire que comme les orbitales occupées sont polarisées dans ce sens, alors les orbitales virtuelles, en particulier la BV, le sont dans le sens opposé, ce qui induit donc une augmentation du coefficient sur le carbone  $C_1$  et une diminution sur  $C_2$  (voir figure : C.1 pour la numérotation), d'où une tendance plus grande de  $C_1$  à "réagir" en solution par rapport à  $C_2$  (se reporter aux tableaux C.5, C.6 et 2.7 rassemblant les longueurs et ordres de liaisons ainsi que les asynchronicités associées).

De la même manière que dans l'étude sur l'acrylate de méthyle, l'augmentation de l'asynchronicité peut être interprétée dans l'espace des degrés de liberté internes de la molécule. Qualitativement, si on considère la coordonnée de réaction à l'état de transition, celle-ci est construite principalement sur les deux liaisons en formation (CC1 et CC2), or lorsqu'on passe du vide au solvant, l'augmentation de l'asynchronicité se traduit par le déplacement de l'état de transition vers un point de l'hypersurface de potentiel où, par rapport au vide, CC1 est allongée tandis que CC2 est raccourcie. On

diénophile	vide	solvant	variation
s-cis syn	2,21	2,28	0,07
s-cis anti	2,58	2,67	0,09
s-trans syn	2,73	2,89	0,16
s-trans anti	3,30	3,49	0,19

TAB. 2.9 – moments dipolaires en Debye dans le vide et le solvant des différents isomères de l’acrylate de (-)-menthyle et différences associées

voit alors que le déplacement de la coordonnée de réaction quand on passe du vide au solvant se fait suivant une direction qui lui est orthogonale.

## 2.4 Conclusion

Sur la question de la sélectivité, notre méthode nous a permis à la fois de retrouver des résultats donnés par de précédentes études mais aussi de nuancer ces derniers grâce à la prise en compte simultanée de deux phénomènes distincts (présence de l’auxiliaire chiral et influence du solvant) mais agissant sur les mêmes facteurs. Nous avons en effet montré que la préférence pour la forme ENDO induite par le solvant n’était vraie que dans les cas où l’encombrement stérique est réduit, ce qui est impossible à prévoir en considérant une molécule modèle d’acrylate de méthyle comme diénophile.

En ce qui concerne la modification du mécanisme réactionnel par la prise en compte des effets à la fois du solvant et de l’encombrement stérique, on voit que l’augmentation de l’asynchronicité de la réaction est due en partie au fort encombrement stérique du groupement (-)-menthyle mais surtout aux effets électrostatiques du solvant. Comme dans le cas de l’acrylate de méthyle, son influence se traduit par une polarisation de l’acrylate, favorisant ainsi l’asynchronicité du mécanisme.

Ce travail montre que la méthode à trois niveaux est un outil bien adapté à l’étude de la réactivité des systèmes de grande taille en solution puisque une partie des résultats que nous obtenons vient affiner ce qui avait pu être déterminé par des approches faisant des approximations soit sur l’auxiliaire chiral, soit sur la présence du solvant.

Cependant, nous avons dû nous limiter à des interprétations qualitatives, d’abord en raison d’un niveau de théorie insuffisant. De plus, dans cette étude, seize structures de transition ont été prises en compte, or il est apparu que des degrés de libertés ”annexes”

(rotation de l'isopropyle) pouvaient jouer un rôle sur la réaction, en particulier en modifiant de façon significative les effets stériques de l'auxiliaire chiral. Afin de procéder à une étude complète et en réduisant à trois le nombre de conformations possibles de l'isopropyle, il aurait fallu étudier quarante-huit structures de transition. On voit donc que, même sur un système chimique de taille raisonnable (moins de cinquante atomes), la question du sens chimique des configurations figées se pose. En effet, le rôle du groupement isopropyle doit sans doute être pris en compte de façon moyenne du fait de sa liberté de rotation, ce qui pourrait être fait à l'aide d'une approche dynamique, alors encore indisponible au moment de notre étude.

Cette nécessité de travailler sur des grandeurs statistiques nous conduit à envisager, comme développement de la méthode LSCF/MM/SCRF, le couplage avec un algorithme de dynamique moléculaire.

Le deuxième point sur lequel nous avons été obligés de faire des approximations est l'interaction spécifique qui peut exister entre le solvant et le solutés sous forme, notamment, de liaisons hydrogène, phénomène dont le rôle a été mis en évidence dans l'accélération de la vitesse de réaction (réduction de l'énergie d'activation) [101, 102, 103]. Le modèle de solvant utilisé étant implicite, il ne décrit ce genre d'interaction que de façon moyenne. L'amélioration de la description du solvant passe donc par l'introduction de molécules explicites de solvant. Ce point a donc fait l'objet des travaux qui sont exposés dans le chapitre suivant.

Enfin, concernant plus particulièrement notre méthode, d'autres pistes de travail peuvent être envisagées. Nous pouvons ainsi citer la possibilité d'utiliser un champ de force polarisable pour la partie classique pour améliorer son interaction avec le continuum et aussi avec la partie quantique. Il est aussi imaginable de ne pas inclure la partie classique dans la cavité. Cette possibilité peut ainsi servir à modéliser les effets d'un environnement qui posséderait une liaison avec le système quantique comme dans le cas d'un site actif au sein d'une protéine. On peut également imaginer de créer une cavité dans laquelle serait placé les atomes classiques mais aussi un milieu polarisable (soit le contraire de ce que nous avons présenté). Cela pourrait permettre, en utilisant une constante diélectrique adaptée, de décrire la polarisabilité de la partie classique.

# Chapitre 3

## Traitement explicite du solvant

Nous avons montré au chapitre précédent qu'il était possible de rendre compte des effets du solvant sur la réactivité d'espèces chimiques de grande taille par le biais de la méthode LSCF/MM/SCRF. Ce modèle reste cependant limité puisque les interactions spécifiques entre le solvant et le soluté ne sont pas décrites complètement. Ainsi, la liaison hydrogène peut être vue comme la somme de deux contributions : l'une est de nature électrostatique tandis que l'autre correspond au transfert de charge entre le soluté et le solvant. La première est correctement traitée par le modèle de continuum alors que la seconde ne peut pas l'être. De plus, le modèle de solvant implicite ne rend pas compte de l'effet directionnel de la liaison hydrogène. Afin de modéliser ce phénomène qui peut avoir une importance fondamentale en chimie, il est possible d'inclure dans les systèmes d'intérêt une ou plusieurs molécules de solvant. C'est ce point que nous avons choisi d'étudier, tout en restant dans le cadre des méthodes hybrides QM/MM. Ce formalisme laisse le choix entre les possibilités d'ajouter ces molécules soit au niveau classique, soit au niveau quantique. La seconde possibilité conduira bien sûr aux résultats les plus fiables. Le prix, en terme de moyens de calculs, peut par contre constituer un facteur limitant en fonction de la nature et du nombre de molécules de solvant.

Outre un rapide rappel sur les approches QM/MM de type "soluté-solvant", ce chapitre est construit en trois parties principales : dans un premier temps, une étude systématique est présentée sur divers systèmes chimiques simples comportant une liaison hydrogène avec une molécule d'eau (solvant auquel on s'intéresse ici) afin de cerner la capacité et les limites des approches QM/MM pour décrire un tel phénomène. A partir des observations faites dans cette étude, il a ensuite été décidé de définir une

approche adaptée pour décrire d'une manière la plus juste possible les liaisons hydrogène entre un soluté quantique (ou hybride) et une ou plusieurs molécules d'eau classiques. Cette approche passe notamment par la détermination de paramètres classiques spécifiques. La dernière partie de ce chapitre concerne enfin l'application de notre modèle dans l'étude des effets du solvant sur la réaction de Claisen.

### 3.1 Les méthodes hybrides de type soluté-solvant

Comme nous l'avons dit dans le premier chapitre, La méthodologie QM/MM (au sens large) a deux domaines d'application principaux. Le premier, dont nous avons déjà donné le détail concerne la description de systèmes moléculaires de grande taille. Le second, vise à décrire un soluté de taille raisonnable, traité au niveau quantique, sous l'influence d'un grand nombre de molécules classiques de solvant [110]. Dans cette approche, contrairement au cas précédent, aucune coupure de liaison covalente n'est nécessaire.

L'hamiltonien d'interaction QM/MM  $H_{QM/MM}$  (équation 1.24) se réduit donc aux termes non-liés, à savoir les interactions de van-der-Waals et les interactions électrostatiques. Cette application du formalisme hybride QM/MM est assez ancienne, c'est d'ailleurs dans le cadre de cette approche que les premières études hybrides ont été menées. Ce formalisme a depuis fait l'objet de nombreux développements et applications. Nous citerons par exemple les travaux de Gao qui a utilisé ce formalisme aux niveaux AM1/TIP3P [111, 112] et 3-21G/TIP3P [113] et pour lesquels il a été amené à développer des paramètres classiques spécifiques pour les atomes quantiques.

A Nancy, l'approche de type soluté-solvant a également été développée, en particulier par M.F Ruiz-López [114, 115, 116] en utilisant un hamiltonien de Kohn-Sham pour la partie quantique.

L'absence de frontière intra-moléculaire rend aisé le calcul analytique des forces. Ceci permet donc d'utiliser cette approche avec des méthodes statistiques Monte-Carlo ou de dynamique moléculaire. Nous citerons donc en particulier les travaux basés sur le logiciel DFMM et concernant l'étude de l'hydrolyse du formamide dans l'eau [117].

Dans l'approche de Jorgensen *et al.*, le chemin réactionnel étudié est d'abord déterminé en phase gazeuse. Certaines configurations en sont extraites et sont utilisées, au niveau classique, pour mener des simulations de Monte-Carlo au sein d'un liquide. Cette approche a été utilisée pour l'étude de l'influence du solvant dans plusieurs réactions chimiques [101, 102, 118] telles que la cycloaddition de Diels-Alder ou la réaction de Claisen.

L'un des derniers développements de l'approche soluté/solvant consiste à coupler cette approche avec un modèle de continuum. Cette approche a été explorée par Q. Cui [119] ou au laboratoire par S. Chalmet *et al.* [120] dans le cas de solutés complètement quantiques.

Nous nous positionnons différemment puisque nous choisissons d'une part, de pou-

voir décrire le soluté dans l'approche QM/MM et d'autre part de n'inclure que les molécules de solvant en interaction directe avec ce dernier, les effets à plus longue portée pouvant être décrits à l'aide du continuum.

Cet état des lieux n'est bien entendu pas exhaustif mais reste représentatif des différentes techniques existant aujourd'hui.

## 3.2 Description de la liaison hydrogène en QM/MM

### 3.2.1 Protocole

Dans l'étude que nous présentons ici, nous avons choisi d'explorer les différentes possibilités offertes en terme de partition QM/MM. C'est pourquoi, outre l'habituelle coupure entre soluté et solvant, nous envisageons aussi les effets d'une coupure intramoléculaire du soluté sur une liaison hydrogène traitée au niveau quantique. L'approches QM/MM envisagée est LSCF/MM [55, 36, 58]. Les résultats obtenus à l'aide des approches hybrides seront comparés à leurs homologues quantiques et classiques.

#### systemes chimiques pris en compte

Afin de nous focaliser sur l'interaction entre le soluté et l'eau, nous avons décidé de prendre en compte uniquement des systèmes comportant une molécule de soluté en interaction avec une seule molécule de solvant. Les différents solutés ont été choisis afin d'offrir un ensemble représentatif des fonctions chimiques pouvant former une liaison hydrogène avec une molécule d'eau. Ils sont donc constitués d'une molécule d'éthane mono substitué avec huit groupements différents (-COOH, -COO<sup>-</sup>, -NH<sub>2</sub>, -NH<sub>3</sub><sup>+</sup>, -OH, -O<sup>-</sup>, -SH et -S<sup>-</sup>) conduisant à dix complexes représentés dans la figure 3.1.

#### Méthodes mises en œuvre

Afin de connaître la capacité des méthodes hybrides à décrire la liaison hydrogène, nous choisissons de comparer les résultats obtenus avec une approche complètement quantique. La méthode ainsi que la base choisie pour la référence seront par ailleurs utilisées pour tous les calculs quantiques, y compris dans le cadre des méthodes hybrides. Nous cherchons ainsi à minimiser les facteurs de différences entre les approches étudiées et les références.

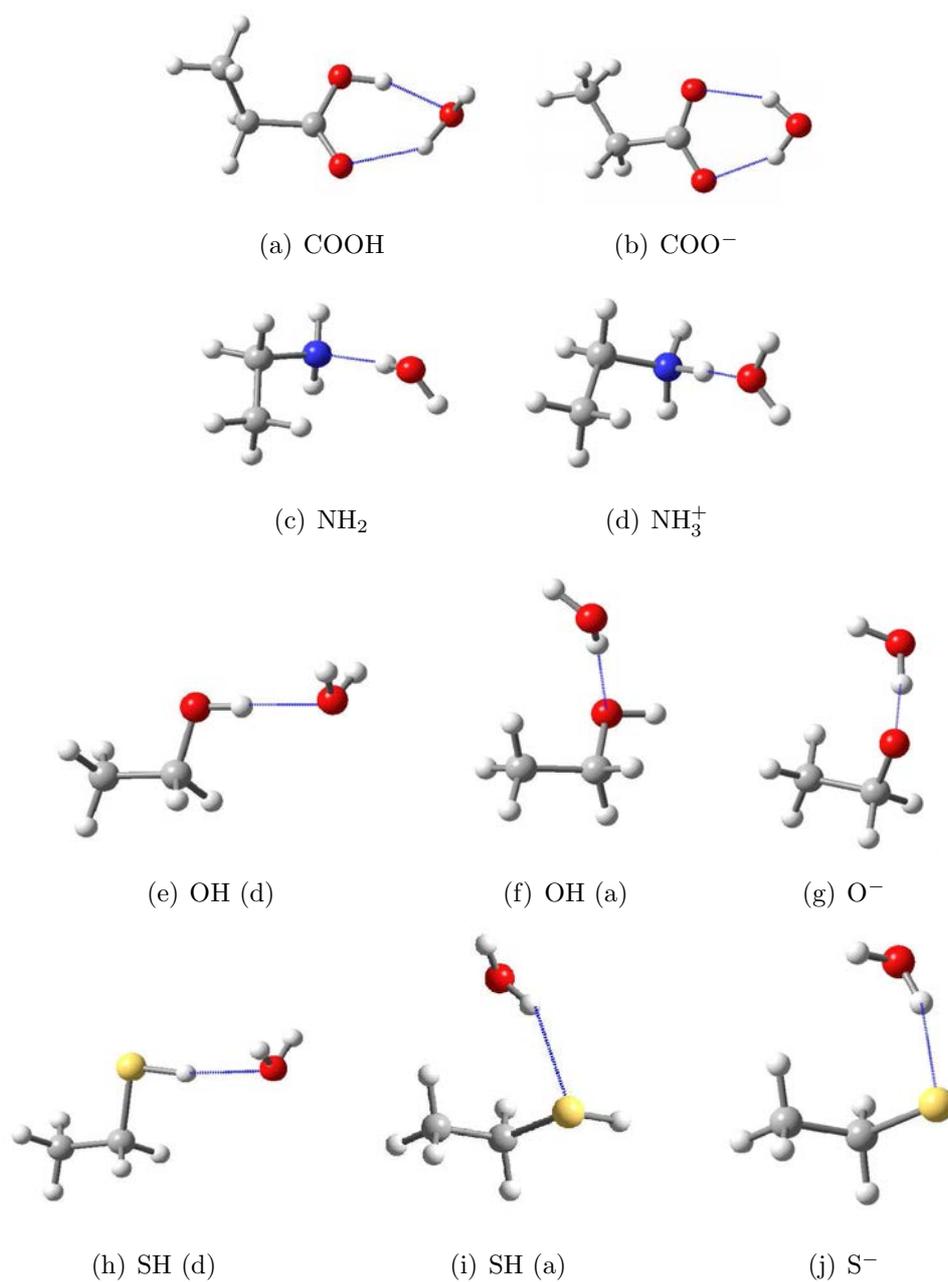


FIG. 3.1 – Structures utilisées dans l'étude. Les traits bleus signalent les liaisons hydrogène prises en compte. Les structures sont désignées par la nature du groupement en interaction avec l'eau.

Comme nous nous intéressons à des systèmes comportant deux molécules, il est important d'avoir une bonne description des interactions intermoléculaires (lorsque celle-ci est entre deux éléments traités de façon quantique). Il faut donc prendre en compte la corrélation électronique. Nous avons choisi de travailler avec la méthode MP2 [14], en incluant, pour le calcul de l'énergie de corrélation tous les niveaux occupés.

Même si nous étudions des espèces anioniques, pour lesquelles l'utilisation d'une base comportant des fonctions diffuses améliore la description, nous choisissons une base n'en comportant pas. Nous préférons augmenter le nombre de fonctions de polarisation. Ce choix a été fait en considérant le phénomène mis en évidence par Surjan et Ángyán [121]. Il s'agit de la "surpolarisation" de la fonction d'onde induite par la proximité de charges ponctuelles classiques, polarisation qui n'est pas compensée par un terme de répulsion (répulsion de Pauli) au niveau quantique puisque les termes QM/MM de van der Waals sont décrits au niveau classique. Ce phénomène a par ailleurs été signalé par Salahub *et al.* [122] ainsi que par Ben-Nun et Martínez [123]. La présence d'orbitales diffuses, de grande extension spatiale et polarisables pouvant faciliter un tel phénomène, la base d'orbitale avec laquelle nous travaillons est donc 6-311G(2d,2p).

Dans le cas de calculs LSCF/MM impliquant la coupure d'une liaison covalente (toujours entre deux atomes de carbone), l'orbitale frontière a été déterminée sur la molécule d'éthane relaxée et avec le critère de Weinstein-Pauncz [62]. Le troisième atome, nécessaire à la détermination du trièdre de référence utilisé pour la rotation de la SLBO, est l'un des atomes d'hydrogène porté par le carbone portant aussi le groupement chimique en interaction avec l'eau.

En ce qui concerne la partie classique, nous avons choisi le champ de forces Amber, dans sa version de 1995 [20]. Les molécules étudiées n'étant pas des acides aminés ou des nucléotides, aucun jeu de charges standard n'existe dans Amber pour nos molécules. Afin de déterminer un jeu de charges cohérent avec le champ de forces mis en œuvre, nous avons utilisé la méthode RESP [124], en considérant un niveau de calcul MP2/6-31G\*. Pour n'avoir qu'un seul jeu de charges par complexe durant toute l'étude, à la fois pour les calculs classiques et hybrides, deux contraintes sont imposées lors de la détermination des charges. D'une part, les charges portées par le groupement méthyle de chaque molécule sont déterminées pour que leur somme soit nulle. D'autre part, les atomes chimiquement équivalents (les trois hydrogènes du méthyle et les deux de l'autre atome de carbone) ont des charges égales.

Les charges ainsi calculées sont rassemblées en annexe dans le tableau D.1.

L'interaction calculée entre deux parties décrites par la mécanique quantique implique la possibilité d'une erreur de superposition de base (BSSE) dont nous devons calculer la valeur. Ce calcul est trivial et même automatisé dans le cas d'approches complètement quantiques mais n'est pas défini pour les approches hybrides dans le programme que nous utilisons [2]. Nous précisons donc ici rapidement les caractéristiques de ce calcul, et les adaptations à la méthode LSCF/MM.

Pour un système donné, l'énergie calculée s'écrit  $E_X^{Y,Z}$  où l'indice X correspond au système chimique, Y se rapporte à la géométrie et Z à la base d'orbitales. Dans le cas de deux fragments  $a$  et  $b$ , l'énergie d'interaction s'écrit :

$$\Delta E_{int.} = E_{ab}^{ab,ab} - E_a^{a,a} - E_b^{b,b} \quad (3.1)$$

Avec  $E_{ab}^{ab,ab}$  l'énergie du complexe et  $E_a^{a,a}$   $E_b^{b,b}$  les énergies respectives des fragments  $a$  et  $b$  dans leur géométries relaxées et avec leur base propre.

Pour corriger les effets de superposition de base, on calcule la différence entre les énergies des monomères avec leur propre base et celle du complexe (base de  $a$  plus base de  $b$ ), mais avec la géométrie qu'ils ont dans le complexe. La correction s'écrit :

$$\delta_{BSSE} = (E_a^{ab,ab} + E_b^{ab,ab}) - (E_a^{ab,a} + E_b^{ab,b}) \quad (3.2)$$

D'où l'énergie d'interaction corrigée :

$$\Delta E_{int.}^{corr.} = \Delta E_{int.} - \delta_{BSSE} \quad (3.3)$$

Pour calculer cette correction dans le cas des approches hybrides, la procédure est équivalente mais avec quelques adaptations sont nécessaires en raison de la présence d'une frontière intra-moléculaire. Le calcul des énergies pour les fragments avec leur propres bases ( $E_a^{ab,a}$  et  $E_b^{ab,b}$ ,  $b$  est le fragment hybride) se font de la même façon que précédemment. Lorsqu'on s'intéresse aux énergies des fragments dans la base complète ( $E_a^{ab,ab}$  et  $E_b^{ab,ab}$ ), le calcul n'est modifié que pour le fragment  $a$  (quantique) et dont la base comprend l'orbitale gelée du fragment  $b$ . Dans ce cas, On impose à cette SLBO d'être inoccupée. Pour le fragment  $b$ , cette orbitale reste naturellement occupée.

Le dernier point à considérer est le modèle d'eau classique dont nous souhaitons étudier l'interaction avec le soluté quantique ou classique. Afin de rester cohérent avec le champ de forces Amber, nous utilisons le modèle TIP3P[68] dans une version modifiée qui rend ce dernier flexible [24]. Les caractéristiques des potentiels intermoléculaires

(charges, potentiel de Lennard-Jones sur l'oxygène uniquement) restent cependant identiques au modèle original. Il faut enfin noter que les paramètres de ce modèle ont été ajustés pour reproduire les caractéristiques de l'eau liquide. On notera en particulier que les charges ponctuelles portées sur les atomes ont des valeurs qui permettent de reproduire le moment dipolaire de la molécule d'eau au sein du liquide. Outre le moment intrinsèque de la molécule, on retrouve aussi le moment induit par l'environnement. Le moment total valant alors 2,34 Debye pour une molécule d'eau TIP3P isolée dans sa géométrie d'équilibre.

### 3.2.2 Résultats, discussion

Avant de passer au traitement des résultats obtenus avec les méthodes hybrides, il est d'abord intéressant de se pencher sur ceux obtenus avec la mécanique moléculaire, en comparaison avec la référence quantique.

Nous choisissons ensuite de nous attarder sur les effets d'une coupure intramoléculaire du soluté, avec le méthyle classique et le reste quantique.

