

NOTE TO USERS

This reproduction is the best copy available.

UMI[®]



Université d'Ottawa • University of Ottawa



Université d'Ottawa - University of Ottawa

FACULTÉ DE ÉTUDES SUPÉRIEURES
ET POSTDOCTORALES

FACULTY OF GRADUATE AND
POSTDOCTORAL STUDIES

Étienne PARADIS

AUTEUR DE LA THÈSE - AUTHOR OF THESIS

Ph.D. (Chimie)

GRADE - DEGREE

Département de Chimie

FACULTÉ, ÉCOLE, DÉPARTEMENT - FACULTY, SCHOOL, DEPARTMENT

TITRE DE LA THÈSE - TITLE OF THE THESIS

Développement de la fonctionnalité QM/MM dans DeFT : Application aux
calculs des énergies libres

A. St-Amant

DIRECTEUR DE LA THÈSE - THESIS SUPERVISOR

CO-DIRECTEUR DE LA THÈSE - THESIS CO-SUPERVISOR

EXAMINATEURS DE LA THÈSE - THESIS EXAMINERS

J. Giorgi

N. Goto

P-N. Roy

J. Wright

J.-M. De Koninck, Ph.D.

LE DOYEN DE LA FACULTÉ DES ÉTUDES
SUPÉRIEURES ET POSTDOCTORALES

DEAN OF THE FACULTY OF GRADUATE
AND POSTDOCTORAL STUDIES

Développement de la fonctionnalité QM/MM dans DeFT
Application aux calculs des énergies libres

Étienne Paradis

Thèse présentée à la Faculté des études supérieures
et postdoctorales de l'Université d'Ottawa
en vue de l'obtention du
Doctorat en Philosophie en Chimie

Département de chimie
Faculté des sciences
Université d'Ottawa

Avril 2004

© Étienne Paradis, Ottawa, Canada, 2004



Library and
Archives Canada

Bibliothèque et
Archives Canada

Published Heritage
Branch

Direction du
Patrimoine de l'édition

395 Wellington Street
Ottawa ON K1A 0N4
Canada

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

ISBN: 0-494-01748-1

Our file Notre référence

ISBN: 0-494-01748-1

NOTICE:

The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protègent cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.


Canada

Sommaire

Depuis le début des années 1990, il y a eu beaucoup d'articles publiés sur le développement de la méthodologie QM/MM. Ces potentiels hybrides ont permis l'étude de la formation ou de la rupture de liaison pour des réactions en solution et pour des réactions enzymatiques.

Dans cette thèse, nous avons interfacé le programme DeFT, utilisant la fonctionnelle de la densité, avec le module Roar de la suite AMBER. Dans notre fonctionnalité QM/MM, nous avons utilisé une contrainte sur l'atome connecté (*link atom*) pour pouvoir redistribuer le gradient de l'atome connecté sur les atomes réels du système. Ainsi, nous sommes moins affectés par les degrés de liberté additionnels qui sont présent dans le système, lors de l'optimisation de géométrie et de simulation de dynamique moléculaire, que lorsqu'il n'y a aucune contrainte sur les atomes connectés. Pour valider cette fonctionnalité, nous avons optimisé une série de molécules pour examiner l'effet de la présence d'un atome connecté sur la distance C-C. Nous avons démontré que l'utilisation de l'atome connecté apportait une déviation moyenne de 0.020 Å. Nous avons également évalué l'affinité de proton pour une série de bases organiques. Ces mesures montrent bien qu'il soit important d'avoir plus qu'un groupement méthyle après le groupement fonctionnel pour obtenir une bonne comparaison avec des calculs complètement quantiques. Par la suite, nous

avons regardé la stabilité de cinq conformères de la dipeptide Alanine auxquelles nous avons ajouté une hélice alpha de 20 glycines.

Dans la deuxième partie de cette thèse, nous avons implémenté dans DeFT la possibilité de faire des trajectoires de dynamique moléculaire qui sont contraintes à suivre un chemin réactionnel prédéfini. Ceci nous a permis d'obtenir le profil d'énergie libre le long du chemin réactionnel en utilisant la méthode de croissance lente. Cette fonctionnalité a été validée avec une série de réactions électrocycliques. Ces différentes simulations ont montré que l'on pouvait obtenir un profil d'énergie libre très lisse mais avec un coût de calcul très élevé. Dans la dernière partie, nous avons optimisé une méthode hybride (AM1/DFT) pour obtenir le profil d'énergie libre d'une réaction. En utilisant la méthode de perturbation thermodynamique non-Boltzmann, il est possible d'obtenir un profil d'énergie libre de niveau de la fonctionnel de la densité en diminuant le temps de calcul par un facteur d'environ 12.

Remerciements

J'aimerais dans un premier temps remercier mon directeur de recherche le Dr Alain St-Amant pour avoir accueilli un expérimentateur physicien dans son équipe.

Je tiens à remercier mes compagnons de travail, c'est-à-dire Michelle Shaw, Sébastien Osborne et Delphine Courmier.

Un gros merci va à ma femme Marie-Josée Gour qui sait toujours comment m'encourager.

à ma femme

Table des matières

Sommaire.....	i
Remerciements	iii
Table des matières	v
Liste des tableaux	viii
Liste des figures.....	ix
Liste des figures.....	ix
Abréviation	xi
Introduction	1
Chapitre 1: Méthodologie.....	8
1.1 La mécanique quantique.....	8
1.1.1 L'approximation de Born-Oppenheimer	9
1.2 La méthode de Hartree-Fock	11
1.3 La fonctionnel de la densité.....	14
1.3.1 Équations de Kohn-Sham	17
1.3.2 La fonctionnel d'échange et de corrélation	19
1.4 Les fonctions de bases	21
1.5 Les méthodes semi-empiriques	24
1.6 Les champs de force	25
1.7 Dynamique moléculaire	27
1.7.1 Intégration des trajectoires	28
1.7.2 Contrôle de la température	30
1.8 Variation d'énergie libre	32
Chapitre 2 : fonctionnalité QM/MM	39
2.1 Concept de base.....	39

2.2 Traitements des atomes connectés dans Roar/DeFT	44
Chapitre 3: Validation de la fonctionnalité QM/MM.....	48
3.1 Éthane.....	49
3.2 Influences de l'atome connecté sur les liaisons σ C-C	54
3.3 Affinité de proton	56
3.3 Analyse conformationnelle de la dipeptide alanine	61
3.4 Discussion	67
Chapitre 4 : Énergie libre le long du chemin réactionnel (IRC)	69
4.1 Introduction	69
4.2 Algorithme pour le calcul de l'énergie libre	72
4.3 Procédure pour les simulations.....	76
4.4 Résultats	78
4.4.1 Ouverture du cycle pour le cyclobutène en 1,3-butadiène	79
4.4.2 Réaction de l'ouverture du cycle de l'oxetène	85
4.4.3 Réaction de cyclisation de l'hexatriène.....	89
4.4.4 Migration 1,3 oxygène-oxygène du méthyle formate et du trifluorométhyle formate.....	94
4.4.5 Isomérisation du cyanure d'hydrogène	100
4.5 Discussion	103
4.6 Conclusion.....	107
Chapitre 5 : Développement d'une méthode hybride AM1/DFT pour le calcul de l'énergie libre	108
5.1 Introduction	108
5.2 Méthode des perturbations thermodynamiques non-Boltzmann ⁷³	111
5.3 Protocole pour les simulations	114
5.4 Test pour diminuer l'erreur dans le calcul de l'énergie libre	116
5.5 Choix du potentiel de référence.....	120