Enfin, nous explorerons aussi la possibilité de placer la coupure entre les deux molécules, comme dans l'approche habituelle soluté-solvant.

Tous les raisonnements seront illustrés par des graphiques comparatifs, à la fois pour les aspects géométriques (longueur de la liaison H) et aussi pour les énergies d'interaction. Les structures précisément évoquées sont indiquées sur ces graphiques. Les valeurs calculées sont quant à elles reportées en annexe.

#### Traitement par la mécanique moléculaire

Outre un point de comparaison pour les calculs hybrides, les résultats obtenus par mécanique moléculaire uniquement nous permettent aussi de vérifier la qualité des jeux de charges utilisés pour les solutés. En effet, bien que nous ayons utilisé le formalisme adéquat (RESP), les contraintes importantes (neutralité du méthyle, charges équivalentes pour les hydrogènes chimiquement équivalents) sont susceptibles d'induire des artefacts dans les résultats.

On reporte les énergies d'interaction obtenues pour les complexes optimisés avec Amber en fonction de des énergies de référence dans la figure 3.2(a). Sur cette dernière, les deux droites (vertes et bleues) parallèles à la diagonale, marquent une erreur de plus ou moins 3 kcal.mol<sup>-1</sup> par rapport à la référence quantique. De la même manière, la longueur des liaisons hydrogène formées entre les solutés et les molécules d'eau sont

reportées sur la figure 3.2(b), ou la marge d'erreur marquée est de plus ou moins 0,1 Å. On peut remarquer cependant que pour dix systèmes traités, douze valeurs sont reportées, les solutés COOH et COO<sup>-</sup> formant deux liaisons hydrogène avec la molécule d'eau.

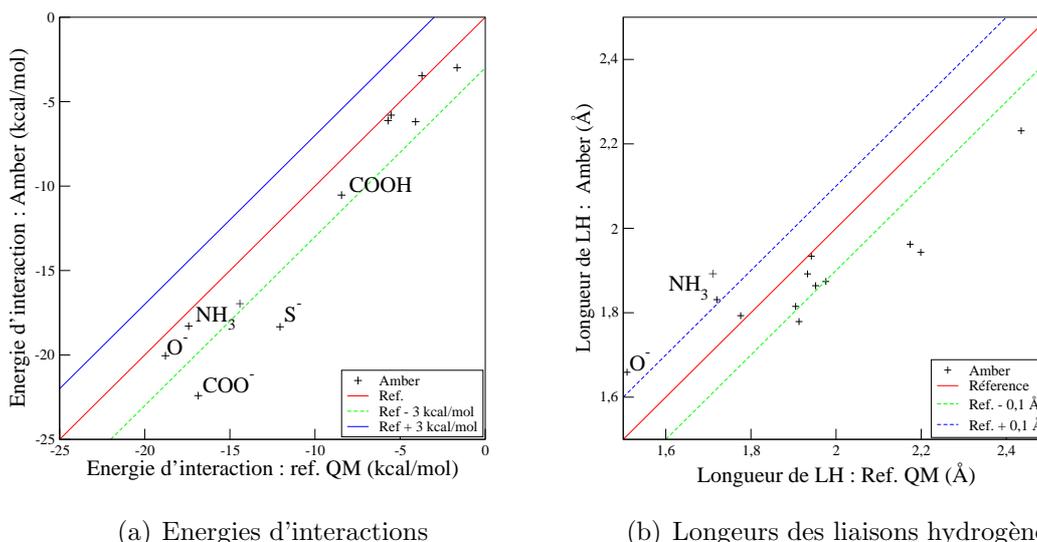


FIG. 3.2 – Valeurs géométriques et énergétiques des calculs menés avec Amber comparées aux références quantiques

En terme d'énergies d'interaction, on peut voir que le modèle totalement classique tend à surestimer les valeurs par rapport à la référence quantique. On peut cependant séparer notre ensemble de molécules en deux groupes : l'un comprenant les espèces chargées et l'autre les espèces neutres. Dans le cas regroupant les solutés neutres, la déviation moyenne par rapport à la référence est d'environ 1 kcal.mol<sup>-1</sup>, avec un maximum de 2,1 kcal.mol<sup>-1</sup> pour le complexe formé avec l'acide propanoïque (COOH). Si on considère maintenant les espèces chargées (COO<sup>-</sup>, NH<sub>3</sub><sup>+</sup>, O<sup>-</sup> et S<sup>-</sup>), la différence moyenne passe à près de 3,5 kcal.mol<sup>-1</sup>, comprise entre 0,87 kcal.mol<sup>-1</sup> pour l'éthylammonium et 6,52 kcal.mol<sup>-1</sup> pour S<sup>-</sup>.

Dans le cas des espèces chargées, il est possible que l'erreur importante par rapport aux résultats de référence provienne de ces derniers. En effet, nous avons fait le choix de ne pas utiliser, dans les calculs quantiques, de jeu de fonctions diffuses. L'absence de ces dernières peut alors induire des valeurs de référence de mauvaise qualité.

En ce qui concerne l'aspect géométrique on s'intéresse aux longueurs des liaisons hydrogène. La figure 3.2(b) montre que, en comparaison des résultats quantiques, les

calculs classiques sous-estiment ces dernières, exception faites des structures  $\text{NH}_3^+$  et  $\text{O}^-$ .

Outre ces deux aspects, on note une différence systématique dans l'arrangement de la molécule d'eau par rapport au soluté, en particulier lorsque ce dernier est le donneur de proton. En effet, dans les calculs quantiques, l'oxygène aqueux est placé dans une sorte d'environnement électronique tétrahédrique, répondant ainsi à l'image habituelle de l'interaction de l'hydrogène avec un des doublets libres de l'oxygène (figure 3.3(a)). Les calculs classiques donnent aux quatre atomes un arrangement plan (figure 3.3(b)). Ce comportement doit sans doute être imputé à l'utilisation d'un modèle décrivant le potentiel électrostatique créé par l'eau à l'aide d'un jeu de trois charges ponctuelles.

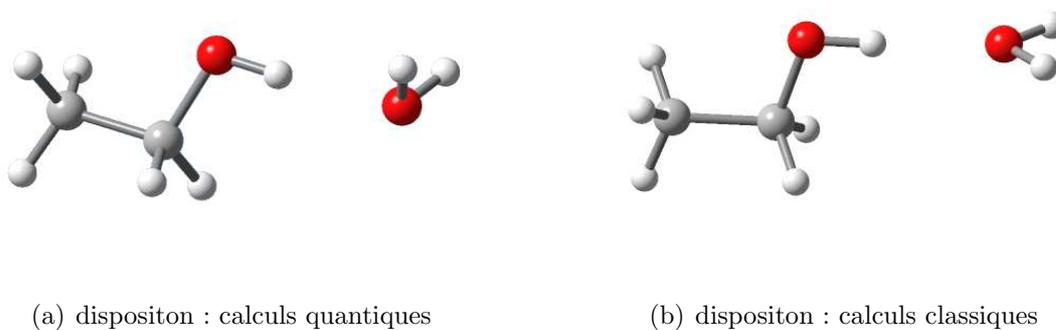


FIG. 3.3 – Différence de disposition de la molécule d'eau par rapport au soluté entre calculs quantiques et classiques

Dans cette approche totalement classique, on doit principalement noter que les énergies d'interaction sont surévaluées par rapport à la référence quantique. Les longueurs des liaisons hydrogène sont généralement trop courtes. Ces faits montrent une surestimation de l'interaction entre le soluté et le solvant. Ce constat reste malgré tout cohérent si on considère que les charges ponctuelles du modèle TIP3P confèrent à la molécule d'eau un moment dipolaire plus important que celui de la molécule isolée (1,86 Debye). Ceci peut effectivement conduire à une interaction trop forte. Malgré cela, les structures en interaction restent proches de leurs homologues de référence.

En dépit des contraintes imposées lors de la détermination des jeux de charges des différents solutés (neutralité du méthyle et charge équivalente pour des atomes chimiquement équivalents), les résultats obtenus par la mécanique moléculaire sont en accord avec les références quantiques, particulièrement concernant les espèces neutres. Nous choisissons donc de conserver les jeux de charges choisis dans la suite de notre

étude.

### Effets d'une frontière dans le soluté

Le point que nous abordons maintenant est l'étude des effets d'une coupure QM/MM au sein du soluté. Dans ce cas, la partie classique n'est constituée, à chaque fois, que du méthyle du soluté portant les charges ponctuelles décrites dans la partie précédente. Lorsqu'on utilise la méthode LSCF/MM, deux facteurs peuvent modifier la description de l'interaction étudiée. Il s'agit d'une part du remplacement du méthyle par sa représentation classique mais aussi de la présence de la liaison gelée. Pour pouvoir séparer les effets de ces deux perturbations, nous présentons d'abord les résultats obtenus avec une approche totalement quantique en introduisant la SLBO, telle qu'elle l'est dans les calculs hybrides.

#### – Influence de la SLBO

Les résultats obtenus après optimisation de géométrie sont présentés dans la figure 3.4, avec les mêmes caractéristiques que précédemment pour la représentation graphique des marges d'erreur.

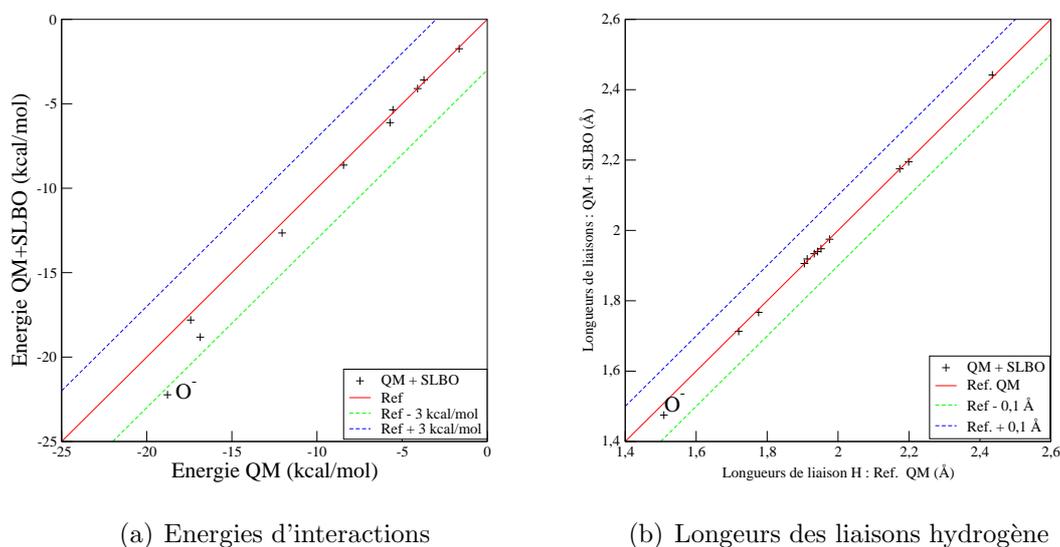


FIG. 3.4 – Valeurs géométriques et énergétiques des calculs menés avec une SLBO, comparés aux références quantiques

En terme d'énergies d'interaction (figure 3.4(a)), les valeurs prédites sont proches de la référence : la déviation la plus importante étant pour le système  $O^-$  avec un surestimation de  $3,4 \text{ kcal.mol}^{-1}$ , ce qui amène la déviation moyenne (moyenne des différences en valeur absolue) à  $0,81 \text{ kcal.mol}^{-1}$ .

Du point de vue géométrique, les longueurs des liaisons H sont presque identiques à leurs références puisque la différence moyenne est inférieure à  $10^{-2} \text{ \AA}$ . L'erreur maximale est obtenue avec la structure de l'éthanolate, dans laquelle la liaison est raccourcie de  $3,4.10^{-2} \text{ \AA}$ .

Afin de déterminer l'influence de l'introduction de la SLBO sur la distribution électronique du système, nous nous intéressons au transfert de charges entre les deux parties du soluté en considérant les charges de Mulliken [100], dans le soluté seul et en interaction avec l'eau. Les charges prises en compte sont celles déterminées en prenant en compte la fonction d'onde obtenue après un calcul Hartree-Fock.

Les sommes des charges de Mulliken pour le méthyle sont présentées dans le tableau 3.1. A chaque fois, la charge totale de la molécule d'eau et aussi donnée afin de déterminer l'influence de la SLBO sur le transfert de charge entre les deux molécules.

On remarque que la charge du groupement méthyle est beaucoup moins importante lorsqu'on introduit une orbitale gelée entre les deux carbones. Comme on pouvait s'y attendre, la SLBO a pour effet de limiter le transfert de charges entre les deux fragments. Ce constat est vérifié dans les deux cas (molécule seule et interagissant avec l'eau).

Ce phénomène est d'autant plus important que la molécule est chargée et qu'elle est de petite taille (c'est à dire que la SLBO est proche du groupement d'interaction). En effet, dans le cas de  $NH_3^+$  et de  $O^-$ , le transfert est réduit respectivement de 0,14 et 0,13 électrons lorsque le soluté est seul. Cette limitation du transfert électronique fait que la charge (positive ou négative) reste localisée sur le groupement en interaction avec la molécule d'eau et conduit ainsi à la surestimation de l'énergie d'interaction constatée dans les deux cas pris en compte ici.

Cette surestimation est par ailleurs marquée, dans le cas de l'éthanolate, par un raccourcissement de la liaison hydrogène formée puisque cette dernière diminue de  $3,4.10^{-2} \text{ \AA}$ . Cette erreur est au moins trois fois supérieure à celle commise dans les autres cas.

Les valeurs des charges totales des molécules d'eau nous indiquent que l'introduction d'une SLBO dans un système quantique ne modifie pas ou très peu le transfert de charge

Complexe	-CH <sub>3</sub>				H <sub>2</sub> O	
	soluté seul		soluté en interaction		H <sub>2</sub> O en interaction	
	QM	QM+SLBO	QM	QM+SLBO	QM	QM+SLBO
COOH	0,054	-0,001	0,050	-0,002	0,011	0,012
COO <sup>-</sup>	-0,038	-0,007	-0,016	-0,001	-0,083	-0,084
NH <sub>2</sub>	0,028	-0,010	0,036	-0,023	-0,017	-0,018
NH <sub>3</sub> <sup>+</sup>	0,172	0,032	0,145	0,015	0,072	0,073
OH (d)	0,017	-0,014	0,005	-0,015	0,033	0,033
OH (a) <sup>a</sup>			0,029	-0,027	-0,006	-0,007
O <sup>-</sup>	-0,206	-0,074	-0,170	-0,070	-0,115	-0,126
SH (d)	0,031	-0,004	0,021	-0,009	0,034	0,035
SH (a) <sup>a</sup>			0,038	-0,013	0,021	0,021
S <sup>-</sup>	-0,061	-0,003	-0,048	0,004	-0,017	-0,017

TAB. 3.1 – Somme des charges de Mulliken des atomes dans le groupement méthyle des différentes structures et charge totale de la molécule d'eau. <sup>a</sup> : les solutés seuls sont les mêmes que les précédents (OH (d) et SH(d)), seul l'arrangement avec la molécule d'eau change.

entre les deux molécules.

Les effets de la présence d'une orbitale de liaison gelée sont à relier directement à l'utilisation d'une SLBO déterminée sur l'éthane. Du fait de sa symétrie, celle-ci ne peut pas rendre compte de la polarisation de la liaison  $\sigma$  qu'implique la proximité d'un groupement polaire, conduisant ainsi à la diminution (et presque l'annulation) du transfert de charge. L'amélioration de la description peut donc passer par la prise en compte d'une SLBO adaptée au système chimique.

Afin d'illustrer ce propos, nous prenons le cas de l'éthanolate pour lequel les résultats sont les plus éloignés de la référence. En utilisant une SLBO directement déterminée sur l'éthanolate, les résultats présentés précédemment sont sensiblement améliorés. En effet, la longueur de liaison hydrogène passe de 1,475 à 1,494 Å (1,509 Å pour le calcul quantique de référence).

De même, l'erreur commise sur l'énergie d'interaction est elle aussi réduite (de 3,4 à 2 kcal.mol<sup>-1</sup>). En ce qui concerne les charges des différents fragments, on remarque que celle du méthyle est plus importante que celle de référence. On se situe en effet respectivement à -0,2813 charges élémentaires pour l'éthanolate seul et -0,2857 lorsqu'il est en interaction avec H<sub>2</sub>O. Le dernier point concerne le transfert de charge intermoléculaire où la valeur calculée de 0,1193 (charge de l'eau) est en bon accord avec la valeur de référence (0,1154).

On voit donc que l'introduction d'une SLBO adaptée permet de diminuer la population électronique sur la partie de la molécule en interaction avec l'eau. Ceci conduit alors à un transfert de charge eau-éthanol moins important et donc une énergie d'interaction plus faible. Par ailleurs, on remarque que le transfert intramoléculaire surestimé par une SLBO "polarisée" s'accompagne d'une énergie d'interaction sous-estimée.

Les résultats obtenus lors de l'introduction d'une SLBO symétrique restent en bon accord avec la référence sauf pour deux cas. Nous considérerons alors dans la suite, où le soluté sera traité de façon hybride avec l'approche LSCF/MM, que les différences avec la référence seront essentiellement dues à la description classique du groupement méthyle. Cette assertion est d'autant plus vraie si on considère que l'utilisation d'une SLBO adaptée au système chimique permet une amélioration sensible du résultat.

Maintenant que les effets de la SLBO seule ont été soulignés, nous pouvons désormais considérer la description hybride elle-même.

- influence du méthyle classique

Les résultats concernant les énergies d'interaction et les longueurs de liaison hydrogène sont reportées dans la figure 3.5, les valeurs correspondantes sont données en annexe.

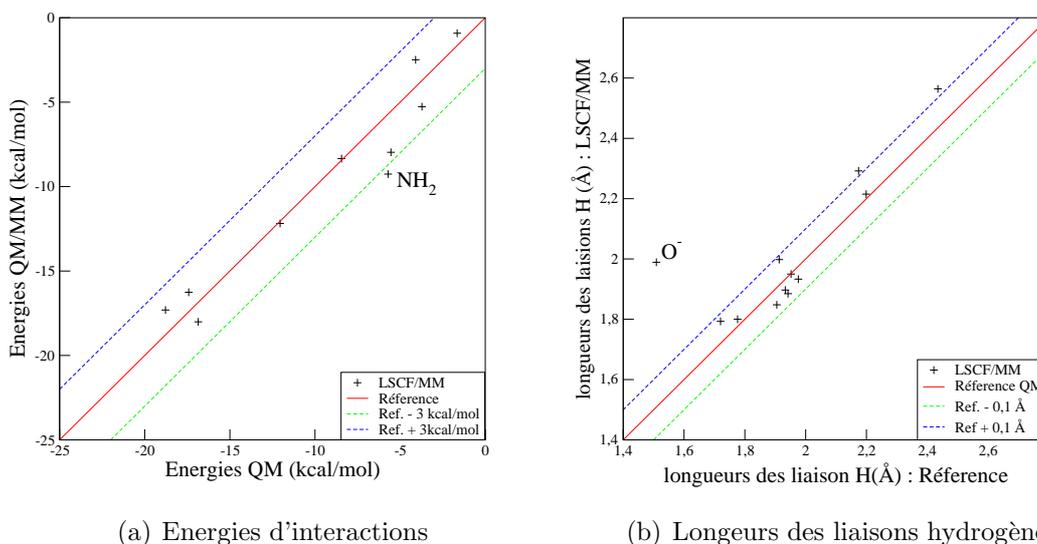


FIG. 3.5 – Valeurs géométriques et énergétiques des calculs menés avec LSCF/MM comparées aux références quantiques. La frontière se situe au sein du soluté

L'erreur moyenne en valeur absolue commise par rapport à la référence est de  $1,81 \text{ kcal.mol}^{-1}$  avec un maximum de  $3,54 \text{ kcal.mol}^{-1}$  pour la structure impliquant l'éthylamine ( $\text{NH}_2$ ). Les longueurs de liaisons associées sont en assez bon accord puisque la différence absolue moyenne est de  $1.10^{-2} \text{ \AA}$ . On notera cependant que dans la structure eau-éthanolate, la liaison hydrogène est allongée de  $0,48 \text{ \AA}$  par rapport à la référence alors que l'énergie d'interaction n'est que légèrement diminuée.

Si on exclut ces deux résultats, les énergies et longueurs de liaisons sont en assez bon accord avec la référence.

Pour expliquer le phénomène constaté dans l'éthanolate, il faut considérer la géométrie complète du complexe incriminé. Avec la méthode LSCF/MM, la distance est surestimée alors que l'énergie reste du même ordre que celle de référence, on remarque par ailleurs que le deuxième atome d'hydrogène de l'eau se trouve à une distance de  $1,4 \text{ \AA}$  de l'hydrogène du méthyle classique le plus proche. Cette configuration pontée du complexe (figure 3.6) n'a aucune réalité physique et existe en raison des charges ponctuelles de  $-0,223$  électrons portées par les hydrogènes classiques.

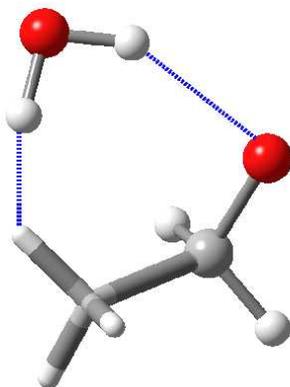


FIG. 3.6 – Structure pontée de l'éthanolate dans l'approche QM/MM : la partie où seules les liaisons sont représentées est traitée au niveau classique

Le phénomène principal mis en évidence dans cette partie est spécifique à l'approche LSCF/MM. Il concerne les effets de la présence d'une SLBO à proximité du centre d'interaction du soluté. En effet nous avons pu voir que cette dernière gelait effectivement le transfert de charge entre deux sous-parties quantiques. Le choix d'utiliser une SLBO déterminée sur la plus petite molécule contenant la liaison coupée (l'éthane) ne permet pas de rendre compte de l'influence induite par la proximité des groupements chimiques les plus polarisants tels que l'alcoolate. Nous avons cependant montré que l'utilisation d'une SLBO adéquate diminuait l'erreur commise.

D'une façon générale, la description de la liaison hydrogène au niveau quantique à proximité d'une coupure QM/MM est envisageable à condition de respecter une certaine distance entre le site concerné et la coupure. Dans le cas spécifique de LSCF/MM, les résultats pourront être améliorés, bien entendu, en utilisant une SLBO qui reproduit la polarité de la liaison qu'elle représente. Enfin une attention particulière devra être portée au jeu de charges ponctuelles, dont le champ créé à longue distance peut reproduire celui de la molécule quantique mais dont les propriétés à courte distance conduisent à des résultats erronés. Ceci nous conduit donc à considérer qu'il est préférable de positionner une coupure sur une liaison faiblement polarisée et à une distance suffisante du centre réactif, surtout si celui-ci est chargé.

### Soluté quantique et molécule d'eau classique

Dans cette partie, nous envisageons une partition plus habituelle puisque c'est celle utilisée dans les calculs de type soluté-solvant avec un soluté totalement quantique et un solvant classique, ici une molécule d'eau TIP3P [68].

On peut par ailleurs mentionner que, puisqu'aucune liaison covalente n'est partagée par les sous systèmes quantiques et classiques, la spécificité liée à l'approche LSCF/MM n'est pas utilisée et les résultats obtenus par cette méthode sont ceux d'une approche soluté-solvant habituelle, impliquant une polarisation directe de la fonction d'onde par l'environnement classique.

Ce mode de description implique deux approximations qu'il convient de rappeler. D'une part, la molécule d'eau étant classique, le transfert de charge avec le soluté ne sera pas décrit explicitement. D'autre part, l'interaction électrostatique entre deux densités électroniques est remplacée par une interaction entre une densité et un jeu de charges ponctuelles fixées, ce qui ne rend possible la polarisation que pour le soluté.

Les valeurs des énergies et longueurs de liaisons hydrogène sont reportés dans la figure 3.7, comme précédemment, les valeurs sont comparées à la référence quantique.

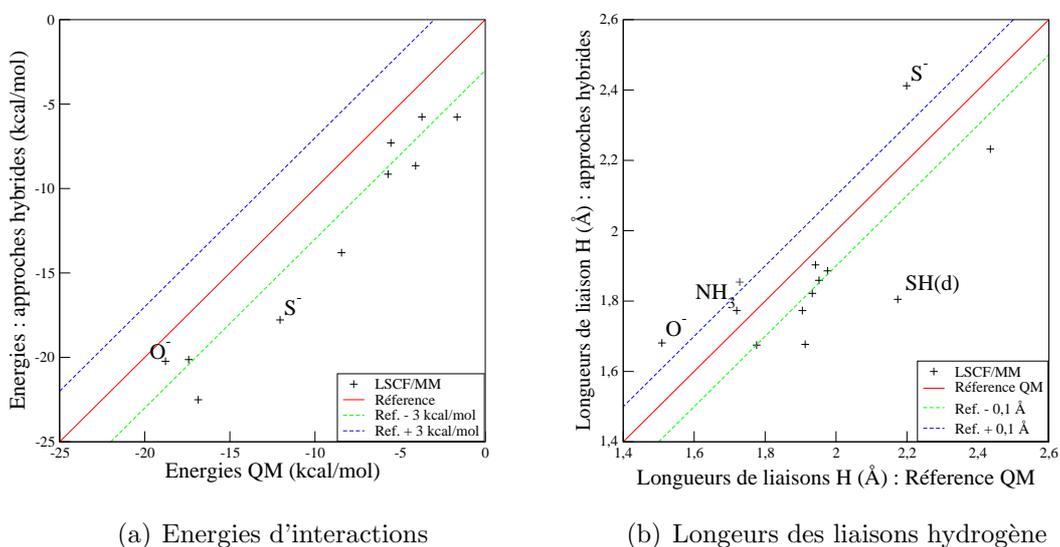


FIG. 3.7 – Valeurs géométriques et énergétiques des calculs menés avec LSCF/MM comparées aux références quantiques. Le soluté est décrit par la mécanique quantique et l'eau avec le modèle TIP3P

Le premier point remarquable concerne les énergies d'interaction que notre approche

a tendance à systématiquement surestimer. L'erreur moyenne est de  $3,7 \text{ kcal.mol}^{-1}$ , avec des valeurs comprises entre  $1,45 \text{ kcal.mol}^{-1}$  pour le complexe de l'éthanolate et  $5,73 \text{ kcal.mol}^{-1}$  pour celui formé avec le complexe  $\text{S}^-$ .

Du point de vue géométrique, cette approche conduit à des structures proches de la référence quantique avec une déviation moyenne de  $0,15 \text{ \AA}$ , l'erreur maximale étant commise sur le complexe SH avec  $0,37 \text{ \AA}$ . On remarque par ailleurs que l'échantillon peut être divisé en deux groupes : le premier, pour lesquels les liaisons hydrogène sont raccourcies comprend toutes les espèces neutres ainsi que la structure  $\text{COO}^-$ . Le second comprend les trois espèces chargées  $\text{NH}_3^+$ ,  $\text{O}^-$  et  $\text{S}^-$ . On peut voir dans cette séparation possible un effet de résultats de référence obtenus avec une base non adaptée à la description de molécules chargées.

Afin de montrer le comportement général dans cette approche, nous choisissons de déterminer l'équation de la droite de régression linéaire à partir des valeurs de l'énergie LSCF/MM en fonction des références, ce qui conduit à une droite de coefficient directeur de  $0,98$  et une ordonnée à l'origine de  $-3,7 \text{ kcal.mol}^{-1}$ . Le coefficient de régression étant de  $0,98$ . Ceci montre qu'on peut ramener les modifications apportées par le remplacement d'une molécule d'eau quantique par son modèle classique à une surestimation moyenne. Ce déplacement est par ailleurs assez constant puisque l'écart quadratique moyen des valeurs calculé par rapport à cette droite de régression est de  $1,5 \text{ kcal.mol}^{-1}$ , ce qui signifie que le phénomène se répète de façon comparable pour les systèmes pris en compte ici.

De la même façon que dans le paragraphe précédent, nous nous intéressons enfin au transfert de charge au sein du soluté. Le tableau 3.2 rassemble la charge totale du groupement méthyle (basée sur une analyse de Mulliken, prenant en compte la fonction d'onde SCF) pour les dix systèmes en interaction avec  $\text{H}_2\text{O}$ . Le transfert de charge entre les deux molécules, dans ce type d'analyse, est bien sûr nul.

Ces valeurs permettent de voir que dans tous les cas, le transfert de charge est plus important dans l'approche hybride que dans la référence. Ce phénomène est particulièrement marqué dans le cas de  $\text{O}^-$ . Par contre, dans le cas de  $\text{OH(d)}$ , on note une inversion de ce transfert. Cette augmentation du transfert de charge vers le méthyle vient de l'absence de transfert vers la molécule d'eau.

Dans le cas d'une coupure intermoléculaire, le phénomène principal mis en évidence est la forte surestimation de l'interaction entre les deux sous-parties, qui est logique en regard de la nature du modèle d'eau utilisé. Ce phénomène peut d'autant plus

Complexe	QM	QM/MM
COOH	0,050	0,052
COO <sup>-</sup>	-0,016	-0,019
NH <sub>2</sub>	0,036	0,052
NH <sub>3</sub> <sup>+</sup>	0,145	0,159
OH (d)	0,005	-0,003
OH (a)	0,029	0,032
O <sup>-</sup>	-0,170	-0,201
SH (d)	0,021	0,022
SH (a)	0,038	0,050
S <sup>-</sup>	-0,048	-0,051

TAB. 3.2 – Somme des charges de Mulliken des atomes dans le groupement méthyle des différentes structures dans les approches QM et QM/MM.

être rattaché à cette molécule puisqu'il a un caractère constant pour tous les solutés envisagés ici.

### 3.2.3 Conclusions de l'étude

L'étude menée ici nous a permis de mettre en avant quelques points-clés concernant l'utilisation des méthodes hybrides pour décrire des systèmes interagissant avec une molécule d'eau par le biais d'une liaison hydrogène.

Il est primordial de bien choisir la position de la frontière quand celle-ci est au sein d'une molécule, en particulier en l'éloignant suffisamment de la zone d'interaction. Dans le cas de l'approche LSCF/MM, nous avons montré que les résultats pouvaient être sensiblement modifiés par l'utilisation d'une SLBO adaptée au système étudié.

Nous avons par ailleurs montré que la méthode est sensible au jeu de charges utilisé.

Un autre aspect mis en évidence dans cette étude est l'importance particulière du transfert de charge entre l'eau et le soluté lorsque ce dernier est chargé. Ce facteur nous amène donc à considérer que l'approche qui consiste à décrire l'eau par la mécanique moléculaire est inadaptée dans ce cas précis, le transfert de charge n'étant pas reproductible. Une modélisation de bonne qualité de tels systèmes ne peut se faire que par l'utilisation de molécules d'eau quantiques.

Le dernier point concerne plus particulièrement la molécule d'eau classique. Nous

avons ainsi mis en évidence que l'utilisation d'un modèle standard d'eau (ici TIP3P) était inadaptée dans notre approche puisque les énergies d'interaction calculées sont systématiquement surestimées en comparaison avec les valeurs de référence. Cette tendance doit être rattachée à l'utilisation de paramètres pour l'eau destinés à reproduire les caractéristiques du liquide et donc mal adaptés à notre approche.

Le fait que l'erreur commise ait un caractère constant nous permet alors d'envisager une possible amélioration de la description de la liaison hydrogène. Cette amélioration peut se faire en déterminant des paramètres classiques spécifiques pour la molécule d'eau. La suite de nos travaux est consacrée à cette voie.

### 3.3 Définition d'une approche soluté-solvant adaptée

Grâce à l'étude précédente, nous avons pu voir que le remplacement d'une molécule d'eau quantique par le modèle classique TIP3P conduisait à une surestimation de l'énergie d'interaction. Le caractère constant de cette erreur nous permet alors d'envisager une possible correction des paramètres classiques utilisés pour rendre compte de l'interaction entre parties quantique et classique.

Différents travaux ont été menés pour corriger ce type d'erreur. Nous citerons particulièrement ceux de Gao [110, 113] qui a proposé un jeu spécifique de paramètres de Lennard-Jones pour les atomes de la partie quantique, conduisant à une amélioration sensible de la description des interactions entre les deux sous-parties. Tu et Laaksonen [125] ont mené une détermination de ces paramètres pour l'eau à partir de deux configurations d'un dimère en reproduisant des calculs au niveau HF/6-311G(d,p). Ils ont par ailleurs montré que ce type de paramètre était difficilement transférable en phase liquide (un grand nombre de molécules classiques et des conditions périodiques aux limites). On notera alors l'approche de M.F. Ruiz-López *et al.* [116, 126, 127], visant à donner des paramètres de Lennard-Jones adaptés à une approche QM/MM dynamique dans l'état liquide.

Outre les paramètres de Lennard-Jones, on peut aussi envisager la modification des charges classiques entrant dans l'expression de l'hamiltonien du soluté quantique, comme l'ont fait Tu et Laaksonen [128]. Ces auteurs ont proposé d'appliquer un facteur d'échelle de 0,84 aux charges classiques entrant dans l'expression de l'hamiltonien de la partie quantique. Cette approche a été utilisée dans des conditions similaires à leur étude précédemment citée et a conduit à des résultats d'assez bonne qualité.

Comme nous l'avons déjà précisé dans ce chapitre, notre approche est assez différente puisque notre but n'est pas d'introduire un grand nombre de molécules de solvant mais quelques unes à proximité du centre réactif pour polariser ce dernier. De ce point de vue, nous restons dans le cadre de systèmes de type complexe 1 :1 (une molécule de soluté et une, voire plusieurs, molécule(s) d'eau). Cette particularité conditionne donc notre choix de procédure pour déterminer un schéma spécifique adapté à notre approche QM/MM.

### 3.3.1 Schéma général

Le principe que nous adoptons est donc de ne modifier que les interactions entre la partie quantique du soluté et les molécules de solvant, les paramètres classiques (entre molécules d'eau et entre l'eau et la partie classique du soluté) étant cohérents entre eux. Cette prise en compte différenciée se fera par le biais de la modification des paramètres de la molécule d'eau. Le tableau 3.3 illustre ce propos. Le choix de ne modifier que les paramètres de la molécule d'eau a été fait en considérant plusieurs facteurs. En effet, il aurait été aussi possible de modifier les paramètres de Lennard-Jones de la partie quantique, comme l'a fait Gao. Cela implique alors un plus grand nombre de paramètres à optimiser, ce qui nécessite un panel de solutés important. Se pose aussi le problème de la transférabilité d'un soluté à l'autre. De plus, le choix de s'intéresser à la molécule d'eau permet aussi de modifier les charges ponctuelles portées par cette dernière.

interaction entre	QM	MM (soluté)	MM (solvant)
QM	normale		
MM (soluté)	normale	normale	
MM (solvant)	<b>modifiée</b>	normale	normale

TAB. 3.3 – Table de vérité des interactions différenciées

### 3.3.2 Détermination des paramètres adaptés

Les paramètres que nous avons choisi de modifier sont ceux portés par la molécule d'eau. Pour cela nous choisissons de nous baser sur le modèle TIP3P [68], dont nous pourrions modifier les deux paramètres  $\sigma$  et  $\varepsilon$  du seul potentiel de Lennard-Jones porté

par l'atome d'oxygène. Les charges portées par les trois atomes pourront aussi être modifiées par le biais d'un facteur d'échelle. Le but est de garder à la fois l'électroneutralité de la molécule ainsi que la symétrie des charges. Nous nous ramenons ainsi à la détermination de trois paramètres.

Ces derniers doivent répondre à deux exigences. Ils devront en effet permettre de reproduire, dans un calcul hybride, l'énergie d'interaction entre un soluté et une molécule d'eau calculée par une approche quantique avec un haut niveau de théorie. L'autre critère est d'obtenir une géométrie proche de cette même référence, en particulier pour la longueur de liaison hydrogène.

Pour cela, nous choisissons de minimiser la différence entre les énergies d'interaction soluté/eau obtenues par une approche hybride et celles données par un calcul quantique sur des complexes 1 :1 équivalents (mêmes molécules). Seuls des complexes faisant interagir des solutés quantiques avec une seule molécule d'eau sont utilisés, puisque c'est l'interaction entre ces deux types de parties que nous voulons améliorer. Parmi les systèmes étudiés précédemment, nous n'utilisons que les six impliquant des solutés neutres (tab 3.1). L'étude précédente a en effet montré qu'une approche hybride soluté/solvant était mal adaptée à la description de la liaison hydrogène entre un soluté chargé décrit au niveau quantique et une molécule d'eau classique. Le transfert de charge entre les deux molécules pouvant être important.

Plutôt que d'augmenter le nombre de systèmes chimiques servant à modifier les paramètres, nous choisissons de garder les six systèmes cités. Nous voulons par ailleurs déterminer des paramètres qui reproduisent la forme de la surface d'énergie potentielle de référence au voisinage de la distance d'équilibre. Pour cela, il existe deux possibilités. La première consiste à calculer l'énergie et son gradient par rapport aux coordonnées nucléaires pour chacune des six structures. En requérant que non seulement la différence énergétique avec la référence soit minimale mais que le gradient soit aussi proche de zéro que possible, il est possible d'obtenir de bonnes énergies correspondant à des minima. Nous avons cependant opté pour la solution qui consiste, pour chaque système, à décrire trois géométries : La première avec la longueur de liaison H de la référence quantique, la deuxième avec une longueur plus courte de 0,1 Å et enfin la troisième avec une longueur plus longue de 0,1 Å. Ce choix permet aussi de tenir compte de la forme de la surface d'énergie potentielle tout en ne calculant que des énergies.

Nous avons déjà montré l'impossibilité de reproduire exactement une géométrie quantique avec un modèle hybride en raison du modèle de solvant. Nous cherche-

rons donc à ne reproduire que la longueur de la liaison hydrogène. C'est pourquoi les systèmes hybrides pris en compte ont été optimisés au niveau LSCF/MM avec une molécule d'eau TIP3P mais sous la contrainte de garder une longueur de liaison hydrogène égale à celle dans le complexe de référence.

En ce qui concerne les systèmes hybrides, nous utilisons toujours Amber pour les paramètres classiques.

Par contre, puisque nous ajustons des paramètres, nous prenons le parti de vouloir reproduire des résultats de haut niveau à l'aide d'une approche moyenne. Nous choisissons donc pour référence les énergies d'interaction (corrigées de la BSSE) des géométries optimisées au niveau MP2/6-311++G(2d,2p) et prenons comme niveau de théorie pour la partie quantique des calculs hybrides MP2/6-31G\*.

### Minimisation

Comme nous souhaitons minimiser la différence d'énergie entre une approche dépendant de trois paramètres et une référence, il est difficile de définir le gradient de cette grandeur en fonction de ces paramètres. Nous choisissons donc la méthode simplex qui ne nécessite pas le calcul des dérivées de la grandeur à minimiser en fonction des variables (référence [23], page 258).

Nous souhaitons reproduire l'énergie d'interaction des références quantiques au niveau hybride. Cela revient à minimiser la différence entre les énergies d'interactions obtenues avec les deux méthodes. Pour cela nous choisissons comme grandeur à minimiser (score, noté  $S$ ) la somme des carrés des différences entre les énergies d'interaction aux niveaux quantique et hybride pour un même système. Ce score s'écrit comme dans l'équation 3.4.

$$S(\sigma, \varepsilon, Q) = \sum_{i=1}^{18} (E_i^{QM/MM} - E_i^{ref.})^2 \quad (3.4)$$

Où  $E_i^{QM/MM}$  est l'énergie du système  $i$  calculé avec l'approche hybride, en fonction de  $\sigma$ ,  $\varepsilon$  et  $Q$  tandis que  $E_i^{ref.}$  est l'énergie de la référence quantique pour le système  $i$ .

Nous choisissons de modifier simultanément les trois paramètres, ce qui nous ramène à l'exploration d'une hypersurface à trois dimensions. Le simplex demande, pour un problème à  $N$  dimensions,  $N+1$  point (ici quatre) dont nous notons les scores respectifs  $S_1, S_2, S_3, S_4$ .

Le mode de fonctionnement consiste à déterminer, parmi ces quatre valeurs laquelle

est la plus grande. On détermine ensuite un nouveau point (triplet  $\sigma$ ,  $\varepsilon$  et  $Q$ ) pour lequel on calcule un nouveau score. Si le nouveau point donne un score meilleur (plus bas), alors il remplace l'ancien. Ce processus est reconduit avec les trois anciens points et le nouveau jusqu'à la convergence.

Pour déterminer les coordonnées du nouveau point, on commence par placer le barycentre des trois points dont le score est le plus bas. Les nouvelles coordonnées sont obtenues par réflexion du point dont le score est le plus haut par rapport au barycentre. Si le score du nouveau point est meilleur, il vient remplacer le point le plus mauvais. Dans le cas contraire, un nouveau point est déterminé comme le milieu du segment reliant le point à remplacer et le barycentre (contraction unidimensionnelle). Si ce point ne conduit toujours pas à un score inférieur, on effectue une contraction autour du point de score minimum. Cela consiste à déterminer le point de score le plus bas, et rapprocher tous les autres points de ce dernier.

Comme critère de convergence, plutôt que de se baser sur une différence de score entre pas successifs, nous avons considéré la différence pour chacun des paramètres entre les quatre points. Cela signifie que le processus s'achève lorsque les points sont suffisamment proches les uns des autres. Ce choix qui rend plus lente la convergence a été fait en raison du fort risque d'avoir une hypersurface très plate, pour laquelle des points aux coordonnées très éloignées risquent de donner des scores suffisamment proches pour arrêter le processus alors que les quatre points ont des coordonnées très différentes.

L'optimisation des paramètres est terminée lorsque les trois expressions suivantes sont satisfaites simultanément :

$$\sum_{i=1}^4 \sqrt{\left(\sigma_i - \frac{1}{4} \sum_{j=1}^4 \sigma_j\right)^2} < \delta_\sigma \quad (3.5)$$

$$\sum_{i=1}^4 \sqrt{\left(\varepsilon_i - \frac{1}{4} \sum_{j=1}^4 \varepsilon_j\right)^2} < \delta_\varepsilon \quad (3.6)$$

$$\sum_{i=1}^4 \sqrt{\left(Q_i - \frac{1}{4} \sum_{j=1}^4 Q_j\right)^2} < \delta_Q \quad (3.7)$$

avec  $\delta_\alpha$  ( $\alpha = \sigma, \varepsilon$  ou  $Q$ ) une constante avec les unités respectives des grandeurs associées. La valeur de cette dernière est la précision avec laquelle on veut déterminer les paramètres.

## Résultats

Les valeurs sont données avec une précision de  $10^{-4}$  à chaque fois et dans les unités respectives des paramètres ( $\text{\AA}$  pour  $\sigma$ ,  $\text{kcal.mol}^{-1}$  pour  $\varepsilon$  et sans unité pour  $Q$ , le facteur d'échelle sur les charges). Les paramètres après optimisation sont 2,0627 pour  $\sigma$ , 0,0245 pour  $\varepsilon$  et un facteur d'échelle de 0,8302 sur les charges ponctuelles. Ces valeurs ont été confirmées en calculant les points voisins de  $\pm 10^{-4}$  unités respectives pour chacun des trois paramètres (six points en tout) qui donnent tous une valeur supérieure à celle obtenue avec les paramètres optimisés. Cela confirme ainsi que les paramètres retenus sont les plus adaptés.

Si on compare ces valeurs avec celles caractérisant TIP3P (tableau 3.4), on voit que le rayon de van der Waals du paramètre de Lennard-Jones est plus grand que celui du modèle TIP3P (1,7683), ce qui va dans le sens d'un allongement de la distance d'équilibre entre l'oxygène et les autres atomes. D'autre part,  $\varepsilon$  est beaucoup moins important (TIP3P : 0,1520) ce qui induit un "puits" moins profond. Cela va dans le sens d'une diminution de l'énergie d'interaction mais implique un potentiel plus "plat" dans la région de la distance d'équilibre, ce qui peut être une conséquence d'une détermination faite avec plusieurs systèmes différents. Enfin, la valeur de 0,8302 du facteur d'échelle conduit à une molécule d'eau, dans sa géométrie relaxée, avec un moment dipolaire de 1,95 Debye. Cette grandeur doit être comparée à la valeur expérimentale de 1,86 en phase gazeuse, conditions dans lesquelles se trouvent les complexes utilisés pour la détermination. Cette valeur peut aussi être comparée aux 2,06 Debye obtenus pour une molécule d'eau optimisée au niveau de théorie MP2/6-311++G(2d,2p), utilisé pour calculer les énergies de référence.

modèle	$\sigma$	$\varepsilon$	$Q$
TIP3P	1,7683	0,1520	1,0000
Nos paramètres	2,0627	0,0245	0,8302
Laaksonen	2,1700	0,0410	1,0000

TAB. 3.4 – Paramètres des différents modèles d'eau.  $\sigma$  en donné en  $\text{\AA}$  et  $\varepsilon$  en  $\text{kcal.mol}^{-1}$ .