5.6 Exemples d'utilisation de la méthode de perturbation thermodynamique non-Boltzman	126
5.6.1 Ouverture du cycle pour le cyclobutène	127
5.6.2 Ouverture du cycle pour l'oxetène	129
5.6.3 Transfert de proton dans la Glycine	132
5.6.4 Comparaison entre les méthodes pour le calcul de l'énergie libre	139
5.7 Conclusion	144
Conclusion	145
Originalité du travail	148
Bibliographie	149

Liste des tableaux

Tableau I : Résultats pour les distances C-C pour une série de molécules.	55
Tableau II. Calculs d'affinités de proton sur les bases organiques	58
Tableau III. Affinité de proton avec la présence d'une charge ponctuelle externe	60
Tableau IV. Résultats : coupure est entre CH ₂ -CO.	64
Tableau V. Résultats : coupure est entre NH-CH ₂	64
Tableau VI. Résultats pour les cinq conformations de la dipeptide alanine.	65
Tableau VII. Paramètre pour les géométries de cyclobutène, 1,3-butadiène et de l'état de transition	82
Tableau VIII. Paramètre des géométries d'oxetène, d'acrolein et de l'état de transition	86
Tableau IX. Paramètres pour les géométries de cyclohexadiène, d'hexatriène et de l'état de transition	92
Tableau X. Paramètres des états de transition pour la migration entre l'oxygène 1 et 3 pour le méthyle et le trifluorométhyle méthyle de formate.	95
Tableau XI. Paramètre pour l'isomérisation du cyanure d'hydrogène	101
Tableau XII. Comparaison des résultats d'énergie libre entre la dynamique moléculaire et l'analyse vibrationnelle.....	106
Tableau XIII. Test pour le nombre d'itération pour équilibrer et échantillonner le système.	118
Tableau XIV. Erreur à chaque fenêtre entre les résultats de la méthode mixte et la méthode complètement DFT.	130
Tableau XV. Comparaison entre la méthode de perturbation thermodynamique complète quantique et mixte (semi-empirique / quantique).....	140

Liste des figures

Figure 1. Définition de la direction avant et arrière pour une fenêtre.	35
Figure 2. Représentation de l'interface entre la partie quantique et classique.	43
Figure 3. Modèle pour la molécule d'éthane où l'atome L représente l'atome connecté.	50
Figure 4. Potentiel d'étirement du lien C-C dans l'éthane.	52
Figure 5. Potentiel de l'angle de torsion H-C-C-H dans l'éthane.	53
Figure 6. Les différentes structures pour le propanol + Na ⁺	59
Figure 7 : Conformations de la dipeptide alanine.	66
Figure 8. Figure montrant l'hystérésis entre la trajectoire vers l'avant et la trajectoire vers l'arrière.	70
Figure 9. Paramètre de contrôle le long de la trajectoire de dynamique moléculaire	77
Figure 10 : Schéma du mouvement conrotatoire et disrotatoire	80
Figure 11. Schéma pour la réaction de l'ouverture du cycle du cyclobutène (1).....	83
Figure 12. Profil d'énergie et d'énergie libre pour la réaction d'ouverture du cycle du cyclobutène le long du chemin réactionnel.	84
Figure 13. Schéma pour la réaction d'ouverture du cycle de l'oxetène (1)	87
Figure 14. Profil d'énergie et d'énergie libre pour la réaction d'ouverture du cycle de l'oxetène le long du chemin réactionnel.....	88
Figure 15. Schéma pour la réaction de cyclisation du hexatriène (3) en cyclohexadiène (1)	90
Figure 16. Profil d'énergie et d'énergie libre pour la réaction de cyclisation de l'hexatriène le long du chemin réactionnel.	93

Figure 17. Schéma représentant la réaction de migration du méthyle formate entre l'oxygène 1 et 3.	96
Figure 18. Profil d'énergie libre pour la migration 1,3 oxygène-oxygène du méthyle formate le long du chemin réactionnel.	98
Figure 19. Profil d'énergie libre pour la migration 1,3 oxygène-oxygène pour le trifluorométhyle formate le long du chemin réactionnel.	99
Figure 20. Profil d'énergie libre pour la réaction d'isomérisation du cyanure d'hydrogène le long du chemin réactionnel.	102
Figure 21. Simulation où λ passe de $0 \rightarrow 1$ et de $1 \rightarrow 0$ pour la réaction d'ouverture du cyclobutène en utilisant le potentiel AM1.	119
Figure 22. Variation de l'énergie le long du chemin réactionnel pour la réaction entre le NH_3 et H_2CO	122
Figure 23. Résultats pour les tests I et II.	123
Figure 24. Résultats pour le test III.	124
Figure 25. Profil d'énergie libre pour la réaction d'ouverture du cyclobutène en utilisant la méthode de perturbation thermodynamique.	128
Figure 26. Profil d'énergie libre pour la réaction d'ouverture de l'oxetène en utilisant la méthode de perturbation thermodynamique.	131
Figure 27 : Structure pour la forme neutre et la forme de l'ion dipolaire	134
Figure 28. Profil d'énergie libre pour la réaction du transfert de proton dans la glycine en utilisant la méthode de perturbation thermodynamique.	137
Figure 29 . Profil d'énergie pour la réaction de transfert du proton dans la glycine en utilisant un modèle de solvation.	138
Figure 30. Profil d'énergie pour la réaction SN_2 ($\text{Cl CH}_3\text{Cl}$).	143

Abréviation

AMBER	Assisted Model Building with Energy Refinement
B3LYP	Combinaison des fonctionnelles de Becke 1988 pour l'échange et LYP pour la corrélation
BP86	Combinaison des fonctionnelles de Becke 1988 pour l'échange et P86 pour la corrélation
DFT	Théorie de la fonctionnelle de la Densité
HOMO	Orbitale moléculaire occupée de plus haute énergie
IRC	Intrinsic Reaction Coordinate
KS	Kohn-Sham
LUMO	Orbitale moléculaire inoccupée de plus basse énergie
LSDA	Approximation de la Densité Locale de Spin
MM	Mécanique Moléculaire
QM	Mécanique Quantique

QM/MM	Mécanique Quantique / Mécanique Moléculaire
SCF	Self-Consistent Field
VWN	Fonctionnelle de corrélation de Vosko, Wilk et Nusair

Introduction

Depuis quelques années, la chimie informatique joue un rôle de plus en plus important en chimie. Ce rôle s'est accentué avec l'augmentation rapide de la puissance et la rapidité des ordinateurs qui nous permet maintenant de modéliser des systèmes moléculaires de plus en plus sophistiqués. Il est maintenant possible de calculer les propriétés structurales, électroniques et dynamiques de systèmes complexes, tels que les protéines, les acides nucléiques et les matériaux avec une exactitude raisonnable. Ces systèmes étaient généralement étudiés avec des logiciels de modélisation moléculaires comme AMBER¹, GROMOS² et CHARMM³. Par contre, la modélisation moléculaire utilise des modèles empiriques qui ne permettent pas d'étudier les ruptures ou la formation de liaisons chimiques.