Nous pouvons par ailleurs comparer nos résultats à ceux obtenus par Tu et Laaksonen (tableau 3.4). Dans une première étude, ils ont commencé par ne modifier que les paramètres de Lennard-Jones pour reproduire l'énergie d'interaction d'un dimère de l'eau mais sans changer les charges [125]. Les paramètres ainsi obtenus sont :  $\sigma =$

2,17 et  $\varepsilon = 0,041$ . Ces valeurs sont de l'ordre de celles obtenues dans notre cas. Une autre étude des mêmes auteurs [128] ne concernait que la modification d'un facteur d'échelle sur les charges entrant dans l'hamiltonien de la partie classique. La valeur alors déterminée pour reproduire de bons résultats pour une molécule d'eau quantique dans une boîte de molécules classiques est de 0,84, proche de celle que nous proposons.

### 3.3.3 Test des paramètres

#### Test sur les complexes de détermination

Afin de tester notre jeu de paramètres adapté à une approche QM/MM, nous commençons naturellement par l'utiliser pour calculer les énergies d'interaction dans les systèmes utilisés pour la détermination des paramètres. Dans un premier temps, nous prenons donc en compte les six complexes neutres dont nous optimisons la géométrie au niveau MP2/6-31G\*. Ce dernier est celui utilisé dans la détermination des paramètres. Les valeurs obtenues (énergies d'interaction et longueurs de liaison hydrogène) sont comparées à celles de référence (uniquement celles dont la géométrie est optimisée) alors prises en compte (niveau MP2/6-311++G(2d,2p)), les valeurs sont données en annexe (tableaux D.4 et D.5).

L'erreur moyenne absolue concernant l'énergie est de 0,48 kcal.mol<sup>-1</sup> avec un maximum de 1,13 kcal.mol<sup>-1</sup> pour le complexe OH(d) de l'éthanol. De même, si on prend en compte les sept liaisons hydrogène créées (deux pour le complexe de l'acide carboxylique), l'erreur moyenne absolue est de 0,104 Å, avec un maximum de 0,23 pour le complexe SH(d) de l'éthanthiol.

Ces valeurs montrent que les paramètres que nous avons déterminé conduisent à des valeurs de l'énergie d'interaction très proches de la référence quantique, ce qui est le facteur que nous avons choisi de privilégier pour obtenir les nouveaux paramètres. L'accord est cependant en deçà de nos attentes en ce qui concerne les distances.

Dans les mêmes conditions, nous appliquons aussi notre modèle dans le cas de solutés chargés, même si nous considérons que nous sortons du domaine d'application du modèle d'eau adapté, les valeurs sont aussi données en annexe (tableaux D.4 et D.5).

Avant de passer aux résultats proprement dit, on peut remarquer que les énergies de référence calculées à l'aide d'une base contenant des orbitales diffuses sont assez différents des valeurs obtenues avec la base 6-311G(2d,2p) (tableau D.2). On observe

en particulier une différence de plus de 3 kcal.mol<sup>-1</sup> dans le cas de l'éthanolate, venant ainsi confirmer l'influence de l'utilisation d'orbitales diffuses.

Pour les espèces chargées, non utilisées pour la détermination, la différence moyenne en terme d'énergie d'interaction par rapport à la référence est de 2,98 kcal.mol<sup>-1</sup>. Le maximum est atteint avec l'éthanolate pour lequel on sous-estime cette énergie de plus de 7 kcal.mol<sup>-1</sup>. Du point de vue géométrique, les longueurs de liaisons H associées sont prédites plus longues en moyenne de 0,107 Å comparées aux valeurs de référence.

Bien qu'il soit préférable de ne pas utiliser une approche hybride dans le cas d'un soluté chargé, on note que nos paramètres conduisent quand même à l'amélioration des résultats, en particulier dans le cas du complexe S<sup>-</sup>.

Cette première série de tests vient donc confirmer que le jeu de paramètres que nous avons déterminé permet d'obtenir avec une bonne précision l'énergie d'interaction entre un soluté quantique neutre et une molécule d'eau traitée par la mécanique moléculaire.

### **Transférabilité de base et méthode.**

Nous avons pu voir que les paramètres adaptés permettaient d'obtenir des énergies d'interaction d'assez bonne qualité dans les cas des complexes précédemment étudiés. Pour déterminer ces paramètres, nous avons choisi de travailler avec une base de niveau moyen pour reproduire des résultats de haut niveau afin de jouer, entre autres, dans le sens de la transférabilité d'une base à l'autre. Pour tester cette propriété, nous prenons en compte un système comportant une molécule d'éthanol quantique en interaction avec une (correspondant au complexe OH (a)) ou deux molécules classiques d'eau (figure 3.8), faisant alors intervenir la différenciation dans le calcul des interactions. L'interaction entre molécules d'eau se fait avec les paramètres TIP3P alors que nos paramètres sont utilisés pour l'interaction entre le soluté et les molécules d'eau.

La géométrie de ce système est optimisée en prenant en compte, pour la partie quantique, les approches Hartree-Fock, B3LYP et MP2 à chaque fois avec quatre bases d'orbitales différentes : 3-21G, 6-31G\*, 6-311G(d,p) et 6-311++G(2d,2p). Les valeurs des énergies calculées sont reportées en annexe (tableaux D.6 et D.7) et présentées dans la figure 3.9. Elles y sont comparées à l'énergie de référence (corrigée de l'erreur de superposition de base) obtenue après optimisation de la géométrie des complexes pris en compte au niveau quantique MP2/6-311++G(2d,2p). A titre de comparaison, les résultats obtenus dans les mêmes conditions mais avec les paramètres provenant de TIP3P sont aussi présentés.

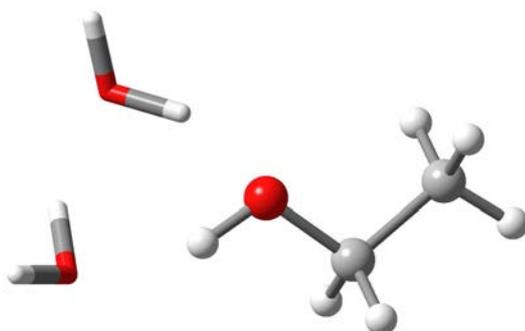
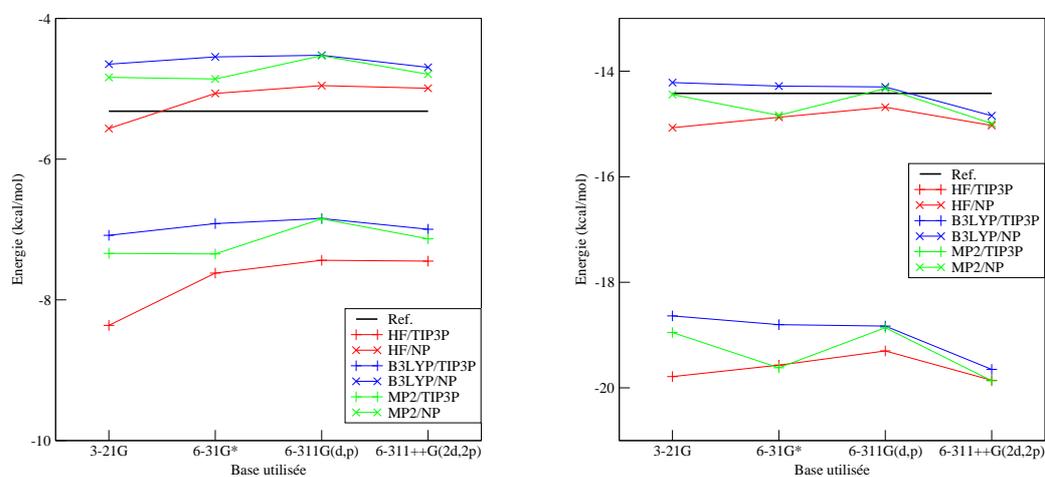


FIG. 3.8 – Complexe de l'éthanol avec deux molécules d'eau



(a) Ethanol en interaction avec une molécule classique d'eau  
 (b) Ethanol en interaction avec deux molécules classiques d'eau

FIG. 3.9 – Energies d'interaction pour l'éthanol quantique en interaction avec une (3.9(a)) ou deux (3.9(b)) molécules d'eau classiques. Le référence est MP2/6-311++G(2d,2p), NP se rapporte aux "Nouveaux Paramètres"

Les deux figures 3.9(a) et 3.9(b) montrent plusieurs aspects intéressants des effets de l'introduction de paramètres spécifiques. Dans un premier temps, on peut voir que ces derniers corrigent les effets de ceux de TIP3P puisque dans tous les cas, l'énergie calculée est plus proche de la référence pour un niveau de théorie identique (méthode et base). Cependant le phénomène le plus remarquable, et ce pour les deux types de paramètres, est la constance des résultats, quelles que soient la méthode et la base. En effet, l'écart maximum entre deux résultats pour un même jeu de paramètres, en prenant en compte les trois méthodes et les quatre types de bases ne dépasse pas  $1,5 \text{ kcal.mol}^{-1}$  dans les deux systèmes. Ceci nous permet donc de dire que les paramètres classiques ont de bonnes qualités de transférabilité d'une base à une autre mais aussi d'une méthode à l'autre. Sur ce sujet, nous pouvons par ailleurs remarquer que, dans les deux systèmes étudiés, les différences d'énergies pour un même jeu de paramètres classiques sont moins importantes lorsqu'on utilise les nôtres. On peut en déduire que ces derniers ont une meilleure transférabilité que ceux de TIP3P. La dernière tendance générale qu'on peut noter concerne les différentes valeurs de l'énergie pour un même jeu de paramètres et une même base (c'est à dire ou seule la méthode quantique change). En effet, on constate que plus la base est importante, plus les résultats sont proches lorsqu'on compare les méthodes entre elles. Concernant ces mêmes valeurs, on peut par ailleurs voir que l'énergie baisse systématiquement lorsqu'on passe d'une base doublement polarisée à la base contenant des orbitales diffuses. Cet effet lié à la base peut sans doute être relié au phénomène déjà évoqué de "surpolarisation" par la présence d'une charge classique à proximité d'un soluté quantique.

Concernant les valeurs des énergies d'interaction, on peut remarquer, dans le cas du complexe avec une seule molécule d'eau, une tendance à sous-estimer cette énergie, comme il avait été constaté sur le même système précédemment (complexe OH(d)). Seule l'approche HF/3-21G conduit à une énergie plus grande. L'erreur maximale étant commise avec l'approche B3LYP/6-311G(d,p) avec une valeur de  $0,8 \text{ kcal.mol}^{-1}$ .

Par contre, dans le cas de l'éthanol en interaction avec deux molécules d'eau, l'accord est bon alors qu'on pourrait s'attendre à retrouver une telle sous-estimation. On compte ici trois liaisons hydrogène : deux entre l'éthanol et les deux molécules d'eau mais aussi une troisième entre ces deux dernières. En raison de l'approche à interaction différenciée adoptée, l'interaction entre les deux molécules d'eau est calculée en prenant en compte les paramètres TIP3P standards. Or l'énergie d'interaction du dimère d'eau dans ce cas est de  $-6.74 \text{ kcal.mol}^{-1}$ , soit une surestimation

d'environ  $1,3 \text{ kcal.mol}^{-1}$  par rapport à la valeur expérimentale généralement retenue ( $5,4 \pm 0,7 \text{ kcal.mol}^{-1}$ ) et  $1,8 \text{ kcal.mol}^{-1}$  par rapport aux valeurs théoriques les plus précises ( $5,0 \pm 0,7 \text{ kcal.mol}^{-1}$ ) [125, 129, 130].

Cette surestimation de l'interaction entre deux molécules d'eau TIP3P peut donc venir compenser l'erreur soulignée pour le cas du monomère.

Cette application nous montre donc que nos paramètres ont pour qualité de conduire à des résultats constants quelles que soient la base ou la méthode utilisées pour la partie quantique. Par ailleurs, dans les cas présentés ici, si on compare cette transférabilité à celle des paramètres TIP3P, on se rend compte que les résultats sont moins différents de la référence.

La dernière question qui se pose concerne la transférabilité d'un champ de forces à un autre. En effet, depuis le début, nous travaillons avec Amber [20], dont les paramètres d'interactions intermoléculaires sont compatibles, par construction, avec le modèle TIP3P. Du fait que les autres champs de forces présents dans Gaussian (UFF et Dreiding) n'incorporent pas directement ce modèle, il est difficile de le tester. La seule supposition quant à l'utilisation de notre modèle est que celui-ci est compatible avec les champs de force incorporant le modèle TIP3P, tels que Charmm ou OPLS [131].

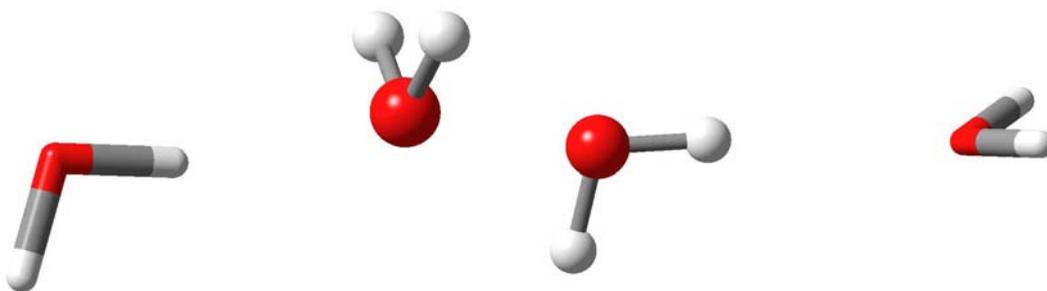
### 3.3.4 Application au dimère de l'eau

Nous avons pu voir que les paramètres adaptés permettaient de reproduire les énergies d'interaction dans les cas des systèmes avec lesquels ils ont été déterminés. Nous procédons au test de ces derniers sur un système avec lequel nos paramètres n'ont pas été ajustés, à savoir le dimère de l'eau. L'intérêt de ce système est double. Le premier réside dans le nombre de résultats existants, à la fois expérimentaux et théoriques notamment au niveau quantique mais aussi dans le cadre d'ajustements de paramètres. Nous souhaitons en particulier comparer nos résultats à ceux obtenus à l'aide du jeu de paramètres déterminé par Tu et Laaksonen [125] dans le but de reproduire les caractéristiques du dimère.

D'autre part, même si nous travaillons dans l'optique de solutés hybrides, la bonne description de l'interaction entre une molécule d'eau classique et une autre quantique est importante. On peut en effet être amené à introduire dans un système une ou plusieurs molécules d'eau quantiques pouvant intervenir directement (par le biais d'un échange protonique, par exemple) dans le processus réactionnel. L'interaction entre les

deux descriptions de  $\text{H}_2\text{O}$  devra alors être de bonne qualité.

Afin de comparer directement nos résultats à ceux de la référence [125], nous travaillons au même niveau, à savoir HF/6-311G(d,p) pour la partie quantique du système. De la même façon, en raison de la présence d'une molécule classique et d'une molécule quantique, il existe deux cas suivant que la molécule quantique est donneuse ou acceptrice de proton. Ces deux cas sont représentés dans la figure 3.10. On note par ailleurs que les paramètres classiques non liés (Lennard Jones) pour les atomes quantiques sont, comme dans le cas de l'étude de référence (référence [125], table II) ceux du modèle TIP3P.



(a) Dimère avec molécule quantique acceptrice de proton (b) Dimère avec molécule quantique donneuse de proton

FIG. 3.10 – Les deux formes du dimère de l'eau en QM/MM : la molécule quantique est celle qui est représentée en "boules et bâtons".

On donne les énergies d'interaction et les données géométriques calculés avec nos paramètres ainsi que ceux de la littérature dans le tableau 3.5. Pour mémoire, les valeurs de l'énergie d'interaction de référence sont respectivement  $5,4 \pm 0,7 \text{ kcal.mol}^{-1}$  [132] pour la grandeur expérimentale et  $5,0 \pm 0,7 \text{ kcal.mol}^{-1}$  pour la valeur théorique la plus précise [125, 129, 130]. La distance inter atomique expérimentale entre les deux oxygènes retenue est  $2,952 \text{ \AA}$  [133], les calculs théoriques poussés déjà mentionnés donnant une valeur similaire ( $2,925 \text{ \AA}$  dans la référence [130]).

On peut voir que, par rapport à une approche visant à reproduire les caractéristiques du dimère de l'eau, nos paramètres donnent des résultats comparables. En effet, dans les deux cas, l'énergie calculée diffère de moins de  $0,2 \text{ kcal.mol}^{-1}$  de celle citée. De même, les angles ont presque la même valeur. Le seul point sur lequel nos paramètres sont moins bons que ceux de la littérature concerne la distance  $\text{O} \cdots \text{O}$  qui est sous-estimée dans notre cas de plus de  $0,1 \text{ \AA}$  par rapport à la référence ou à la valeur expérimentale.

	$E_{int.}$ (kcal.mol <sup>-1</sup> )		distance O...O (Å)		angle O...H-O (degrés)	
	Laaksonen	N. P.	Laaksonen	N. P.	Laaksonen	N. P.
QM donneur	5,32	5,19	2,93	2,79	177,3	178,5
QM accepteur	5,13	4,94	2,95	2,82	175,6	177,0

TAB. 3.5 – Energies d’interactions et paramètres géométriques du dimère de l’eau traité en QM/MM. Laaksonen correspond aux valeurs obtenues par cet auteur avec le jeu de paramètres déjà donné (tableau 3.4) et N.P. est pour nos paramètres.

Bien qu’il n’ait pas été spécifiquement développé pour ce type de système, notre jeu de paramètres conduit quand même à des résultats assez satisfaisants. En effet, les énergies d’interaction sont en très bon accord avec les résultats obtenus à l’aide d’un modèle développé explicitement pour reproduire les caractéristiques du dimère de l’eau [125].

Ce court exemple nous permet par ailleurs de confirmer les tendances de notre approche (au moins en phase gazeuse). Nous avons d’une part une bonne reproduction de l’énergie d’interaction entre un soluté quantique et une molécule d’eau classique et d’autre part une sous-estimation de la longueur de la liaison hydrogène.

Nous pouvons donc déduire que dans le cas d’un système incluant aussi une ou quelques molécules d’eau traitées au niveau quantique, notre modèle pourra être utilisé.

Avec les différentes applications présentées, nous avons pu montrer que nos paramètres permettaient de décrire de façon satisfaisante la liaison hydrogène, au moins du point de vue énergétique. Cependant, notre but est plutôt de modéliser ses effets sur la réactivité d’un soluté quantique ou d’un soluté décrit par une approche hybride QM/MM. C’est à cet aspect que nous nous intéressons maintenant.

### 3.4 Effets du solvant sur le réarrangement de Claisen

Afin d’illustrer l’utilisation de notre approche, nous proposons de la mettre en œuvre dans l’étude des effets du solvant sur le mécanisme du réarrangement de Claisen dans le cas de l’allyle vinyle ether qui est l’exemple le plus simple pour cette réaction (figure 3.11).

Les deux principaux points d’intérêt pour ce mécanisme sont d’une part son appli-

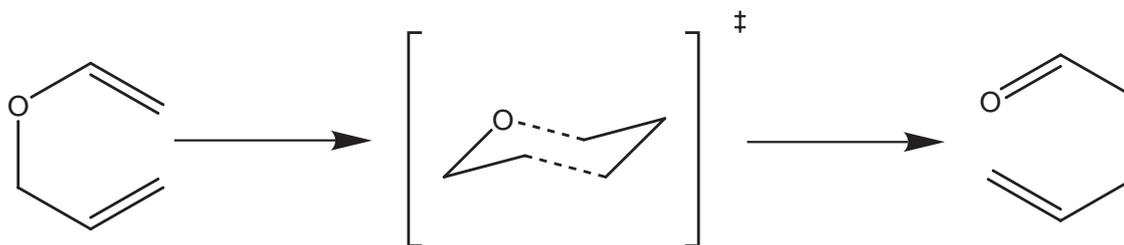


FIG. 3.11 – réarrangement de Claisen pour l'allyle vinyle éther

cation en synthèse organique mais aussi dans le domaine biologique. En effet, c'est par un tel mécanisme que l'enzyme chorismate-mutase catalyse la transformation du chorismate en préphenate [118]. L'aspect qui nous intéresse particulièrement ici concerne la modification du mécanisme réactionnel induit par la présence du solvant. Il est généralement admis que, d'une façon analogue à ce qui a été observé dans le cas de la réaction de Diels-Alder [95], le solvant a pour effet de déplacer l'état de transition par rapport à sa position lorsque la réaction se produit en phase gazeuse. Dans le cas de l'allyle vinyle éther, l'eau a pour effet de déplacer la structure de transition vers la forme aldéhyde (mieux solvatée que la forme éther), c'est-à-dire le long de la coordonnée de réaction.

En phase gazeuse (dans le cas d'une inter conversion thermique), l'état de transition entre les deux formes a été déterminé comme étant "aromatique", la liaison C-O et la création de la liaison C-C se font de manière concertée. A l'opposé, un autre mécanisme, impliquant un intermédiaire à séparation de charges (zwitterionique) est favorisé par la présence d'un solvant polaire. Dans ce cas, la structure de transition est déplacée, dans l'espace des degrés de liberté internes dans une direction orthogonale à la coordonnée de réaction [134].

Expérimentalement, on constate que le réarrangement de Claisen est plus de deux cents fois plus rapide dans l'eau que dans le cyclohexane [135]. D'après cette même référence, deux facteurs peuvent expliquer cet effet : d'une part, l'hydrophobicité, et d'autre part la présence de liaisons hydrogène entre une ou plusieurs molécules de solvant (l'eau) et l'oxygène.

Le premier facteur a été étudié théoriquement grâce à l'utilisation de modèles implicites de solvant. Nous citerons ainsi les travaux de Hillier *et al.* [136, 137]. Cet auteur a aussi conduit une étude analogue à celle présentée ici [138] puisque le système étudié était composé du soluté en interaction avec deux molécules d'eau, l'ensemble étant

plongé dans un continuum. Cette étude a d'ailleurs permis de montrer que le facteur principal d'abaissement de la barrière d'activation de la réaction de Claisen provenait principalement de la formation de liaisons hydrogène avec des molécules d'eau.

En ce qui concerne la formation de liaisons hydrogène, Severance et Jorgensen [139] ont étudié cette réaction en modélisant de façon explicite le solvant et en utilisant une approche statistique de type Monte-Carlo. Ils ont entre autre calculé le nombre moyen de liaisons hydrogène formées au cours de la réaction (de l'éther vers l'aldéhyde). Ce dernier passe de 0,9 pour l'éther à 1,9 pour l'état de transition et vaut 1,7 pour le réactif.

De façon générale, la plus importante polarisation de la structure de transition et du produit par rapport au réactif tend à accélérer le processus de conversion.

Nous citerons aussi les travaux de Sehgal et Gao [140] visant à déterminer les effets de la présence de différents substituants sur ce mécanisme dans une approche de type soluté quantique/eau classique.

Cette réaction impliquant un soluté de taille réduite en interaction avec un nombre restreint de molécules d'eau est donc un candidat de choix pour tester notre approche. Afin d'en vérifier la validité, nous considérons ici le réarrangement en présence d'une seule molécule d'eau classique. Les résultats obtenus seront comparés à une approche complètement quantique. De même, cette comparaison pourra se faire en prenant en compte les effets du solvant à l'aide du modèle de continuum. Notre but n'est pas ici l'étude approfondie des effets du solvant sur cette réaction mais plutôt, encore une fois, la capacité de notre modèle d'eau explicite à reproduire les effets d'une molécule quantique. Nous procédons donc ici principalement à la comparaison des modifications sur la réactivité de l'allyle vinyle ether induites par les deux approches et en prenant les données des calculs quantiques comme référence.

### 3.4.1 Détails calculatoires

Notre approche est de type soluté/solvant et ne nécessite pas, ici, de détermination d'orbitale frontière. Pour la partie quantique, nous choisissons de travailler au niveau B3LYP/6-31G\*\*, déjà utilisé par Hammes-Schiffer *et al.* [141]. La fonctionnelle B3LYP a montré dans plusieurs études sur le réarrangement de Claisen un bon accord avec des calculs au niveau MP2 [141, 142].

Pour les paramètres classiques, nous utilisons ceux qui ont été détaillés dans ce chapitre pour l'eau ainsi que les paramètres non liés de Amber [20] pour les atomes

quantiques.

Pour le modèle de solvant implicite, nous choisissons d'utiliser la constante diélectrique relative de l'eau ( $\epsilon_0 = 78,4$ ). La cavité sera de forme adaptée à la géométrie moléculaire et le développement multipolaire, poussé jusqu'à l'ordre six sera fait en un seul point. Enfin, nous choisissons de multiplier les rayons de van der Waals utilisés pour le calcul du volume de la cavité par un facteur deux au lieu de la valeur par défaut. Ce choix est fait pour assurer une convergence plus facile du développement multipolaire et des optimisations de géométrie. Cependant, les effets du solvant tels que décrits par le continuum seront sous-estimés.

Cette approche nous permet par ailleurs de vérifier la compatibilité de nos paramètres classiques avec le modèle de continuum.

Nous envisageons deux types de complexes différents : le premier ne comporte que le système réactif alors que dans le deuxième, une molécule d'eau est ajoutée et forme une liaison hydrogène avec l'oxygène de la molécule. En ce qui concerne les états de transition, nous choisissons de ne prendre en compte que ceux passant par la configuration "chaise", décrite comme la plus stable [143].

### 3.4.2 Résultats, discussion

#### Description de la liaison hydrogène

Comme dans les études précédentes, nous voulons d'abord vérifier que notre modèle reproduit suffisamment bien les résultats de référence obtenus avec une approche complètement quantique. Pour cela, nous étudions les énergies d'interaction (tableau 3.6) mais uniquement dans le vide, le calcul de correction d'erreur de superposition de base dans le continuum n'étant pas défini. De plus, techniquement, il est encore impossible d'étudier un soluté classique seul dans le continuum. Les longueurs de liaisons hydrogène formées à la fois dans le vide et dans le continuum sont aussi données dans le tableau 3.7.

Les énergies d'interaction de liaison hydrogène (corrigées de l'erreur de superposition de base pour la référence quantique) sont très bien reproduites lorsqu'on utilise notre approche hybride. L'erreur maximum étant commise sur la structure de l'éther. Ces valeurs viennent donc encore une fois confirmer la bonne capacité de notre modèle à reproduire les énergies d'interaction de la liaison hydrogène.

Nous pouvons remarquer que, comme noté précédemment, notre approche tend à donner des liaisons un peu plus courtes que celles de référence, à la fois dans le vide et

$E_{int.}$	H <sub>2</sub> O QM	H <sub>2</sub> O MM	déviaton
Ether	-2,731	-3,269	-0,538
TS	-5,116	-5,032	0,085
aldéhyde	-4,186	-4,199	-0,012

TAB. 3.6 – Energies d’interaction en kcal.mol<sup>-1</sup> dans le vide entre le réactif et une molécule d’eau suivant l’approche quantique ou hybride. La différence d’énergie calculée entre les deux approches est aussi donnée.

liaison H	ether		TS		aldéhyde	
	QM	QM/MM	QM	QM/MM	QM	QM/MM
dans le vide	2,026	1,899	1,869	1,826	1,972	1,865
dans le continuum	1,991	1,913	1,868	1,824	1,928	1,870
différence	-0,035	0,014	-0,001	-0,002	-0,044	0,005

TAB. 3.7 – Longueurs des liaisons hydrogène en Å formées suivant l’approche (quantique ou hybride), dans le vide et avec le continuum.

dans le continuum. Par contre, l’effet du passage dans le continuum (ligne ”différence” du tableau) est différent suivant l’approche (soluté seul, H<sub>2</sub>O quantique ou classique). Si on regarde la variation de la longueur de liaison hydrogène, on voit que celle ci est raccourcie pour le réactif et le produit lorsque la molécule d’eau est quantique alors que le phénomène inverse se produit pour une molécule classique. Néanmoins, ces variations sont minimales et même difficilement interprétables dans le cas des états de transition.

L’aspect qui nous intéresse plus particulièrement dans cette étude concerne l’influence du solvant sur la réactivité du soluté. Pour cela, nous prenons en compte à la fois les énergies (de réaction et d’activation), reportées dans le tableau 3.8, ainsi que les aspects géométriques avec, en particulier, les longueurs de liaisons en cours de modification. Nous mesurons enfin les effets du solvant sur la séparation de charge induite à l’aide des charges de Mulliken.

### Aspect énergétique

Expérimentalement, Hall *et al.* rapportent un abaissement de la barrière d’activation en passant du vide dans l’eau d’environ 14 à 19 kJ.mol<sup>-1</sup> (3,5 à 4,7 kcal.mol<sup>-1</sup>

environ) [137, 138]. Dans leur étude au niveau BLYP/6-31G\* et utilisant uniquement un modèle de continuum, ces mêmes auteurs calculent une valeur de 1,2 kJ.mol<sup>-1</sup> (0,29 kcal.mol<sup>-1</sup>). Par contre dans une étude ultérieure du même auteur, où deux molécules d'eau sont ajoutées au soluté au sein du continuum[138], l'abaissement total calculé vaut 5,5 kcal.mol<sup>-1</sup> au niveau MP2/6-31G\*, en tenant compte, comme ici, des seules énergies internes. Dans notre cas, on peut s'attendre à un abaissement moins important en raison de l'inclusion d'une seule molécule d'eau.

Les valeurs reportées dans le tableau 3.8 montrent que lorsqu'on somme les effets du continuum et de l'ajout de la molécule d'eau quantique, on obtient un abaissement de la barrière de 3,94 kcal.mol<sup>-1</sup>. Cette valeur est du même ordre que celle déjà mentionnée mais est effectivement inférieure. Il est ensuite difficile de comparer directement nos résultats à ceux pris en référence, les systèmes chimiques ainsi que les méthodes (niveaux de théories et modèles de continuum) étant différents dans les deux cas.

	$E_{int.}$ kcal.mol <sup>-1</sup>	sans molécule d'eau	H <sub>2</sub> O quantique	H <sub>2</sub> O classique
Dans le vide	$\Delta E_a$	27,69	24,65	25,93
	$\Delta E_r$	-17,86	-19,42	-18,79
Dans le continuum	$\Delta E_a$	27,26	23,75	25,58
	$\Delta E_r$	-18,40	-19,54	-18,99

TAB. 3.8 – Energies d'activation et de réaction pour les différentes approches : dans le vide, eau traitée au niveau quantique ou au niveau classique. les valeurs analogues sont aussi reportées pour les calculs conduits dans le continuum

Lorsqu'on étudie séparément les effets de chacune des description du solvant (implicite et explicite), on voit que c'est l'ajout d'une molécule d'eau qui est le principale facteur d'abaissement de la barrière d'activation. Ainsi en partant du soluté dans le vide et en le plaçant dans le continuum, on observe un gain énergétique de 0,42 kcal.mol<sup>-1</sup>, du même ordre que celui décrit dans la référence [137]. Par contre, la barrière est abaissée de 3,041 kcal.mol<sup>-1</sup> lorsqu'on étudie le soluté dans le vide en interaction avec une molécule d'eau. Ces données sont donc en accord avec les différentes études évoquées selon lesquelles le facteur majoritaire d'abaissement de la barrière d'activation provient de la formation d'une (ou plusieurs) liaison(s) hydrogène avec l'oxygène. Ces résultats montrent donc que la référence que nous avons choisi est de bonne qualité.

Si on s'intéresse maintenant aux résultats obtenus à l'aide de notre modèle hybride, on remarque que la barrière est effectivement globalement abaissée mais de seulement  $1,68 \text{ kcal.mol}^{-1}$ . Le passage du soluté seul à un soluté en interaction avec l'eau et dans le continuum peut se faire de deux façon différentes : soit en ajoutant une molécule d'eau au soluté déjà placé dans le continuum, soit en plongeant le soluté en interaction avec une molécule d'eau dans le continuum. On remarque que dans le deuxième cas, la deuxième étape (mise dans le continuum du complexe soluté+eau) fait légèrement augmenter la barrière d'activation de la réaction par rapport à ce qu'on observe dans le vide.

En comparaison avec les résultats de référence, notre modèle modifie la grandeur étudiée dans le bon sens mais ne reproduit pas entièrement la modification. En effet même si on prend en compte l'erreur sur l'énergie d'interaction de l'éther avec l'eau (de  $-0,54 \text{ kcal.mol}^{-1}$ ), on ne parvient à retrouver la hauteur de barrière calculée avec une molécule d'eau quantique.

En ce qui concerne l'énergie de la réaction, exothermique de  $17,85 \text{ kcal.mol}^{-1}$  dans le vide, on obtient une valeur légèrement plus importante dans les deux cas impliquant la présence d'eau. De nouveau, l'introduction d'une molécule classique conduit à un abaissement plus faible. Cette différence est de l'ordre de  $0,63 \text{ kcal.mol}^{-1}$  dans le vide et  $0,55 \text{ kcal.mol}^{-1}$  dans le continuum. Ces valeurs sont du même ordre que l'erreur commise sur l'énergie d'interaction soluté-eau entre les approches quantiques et hybrides pour le réactif. Le signe de cette erreur implique par ailleurs un abaissement de l'énergie de l'éther, ce qui conduit alors à une énergie de réaction libérée moins importante. On peut donc estimer que, dans la marge d'erreur fixée par les énergies d'interaction, l'énergie de réaction peut être considérée comme assez bien reproduite par l'approche hybride.

### Aspect géométrique

Pour étudier l'influence du solvant sur le mécanisme réactionnel, on s'intéresse aux liaisons en cours de modification. On reporte donc dans les tableaux 3.9 et 3.10 les valeurs des longueurs de liaisons des différents produits, états de transition et réactifs. Les distances carbone-carbone dans le réactif et carbone-oxygène dans le produit sont présentés à titre indicatif (on ne peut pas parler de liaison). Elles montrent seulement que les géométries prises en compte pour les réactifs et les produits sont analogues pour les différentes approches.

	liaison	Réactif	TS	Produit
soluté	C...O	1,424	1,914	3,536
seul	C...C	3,674	2,322	1,537
H <sub>2</sub> O	C...O	1,433	1,949	3,539
QM	C...C	3,654	2,392	1,537
H <sub>2</sub> O	C...O	1,428	1,935	3,532
MM	C...C	3,683	2,359	1,537

TAB. 3.9 – longueurs des liaisons (Å) modifiées au cours de la réaction de Claisen pour les différentes approches (sans molécule d'eau, avec une molécule quantique ou avec une classique) sans le continuum

	liaison	Réactif	TS	Produit
soluté	C...O	1,425	1,920	3,564
seul	C...C	3,675	2,333	1,537
H <sub>2</sub> O	C...O	1,433	1,950	3,564
QM	C...C	3,671	2,397	1,537
H <sub>2</sub> O	C...O	1,428	1,937	3,521
MM	C...C	3,653	2,361	1,537

TAB. 3.10 – longueurs des liaisons (Å) modifiées au cours de la réaction de Claisen pour les différentes approches (sans molécule d'eau, avec une molécule quantique ou avec une classique) avec influence du continuum

D'abord concernant la réaction sans influence du continuum, on remarque que les longueurs de liaisons sont très proches à la fois pour les produits et les réactifs quelque soit le système envisagé. Ainsi concernant le produit, on ne note aucune différence entre les trois liaisons carbone-carbone formées. Les variations de la longueur de la liaison C...O sont aussi du même ordre.

L'effet de la présence d'une molécule d'eau se fait par contre sentir sur la structure de transition. En effet, la distance carbone-oxygène est allongée de 0,035 Å avec la molécule quantique alors qu'elle ne l'est que de 0,021 avec la molécule classique. Le phénomène est analogue avec la liaison carbone-carbone (respectivement 0,07 et 0,037 Å). Encore une fois, les modifications induites par notre modèle se font dans le bon sens mais ce dernier ne reproduit pas complètement les effets d'une molécule d'eau quantique.

En ce qui concerne l'effet du continuum, on peut voir que ce dernier est assez faible si on compare les deux tableaux ligne à ligne. En effet, par exemple, l'allongement de la liaison C-O dans l'état de transition du soluté seul est allongée de 0,006 Å. Cela montre une fois encore que le principal facteur de modification du chemin réactionnel est la formation d'une liaison hydrogène avec l'oxygène.

Si on compare de nouveau les résultats pour les deux descriptions de la molécule d'eau, on voit que, dans le cas du modèle d'eau classique, les longueurs des liaisons considérées à l'état de transition ont une valeur intermédiaire entre l'approche complètement quantique et celle sans molécule d'eau. L'influence du modèle d'eau classique est donc analogue dans le solvant et dans le vide.

### Aspect électronique

Nous choisissons enfin de comparer les effets de l'introduction de molécules d'eau sur la fonction d'onde. Les données de la littérature indiquent que le solvant a pour effet de favoriser les structures de transition où la séparation de charge entre les deux fragments (de part et d'autre des deux liaisons en modification) est plus importante. Nous comparons donc les charges des fragments pour les trois états de transition en se basant sur les charges atomiques de Mulliken [100] (tableau 3.11). Dans le cas où la molécule d'eau est traitée au niveau quantique, il existe un léger transfert de charge vers les molécules d'eau. La charge de la molécule d'eau vaut -0,035 électrons dans le vide et -0,037 dans le continuum. Afin de ne pas prendre en compte ce dernier, nous choisissons de présenter la valeur de la charge du fragment qui ne se trouve pas en interaction avec

la molécule d'eau. nous faisons ainsi l'approximation que le transfert de charge vers la molécule d'eau provient uniquement du fragment qui contient l'oxygène.

	vide	H <sub>2</sub> O QM	H <sub>2</sub> O MM
Dans le vide	0,2154	0,2624	0,2484
Dans le continuum	0,2273	0,2587	0,2535

TAB. 3.11 – Charges du fragment ne contenant pas l'atome d'oxygène dans les trois état de transition pris en compte et calculées à partir des charges atomiques de Mulliken.

D'après ces valeurs, on voit de nouveau que la modification principale provient de l'introduction d'une molécule d'eau plutôt que du continuum. Ces charges montrent par ailleurs que la molécule d'eau classique a tendance à moins polariser le soluté. On remarque aussi que les effets du continuum et d'une molécule d'eau classique sont cumulatifs alors que le contraire se passe lorsque la molécule est quantique puisque la séparation de charge est moins importante dans le continuum que dans le vide. On peut interpréter cette différence comme provenant de la possibilité qu'a la molécule d'eau quantique à être polarisée, à la fois par le soluté et par le continuum, contrairement au modèle classique.

### 3.4.3 Conclusion sur l'étude du réarrangement de Claisen

Nous avons voulu, par le biais de cette étude, vérifier les capacités de notre modèle à reproduire l'influence d'une molécule d'eau quantique sur une réaction chimique. Les résultats que nous obtenons sont assez encourageants. En effet, les énergies d'interaction de la liaison hydrogène sont très bien reproduites, bien que les autres facteurs tels que la hauteur de barrière d'activation ou le déplacement de l'état de transition sur la surface d'énergie potentielle ne soient que partiellement reproduits. En d'autres termes, notre modèle ne rend correctement compte que d'une partie des interactions modifiant le chemin réactionnel. Cette remarque s'applique de la même façon aux résultats obtenus pour les calculs menés dans le solvant. Nous devons, après cette application, conclure que le modèle d'eau classique que nous avons développé remplit le rôle pour lequel il a été créé, à savoir reproduire les modifications que subit une réaction chimique lorsque les réactifs interagissent par le biais de liaisons hydrogène avec les molécules du solvant. Cependant, les effets de la molécule d'eau sur la réactivité ne sont décrits qu'en partie :

l'usage d'une molécule classique introduit des erreurs non corrigibles (avec la forme du champ de forces actuel).

### 3.5 Conclusion

Dans ce chapitre, nous avons choisi d'explorer la voie qui consiste à introduire une ou quelques molécules explicites de solvant, ici l'eau. Nous nous sommes particulièrement intéressés aux effets de la frontière entre les sous-parties quantiques et classiques sur la description de la liaison hydrogène.

De façon générale, nous avons pu constater, après une étude comparative, que le modèle standard choisi (TIP3P) n'est pas adapté à l'approche choisie. On remarque en effet que ce dernier, tend à surestimer l'interaction entre le soluté et le solvant explicite. Couper une molécule à proximité d'une liaison hydrogène peut modifier notablement les caractéristiques de cette dernière. Ceci est particulièrement vrai pour les espèces chargées et de petite taille comme l'éthanol. De ces constatations, nous pouvons dégager quelques remarques d'ordre général. D'une part, une attention particulière devra être prise quant à la position de la coupure, comme il est généralement constaté, il est préférable d'éloigner au plus cette dernière du centre réactif. Dans le cas particulier de l'approche LSCF/MM, la SLBO est aussi un facteur qui pourra influencer le centre réactif, sa détermination devra alors faire l'objet d'une attention particulière.

Ensuite, l'introduction de molécules d'eau au niveau classique pourra se faire mais à condition que ce modèle soit adapté au problème à traiter, à savoir l'interaction entre une molécule classique et un soluté quantique ou hybride.

Cette possibilité nous a donc mené à la mise en place d'un schéma adapté. Ce dernier nécessitant alors la détermination de paramètres spécifiques permettant une bonne description des interactions entre l'eau classique et la partie quantique du soluté. Les interactions entre molécules de solvant ou avec la partie classique du soluté restent inchangées. De nouveaux paramètres spécifiques ont donc été mis au point pour reproduire les caractéristiques d'énergie et de longueurs de la liaison hydrogène entre l'eau et des solutés hydrophiles. Nous avons pu voir par différents exemples que l'utilisation de ces paramètres conduisait à des résultats satisfaisants concernant l'énergie d'interaction dans le vide.

L'étude finale nous a montré que l'utilisation de ces paramètres rendait partiellement compte de l'influence du solvant sur la modification du chemin réactionnel

provoquée par le solvant. En effet, le comportement global suit le même sens que la référence. On voit par contre que tous les résultats obtenus sont intermédiaires entre ceux obtenus pour un soluté seul et un soluté en interaction avec une molécule d'eau quantique. Pour résumer, notre modèle reproduit une partie de l'influence due à la présence d'une molécule d'eau.

Les raisons de cette limitation peuvent être recherchées dans deux directions. La première concerne directement les paramètres : on peut en effet considérer que ces derniers sont simplement mal adaptés à la modélisation des effets du solvant sur les réactions chimiques. Cependant, c'est sans doute une seconde piste qui doit être évoquée. Cette dernière concerne la nature même du modèle. Celui-ci, ne peut, en effet, ni se polariser ni échanger de charges avec le soluté.

Pour résumer, nous pouvons dire que le modèle que nous avons développé permet de reproduire certaines caractéristiques de la liaison hydrogène, en particulier l'énergie d'interaction.



# Conclusion générale

L'objectif principal de ce travail était le développement d'une méthode dans le but de modéliser les effets du solvant sur la réactivité des systèmes de grande taille.

Pour cela, deux aspects compatibles ont été envisagés.

Dans un premier temps, nous avons été amenés à coupler l'approche hybride LSCF/MM permettant la description des systèmes moléculaires de grande taille avec le modèle implicite de solvant désigné SCRF. Nous avons ainsi défini l'approche à trois niveaux LSCF/MM/SCRF dans laquelle un soluté hybride (QM/MM) ressent l'influence du solvant. Cette méthode a permis, lors de l'étude de la cycloaddition du cyclopentadiène avec l'acrylate de (-)-menthyle, d'observer les influences simultanées du solvant et de l'auxiliaire chiral (le groupement menthyle). Ces effets auraient été impossibles à modéliser sans l'utilisation de notre approche à trois niveaux.

Dans un deuxième temps, afin d'affiner la description des effets du solvant sur la réactivité chimique, nous avons choisi d'explorer la voie qui consiste à introduire des molécules de solvant explicites au voisinage du centre réactif. Pour cela, nous avons commencé par étudier les effets d'une frontière QM/MM à proximité d'une liaison hydrogène qui est l'interaction spécifique sur laquelle nous nous sommes focalisés. Cette étude nous a montré d'une part qu'il était important de bien décrire la frontière lorsque celle-ci est à l'intérieur du soluté et que la liaison hydrogène est modélisée entre des atomes décrits par la mécanique quantique. Nous avons par ailleurs montré, lorsqu'on s'intéresse à un système où le soluté est quantique et la molécule d'eau classique, que les paramètres classiques standards étaient inadaptés à la description d'une interaction d'un soluté quantique avec un solvant classique. Cette étude nous a donc mené vers la détermination de paramètres spécifiques pour la molécule d'eau classique afin de décrire la liaison hydrogène entre un soluté quantique (ou hybride) et une ou quelques molécules d'eau classique. Par le biais de quelques tests, nous avons montré que ces paramètres reproduisaient bien les énergies d'interaction obtenues à un niveau quantique poussé.

Cependant, bien que la liaison hydrogène soit assez bien modélisée, l'influence d'une molécule d'eau classique sur la réactivité chimique (dans le cas de la réaction de Claisen) n'est pas complètement décrite par notre modèle.

Le modèle de continuum ne peut décrire que les interactions à longue portée entre le solvant et le soluté ainsi que les interactions spécifiques de façon moyenne. En ce qui concerne l'introduction de molécules classiques, même si elles ont des paramètres adaptés, ces dernières ne rendent compte que d'une partie des effets du solvant sur la réactivité du soluté. Seule l'introduction de molécules d'eau quantique permet de rendre compte de la polarisation du solvant par le soluté ainsi que du transfert de charge qui, dans le cas d'un soluté chargé, peut être important.