L'utilisation de la mécanique quantique permet par contre d'étudier la formation et la rupture des liaisons. On regroupe les méthodes quantiques en deux classes : les modèles semi-empiriques et les modèles ab-initio. Les modèles semi-empiriques utilisent un certain nombre de paramètres empiriques. L'utilisation de modèles semi-empiriques permet d'obtenir une bonne approximation des propriétés physiques que l'on désire. Ces dernières années, comme pour le développement des ressources informatiques, les programmes qui utilisent des modèles ab-initio sont devenus de plus en plus sophistiqués et beaucoup plus rapides. Les méthodes ab-

initio (Hartree-Fock, MP2⁴, CCSD⁵, DFT, ...) permettent d'obtenir des descriptions précises de systèmes moléculaires à 0 K dans la phase gazeuse.

Typiquement, pour les méthodes ab-initio le temps de calcul augmente par N^4 où le N représente le nombre d'atomes du système et augmente par K^n avec la grosseur de la base qui est utilisée. Pour diminuer ce temps de calcul, la méthode *divide-and-conquer* a été développée par Yang⁶ pour la fonctionnelle de la densité. Cette méthode permet de ramener le temps de calcul d'une façon linéaire avec le nombre d'atomes du système. Cette méthode est également disponible pour l'approche semi-empirique⁷. Mais pour un système enzymatique, le nombre d'atomes étant très grand, le temps de calcul et la capacité de stockage des données deviendraient trop importants. Donc, l'utilisation d'une méthode quantique seule pour ce genre de systèmes est presque impensable pour le moment. Il est également très utile de connaître le comportement des molécules en solution. Des modèles de solvation (Onsager⁸, PCM⁹, COSMO¹⁰) ont été développés pour étudier ces systèmes au niveau quantique. Mais, il est très souvent important de traiter explicitement les molécules d'eaux qui sont directement autour de la molécule.

Le développement de potentiel hybride mécanique quantique/mécanique moléculaire (QM/MM)^{11,12,13} permet maintenant d'exécuter ce genre de simulations. C'est ainsi que les atomes, faisant partie du système réactionnel, ont été traités par la

mécanique quantique tandis que les autres atomes, qui représentent les contraintes de l'environnement, sont traités par la mécanique moléculaire.

Les concepts de bases, sur l'utilisation des potentiels hybrides, ont été publiés en 1976 par Warshel et Levitt¹¹ avec leur travail sur la lysozyme. Ils ont introduit une méthode qui traite une petite portion du système réactionnel par des calculs semi-empiriques, et en même temps, la protéine et le solvant ont été traités par des fonctions d'énergie empirique. Les parties classiques et quantiques de l'énergie sont couplées en tenant compte des changements des charges QM dues au potentiel électrostatique provenant des charges ponctuelles et des dipôles induits dans la région classique. Ils ont également considéré les interactions électrostatiques et van der Waals entre les atomes quantiques et les régions classiques. La liaison de frontière, σ_{Q-M} , est décrite par deux orbitales hybrides. Cette liaison, σ_{Q-M} , est incluse dans le calcul de la partie quantique en considérant une orbital hybride supplémentaire situé sur l'atome classique. Par la suite, Singh et Kollman¹⁴ ont publié un article utilisant la méthode QM/MM pour l'étude d'une réaction SN2. Dans cet article, les auteurs ont introduit la notion d'atome *dummy* pour compléter la valence de la partie quantique. Le grand intérêt de l'utilisation de la méthode QM/MM est surtout venu après l'article publié par Field, Bash et Karplus¹². C'est dans cet article que les auteurs ont introduit le concept des atomes connectés (link atom).

L'utilisation d'un atome connecté consiste à compléter la valence de la partie quantique avec un atome monovalent. Généralement, l'atome qui est utilisé est un atome d'hydrogène. Parfois des atomes d'halogènes sont utilisés. L'utilisation d'un atome connecté amène généralement un lien σ_{Q-M} trop grand. Ce lien est trop grand principalement parce que les interactions électrostatiques autour de la frontière ne sont pas très bien décrites. Ce problème est résolu en annulant les charges sur les atomes classiques, mais pose des problèmes si les charges proches de la frontière sont importantes. Dans le même esprit que l'approche d'orbitale hybride proposé par Warshel et Levitt¹¹, Rivail et al.¹⁵ ont développé une méthode, le champ auto-cohérent local, dans laquelle des orbitales hybrides statiques sont utilisées pour saturer la valence de la liaison de frontière entre la partie quantique et la partie classique. La population des orbitales locale est obtenue en optimisant, par calcul quantique, sur un modèle simple. Ce modèle est considéré transférable à un système plus complexe. Par contre, cette paramétrisation doit être refait si on change de système. Cette méthode a déjà été implémentée dans des programmes utilisant les méthodes empiriques, ab-initio et utilisant la théorie de la fonctionnelle de la densité. Une autre approche développée par Antes et Thiel¹⁶, pour les méthodes semi-empiriques, consiste à utiliser un pseudo atome dont les paramètres sont ajustés pour simuler l'influence de l'environnement MM. Ce pseudo-atome représente généralement un groupement alkyle. Cette méthode a également été implantée pour

la DFT. Dans cette méthode, les paramètres du pseudo-atome sont également ajustés en utilisant un pseudo-potentiel qui est ajusté avec des systèmes modèles.

Généralement, on classe les programmes QM/MM en quatre grandes classes¹³ :

- A. Le modèle QM/MM le plus simple, les interactions QM/MM ne tiennent pas en compte de la polarisation des atomes QM par l'environnement MM.
- B. Le modèle permet la polarisation du système QM par le système MM, y compris les charges MM dans le calcul QM. Le calcul électronique inclut les effets du système MM.
- C. Ce modèle dépasse le modèle B en incluant la polarisation de la région MM également (ex : interaction de dipôle).
- D. Ce modèle inclut la polarisation auto-cohérente de la région MM par un processus itératif.

Les calculs d'un seul cycle auto-cohérent (*single point*) et les optimisations des géométries peuvent apporter une multitude d'informations. Généralement, on retire les énergies et les enthalpies d'activation et de la réaction. Ces informations sont importantes, mais la quantité d'intérêt pour beaucoup de chimistes est l'énergie libre. Un des principaux intérêts dans les simulations de systèmes chimiques ou biologiques est de connaître le changement d'énergie libre le long des coordonnées

d'une réaction. On obtient ce changement d'énergie libre en calculant la probabilité de présence du système le long du chemin réactionnel. Les applications conventionnelles de la mécanique quantique (semi-empirique, ab-initio, DFT) peuvent fournir des énergies libres approximatives en utilisant l'analyse des modes normaux. Ces résultats sont seulement valides si l'approximation harmonique est valide. Cette approximation n'est pas très efficace si les systèmes sont très grands. Donc, si l'approximation harmonique ne peut être utilisée, nous devons balayer la surface d'énergie potentielle d'un système moléculaire. Alors, le prélèvement à proximité immédiate des points stationnaires de la surface d'énergie potentielle est insuffisant. Pour obtenir des énergies libres précises, nous utilisons généralement la méthode des perturbations thermodynamiques¹⁷ ou la méthode d'intégration thermodynamique¹⁸. Avec ces méthodes, nous devons échantillonner suffisamment la surface d'énergie potentielle pour bien représenter l'ensemble des états du système. Ces techniques peuvent être améliorées en utilisant un échantillonnage préférentiel (Umbrella Sampling)¹⁹ pour mieux représenter l'espace des phases que nous désirons étudier. Ainsi, en combinant une des méthodes pour le calcul de l'énergie libre avec un potentiel hybride QM/MM, il est possible d'évaluer le profil d'énergie libre d'une réaction en solution.

L'objectif principal de cette thèse est d'introduire la fonctionnalité QM/MM dans le logiciel DeFT qui utilise la fonctionnelle de la densité. Cette fonctionnalité

sera introduite en interfaçant le module Roar de AMBER avec le logiciel DeFT. Cette nouvelle fonctionnalité nous permettra d'étudier l'énergie libre pour des réactions en solution. Le deuxième objectif de cette thèse est d'introduire des méthodes de calculs pour l'énergie libre dans DeFT. Dans un premier temps, le calcul de l'énergie libre le long du chemin réactionnel sera introduit et dans un deuxième temps, nous utiliserons la méthode d'échantillonnage non-Boltzmann. Cette méthode va permettre de diminuer le temps de calculs.

Dans le chapitre 1, nous présenterons les outils théoriques utilisés dans cette thèse. Par la suite, le chapitre 2 présentera l'interface QM/MM et le traitement que nous avons utilisé pour couper les liens covalents entre la partie quantique et la partie classique. Dans le chapitre 3, nous présenterons les résultats pour la validation de la fonctionnalité QM/MM. Le chapitre 4 introduira la méthode de calcul de l'énergie libre le long de la réaction pour des calculs complètement quantiques. Le chapitre 5 sera consacré au développement de la méthode qui utilise un échantillonnage non-Boltzmann pour le calcul de l'énergie libre d'une réaction. Cette méthode nous permettra d'obtenir des profils d'énergie libre le long d'une coordonnée réactionnelle comparable à ceux obtenus en utilisant seulement des calculs de niveau de la fonctionnelle de la densité pour une fraction du temps de calculs.

Chapitre 1: Méthodologie

1.1 La mécanique quantique²⁰

L'équation principale de la mécanique quantique est l'équation de Schrödinger. Pour les états stationnaires, elle s'écrit comme :

$$\hat{H}\Psi = E\Psi \quad (1.1)$$

Dans cette équation, H est l'Hamiltonien, E correspond à l'énergie et Ψ est la fonction d'onde du système. Pour une molécule simple, l'Hamiltonien, H, peut s'écrire :

$$\hat{H} = -\frac{\hbar^2}{2} \sum_{\alpha=1} \frac{1}{M_{\alpha}} \nabla_{\alpha}^2 - \frac{\hbar^2}{2m_e} \sum_{i=1}^n \nabla_i^2 + \sum_{\beta \neq \alpha}^N \sum \frac{Z_{\alpha} Z_{\beta} e^2}{R_{\alpha\beta}} - \sum_{\alpha} \sum_i^n \frac{Z_{\alpha} e^2}{R_{i\alpha}} + \sum_{i \neq j}^n \sum \frac{e^2}{r_{ij}} \quad (1.2)$$

où

- α et β représentent les noyaux;
- i et j représentent les électrons;
- M_{α} sont les masses des noyaux ;
- $R_{\alpha\beta}$ est la distance entre les noyaux α et β ;
- Z_{α} est la charge du noyau α ;
- r_{ij} est la distance entre les électrons i et j ;

- $R_{i\alpha}$ est la distance entre l'électron i et le noyau α ;
- ∇_{α}^2 est l'opérateur gradient pour les coordonnées des noyaux;
- ∇_i^2 est l'opérateur gradient pour les coordonnées électroniques.

Le premier terme est l'énergie cinétique de chaque noyau. Le deuxième terme est l'énergie cinétique des électrons. Le troisième terme correspond à la répulsion électrostatique entre les noyaux. Le quatrième terme est l'attraction électrostatique entre l'électron i et le noyau α_A . Le dernier terme est la répulsion électrostatique des électrons. Cet Hamiltonien peut se résoudre seulement pour un nombre restreint de cas car celui-ci est fonction non seulement des coordonnées électroniques mais également des coordonnées nucléaires. Ainsi, nous devons réduire la complexité de l'équation en séparant les mouvements nucléaires et électroniques. Il est alors nécessaire d'effectuer des approximations pour résoudre cette équation. L'approximation qui sera utilisée est l'approximation de Born-Oppenheimer. Le couplage spin-orbitale sera également négligé dans l'Hamiltonien moléculaire.

1.1.1 L'approximation de Born-Oppenheimer

L'approximation principale consiste à séparer les mouvements des noyaux de ceux des électrons. Cette approximation porte le nom d'approximation de Born-

Oppenheimer²¹. Les électrons étant beaucoup plus légers que les noyaux, ceux-ci peuvent rapidement ajuster leur distribution si la position des noyaux est changée. Cette approximation amène ainsi à supposer que la fonction d'onde totale puisse s'exprimer sous la forme du produit des fonctions nucléaire et électronique.

$$\Psi(r, R) = \Psi_{el}(r, R) \Psi_{nucléaire}(R) \quad (1.3)$$

L'équation 1.1 peut ainsi s'écrire sous la forme suivante :

$$\hat{H}_{el} \Psi_{el}(r, R) = E_{eff}(R) \Psi_{el}(r, R) \quad (1.4)$$

où H_{el} est l'Hamiltonien électronique, $\Psi_{el}(r; \mathbf{R})$ la fonction d'onde électronique qui dépend des coordonnées des électrons, $\mathbf{r}$, ainsi que des coordonnées des noyaux, $\mathbf{R}$, finalement E_{eff} est l'énergie effective qui est une fonction de la position des noyaux $\mathbf{R}$.

Dans cette équation, l'Hamiltonien électronique, H_{el} , s'écrit, lorsque nous utilisons les unités atomiques :

$$H_{el} = -\sum_{i=1}^n \frac{1}{2} \nabla_i^2 - \sum_{i=1}^n \sum_{\alpha=1}^N \frac{Z_{\alpha}}{r_{i\alpha}} + \sum_{i<j} \frac{1}{r_{ij}} \quad (1.5)$$

où le premier terme correspond à l'énergie cinétique des électrons, le deuxième terme à l'attraction électrons noyaux et le troisième terme à la répulsion électrons-

électrons. Dans l'équation 1.5, la fonction d'onde ψ_{el} est représentée par une fonction de n coordonnées d'électrons par n fonctions monoélectroniques. Pour tenir en compte du principe de Pauli, la fonction d'onde ψ_{el} sera une fonction monoélectronique dépendant des variables spatiales et des variables de spin. Cette fonction sera appelée spinorbitale. Ainsi, puisque les électrons sont indiscernables, les fonctions d'ondes sont construites avec les spins orbitaux. Les spins orbitaux sont écrits sous la forme d'un déterminant de $n^{ième}$ ordre du déterminant de Slater²² :

$$\psi = \frac{1}{\sqrt{n!}} \begin{vmatrix} \varphi_1(1) & \varphi_2(1) & \varphi_n(1) \\ \vdots & \dots & \vdots \\ \varphi_1(n) & \varphi_2(n) & \varphi_n(n) \end{vmatrix} \quad (1.6)$$

où $1/\sqrt{n!}$ est le facteur de normalisation. Ce facteur sera valable seulement si les spins orbitaux sont également normés. La permutation de deux lignes du déterminant correspond à la permutation des coordonnées spatiale et de spin de deux électrons. Si l'on permute deux lignes ou deux colonnes, le signe du déterminant changera, ce qui répond au principe d'antisymétrie.