L'une des idées sous-jacentes à la définition d'un modèle adapté était son utilisation dans des simulations dynamiques de réactions chimiques, nécessitant ainsi un modèle fiable mais peu coûteux.

Afin de répondre à cet objectif, il doit être envisageable de définir une approche où l'introduction de molécules de solvant serait peu coûteuse (en terme de moyens de calculs) mais dont le modèle choisi pourrait au moins subir l'effet de son environnement (continuum ou autres molécules explicites de solvant) ainsi que du soluté. Une piste se trouve peut-être dans le passage à une approche où le système serait traité en mécanique quantique avec une méthode *ab initio* et le reste avec une approche semi-empirique (méthode QM/QM'). Dans ce cas, toutes les parties du systèmes sont capables d'être polarisées.

En plus de son objectif initial, ce travail de thèse a aussi été l'occasion d'apporter quelques compléments à l'approche LSCF/MM. Nous évoquerons ainsi la méthode présentée dans le premier chapitre et qui consiste à calculer le hessien d'un système traité par LSCF/MM en blocs. Nous avons aussi été amenés à "porter" le code LSCF/MM dans le logiciel Gaussian03 [2], nous obligeant ainsi à y apporter quelques adaptations. Ce travail a par ailleurs permis de bénéficier de certaines des nouvelles possibilités du programme au profit de notre méthode. Après un nombre restreint d'adaptations, il est, par exemple, désormais possible de conduire des calculs de dynamique moléculaire en utilisant une approche LSCF/MM. Cet aspect, s'il continue à être développé, pourra permettre l'accès aux grandeurs statistiques pour l'étude complète de chemins réactionnels impliquant des systèmes de grande taille.

# 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, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, J. L. Andres, C. Gonzalez, M. Head-Gordon, E. S. Replogle, , and J. A. Pople. *Gaussian 98, Revision A.9*. Gaussian, Inc., Pittsburgh PA, 1998.
- [2] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, and J. A. Pople. *Gaussian 03*,

- Revision B05*. Gaussian, Inc., Pittsburgh PA, 2003.
- [3] J.L. Rivail. *Elements de Chimie Quantique à l'Usage des Chimistes*. CNRS éditions, Paris, troisième édition, 1999.
- [4] A. Szabo and N. Ostlund. *Modern Quantum Chemistry*. McGraw-Hill, New York, deuxième édition, 1989.
- [5] M. Born and J.R. Oppenheimer. *Ann. Physik*, 84 :457, 1927.
- [6] W. Pauli Jr. *Z. Physik*, 31 :765, 1925.
- [7] J.C. Slater. *Phys. Rev.*, 34 :1293, 1929.
- [8] D.R. Hartree. *Proc. Cambridge Phil. Soc.*, 244 :89, 1928.
- [9] V. Fock. *Z. Physik*, 61 :126, 1930.
- [10] C.C.J. Roothaan. *Rev. Mod. Phys.*, 23 :69, 1951.
- [11] J.A. Pople and R.K. Nesbet. *J. Chem. Phys.*, 22 :571, 1954.
- [12] G. Berthier. *C.R. Acad. Sci. Paris*, 238 :91, 1954.
- [13] R.S. Mulliken. *J. Am. Chem. Soc.*, 74 :811, 1952.
- [14] C. Møller and M.S. Plesset. *Phys. Rev.*, 46 :618, 1934.
- [15] P. Hohenberg and W. Kohn. *Phys. Rev. B*, 136 :864, 1964.
- [16] W. Kohn and L.J. Sham. *Phys. Rev. A*, 140 :1133, 1965.
- [17] A.D Becke. *J. Chem. Phys.*, 98 :1372, 1993.
- [18] A.D Becke. *Phys. Rev. A*, 38 :3098, 1988.
- [19] C. Lee, W. Yang, and R.G. Parr. *Phys. Rev. B*, 37 :785, 1988.
- [20] W.D. Cornell, P. Cieplak, C.I. Bayly, I. R. Gould, K. M. Merz Jr., D.M. Ferguson, D.C. Spellmeyer, T. Fox, J.W. Caldwell, and P.A. Kollman. *J. Am. Chem. Soc.*, 117 :5179, 1995.
- [21] S.L. Mayo, B.D. Olafson, and W.A. Goddard. *J. Phys. Chem.*, 94 :8897, 1990.
- [22] A.K. Rappé, C.J. Casewit, K.S. Colwell, W.A. Goddard III, and W.M. Skiff. *J. Am. Chem. Soc.*, 114 :10024, 1990.
- [23] A. R. Leach. *Molecular Modelling Principles and Applications*. Pearson Education Limited, Harlow, England, deuxième édition, 2001.
- [24] M.J. Field, P.A. Bash, and M. Karplus. *J. Comp. Chem.*, 11 :700, 1990.
- [25] P.L. Cummins and J.E. Gready. *J. Phys. Chem. B*, 104 :4503, 2000.

- [26] S. Ranganathan and J.E. Gready. *J. Phys. Chem. B*, 101 :5614, 1997.
- [27] I. Antes and W. Thiel. *J. Phys. Chem. A*, 103 :9290, 1999.
- [28] Y.K. Zhang, H. Liu, and W.T. Yang. *J. Chem. Phys.*, 112 :3483, 1999.
- [29] Y.K. Zhang, T.S. Lee, and W.T. Yang. *J. Chem. Phys.*, 110 :46, 1999.
- [30] G. A. DiLabio, M. M.Hurley, and P. A.Christiansen. *J. Chem. Phys.*, 116 :9578, 2002.
- [31] M. Swart. *Int. J. Quant. Chem.*, 91 :177, 2002.
- [32] D. Das, K. P. Eurenus, E. M Billings, P. Sherwood, D. C. Chatfieldand, M. Hodoscekand, and B. R. Brook. *J. Chem. Phys.*, 117 :10534, 2002.
- [33] A. Warshel and M. Levitt. *J. Mol. Biol.*, 103 :227, 1976.
- [34] V. Théry, D. Rinaldi, J.-L Rivail, B. Maignret, and G.G Ferenczy. *J.Comp.Chem.*, 15 :269, 1994.
- [35] G. Monard, M. Loos, V. Théry, K. Baka, and J.L. Rivail. *Int. J. Quant. Chem.*, 58 :153, 1996.
- [36] N. Ferré, J.-L. Rivail, and X. Assfeld. *J.Comp.Chem.*, 23 :610, 2002.
- [37] J. Gao, P. Amara, C.Alhambra, and M.J. Field. *J. Phys Chem. A*, 12 :4714, 1998.
- [38] P. Amara, M.J. Field, C.Alhambra, and J. Gao. *Theor. Chem. Acc.*, 104 :336, 2000.
- [39] J. Pu, J. J. Gao, and D. G. Truhlar. *J. Phys. Chem.*, 108 :632, 2004.
- [40] R. Murphy, D. Philipp, and R. Friesner. *Chem. Phys. Lett.*, 321 :113, 2000.
- [41] R. Murphy, D. Philipp, and R. Friesner. *J. Comp. Chem.*, 21 :1442, 2000.
- [42] H. Lin and D. G. Truhlar. *J. Phys. Chem. A*, 109 :3991, 2005.
- [43] N. Reuter, A. Dejaegere, B. Maignret, and M. Karplus. *J. Phys. Chem. A*, 104 :1720, 2000.
- [44] M.A. Thompson and G.K. Schenter. *J. Phys. Chem.*, 100 :10580, 1996.
- [45] A. Laio, V. Vondele, and U. Rothlisberger. *J. Chem. Phys.*, 116 :6941, 2002.
- [46] G. Monard and K. H. Merz Jr. *Acc. Chem. Res.*, 32 :904, 1998.
- [47] R. Hall, S. A. Hindle, N. A. Burton, and I. A. Hillier. *J. Comp. Chem.*, 21 :1433, 2000.

- [48] G. Monard, X. Prat-Resina, A. González-Lafont, and J. M. Lluch. *Int. J. Quant. Chem.*, 93 :229, 2003.
- [49] P. Amara and M. J. Field. *Theor. Chem. Acc.*, 109 :43, 2003.
- [50] *Journal of Molecular Structure : THEOCHEM, Special Issue. Combined QM/MM Calculations in Chemistry and Biochemistry*, volume 632. Elsevier, Amsterdam, 2003.
- [51] F. Maseras and K. Morokuma. *J. Comp. Chem.*, 16 :1170, 1995.
- [52] S. Humbel, S. Sieber, and K. Morokuma. *J. Chem. Phys.*, 105 :1959, 1996.
- [53] M. Svensson, S. Humbel, R. D. J. Froese, T. Matsubara, S. Sieber, and K. Morokuma. *J. Phys Chem.*, 100 :19357, 1996.
- [54] N. Rega, S.S. Iyengar, G.A. Voth, H.B. Schlegel, T. Vreven, and M.J. Frisch. *J. Phys. Chem. B*, 108 :4210, 2004.
- [55] X. Assfeld and J.L. Rivail. *Chem. Phys. Lett.*, 263 :100, 1996.
- [56] P.O. Lowdin. *J. Chem. Phys.*, 18 :365, 1950.
- [57] N. Ferré. *Développements de la méthode LSCF ab initio et d'un champs de forces hybride QM/MM/Madelung pour l'étude de la réactivité des zéolithes*. Thèse de doctorat, Université Henri Poincaré Nancy I, 2001.
- [58] N. Ferré, X. Assfeld, and J.L. Rivail. *J. Chem. Phys.*, 117 :4119, 2002.
- [59] J. Foster and S. Boys. *Rev. Mod. Phys.*, 32 :300, 1960.
- [60] C. Edmiston and K. Ruedenberg. *Rev. Mod. Phys.*, 35 :457, 1963.
- [61] J. Pipewk and P. Mezey. *J. Chem. Phys.*, 90 :4916, 1989.
- [62] H. Weinstein, R. Pauncz, and M. Cohen. *dans Advances in Atomic and Molecular Physics*. Academic Press, New York, 1971.
- [63] M. Sironi and A. Famulari. *Theor. Chem Acc.*, 103 :417, 2000.
- [64] A. Fornili, Y. Moreau, M. Sironi, and X. Assfeld. *en préparation*, 2005.
- [65] X. Prat-Resina, M. Garcia-Viloca, G. Monard, A. González-Lafont, J. M. Lluch, J. M. Bofill, and J. M. Anglada. *Theor. Chem. Acc.*, 107 :147, 2002.
- [66] J. Pu and D. G. Truhlar. *J. Chem. Theory Comput.*, 1 :54, 2005.
- [67] B.H. Besler, K. M. Merz Jr., and P.A. Kollman. *J. Comp. Chem.*, 11 :431, 1990.
- [68] W.L. Jorgensen, J. Chandreskhar, J. D. Madura, R. W. Impey, and M. L. Klein. *J. Chem. Phys.*, 79 :926, 1982.

- [69] Y. Moreau, P. F. Loos, and X. Assfeld. *Theor. Chem Acc.*, 112 :228, 2004.
- [70] D. Rinaldi and J.L. Rivail. *Theor. Chim Acc.*, 32 :57, 1973.
- [71] M.Z. Born. *Physik*, 1 :45, 1920.
- [72] J.G. Kirkwood. *J. Chem. Phys.*, 7 :911, 1939.
- [73] L. Onsager. *J. Am. Chem. Soc.*, 58 :1486, 1936.
- [74] S. Miertus, E. Scorocco, and J. Tomasi. *Chem. Phys.*, 55 :1174, 1981.
- [75] C. Amovilli, V. Barone, R. Cammi, E. Cancès, R. Cossi, B. Mennucci, CS. Pomelli, and J. Tomasi. *Adv. Quantum Chem.*, 32 :227, 1998.
- [76] J. Tomasi, M. Mennucci, R. Cammi, and M. Cossi. *Quantum mechanical models for reaction in solution. in Chemistry and Biochemistry. In : G. Náray-Szabo , A. Warshel (Eds.) Computational approaches to biochemical reactivity.* Kluwer, Dodrecht, 1997.
- [77] C. J. Cramer and D. G. Truhlar. *J. Am. Chem. Soc.*, 113 :8305, 1991.
- [78] C. J. Cramer and D. G. Truhlar. *Chem. Rev.*, 99 :2166, 1999.
- [79] J. D. Thompson, C. J. Cramer, and D. G. Truhlar. *J. Phys. Chem. A*, 108 :6832, 2004.
- [80] W. C. Still, A. Tempczyk, R. C. Hawley, and T. Hendrickson. *J. Am. Chem. Soc.*, 112 :6127, 1990.
- [81] F. Eckert and A. Klamt. *AIChE J.*, 48 :369, 2002.
- [82] J. Tomasi. *Theor. Chem. Acc.*, 112 :184, 2004.
- [83] D. Rinaldi and J.L. Rivail. *Chem. Phys.*, 18 :233, 1976.
- [84] D. Rinaldi, M.F. Ruiz López, and J.L. Rivail. *J. Chem. Phys.*, 78 :834, 1983.
- [85] D. Rinaldi, J.L. Rivail, and M.F. Ruiz López. dans “*Theoretical and Computational Models for Organic Chemistry*”, volume 79. Kluwer Academic Publishers, 1991.
- [86] J.L. Rivail and D. Rinaldi. *Computational Chemistry, Review of Current Trends*, volume 139. J. Leszczynski Ed., World Scientific Publishing, 1996.
- [87] V. Dillet, D. Rinaldi, J.G. Ángyán, and J.L. Rivail. *Chem. Phys. Lett.*, 18 :202, 1993.
- [88] V. Dillet, D. Rinaldi, and J.L. Rivail. *J. Phys. Chem.*, 98 :5034, 1994.

- [89] J.G. Kirkwood. *J. Chem. Phys.*, 2 :351, 1934.
- [90] F. M. Richards. *Ann. Rev. Biophys. Biornng.*, 6 :151, 1977.
- [91] D. Rinaldi, J.L. Rivail, and N. Rguini. *J. Comp. Chem.*, 13 :675, 1992.
- [92] V. Dillet, D. Rinaldi, J. Bertran, and J.L. Rivail. *J. Chem. Phys.*, 104 :23, 1996.
- [93] T. Vreven, B. Mennucci, C. O. da Silva, K. Morokuma, and J. Tomasi. *J. Chem. Phys.*, 1 :115, 2001.
- [94] C. Cativiela, J.I. García, J.A. Mayoral, A. J. Royo, L. Salvatella ;, M.F. Ruiz-López, and X. Assfeld. *J.Phys. Org. Chem.*, 5 :230, 1992.
- [95] M.F. Ruiz-López, X. Assfeld, J.I. García, J.A. Mayoral, and L. Salvatella. *J. Am. Chem. Soc.*, 115 :8780, 1993.
- [96] L. Salvatella, A. Mokrane, A. Cartier, and M.F. Ruiz-López. *J. Org. Chem.*, 63 :4664, 1998.
- [97] L. E. Chirlian and M.M. Francl. *J. Comp. Chem.*, 8 :894, 1987.
- [98] C.M. Breneman and K.B. Wiberg. *J. Comp. Chem.*, 11 :361, 1990.
- [99] J.G. Ángyán, C. Chipot, F. Dehez, C. Hättig, G. Hansen, and C. Millot. *OPEP : a tool for the optimal partitionning of electric properties*. Université Henri Poincaré, Vandœuvre-lès-Nancy Cedex, France.
- [100] R.S. Mulliken. *J. Chim. Phys.*, 46 :497, 1949.
- [101] J. F. Blake, D. Lim, and W. L. Jorgensen. *J. Org. Chem.*, 59 :803, 1994.
- [102] J. Chandrasekhar, S. Sharillskul, and W. L. Jorgensen. *J. Phys. Chem. B*, 106 :8078, 2002.
- [103] T.R. Furlani and J. Gao. *J. Org. Chem.*, 61 :5492, 1996.
- [104] S. Kong and J. D. Evanseck. *J.Am. Chem. Soc.*, 122 :10418, 2000.
- [105] G.D. Hawkins, D.J. Giesen, G.C Lynch, C.C Chambers, I. Rossi, J.W. Storer, J. Li, D. Rinaldi, D.A. Liotard, C.J. Cramer, and D.G Truhlar. *AMSOL version 6.5.2*. University of Minnesota, Minneapolis, based in part on AMPAC version 2.1 (D.A. Liotard, E.F. Healey, J.M. Ruiz, M.J.S Dewar), 1997.
- [106] M.J.S. Dewar, E.F. Zoebisch, E.F. Healy, and J.J.P. Stewart. *J.Am. Chem. Soc.*, 107 :3902, 1985.
- [107] C.J. Cramer and D.G. Truhlar. *Science*, 256 :213, 1992.
- [108] J.F Maddaluno, N. Gresh, and C. Giessner-Prettre. *J. Org. Chem.*, 59 :793, 1994.

- [109] I. Mayer. *Chem. Phys. Lett.*, 97 :270, 1983.
- [110] J. Gao. *Acc. Chem. Res.*, 29 :298, 1996.
- [111] J. Gao and X. Xia. *Science*, 258 :631, 1992.
- [112] J. Gao. *J. Phys. Chem.*, 96 :537, 1992.
- [113] M. Freindhorf and J. Gao. *J. Comp. Chem.*, 17 :386, 1995.
- [114] I. Tuñon, M.T.C. Martins-Costa, C. Millot, and M. F. Ruiz-López. *J. Mol. Model.*, 1 :196, 1995.
- [115] I. Tuñon, M.T.C. Martins-Costa, C. Millot, M. F. Ruiz-López, and J. L. Rivail. *J. Comp. Chem.*, 17 :19, 1996.
- [116] S. Chalmet. *Simulation de processus chimiques en phase liquide par des méthodes hybrides DFT/MM*. Thèse de doctorat, Université Henri Poincaré Nancy I, 2000.
- [117] S. Chalmet and M. F. Ruiz-López. *J. Chem. Phys.*, 111 :1117, 1999.
- [118] M.P. Repasky, C.R. Werneck Guimar aes, J. Chandrasekhar, J. Tirado-Rives, and W. L. Jorgensen. *J. Am. Chem. Soc.*, 125 :6663, 2003.
- [119] Q. Cui. *J. Chem. Phys.*, 117 :4720, 2002.
- [120] S. Chalmet, D. Rinaldi, and M. F. Ruiz-López. *Int. J. Quant. Chem.*, 84 :559, 2001.
- [121] P. Surjan and J. Ángyán. *Chem. Phys. Lett.*, 225 :258, 1994.
- [122] D. Wei and D.R. Salahub. *Chem. Phys. Lett.*, 224 :291, 1994.
- [123] M. Ben-Nun and T.J. Marytínez. *Chem. Phys. Lett.*, 290 :289, 1998.
- [124] W. Cornell, P. Cieplak, C. Bayly, and P.A. Kollman. *J. Phys. Chem.*, 97 :10269, 1993.
- [125] Y. Tu and A. Laaksonen. *J. Chem. Phys.*, 111 :7519, 1999.
- [126] S. Chalmet and M. F. Ruiz-López. *Chem. Phys. Lett.*, 329 :154, 2000.
- [127] M.E. Martín, M. A. Aguilar, S. Chalmet, and M. F. Ruiz-López. *Chem. Phys.*, 284(3) :607, 2002.
- [128] Y. Tu and A. Laaksonen. *J. Chem. Phys.*, 113 :11264, 2000.
- [129] A. Halkier, H. Koch, P. Jørgensen, O. Christiansen, I. M. Beck Nielsen, and T. Helgaker. *Theor. Chem. Acc.*, 97 :150, 1997.
- [130] M. Schütz, S. Brdrski, P-O. Widmark, R. Lindh, and G. Karlström. *J. Chem. Phys.*, 107 :4597, 1997.

- [131] W. L. Jorgensen and J. Tirad-Rives. *J. Am. Chem. Soc.*, 110 :1657, 1988.
- [132] L.A. Curtiss, D.J. Frurip, and M. Blander. *J. Chem. Phys.*, 71 :2703, 1979.
- [133] J.G.C.M. van Duijneveldt-van de Rijdt and F.B. van Duijneveldt. *J. Chem. Phys.*, 97 :5019, 1992.
- [134] C. Cramer. *Essential of computational chemistry : theories and models*. Kluwer, Dodrecht, 1997.
- [135] J. J. Gajewski. *Acc. Chem. Res.*, 30 :219, 1997.
- [136] M. M. Davidson, I. H. Hillier, R. J. Hall, and N. A. Burton. *J. Am. Chem. Soc.*, 116 :9294, 1994.
- [137] R. J. Hall, M. M. Davidson, N. A. Burton, and I. H. Hillier. *J. Phys. Chem.*, 99 :921, 1995.
- [138] M. M. Davidson and I. H. Hillier. *J. Phys. Chem.*, 99 :6748, 1995.
- [139] D. L. Severance and W. L. Jorgensen. *J. Am. Chem. Soc.*, 114 :10966, 1992.
- [140] A. Sehgal, L. Shao, and J. Gao. *J. Am. Chem. Soc.*, 117 :11337, 1995.
- [141] H. Hu, M. N. Kobrak C. Xu, and S. Hammes-Schiffer. *J. Phys. Chem. A*, 104 :8058, 2000.
- [142] O. Wiest, K. Black, and K.N. Houk. *J. Am. Chem. Soc.*, 116 :10336, 1994.
- [143] R. L. Vance, N. G. Rondan, K. N. Houk, F. Jensen, W Thatcher Borden, A. Kormornicki, and E. Wimmer. *J. Am. Chem. Soc.*, 110 :2314, 1988.

# Annexe A

## Modifications des logiciels

Dans cette annexe, nous avons choisi de détailler rapidement les différentes modifications nécessaires dans les programmes Gaussian98 [1] et Gaussian03 [2] pour mener nos travaux.

Nous évoquons ici les modifications pour chacune des méthodes programmées lors de nos travaux. Nous avons choisi ici de respecter l'ordre chronologique dans lequel les modifications ont été faites. Nous débutons ainsi par les modifications dans le logiciel Gaussian 98 pour la méthode LSCF/MM/SCRF. Viennent ensuite les détails de l'implantation de LSCF/MM dans Gaussian 03. Nous évoquons ensuite les modifications faites pour la description différenciée des interactions non liées. Le dernier point concerne enfin le calcul par blocs de la matrice hessienne.

## A.1 Le programme Gaussian

Afin d'appréhender les modifications portées dans le programme, nous commençons par en présenter le fonctionnement général.

Le programme est constitué d'un ensemble de programmes appelés links et portant un numéro précis. Les links sont regroupés en overlays suivant leur rôle (tableau A.1). Chaque link porte donc le numéro de l'overlay suivi de son propre numéro au sein de l'overlay. Le link101 est ainsi le numéro 1 dans l'overlay 1.

overlay	rôle
1	lecture du fichier d'entrée, optimisation
2	verification des information moléculaires
3	attribution de la base d'orbitales, détermination des intégrales
4	calcul du guess ; calculs semi-empiriques et classiques
5	calcul SCF (Hartree-Fock et Kohn-Sham)
6	analyse de population électronique
7	calcul des dérivées de l'énergie
8	calcul des intégrales biélectroniques pour les calculs post-HF
9	calcul post-HF
10	Résolution du CPHF
11	calcul des dérivées secondes les méthodes post-Hartree Fock
99	fin du programme

TAB. A.1 – numéro et rôles des overlays

Ces liens sont exécutés successivement et communiquent entre eux. Chacun d'entre eux a un rôle précis. Ainsi le link 502 (1502) a pour but de résoudre les équations de Hartree-Fock ou de Kohn-Sham, connaissant la base d'orbitales et le nombre d'électrons. Dans nos travaux, nous avons eu à intervenir dans les links 101, 103, 106, 118, 301, 302, 347, 402, 447, 502, 601, 602, 701 et 716.

On note par ailleurs que d'une version à l'autre du programme, la structure générale reste la même, chaque link garde son rôle.

Dans le fichier d'entrée, chaque mot clé est traduit pour donner la "route", structure qui donne l'ordre d'appel des links et les options sous forme de nombres. Ainsi pour l'option numéro 199 utilisée dans le link101 et dont la valeur est fixée à 4, on notera l'ioP avec le numero d'overlay, le numero d'option et sa valeur comme : IOP(1/199=4)

Les options utilisées n'étant pas standards dans le programme, nous sommes obligés, pour un calcul QM/MM, de travailler directement avec la route qu'on qualifie de "non-standard".

Les modifications liées à la méthode LSCF/MM que nous n'avons pas introduit ne sont pas détaillées ici et pourront être trouvées dans la référence [57].

Dans un fichier utilisé pour un calcul QM/MM, on trouve :

- La route, qui indique quels sont les links à utiliser et les options
- Le titre
- Les coordonnées atomiques (précédées de la charge et de la multiplicité de spin). Ces coordonnées contiennent aussi les informations quant à la nature des atomes (quantiques ou classiques) ainsi que leur nature et la charge associée pour la construction du champ de forces de la partie classique.
- Les données nécessaires comme la connectivité, les spécifications pour le modèle de solvant et/ou pour les orbitales gelées ou strictement localisées.

## A.2 Méthode LSCF/MM/SCRF

Dans ce cas, le travail a consisté à introduire la méthode LSCF/MM dans le code Gaussian où préexistait déjà le modèle de solvant implicite. Dans un calcul QM/MM, les atomes sont décrits, dans le programme, par un code permettant de connaître sa nature (QM, MM ou frontière). Le type particulier de l'atome frontière fait que dans le cas d'un calcul à trois niveaux, aucun rayon de van der Waals ne lui était attribué. Ce qui rend alors faux la forme et le volume de la cavité considérée. Notre travail a donc consisté, ici, à préciser la nature de l'atome frontière.

Le fichier présenté ici a servi à optimiser la géométrie d'un des seize états de transition pris en compte dans l'étude de la réaction de l'acrylate de (-)-menthyle avec le cyclopentadiène. Ce fichier est prévu pour Gaussian98. Ce fichier a été tronqué au niveau des coordonnées atomiques.

```
%subst 1101 /home/moreau/YG98
%subst 1301 /home/moreau/YG98
%subst 1347 /home/moreau/YG98
%subst 1402 /home/moreau/YG98
%subst 1447 /home/moreau/YG98
%subst 1502 /home/moreau/YG98
```

```

%subst 1601 /home/moreau/YG98
%subst 1701 /home/moreau/YG98
%NPROC=1
%chk=tsendocisbis.chk
%mem=40mw
# nonstd
  1/5=1,10=3,11=1,18=20,38=1,98=1/1,3;
  2/9=110,15=1,17=6,18=5,40=1/2;
  4/20=9,22=1,24=1,67=2/2;
  3/5=1,6=6,7=1,11=1,24=10,25=1,30=1,96=1/1,2,3,47;
  4/7=1,96=2/1,47;
  5/5=2,38=4,96=1,99=1/2;
  6/7=3,28=1/1;
  7/30=1,96=-1/1,2,3,47,16;
  1/5=1,11=1,18=20/3(1);
  99/96=1/99;
  2/9=110,15=1/2;
  4/16=2,20=9,22=1,24=1,67=2/2;
  3/5=7,6=1,7=1,11=1,25=1,30=1,94=2,96=1/1,2,3,47;
  4/5=5,7=1,16=2,96=2/1,47;
  5/5=2,38=4,96=1,99=1/2;
  7/30=1,96=-1/1,2,3,47,16;
  1/5=1,11=1,18=20/3(-6);
  2/9=110,15=1/2;
  6/7=3,19=2,28=1/1;
  99/9=1,96=1/99;

tsendocis QMMM

0 1
C-C_2          -3.778184   .025571   1.957555
  ...
0-0_2          -.482257   -.722208   1.065695
C-C_3-0.0270864  1.277322   -.218216   -.939464      M
  ...
H-H_-0.124685   3.404006  -3.838509  -1.325526      L

```

LPMAX=6 EPS0=78.4 EPSI=2.5 EPSC=1.0 ELLIP=VDW ISHAPE=4 Xvol1=2.

SLB0

11 20 18 17 ROTATE

-.116654 -.294340 .274281

1.243078 -.443233 -.152381

2.083679 -1.127410 .607414

.0023426 .2083777 .2515291 -.0181450 .0017829 .0992273 .0634403 .0095886

.0018211 .0374353 -.0096958 -.0102400 -.0040098 .0000936 -.0003132

-.0586315 .2346656 -.3922652 .0446674 -.0007542 .1938485 -.2523734 .0474052

-.0022385 .0267401 -.0007825 -.0036398 -.0183569 .0013777 -.0001006

\*\*\*\*

### A.3 Implantation de LSCF/MM dans Gaussian03

Comme dit précédemment, la structure des programmes (98 et 03) sont similaires, peu de modifications ont été nécessaires pour le passage de l'un à l'autre de la partie du code concernant le calcul quantique (énergies et forces LSCF). Par contre, la construction du champs de force qui se faisait dans le link 402, juste avant les calculs classiques a été reportée dans le link101, pour n'être faite qu'une seule fois. Notre travail le plus important a donc été d'introduire la constructoin du champ de forces LSCF/MM (choix des paramètres à utiliser, lecture des paramètres spécifiques de liaison frontière) dans ce link et de créer les fichiers de communication permettant la lecture de ces données lors du calcul dans le link402.

Notre travail de transfert a aussi consisté à définir la route compatible avec ce nouveau programme. Nous reportons ci-après une route typique d'une optimisation de géométrie au niveau HF ou DFT, les ... figurant les options.

```

1/.../1,3;
2/.../2;
3/.../1;
4/.../2;
3/.../1,2,3;
4/.../1;
5/.../2;
6/.../1;
7/.../1,2,3,16;
1/.../3(1);
99/.../99;
2/.../2;
3/.../1;
4/.../2;
3/.../1,2,3;
4/.../1;
5/.../2;
7/.../1,2,3,16;
1/.../3(-7);
2/.../2;

```

6/.../1;  
99/.../99;

On remarque, par rapport à la route de Gaussian98 le passage par le link 301 avant le calcul classique. Ce passage a été fait pour reproduire la route utilisée dans le cas d'un calcul classique standard. Le nombre d'options ayant augmenté d'une version à l'autre du logiciel, la numérotation des IOPs spécifiques à LSCF/MM a été refaite (généralement augmentée de 100).

## A.4 Insertion de molécules d'eau discrètes

Les modifications liées à ce travail concernent plusieurs links.

Dans le link 101, on lit dans le fichier d'entrée quels atomes appartiennent à la partie soluté ou solvant, quels paramètres sont modifiés (charges, Lennard-Jones ou les deux) ainsi que les valeurs associées. L'IOP(1/198) a été créé afin de permettre la lecture des paramètres. Sa valeur est 1 pour la lecture du facteur d'échelle, 10 pour la lecture des paramètres de Lennard-Jones spécifiques et 11 pour les deux. Les spécifications sont lues dans le fichier d'entrée après les données de structure.

Dans le link 402, le calcul des interaction non liées est modifié de telle sorte que les paramètres spécifiques ne sont utilisés pour les atomes du solvant que lorsqu'on calcul leur interaction avec un atome du soluté. Les dérivées de l'énergie associée sont calculées en prenant en compte ce traitement.

Dans le link 302, les charges des atomes classiques de la partie "solvant" sont modifiées (par le biais du facteur d'échelle) pour le calcul de l'énergie de répulsion noyaux-atomes classiques mais aussi pour le calcul des intégrales monoélectroniques dans lesquelles elle interviennent. On note par ailleurs que ces charges restent modifiées jusqu'au link 701 (calcul des dérivées de l'énergie) permettant ainsi l'optimisation de géométrie QM/MM avec cette approche différenciée.

Le fichier suivant à été utilisé pour l'optimisation de géométrie du complexe de l'éthanol avec deux molécules classiques d'eau. L'interaction entre le soluté et le solvant se fait avec les paramètres adaptés tandis que l'interaction entre les molécules d'eau se fait avec les paramètres TIP3P. Ce fichier est utilisé dans Gaussian 03 [2]. Le niveau de calcul est ici B3LYP/6-311G(d,p) pour la partie quantique, les paramètres classiques sont ceux d'Amber et les notres.

```
# nonstd
1/18=20,38=1,56=2,57=2,64=3,66=2,198=11,199=4/1,3;
2/9=110,15=1,17=6,18=5,40=1/2;
3/5=4,6=6,7=101,11=9,16=1,25=1,30=1/1;
4/20=11,22=1,24=3,67=2/2;
3/5=4,6=6,7=101,11=9,16=1,25=1,30=1,74=-5/1,2,3;
4//1;
5/5=2,38=5/2;
6/7=2,8=2,9=2,10=2/1;
7/30=1/1,2,3,16;
```

1/18=20,57=2/3(1);  
 99//99;  
 2/9=110,15=1/2;  
 3/5=7,6=1,7=1,11=9,16=1,25=1,30=1/1;  
 4/16=3,20=11,22=1,24=3,67=2,33=3/2;  
 3/5=7,6=1,7=1,11=9,16=1,25=1,30=1,74=-5/1,2,3;  
 4/5=5,16=3/1;  
 5/5=2,38=5/2;  
 7/30=1/1,2,3,16;  
 1/18=20,57=2/3(-7);  
 2/9=110,15=1/2;  
 6/7=2,8=2,9=2,10=2/1;  
 99//99;

MP2 SP

0 1

C-CT-		-.075700	-.242589	.004980	
...					
O-OW--	.834	-.820731	4.610151	-.799840	1
H-HW-	.417	-1.016542	5.491770	-1.120335	1
H-HW-	.417	.140812	4.526862	-.934410	1

1 2 1.0 3 1.0 4 1.0 5 1.0

...

15

SC\_FACT

0.8302

VDW\_PRM

2.0627

0.0245

SOLUTE

1

...

9

SOLVENT

```
10
...
15
****
```

La dernière partie du fichier (après la table de connectivité) permet d'entrer les valeurs des paramètres spécifiques. Ces dernières se présentent sous cette forme : **SC\_FACT** indique le facteur d'échelle appliqué aux charges classiques qui entrent dans l'expression de l'hamiltonien quantique. **VDW\_PRM** indique dans l'ordre les valeurs de  $\sigma$  et  $\epsilon$  utilisées pour calculer l'interaction des oxygènes des molécules d'eau avec des atomes quantiques. Enfin, les sections **SOLUTE** et **SOLVANT** indiquent entre quels atomes les modifications doivent intervenir. (uniquement entre un atome de type soluté et un atome de type solvant). On note par ailleurs que tous les atomes ne doivent pas être obligatoirement dans l'une ou l'autre section. Un tel atome interagira alors avec tous les autres avec les paramètres standards du programme.

Lors d'un tel calcul, le fichier de sortie indique précisément les modifications dès le lien 101 où est effectuée la lecture des spécifications. ci-après, un extrait du fichier de sortie obtenu avec les données précédentes :

```
WARNING :
modified interactions between solute and solvent
scaling Coulomb interaction by a factor .83020
VdW parameters for O(W) are 2.0627 and .02450
```

## A.5 Calcul de la matrice hessienne par blocs

Comme présenté dans ce mémoire, ce type de calcul peut se faire en incluant ou non des charges classiques ajustées sur le potentiel électrostatique de la fonction d'onde de la partie quantique dans le calcul classique.

Dans le link 101, l'IOP(1/199) peut donc recevoir différentes valeurs permettant ainsi de choisir quel type de calcul est effectué. Comme cet iop indique aussi si on effectue un calcul LSCF/MM, (il vaut 4 si on fait un calcul LSCF/MM, 14 si on doit lire un potentiel de liaison frontière et le chiffre des centaines est associé au calcul des constantes de force).

- si iop(1/199)=100, on effectue un calcul numérique complet
- si iop(1/199)=200, on effectue un calcul du hessien par blocs, sans prise en compte des charges dérivées du potentiel électrostatique
- si iop(1/199)=300, on effectue un calcul du hessien par blocs, avec prise en compte des charges dérivées du potentiel électrostatique

Le link 106, dont le rôle est de conduire le calcul numérique a été modifié de telle sorte que seuls les atomes quantiques (y compris l'atome frontière) ne soient concernés.

Le link 402 a été modifié pour prendre en compte les charges électrostatiques dans le calcul des constantes de force de la partie classique, au niveau classique.

Le link 602, a été modifié de telle sorte que le calcul des charges ajustées sur le potentiel électrostatique ne prenne en compte que les atomes de la partie quantique, y compris l'atome frontière. Ces charges sont ensuite conservées pour le calcul du sous bloc du hessien correspondant à la partie classique.

Le link 716 a été modifié pour pouvoir rechercher les différents blocs du hessien et reconstruire ce dernier.

Le link 103, où s'effectue le calcul des fréquences et modes de vibration a été modifié pour assurer que les atomes frontières ont la bonne masse (c'est à dire non nulle), et ainsi permettre la diagonalisation de la matrice hessienne pondérée par les masses.

Dans le fichier d'entrée, seule la route non standard est modifiée. Nous ne présentons donc ici que cette partie. La route présentée ici est utilisée avec le logiciel Gaussian03.

# nonstd

```
1/10=4,30=1,38=11,56=1,64=1,66=2,199=304/1,3,6;
2/15=1,17=6,18=5,40=1/2;
3/5=26,6=2,7=1,11=9,16=1,25=1,30=1/1;
4/20=9,22=1,24=3,33=3/2;
```

$3/5=1, 6=6, 7=1, 11=1, 16=1, 25=1, 30=1/1, 2, 3;$   
 $4/7=1, 11=3, 69=2/1;$   
 $5/5=2, 38=5, 199=1/2;$   
 $6/7=2, 8=2, 9=2, 10=2, 15=8, 20=1, 28=1/1, 2;$   
 $7/30=1/1, 2, 3, 16;$   
 $1/38=10, 57=2/6(3);$   
 $7/8=1, 25=1, 30=1, 44=-1/16;$   
 $1/10=4, 30=1, 38=10, 57=2/3;$   
 $99//99;$   
 $3/5=7, 6=1, 7=1, 11=9, 16=1, 25=1, 30=1/1;$   
 $4/16=2, 20=9, 22=1, 24=3/2;$   
 $3/5=7, 6=1, 7=1, 11=1, 16=1, 25=1, 30=1/1, 2, 3;$   
 $4/5=5, 7=1, 11=3, 16=3, 69=2/1;$   
 $5/5=2, 38=5, 199=1/2;$   
 $7/7=1, 30=1/1, 2, 3, 16;$   
 $1/38=10, 57=2/6(-6);$   
 $3/5=7, 6=1, 7=1, 11=9, 16=1, 25=1, 30=1/1;$   
 $4/16=2, 20=9, 22=2, 24=3/2;$   
 $7/8=1, 25=1, 30=1, 44=-1/16;$   
 $1/10=4, 30=1, 38=10, 57=2/3;$   
 $99//99;$

Dans cet exemple, l'IOP(1/199) vaut 304, indiquant ainsi un calcul LSCF/MM des constantes de forces, avec prise en compte des charges dans le calcul classique des constantes de force de la partie classique. Le passage par le link 602 permet le calcul et le stockage des charges. Le calcul se fait de façon habituelle pour les atomes QM. Un calcul classique sur toute la molécule (champ de force crée au départ plus charges) de l'énergie et des dérivées de celle ci jusqu'à l'ordre deux et ajouté à la fin. L'assemblage du hessien (link716) et le calcul des fréquences (link103) se fait ensuite.

Dans le cas d'un calcul sans prise en compte des charges, le bloc du hessien concernant la sous partie classique est déterminé dans le premier passage du link 402, ensuite, la procédure est la même mais sans le calcul des charges ni le dernier calcul au niveau classique.

## Annexe B

### Données associées au premier chapitre

747,1344 <i>i</i>	24,0480	30,8530	36,3731	69,8681
78,3405	98,1410	136,8335	151,9952	204,1937
236,0339	249,7757	267,3560	270,7924	284,7308
294,7934	318,8833	323,1392	351,2560	374,6738
381,3122	422,9687	427,6185	482,1577	502,7596
511,4527	520,2065	546,7043	554,7525	570,2655
618,2404	626,7812	660,5859	762,4012	807,5254
829,6610	858,7512	863,3807	867,1967	918,4073
919,9988	939,8378	948,8396	974,0208	989,7430
994,0861	1017,3411	1018,5427	1044,3926	1047,3221
1060,5507	1064,1838	1068,1697	1073,6568	1085,5701
1090,6280	1099,7994	1105,6200	1130,4974	1143,9350
1158,8311	1176,3974	1186,3383	1187,7939	1195,7330
1198,3037	1207,2591	1209,8698	1223,3868	1239,0180
1248,6350	1286,1379	1328,6725	1352,2294	1377,8743
1419,5393	1420,8225	1423,7386	1443,5663	1448,8615
1454,0605	1468,0852	1478,1160	1488,4275	1524,8533
1549,0668	1570,0059	1582,3745	1600,9589	1607,3063
1631,9971	1634,4012	1651,5059	1656,1962	1668,8594
1670,9677	1691,9099	1694,3777	1697,7306	1701,3052
1704,5023	1710,8432	1718,2762	1720,4279	1736,9235
1755,8695	1785,1185	1830,1994	1845,4542	1940,9552
2919,3566	2920,4146	2920,7186	2956,9700	2970,5864
2975,6051	2998,4145	3005,2135	3006,0884	3021,3967
3027,9346	3028,1509	3028,8278	3029,7844	3030,4403
3031,9181	3032,4155	3036,3782	3038,1233	3184,9322
3303,8426	3341,0886	3375,3801	3396,6632	3404,6872
3408,1283	3423,3823	3440,7669		

TAB. B.1 – Fréquences de vibration en  $\text{cm}^{-1}$  calculées avec une procédure numérique complète pour l'état de transition dans la réaction de Diels-Alder

747,1315 <i>i</i>	9,6893	27,6761	31,4753	64,5883
77,8914	101,5817	123,2844	151,6639	205,0376
234,1993	249,9202	266,1255	271,3927	286,5343
297,9740	319,2633	323,8216	351,0821	374,5425
382,7506	423,1343	428,5661	485,0813	503,8350
512,0454	521,2735	547,5131	557,3968	571,2754
618,5774	628,1911	661,5963	762,6810	807,6082
829,6624	857,9507	861,4146	867,0025	917,2522
919,8382	939,8052	948,2022	972,7079	988,6209
994,0453	1017,2338	1018,1176	1043,3183	1047,2185
1060,1648	1064,0580	1068,0850	1073,1481	1085,6358
1090,3658	1099,8267	1108,6281	1130,2586	1143,8458
1158,8306	1175,8156	1186,3378	1188,7919	1195,8429
1199,7840	1207,3265	1210,0834	1223,5078	1243,1078
1248,6994	1286,4981	1329,0826	1352,3716	1378,8134
1420,7791	1424,4922	1429,0249	1447,8819	1448,8641
1454,9355	1469,1745	1479,7785	1491,3213	1524,9433
1551,5175	1570,5781	1589,5379	1600,9619	1609,0505
1632,0051	1635,0419	1652,0880	1656,1934	1670,0177
1671,6771	1692,6175	1694,5072	1697,9933	1701,7648
1705,9870	1714,4594	1719,7908	1720,9683	1737,1874
1757,5081	1785,9356	1830,6103	1849,0161	1940,9786
2919,3280	2919,9336	2920,6133	2956,6915	2966,1037
2975,3384	2997,1561	3003,3615	3005,8075	3012,5353
3021,3071	3027,8486	3028,2126	3029,3849	3029,9990
3031,2671	3031,4248	3035,1753	3036,1998	3184,9323
3303,8425	3341,0885	3375,3800	3396,6631	3404,6872
3408,1280	3423,3823	3440,7668		

TAB. B.2 – Fréquences de vibration en  $\text{cm}^{-1}$  calculées sans prise en compte des interactions électrostatiques entre les sous parties quantiques et classiques pour l'état de transition dans la réaction de Diels-Alder

747,2526 <i>i</i>	11,1820	27,9031	31,7710	64,5257
77,9417	101,5989	123,4300	151,7197	205,1084
234,2510	249,9247	266,1252	271,4030	286,5456
297,9865	319,2957	323,8462	351,0748	374,5371
382,7521	423,1390	428,5669	485,1080	503,8383
512,0364	521,2734	547,5129	557,3976	571,2998
618,5813	628,1898	661,6087	762,6911	807,5699
829,7293	857,9841	861,3944	867,0239	917,2468
919,7806	939,8415	948,2052	972,7137	988,6187
994,0430	1017,1083	1018,0298	1043,3149	1047,1889
1060,1768	1064,0733	1068,0922	1073,1311	1085,5962
1090,3617	1099,8266	1108,6265	1130,2586	1143,8458
1158,7375	1175,8156	1186,3260	1188,7947	1195,8880
1199,7863	1207,3459	1210,0895	1223,5841	1243,1037
1248,6339	1286,4981	1329,0767	1352,3533	1378,8135
1420,7450	1424,4924	1429,0274	1447,8825	1448,8905
1454,9367	1469,1750	1479,7786	1491,3220	1525,1086
1551,5166	1570,5699	1589,5396	1600,9614	1609,0506
1631,9921	1635,0419	1652,0880	1656,1582	1670,0177
1671,6771	1692,6170	1694,5008	1697,9933	1701,7648
1705,9859	1714,4589	1719,7907	1720,9682	1737,1874
1757,5080	1785,9356	1830,6103	1849,0169	1940,9619
2919,3280	2919,9336	2920,6133	2956,6915	2966,1037
2975,3384	2997,1561	3003,3615	3005,8075	3012,5354
3021,3071	3027,8486	3028,2126	3029,3849	3029,9990
3031,2671	3031,4248	3035,1753	3036,1998	3184,9573
3303,8503	3341,0755	3375,3831	3396,6601	3404,6690
3408,1115	3423,3893	3440,8008		

TAB. B.3 – Fréquences de vibration en  $\text{cm}^{-1}$  calculées en prenant en compte les interactions électrostatiques entre les sous parties quantiques et classiques pour l'état de transition dans la réaction de Diels-Alder

atome	procédure numérique complète			sans prise en compte des interactions électrostatiques			avec prise en compte des interactions électrostatiques		
	X	Y	Z	X	Y	Z	X	Y	Z
C 1	0,10	-0,07	-0,02	0,10	-0,07	-0,02	0,10	-0,07	-0,02
C 2	-0,07	-0,03	0,09	-0,07	-0,03	0,09	-0,07	-0,03	0,09
H 3	0,00	-0,09	-0,06	0,00	-0,09	-0,06	0,00	-0,09	-0,06
H 4	-0,10	-0,08	-0,02	-0,10	-0,08	-0,02	-0,10	-0,08	-0,02
C 5	0,13	0,30	0,16	0,13	0,30	0,16	0,13	0,30	0,16
C 6	0,25	0,35	0,14	0,25	0,35	0,14	0,25	0,35	0,14
H 7	-0,01	0,06	0,03	-0,01	0,06	0,03	-0,01	0,06	0,03
C 8	-0,07	-0,30	-0,22	-0,07	-0,30	-0,22	-0,07	-0,30	-0,22
C 9	-0,02	0,01	-0,02	-0,02	0,01	-0,02	-0,02	0,01	-0,02
H 10	0,01	-0,03	-0,07	0,01	-0,03	-0,07	0,01	-0,03	-0,07
C 11	-0,34	-0,29	-0,12	-0,34	-0,29	-0,12	-0,34	-0,29	-0,12
H 12	-0,09	-0,03	-0,14	-0,09	-0,03	-0,14	-0,09	-0,03	-0,14
H 13	0,00	-0,15	0,10	0,00	-0,15	0,10	0,00	-0,15	0,10
H 14	0,20	0,20	0,13	0,20	0,20	0,13	0,20	0,20	0,13
H 15	0,02	0,09	0,00	0,02	0,09	0,00	0,03	0,09	0,00
H 16	0,09	0,17	0,15	0,09	0,17	0,15	0,09	0,17	0,15
C 17	-0,01	-0,01	-0,02	-0,01	-0,01	-0,02	-0,01	-0,01	-0,02
O 18	0,01	0,00	0,00	0,01	0,00	0,00	0,01	0,00	0,00
O 19	0,00	0,01	0,00	0,00	0,01	0,00	0,00	0,01	0,00
C 20	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00

TAB. B.4 – Expressions du vecteur de transition suivant la méthode de calcul utilisée dans l'espace des degrés de liberté cartésiens des atomes. Seuls les déplacements liés aux atomes quantiques sont reportés, les valeurs liés aux atomes classiques étant systématiquement nulles.