1.2 La méthode de Hartree-Fock

C'est en 1927 que D. Hartree²³ a proposé la méthode de champ auto-cohérent pour la résolution des problèmes de chimie quantique. Dans cette méthode,

l'interaction avec un champ moyen créé par le noyau et tous les autres électrons remplace l'interaction instantanée entre les électrons de l'atome. Cette méthode permet de remplacer le potentiel de type $1/r_{ij}$ par une expression où l'interaction électronique est une fonction des coordonnées de chaque électron isolé.

Par la suite, Fock²⁴ a perfectionné la méthode de Hartree en introduisant un terme d'échange. Ce terme permet de tenir en compte de l'antisymétrie de la fonction d'onde. Donc, en utilisant l'Hamiltonien électronique et la norme de la fonction d'onde, il est possible d'évaluer l'énergie en utilisant la méthode variationnelle. Si on se limite à des systèmes à couches fermées, les équations de Hartree-Fock vont être simplifiées car on ne prendra compte seulement des orbitales spatiales. On écrit alors les équations de Hartree-Fock sous la forme suivante :

$$\left\{ \frac{-\nabla_1^2}{2} - \sum_{A=1}^M \frac{Z_A}{r_{1A}} + \sum_{b=1}^N [J_b(X_1) - K_b(X_1)] \right\} \chi_a(X_1) = \sum_{b=1}^N \epsilon_{ba} \chi_b(X_1) \quad (1.7)$$

où ϵ_{ba} sont les multiplicateurs de Lagrange. Ces équations peuvent être résolues par une méthode itérative. Ainsi, l'opérateur de Fock est donné par :

$$F(X_1) = \left\{ \frac{-\nabla_1^2}{2} - \sum_{A=1}^M \frac{Z_A}{r_{1A}} + \sum_{b=1}^N [J_b(X_1) - K_b(X_1)] \right\} \quad (1.8)$$

Le terme $J_b(X_1)$ est l'opérateur Coulombien qui prend la forme :

$$J_b(X_1) = \int \frac{\chi_b(X_2)\chi_b(X_2)}{r_{12}} dX_2 \quad (1.9)$$

L'intégrale coulombienne représente l'énergie moyenne de répulsion électrostatique, des électrons occupant les orbitales χ_{bd} .

et le terme $K_b(X_1)$ est l'opérateur d'échange auquel correspond l'intégrale d'échange.

$$\iint \frac{\chi_a(X_1)\chi_b(X_1)\chi_a(X_2)\chi_b(X_2)}{r_{12}} dX_1 dX_2 \quad (1.10)$$

Les équations de Hartree-Fock se résolvent plus facilement en utilisant un système d'équations algébriques proposées par Roothan²⁵.

$$FC = SC\varepsilon \quad (1.11)$$

F est la matrice de Fock, S est la matrice de recouvrement, C est la matrice des coefficients du développement des orbitales et ε est le vecteur des énergies de l'orbitale. Les éléments de la matrice de Fock et les éléments de la matrice de recouvrement s'écrivent sous la forme :

$$F_{\mu\nu} = \int \phi_\mu(r_1)F(r_1)\phi_\nu(r_1)dr_1 \quad (1.12)$$

$$S_{\mu\nu} = \int \phi_\mu(r_1)\phi_\nu(r_1)dr_1 \quad (1.13)$$

Le principal défaut de la méthode de Hartree-Fock est que la méthode ignore la corrélation qui existe entre le mouvement d'un électron et les mouvements des autres. Cette ignorance provient du fait que, dans la méthode Hartree-Fock, l'électron est placé dans un champ moyen.

1.3 La fonctionnel de la densité²⁶

La théorie de la fonctionnelle de la densité a été introduite en 1964 par Hohenberg et Kohn²⁷. Ces auteurs ont établi que l'énergie d'un système dans son état fondamental non dégénéré est une fonctionnelle de la densité électronique de ce système $\rho(\mathbf{r})$ et ainsi toute autre densité $\rho'(\mathbf{r})$ différentes de la densité réelle va conduire nécessairement vers une énergie supérieure. Cette théorie se base sur deux théorèmes :

- 1. Si le potentiel qui agit sur les électrons est local, et le même pour tous les électrons, alors ce potentiel est une fonctionnelle précise de la densité électronique dans l'état fondamental.*

La preuve de ce théorème peut se faire par l'absurde. Montrons que, pour un potentiel V , on obtient une densité $\rho(\mathbf{r})$.

$$H|\Psi\rangle = (\hat{T} + \hat{U} + \hat{V})|\Psi\rangle = E_{\text{fond.}}|\Psi\rangle \quad (1.14)$$

où T est l'énergie cinétique, U est la répulsion coulombienne et V un potentiel cristallin. Si on connaît le potentiel, V, alors :

$$\langle\Psi|\hat{\rho}(\vec{r})|\Psi\rangle = \rho(\vec{r}) \quad (1.15)$$

Maintenant, supposons qu'il existe un V' tel que

$$(\hat{T} + \hat{U} + \hat{V}')|\Psi'\rangle = E'_{\text{fond.}}|\Psi'\rangle \quad (1.16)$$

$$\langle\Psi'|\hat{\rho}(\vec{r})|\Psi'\rangle = \rho(\vec{r}) \quad (1.17)$$

Par le principe de Ritz, on sait que :

$$E' = \langle\Psi'|\hat{H}'|\Psi'\rangle < \langle\Psi|\hat{H}|\Psi\rangle \quad (1.18)$$

$$E' = \langle\Psi'|\hat{H}'|\Psi'\rangle < \langle\Psi|(\hat{T} + \hat{U} + \hat{V})|\Psi\rangle + \langle\Psi'|\hat{V}' - \hat{V}|\Psi'\rangle \quad (1.19)$$

$$E' < E + \int d^3r \rho(\vec{r}) [V'(\vec{r}) - V(\vec{r})] \quad (1.20)$$

De façon analogue, pour l'énergie E, on obtient :

$$E < E' + \int d^3r \rho(\vec{r}) [V'(\vec{r}) - V(\vec{r})] \quad (1.21)$$

Ces deux équations sont contradictoires. Donc, il ne peut exister qu'un seul potentiel correspondant à une densité électronique dans l'état fondamental.

2. *L'énergie de l'état fondamental est minimale par rapport à toute variation de la densité de particules qui conserve le nombre de particules.*

Supposons une variation $\rho'(\mathbf{r}) = \rho(\mathbf{r}) + \delta\rho(\mathbf{r})$ et que H est invariant :

$$E(\rho') = \langle \Psi(\rho') | \hat{H} | \Psi(\rho') \rangle = \langle \Psi(\rho') | (\hat{T} + \hat{U}) | \Psi(\rho') \rangle + \langle \Psi(\rho') | \hat{V} | \Psi(\rho') \rangle \quad (1.22)$$

on sait que :

$$F_{HK}(\rho') = \langle \Psi(\rho') | (\hat{T} + \hat{U}) | \Psi(\rho') \rangle \quad (1.23)$$

$$\int d^3r \rho'(\mathbf{r}) V(\mathbf{r}) = \langle \Psi(\rho') | \hat{V} | \Psi(\rho') \rangle \quad (1.24)$$

Alors,

$$E(\rho') > \langle \Psi(\rho) | \hat{H} | \Psi(\rho) \rangle = F_{HK} + \int d^3r \rho(\mathbf{r}) V(\mathbf{r}) \quad (1.25)$$

$$\left[\frac{\delta E(\rho)}{\delta \rho} \right]_{\rho'(\mathbf{r})} = 0 \quad (1.26)$$

Donc, l'énergie de l'état fondamental est minimale par rapport à n'importe quelle variation de densité électronique.