## Annexe C

### Données associées au deuxième chapitre

	QM	MK	MK×1.1	Chelp	ChelpG	OPEP
$C_1$	-0.354	-0.349	-0.349	-0.346	-0.348	-0.352
$C_2$	-0.250	-0.251	-0.251	-0.249	-0.251	-0.254
$H_3$	0.218	0.221	0.221	0.224	0.222	0.218
$H_4$	0.197	0.193	0.193	0.198	0.194	0.185
$C_5$	0.807	0.791	0.791	0.793	0.792	0.787
$C_6$	0.223	0.224	0.224	0.230	0.225	0.214
$O_7$	-0.584	-0.587	-0.587	-0.572	-0.585	-0.614
$C_8$	-0.672	-0.488	-0.488	-0.494	-0.490	-0.457
$O_9$	0.163	0.271	0.274	0.399	0.296	0.143

TAB. C.1 – charges de Mulliken des atomes de la partie quantique du diénophile, suivant le jeu de charges classiques portées par la partie classique de la même molécule, la numérotation des atomes se fait comme indiqué par la figure C.1

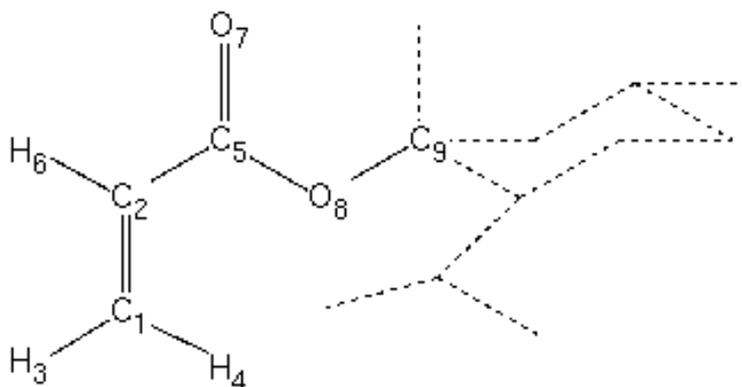


FIG. C.1 – Numérotation des atomes du diénophile

espèce	Vide	Solvant
cyclopentadiène	-192.79172	-192,79209
acrylate de (-)-menthyle s-cis syn	-265,63599	-265,63841
acrylate de (-)-menthyle s-cis anti	-265,62659	-265,62814
acrylate de (-)-menthyle s-trans syn	-265,63476	-265,63637
acrylate de (-)-menthyle s-trans anti	-265,62562	-265,62730

TAB. C.2 – Energies absolues (u.a.) des réactifs dans le vide et dans le solvant

TS	$E_{TS}$ (u.a.)	$E_a$ (kcal.mol <sup>-1</sup> )	$E_{prods}$ (u.a.)	$\Delta E$ (kcal.mol <sup>-1</sup> )
endo s-cis RE syn	-458,372459	34,673	-458,462569	-21,872
endo s-cis RE anti	-458,359942	42,528	-458,450848	-14,517
endo s-cis SI syn	-458,371177	35,478	-458,461493	-21,197
endo s-cis SI anti	-458,362421	40,972	-458,452973	-15,850
endo s-trans RE syn	-458,379067	30,527	-458,460110	-20,328
endo s-trans RE anti	-458,369506	36,527	-458,452600	-15,616
endo s-trans SI syn	-458,370311	36,021	-458,462895	-22,076
endo s-trans SI anti	-458,362421	40,972	-458,451856	-15,149
exo s-cis RE syn	-458,369793	36,346	-458,461507	-21,205
exo s-cis RE anti	-458,361399	41,614	-458,453217	-16,003
exo s-cis SI syn	-458,371597	35,215	-458,463278	-22,316
exo s-cis SI anti	-458,367407	37,843	-458,452802	-15,743
exo s-trans RE syn	-458,369250	36,687	-458,461942	-21,478
exo s-trans RE anti	-458,369055	36,810	-458,453217	-16,003
exo s-trans SI syn	-458,367164	37,996	-458,459632	-20,029
exo s-trans SI anti	-458,359193	42,998	-458,450519	-14,310

TAB. C.3 – Energies absolues des TS et produits et énergies d'activation et de réaction associées dans le vide

TS	$E_{TS}$ (u.a.)	$E_a$ (kcal.mol <sup>-1</sup> )	$E_{prods}$ (u.a.)	$\Delta E$ (kcal.mol <sup>-1</sup> )
endo s-cis RE syn	-458,374028	35,439	-458,464971	-21,628
endo s-cis RE anti	-458,361324	43,411	-458,457363	-16,855
endo s-cis SI syn	-458,372496	36,400	-458,462784	-20,256
endo s-cis SI anti	-458,364042	41,706	-458,454459	-15,032
endo s-trans RE syn	-458,380979	31,077	-458,462436	-20,038
endo s-trans RE anti	-458,371453	37,055	-458,453862	-14,658
endo s-trans SI syn	-458,372429	36,442	-458,464087	-21,074
endo s-trans SI anti	-458,367506	39,532	-458,453197	-14,240
exo s-cis RE syn	-458,370915	36,042	-458,464512	-21,341
exo s-cis RE anti	-458,362901	38,745	-458,454204	-14,872
exo s-cis SI syn	-458,373068	37,392	-458,463560	-20,743
exo s-cis SI anti	-458,368759	42,421	-458,454566	-15,100
exo s-trans RE syn	-458,371335	38,719	-458,460953	-19,107
exo s-trans RE anti	-458,370944	43,400	-458,452194	-13,611
exo s-trans SI syn	-458,368802	37,129	-458,464274	-21,191
exo s-trans SI anti	-458,361341	37,374	-458,454571	-15,102

TAB. C.4 – Energies absolues des TS et produits et énergies d'activation et de réaction associées dans le solvant

	Vide				Solvant			
	TS		Prod		TS		Prod	
	CC1	CC2	CC1	CC2	CC1	CC2	CC1	CC2
endo s-cis RE syn	2,089	2,300	1,555	1,562	2,086	2,302	1,555	1,562
endo s-cis RE anti	2,088	2,313	1,557	1,568	2,086	2,315	1,558	1,566
endo s-cis SI syn	2,080	2,319	1,557	1,569	2,078	2,321	1,557	1,569
endo s-cis SI anti	2,089	2,301	1,556	1,565	2,086	2,304	1,556	1,565
endo s-trans RE syn	2,110	2,275	1,555	1,559	2,100	2,285	1,555	1,559
endo s-trans RE anti	2,102	2,284	1,555	1,553	2,093	2,295	1,555	1,553
endo s-trans SI syn	2,099	2,284	1,555	1,553	2,089	2,295	1,555	1,553
endo s-trans SI anti	2,098	2,300	1,555	1,556	2,090	2,309	1,555	1,556
exo s-cis RE syn	2,102	2,305	1,556	1,572	2,098	2,310	1,556	1,572
exo s-cis RE anti	2,109	2,298	1,556	1,571	2,091	2,324	1,556	1,571
exo s-cis SI syn	2,095	2,325	1,556	1,571	2,091	2,329	1,556	1,571
exo s-cis SI anti	2,104	2,305	1,554	1,555	2,100	2,310	1,553	1,555
exo s-trans RE syn	2,109	2,292	1,554	1,555	2,098	2,302	1,553	1,555
exo s-trans RE anti	2,119	2,294	1,556	1,571	2,102	2,303	1,556	1,571
exo s-trans SI syn	2,106	2,298	1,557	1,570	2,098	2,308	1,557	1,570
exo s-trans SI anti	2,111	2,296	1,557	1,574	2,102	2,308	1,557	1,574

TAB. C.5 – Longueurs des liaisons CC1 et CC2 dans les TS et produits : vide et le solvant

	Vide				Solvant			
	TS		Prod		TS		Prod	
	CC1	CC2	CC1	CC2	CC1	CC2	CC1	CC2
endo s-cis RE syn	0,436	0,326	0,964	0,962	0,438	0,325	0,965	0,962
endo s-cis RE anti	0,429	0,323	0,964	0,961	0,430	0,323	0,964	0,960
endo s-cis SI syn	0,435	0,322	0,964	0,960	0,437	0,321	0,964	0,960
endo s-cis SI anti	0,436	0,327	0,964	0,959	0,438	0,325	0,965	0,959
endo s-trans RE syn	0,426	0,332	0,965	0,967	0,431	0,328	0,965	0,967
endo s-trans RE anti	0,427	0,332	0,964	0,968	0,431	0,327	0,963	0,968
endo s-trans SI syn	0,430	0,333	0,964	0,968	0,436	0,328	0,963	0,968
endo s-trans SI anti	0,425	0,327	0,965	0,967	0,430	0,323	0,965	0,967
exo s-cis RE syn	0,431	0,331	0,965	0,956	0,434	0,328	0,965	0,956
exo s-cis RE anti	0,435	0,321	0,964	0,963	0,436	0,320	0,964	0,963
exo s-cis SI syn	0,429	0,327	0,966	0,960	0,431	0,325	0,966	0,960
exo s-cis SI anti	0,430	0,331	0,965	0,957	0,432	0,328	0,966	0,957
exo s-trans RE syn	0,431	0,333	0,964	0,959	0,435	0,329	0,964	0,959
exo s-trans RE anti	0,425	0,337	0,965	0,956	0,430	0,333	0,965	0,956
exo s-trans SI syn	0,429	0,335	0,963	0,964	0,434	0,331	0,963	0,963
exo s-trans SI anti	0,428	0,336	0,965	0,957	0,432	0,330	0,966	0,957

TAB. C.6 – Ordres des liaisons de Mayer CC1 et CC2 dans les TS et produits : vide et le solvant

	s-trans-syn			s-trans anti		
	vide	solvant	variation	vide	solvant	variation
$C_1$	-0,351	-0,349	0,001	-0,351	-0,350	0,001
$C_2$	-0,249	-0,254	-0,005	-0,254	-0,259	-0,005
$H_3$	0,216	0,220	0,003	0,216	0,220	0,003
$H_4$	0,191	0,198	0,007	0,190	0,197	0,008
$C_5$	0,792	0,795	0,002	0,805	0,808	0,003
$C_6$	0,218	0,215	-0,003	0,216	0,212	-0,004
$O_7$	-0,589	-0,601	-0,012	-0,591	-0,604	-0,013
$C_8$	-0,473	-0,468	0,005	-0,483	-0,47	0,006
$O_9$	0,271	0,272	0,001	0,278	0,279	0,001
	s-cis syn			s-cis anti		
	vide	solvant	variation	vide	solvant	variation
$C_1$	-0,350	-0,353	-0,002	-0,350	-0,353	-0,002
$C_2$	-0,253	-0,253	0,000	-0,255	-0,255	0,000
$H_3$	0,223	0,218	-0,004	0,223	0,218	-0,005
$H_4$	0,191	0,198	0,007	0,191	0,198	0,007
$C_5$	0,793	0,793	0,000	0,803	0,804	0,001
$C_6$	0,217	0,220	0,003	0,215	0,217	0,002
$O_7$	-0,601	-0,610	-0,009	-0,599	-0,608	-0,009
$C_8$	-0,465	-0,460	0,005	-0,479	-0,474	0,005
$O_9$	0,272	0,273	0,001	0,279	0,280	0,001

TAB. C.7 – Charges de Mulliken sur les différents isomères du dièneophile dans le vide et en solvant et variation des charges du au passage du vide au solvant. La numérotation est celle de la figure C.1



## Annexe D

### Données associées au troisième chapitre

Molécule		C <sub>2</sub> H <sub>5</sub> -COOH	C <sub>2</sub> H <sub>5</sub> -COO <sup>-</sup>	C <sub>2</sub> H <sub>5</sub> -NH <sub>2</sub>	C <sub>2</sub> H <sub>5</sub> -NH <sub>3</sub>
Méthyle	C	-0,1361	0,2253	-0,0358	-0,4626
	H	0,0454	-0,0751	0,0119	0,1542
CH <sub>2</sub>	C	0,0623	-0,1659	0,2071	0,3469
	H	0,0155	-0,0201	0,0024	0,0456
Groupe	C : 0,6600	C : 0,6600	C : 0,8731	N : -0,8520	N : -0,5064
	=O : -0,5815	=O : -0,5815	O : -0,8335	H : 0,3201	H : 0,3560
	-O : -0,6105	-O : -0,6105			
	H : 0,4387	H : 0,4387			
Molécule		C <sub>2</sub> H <sub>5</sub> -OH	C <sub>2</sub> H <sub>5</sub> -O <sup>-</sup>	C <sub>2</sub> H <sub>5</sub> -SH	C <sub>2</sub> H <sub>5</sub> -S <sup>-</sup>
Méthyle	C	-0,1239	0,6842	-0,2724	0,2749
	H	0,0413	-0,2281	0,09085	-0,0916
CH <sub>2</sub>	C	0,2766	0,1703	0,1295	-0,0408
	H	-0,0165	-0,1531	0,0258	0,0013
Groupe	O : -0,6305	O : -0,6305	O : -0,8640	S : -0,3558	S : -0,9619
	H : 0,3870	H : 0,3870		H : 0,1748	

TAB. D.1 – Jeux de charges calculés pour les différents solutés envisagés à l'aide de la méthode RESP au niveau MP2/6-31G\*

structure	QM	AMBER	QM +SLBO	soluté QM/MM, eau QM	soluté QM, eau MM
COOH	-8,442	-10,533	-8,631	-8,338	-13,804
COO <sup>-</sup>	-16,859	-22,410	-18,817	-18,015	-22,512
NH <sub>2</sub>	-5,710	-6,117	-6,850	-9,257	-9,147
NH <sub>3</sub> <sup>+</sup>	-17,420	-18,294	-17,813	-16,265	-20,130
OH (d)	-4,094	-6,188	-4,101	-2,488	-8,652
OH (a)	-5,535	-5,797	-5,356	-7,973	-7,304
O <sup>-</sup>	-18,783	-20,048	-22,235	-17,313	-20,233
SH (d)	-1,653	-2,987	-1,737	-0,923	-5,772
SH (a)	-3,715	-3,460	-3,593	-5,274	-5,763
S <sup>-</sup>	-12,054	-18,333	-12,649	-12,175	-17,784

TAB. D.2 – Energies d’interaction soluté/eau suivant les approches considérées en kcal.mol<sup>-1</sup>

structure	QM	AMBER	QM +SLBO	soluté QM/MM, eau QM	soluté QM, eau MM
COOH	1,942	1,934	1,940	1,885	1,903
COOH	1,776	1,793	1,767	1,800	1,675
COO <sup>-</sup>	1,952	1,864	1,948	1,950	1,859
COO <sup>-</sup>	1,976	1,874	1,975	1,933	1,886
NH <sub>2</sub>	1,933	1,892	1,934	1,897	1,822
NH <sub>3</sub> <sup>+</sup>	1,720	1,831	1,713	1,793	1,773
OH (d)	1,913	1,779	1,919	1,998	1,677
OH (a)	1,905	1,815	1,906	1,848	1,773
O <sup>-</sup>	1,509	1,659	1,475	1,989	1,681
SH (d)	2,174	1,962	2,175	2,292	1,805
SH (a)	2,435	2,231	2,442	2,564	2,232
S <sup>-</sup>	2,199	1,943	2,195	2,215	2,412

TAB. D.3 – Longueurs des liaisons hydrogène considérées en Ångstroms

structure	Ref. QM : MP2/ 6-311++G(2d,2p)	MP2/6-31G* paramètres adaptés
COOH	-8,767	-8,555
COO <sup>-</sup>	-18,070	-15,978
NH <sub>2</sub>	-6,938	-6,687
NH <sub>3</sub> <sup>+</sup>	-16,653	-14,167
OH (d)	-4,324	-5,456
OH (a)	-5,321	-4,860
O <sup>-</sup>	-21,831	-14,703
SH (d)	-2,019	-2,735
SH (a)	-4,160	-4,260
S <sup>-</sup>	-14,082	-14,294

TAB. D.4 – Energies d'interaction soluté/eau (kcal.mol<sup>-1</sup>) en utilisant les paramètres adaptés, et références quantiques auxquelles elles sont comparées

structure	Ref. QM : MP2/ 6-311++G(2d,2p)	MP2/6-31G* paramètres adaptés
COOH	1,959	2,055
COOH	1,806	1,795
COO <sup>-</sup>	1,961	1,972
COO <sup>-</sup>	1,990	2,002
NH <sub>2</sub>	1,916	1,880
NH <sub>3</sub> <sup>+</sup>	1,735	1,896
OH (d)	1,943	1,796
OH (a)	1,901	1,860
O <sup>-</sup>	1,476	1,765
SH (d)	2,230	2,004
SH (a)	2,408	2,234
S <sup>-</sup>	2,156	2,220

TAB. D.5 – Longueurs des liaisons hydrogène considérées en Ångstroms en utilisant les paramètres adaptés, et références quantiques auxquelles elles sont comparées

base	Nos paramètres			TIP3P		
	RHF	B3LYP	MP2	RHF	B3LYP	MP2
3-21G	-5,565	-4,650	-4,838	-8,362	-7,082	-7,340
6-31G*	-5,066	-4,547	-4,860	-7,620	-6,916	-7,348
6-311G(d,p)	-4,955	-4,523	-4,529	-7,437	-6,843	-6,849
6-311++G(2d,2p)	-4,992	-4,695	-4,793	-7,449	-6,996	-7,132

TAB. D.6 – Energies d’interaction entre une molécule quantique d’éthanol et une molécule d’eau traitée au niveau classique suivant le jeu de paramètres adopté, la base d’orbitales et la méthode. suivant le jeu de paramètres adopté. La référence quantique MP2/6-311++G(2d,2p) est à  $-5,319 \text{ kcal.mol}^{-1}$

base	Nos paramètres			paramètres issus de TIP3P		
	RHF	B3LYP	MP2	RHF	B3LYP	MP2
3-21G	-14,237	-13,373	-13,602	-19,787	-18,637	-18,951
6-31G*	-14,033	-13,440	-13,996	-19,570	-18,803	-19,620
6-311G(d,p)	-13,842	-13,454	-13,483	-19,301	-18,830	-18,861
6-311++G(2d,2p)	-14,180	-13,989	-14,138	-19,859	-19,650	-19,866

TAB. D.7 – Energies d’interaction ( $\text{kcal.mol}^{-1}$ ) entre une molécule quantique d’éthanol et deux molécules classiques d’eau suivant le jeu de paramètres adopté, la base d’orbitales et la méthode. La référence quantique MP2/6-311++G(2d,2p) est à  $-14,420 \text{ kcal.mol}^{-1}$





## Résumé

L'objectif principal de ce travail est le développement d'outils permettant la prise en compte simultanée de l'influence du solvant et de l'environnement moléculaire dans l'étude de la réactivité chimique.

La méthode LSCF/MM permet de modéliser un système moléculaire complexe (comportant un grand nombre d'atomes) en traitant le centre réactif au niveau quantique et le reste (environnement moléculaire) au niveau classique. Dans l'approche SCRF, le solvant est décrit comme un milieu continu polarisable.

Le couplage des deux conduit à l'approche à trois niveaux qui est utilisée pour étudier la cycloaddition de l'acrylate de (-)-menthyle avec le cyclopentadiène, permettant la description des effets couplés du solvant et du groupement chiral sur la réactivité.

L'introduction de molécules explicites d'eau à proximité du centre réactif est aussi envisagée afin de modéliser les interactions spécifiques soluté/solvant comme la liaison hydrogène. Des paramètres spécifiques sont déterminés pour rendre compte de l'interaction entre eau classique et soluté quantique. Cette approche est utilisée dans l'étude des effets du solvant sur le réarrangement de Claisen.

Mots-clés :

QM/MM, LSCF/MM, solvant, SCRF, liaison hydrogène, réactivité, Diels-Alder, Claisen, hessien