1.3.1 Équations de Kohn-Sham

Bien que Hohenberg et Kohn aient jeté les bases de la théorie de la fonctionnelle de la densité, ce sont les travaux de Kohn et Sham²⁸ (KS) qui ont permis de faire des calculs avec la fonctionnelle de la densité. Kohn et Sham ont montré que la fonction $\rho(\mathbf{r})$ exacte correspondait au minimum de l'énergie. Dans l'approche de KS, l'énergie est divisée en trois termes :

$$E[\rho(\tilde{\mathbf{r}})] = U[\rho(\tilde{\mathbf{r}})] + T[\rho(\tilde{\mathbf{r}})] + E_{xc}[\rho(\tilde{\mathbf{r}})] \quad (1.27)$$

Dans cette équation, le terme $U[\rho(\mathbf{r})]$ correspondent à l'attraction coulombienne électron-noyaux. Ce terme peut s'écrire comme suit :

$$U[\rho(\tilde{\mathbf{r}})] = \left(\sum_A \int \frac{-Z_A \rho(\tilde{\mathbf{r}})}{|\tilde{\mathbf{r}} - \tilde{\mathbf{R}}_A|} d\tilde{\mathbf{r}} \right) + \frac{1}{2} \iint \frac{\rho(\tilde{\mathbf{r}}) \rho(\tilde{\mathbf{r}}')}{|\tilde{\mathbf{r}} - \tilde{\mathbf{r}}'|} d\tilde{\mathbf{r}} d\tilde{\mathbf{r}}' \quad (1.28)$$

Le deuxième terme de l'équation de l'énergie $E[\rho(\mathbf{r})]$ correspond à l'énergie cinétique du système mais dans lequel les électrons n'interagissent pas entre eux. Le dernier terme de cette équation rassemble les contributions liées aux effets d'échanges et de corrélation.

Dans le développement de Kohn et Sham, la densité $\rho(\mathbf{r})$ peut-être défini par la somme des carrés des modules des orbitales moléculaires de Kohn-Sham qui sont orthonormales entre elles et occupées par un seul électron.

$$\rho(\vec{r}) = \sum_i^{N^\alpha} |\psi_i^\alpha(\vec{r})|^2 + \sum_i^{N^\beta} |\psi_i^\beta(\vec{r})|^2 \quad (1.29)$$

Dans cette équation, le premier terme correspond à la densité électronique de spin $1/2$ tandis que le deuxième terme correspond à la densité électronique de spin $-1/2$.

L'énergie cinétique $T[\rho(\mathbf{r})]$ se définit comme :

$$T[\rho(\vec{r})] = \sum_{\sigma=\alpha,\beta} \sum_i^{N^\sigma} \int \psi_i^\sigma(\vec{r}) \left(-\frac{\nabla^2}{2} \right) \psi_i^\sigma(\vec{r}) d\vec{r} \quad (1.30)$$

Comme la fonction $\rho(\mathbf{r})$ correspond au minimum de l'énergie :

$$\frac{\delta E[\rho(\vec{r})]}{\delta \rho^\alpha(\vec{r})} = \frac{\delta E[\rho(\vec{r})]}{\delta \rho^\beta(\vec{r})} = 0 \quad (1.31)$$

On peut maintenant établir l'équation Kohn-Sham à un électron comme étant :

$$\left\{ \frac{-\nabla^2}{2} - \left(\sum_A \frac{Z_A}{|\vec{r} - \vec{R}_A|} \right) + \int \frac{\rho(\vec{r}')}{|\vec{r} - \vec{r}'|} d\vec{r}' + \frac{\delta E_{xc}[\rho(\vec{r})]}{\delta \rho^\sigma(\vec{r})} \right\} \psi_i^\sigma(\vec{r}) = \epsilon_i \psi_i^\sigma(\vec{r}) \quad (1.32)$$

où $\sigma=\alpha,\beta$.

L'obtention des valeurs propres et des vecteurs propres ($\psi_i^\sigma(\mathbf{r})$), va nous permettre de générer $\rho^\alpha(\mathbf{r})$ et $\rho^\beta(\mathbf{r})$ et ainsi résoudre l'équation de Kohn Sham d'une

manière auto-cohérente. Seules l'énergie totale et la densité de charge ont une signification physique. Les valeurs propres et les vecteurs propres ne sont que des intermédiaires de calculs. Le principal problème est la détermination du terme d'énergie d'échange et de corrélation pour lequel il n'existe pas d'expression analytique.

1.3.2 La fonctionnel d'échange et de corrélation

Dans les termes de l'énergie $E[\rho(\mathbf{r})]$, il nous reste un terme d'échange et de corrélation $E_{xc}[\rho(\mathbf{r})]$ à déterminer. Dans la majorité des systèmes, il est impossible d'évaluer directement ce terme. Diverses approximations ont donc été proposées dont l'approximation de la densité de spin locale (LSDA). Cette approximation permet d'introduire la densité de spin en divisant la densité totale avec une contribution des électrons de spin α et les électrons de spin β . Le terme $E[\rho(\mathbf{r})]$ peut être approximé par :

$$E_{xc}[\rho(\vec{r})] = \int \rho(\vec{r}) \epsilon_{xc}(\rho^\alpha(\vec{r}), \rho^\beta(\vec{r})) d\vec{r} \quad (1.33)$$

Le terme ϵ_{xc} peut être évalué à partir de calculs de Monté Carlo quantiques sur un gaz d'électrons de densité uniforme. Cette technique permet de trouver les paramètres à partir des différentes valeurs de la densité. Les fonctionnelles les plus précises ont été

développées par Perdew et Zunger (PZ)²⁹ et Vosko, Wilk et Nusair (VWN)³⁰. L'utilisation de ces fonctionnelles apporte d'excellents résultats pour les géométries à l'équilibre pour les molécules avec des liaisons covalentes. Par contre, les énergies de liaison sont surestimées et ces fonctionnelles donnent également de mauvais résultats pour les molécules contenant des ponts hydrogènes³¹. Ainsi, dans les systèmes réels, la densité n'est pas uniforme. Ceci implique que l'approximation n'est valable que localement. Pour corriger cette situation, le gradient de la densité, a été introduit dans les fonctionnelles. Ces fonctionnelles sont dites corrigées du gradient (GGA, Generalized Gradient Approximations). Cette correction permet de mieux représenter les régions de grande fluctuation. Dans cette thèse, on utilisera la fonctionnelle d'échange de Becke (ΔE_x^{B88})³² et la fonctionnelle de corrélation de Perdew (E_c^{P86})³³. On utilisera l'abréviation BP86 pour désigner cette association.

$$E_x^{B88} = E_x^{LDA} - \beta \sum_{\sigma} \int \rho_{\sigma}^{4/3} \left(\frac{x_{\sigma}^2}{(1 + 6\beta x_{\sigma} \sinh^{-1} x_{\sigma})} \right) d^3r \quad (1.34)$$

où β est une constante et x_{σ} est la densité réduite du gradient pour le spin σ .

$$x_{\sigma}(\vec{r}) = \frac{|\nabla \rho_{\sigma}|}{\rho_{\sigma}^{4/3}(\vec{r})} \quad (1.35)$$

Dans cette fonctionnelle, un comportement asymptotique ($-1/r$) a été imposé. Le paramètre β a été lissé par la méthode des moindres carrés en utilisant les énergies d'échange exactes de Hartree-Fock pour les atomes de gaz rares.

Malgré le fait que les résultats obtenus avec les fonctionnelles GGA soient plus précis, l'échange est toujours traité d'une façon locale. Pour éliminer cette évaluation de l'échange local, Becke³⁴ a introduit les méthodes dites hybrides. Dans les méthodes hybrides, l'énergie d'échange et corrélation s'écrit :

$$E_{xc}^{hybride} = aE_{xc}^{LDA} + (1-a)K^{KS} + b\Delta E_x^{GGA} + cE_c^{VWN} + (1-c)E_c^{GGA} \quad (1.36)$$

où K^{KS} représente la partie d'échange Hartree Fock non local et a,b,c sont des paramètres ajustables qui sont déterminés par un lissage des moindres carré sur un ensemble de données expérimentales. La fonctionnelle la plus utilisée est la fonctionnelle B3LYP.

1.4 Les fonctions de bases

La combinaison linéaire d'orbitales atomiques (LCAO) est utilisée pour définir les orbitales moléculaire, ψ_i . Ces fonctions mono-électroniques, ϕ_μ , sont appelées les fonctions de bases.

$$\psi_i = \sum_{\mu=1}^N c_{\mu i} \phi_\mu \quad (1.37)$$

où l'indice i représente les orbitales moléculaires et l'indice μ les fonctions de base, N est le nombre d'électrons et $c_{\mu i}$ sont les coefficients d'expansion des fonctions de base.

Deux types de fonctions de bases sont généralement utilisés. Premièrement, nous avons les bases obtenues à partir des orbitales de Slater³⁵. Ces orbitales sont des puissances de x, y, z multiplié par $\exp(-\xi r)$. Dans ces orbitales, ξ est une constante qui indique la taille de l'orbitale.

$$\phi_n^{l,m}(\xi, r) = N r^{n-1} e^{-\xi r} Y_l^m \quad (1.38)$$

Deuxièmement, nous avons les bases formées à partir de gaussiennes. Les fonctions sont également des puissances de x, y, z mais multiplié par $\exp(-\alpha r^2)$ où le α est une constante qui représente l'extension radiale de la fonction.

$$\phi_{i,j,k}(\alpha, r) = N x^i y^j z^k e^{-\alpha r^2} \quad (1.39)$$

Les bases de Slater décrivent d'une façon raisonnable les orbitales atomiques tandis que les bases gaussiennes représentent moins bien les orbitales atomiques car elles n'ont pas le comportement exact à l'origine ni aux grandes distances. Par contre, les intégrales des bases gaussiennes peuvent être calculées explicitement et sans recours à une intégration numérique. Ainsi, on utilise des combinaisons linéaires

de gaussiennes pour représenter les fonctions de base car le produit de deux gaussiennes peut-être exprimées en tant qu'une simple gaussienne.

Les bases généralement utilisées sont celles développées par Pople et al³⁶. La base minimale que l'on peut utiliser est la base STO-3G. Cette expression nous indique que trois fonctions gaussiennes ont été utilisées pour représenter les orbitales de type Slater. Un ensemble minimal de base est une représentation qui contient juste le nombre de fonctions exigé pour adapter à toutes les orbitales remplies pour chaque atome. Une autre approche est de doubler le nombre de fonctions employées pour décrire les électrons de valence et d'utiliser une fonction simple pour les couches intérieures. Le raisonnement derrière cette approche, est que les orbitales de cœurs, à la différence des orbitales de valence, n'affectent presque pas les propriétés chimiques et peuvent être considérées comme étant les mêmes d'une molécule à une autre. Ces bases, d'un niveau plus élevé, portent le nom de split-valence et sont représentées par 3-21G, 4-31G et 6-31G. Pour la base 6-31G, on aura six gaussiennes pour l'orbitale 1S, trois gaussiennes pour une des fonctions représentent l'orbitale 2S et 1 gaussienne pour l'autre fonction. Il est également possible d'ajouter des fonctions de polarisation. L'ajout de fonctions de polarisation est représenté par l'ajout d'un astérisque (6-31G*). Donc, pour les bases 6-31G*, on ajoute des fonctions d à tous les atomes (sauf pour l'hydrogène) dans la molécule. Pour la base

6-31G**, des fonctions p ont été ajoutées aux hydrogènes et des fonctions d aux autres atomes.

1.5 Les méthodes semi-empiriques

Les méthodes semi-empiriques utilisent des approximations afin d'augmenter la vitesse des calculs. Ainsi, les méthodes semi-empiriques peuvent être utilisées pour exécuter des calculs sur des gros systèmes avec un coût raisonnable en temps et en ressources informatiques. Ces approximations permettent de réduire le nombre d'intégrales à calculer. Les principales approximations des méthodes semi-empiriques sont :

- Seuls les électrons de valence sont considérés. Les électrons des couches internes sont maintenus fixes.
- Les orbitales de Slater sont utilisées pour les fonctions de bases.
- Les intégrales de recouvrement sont négligées.
- Les intégrales calculées sont déterminées à partir d'une paramétrisation (valeurs expérimentales) et toutes les intégrales biélectroniques, qui met en jeu trois ou quatre centres, sont nulles.

Dans cette thèse, les méthodes AM1³⁷ et PM3³⁸ seront utilisées. Ces méthodes découlent de la méthode NDDO (Neglect of Differential Diatomic Overlap). Dans cette approximation, on néglige le recouvrement AB à chaque point de l'espace si A et B sont centrés sur des atomes différents.

1.6 Les champs de force

Les champs de force ont été développés pour être utilisés avec les programmes de mécanique moléculaire. La mécanique moléculaire est un excellent outil pour l'étude de la conformation des molécules. Pour notre travail, nous avons utilisé le champ de force de AMBER.

Le champ de force AMBER¹(Assited Model Building with Energy Refinement) utilise une représentation empirique de l'énergie potentielle. Ce potentiel possède deux interactions principales, soit les interactions intramoléculaires (liaisons covalentes) et les interactions intermoléculaires (liaisons non covalentes). Les interactions intramoléculaires sont celles qui tiennent compte de la déformation des liaisons chimiques, de la déformation angulaire et de la déformation provenant des dièdres.

$$E_{\text{liantes}} = \sum_{\text{liens}} k_r (r - r_o)^2 \quad (1.40)$$

$$E_{\text{liantes}} = \sum_{\text{angles}} k_{\theta} (\theta - \theta_0)^2 \quad (1.41)$$

$$E_{\text{liantes}} = \sum_{\text{dièdres}} \frac{V_n}{2} [1 + \cos(n\phi - \omega)] \quad (1.42)$$

où K_r et K_{θ} sont les constantes de forces, r_0 est la position à l'équilibre et θ_0 est l'angle à l'équilibre. V_n est la constante de force pour le contrôle de l'amplitude, ω la phase, n la périodicité et ϕ est l'angle de torsion. Les interactions non-liantes correspondent aux interactions de van der Waals et électrostatiques. L'énergie non-liante est représentée par :

$$E_{\text{non-liantes}} = \sum_{j=1}^N \sum_{i>j}^N \epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] + \sum_{j=1}^N \sum_{i>j}^N \frac{q_i q_j}{r_{ij}} \quad (1.43)$$

où σ_{ij} est la séparation entre deux atomes pour lequel le potentiel de van der Waals est nul, ϵ_{ij} est la profondeur du puits, q_i représente la charge ponctuelle d'un atome et r_{ij} est la distance séparant deux atomes i et j .

Pour les interactions non-covalentes, celles-ci impliquent les atomes séparés par au moins trois liaisons. Ces interactions sont de deux types : coulombienne représentée par un terme en $1/r$ et de Van der Waals représenté par un terme en $1/r^6$ et $1/r^{12}$. L'interaction coulombienne augmente avec la polarité des liaisons chimiques et peut être importante dans le cas des molécules contenant des hétéroatomes. Le

terme en $1/r^6$ pour l'interaction de Van der Waals constitue la dispersion attractive de London entre deux atomes tandis que le terme en $1/r^{12}$ est la répulsion due à l'exclusion de Pauli. Ces interactions entre les atomes non-liés sont responsables de l'encombrement stérique.

1.7 Dynamique moléculaire

Cette section décrira les différents aspects des simulations de dynamique moléculaire³⁹. La dynamique moléculaire sera utilisée pour l'étude de réactions. Celle-ci nous permettra d'évaluer les énergies d'activation et de réactions mais également de générer un ensemble statistique pour le calcul de la variation de l'énergie libre d'une réaction.

La dynamique moléculaire est une technique déterministe qui explore la trajectoire de phase d'un système à partir du calcul des interactions des particules qui forment le système que l'on désire étudier. La trajectoire est évaluée en fonction du temps en intégrant les équations du mouvement. Les équations du mouvement sont les équations de Newton où la force peut s'exprimer par :

$$F_i = m_i \frac{d^2 \vec{r}_i}{dt^2} \quad (1.44)$$

L'énergie totale du système correspond la somme de l'énergie cinétique et de l'énergie potentielle. L'énergie potentielle est évaluée par un champ de force ou par la mécanique quantique. Pour s'assurer que la trajectoire de la simulation de dynamique moléculaire modélise bien le système que l'on désire étudier, celle-ci va dépendre des deux conditions suivantes :

1. La qualité de la méthode d'intégration des équations de mouvement ;
2. La capacité du potentiel choisi à bien modéliser les interactions inter et intramoléculaire du système. Ce potentiel utilisé sera dans cette thèse soit le champ de force de AMBER, soit la fonctionnelle de la densité.

1.7.1 Intégration des trajectoires

En connaissant les positions et les vitesses des atomes à un instant t , il est possible d'intégrer les équations du mouvement et ainsi trouver de nouvelles positions et vitesses pour le temps $t + dt$.

Les deux principaux algorithmes d'intégrations des équations du mouvement sont l'algorithme de Verlet⁴⁰ et l'algorithme de leap-frog⁴¹. L'algorithme de Verlet s'obtient à partir de la différence de deux séries de Taylor des positions $\mathbf{r}(t + \Delta t)$ et $\mathbf{r}(t - \Delta t)$. On obtient ainsi :

$$r_i(t + \Delta t) = r_i(t) + v_i(t)\Delta t + \frac{F_i(t)}{2m_i}(\Delta t)^2 + \dots \quad (1.45)$$

$$r_i(t - \Delta t) = r_i(t) - v_i(t)\Delta t + \frac{F_i(t)}{2m_i}(\Delta t)^2 - \dots \quad (1.46)$$

En sommant les deux équations, nous obtenons :

$$r_i(t + \Delta t) = 2r_i(t) - r_i(t - \Delta t) + \frac{F_i(t)}{m_i}(\Delta t)^2 \quad (1.47)$$

$$v_i(t) = \frac{r_i(t + \Delta t) - r_i(t - \Delta t)}{2\Delta t} \quad (1.48).$$

Pour l'algorithme de leap-frog, on utilise également la différence de deux séries de Taylor des positions et des vitesses. L'algorithme s'écrit alors :

$$v_i\left(t + \frac{\Delta t}{2}\right) = v_i\left(t - \frac{\Delta t}{2}\right) + F_i(t)\Delta t \quad (1.49)$$

$$v_i(t) = \frac{1}{2}\left(v_i\left(t + \frac{\Delta t}{2}\right) + v_i\left(t - \frac{\Delta t}{2}\right)\right) \quad (1.50)$$

$$r_i(t + \Delta t) = r_i(t) + v_i\left(t + \frac{\Delta t}{2}\right)\Delta t \quad (1.51)$$

Le principal problème de l'algorithme de Verlet est qu'il faut additionner un très petit terme à une grande différence. L'algorithme de leap-frog corrige ce fait, par

contre, il est impossible de calculer simultanément l'énergie potentielle et l'énergie cinétique. Dans l'intégration des équations du mouvement, l'intervalle de temps (δt) doit être choisi d'une façon justifiable. Cet intervalle de temps doit être inférieur à la période de mouvement de plus haute fréquence (généralement les vibrations X-H) présente à l'intérieur du système. Habituellement, le pas d'intégration utilisé est de l'ordre de 10^{-15} s.

1.7.2 Contrôle de la température

La température T d'un système est reliée à l'énergie cinétique par :

$$T \equiv \frac{2}{k_B(3N - N_c)} \sum_{i=1}^N \frac{1}{2} m_i |v_i|^2 \quad (1.52)$$

où v_i est la vitesse de l'atome i , $(3N - N_c)$ est le nombre de degré de liberté, m_i la masse de l'atome i et k_B la constante de Boltzmann. Une méthode simple pour le contrôle de la température est la méthode proposée par Berendsen⁴². Dans cette méthode, les vitesses sont ajustées à chaque pas de dynamique moléculaire par le facteur de multiplicatif suivant :

$$\lambda = \left[1 + \frac{\Delta t}{\tau_T} \left(\frac{T_o}{T} - 1 \right) \right]^{\frac{1}{2}} \quad (1.53)$$

où T_0 est la température de référence et τ_T le temps de relaxation de la température. Si τ_T est grand, le couplage avec le système sera faible. Si τ_T est petit, le couplage avec le système sera fort. Avec un temps de relaxation égale au temps d'intégration, l'algorithme de Berendsen devient équivalent à appliquer un simple facteur multiplicatif. Pour un temps d'intégration de 1 fs, il est suggéré⁴² d'utiliser un temps de relaxation de 0.4 ps. La méthode de contrôle de la température de Berendsen a été implémentée dans DeFT pour les simulations de dynamique moléculaire complètement quantique.

Pour les simulations de dynamique moléculaire utilisant l'interface entre Roar et DeFT, nous avons utilisé la méthode de contrôle thermique de Nosé-Hoover^{43,44}. Dans la méthode de Nosé-Hoover, un degré de liberté additionnel, s , a été introduit pour agir en tant que bain thermique. Pour contrôler la température, le système est placé dans un grand bain thermique qui est à une température constante. Si on utilise un coefficient de friction ξ , les équations de mouvement s'écrivent alors :

$$\vec{\dot{r}}_i = \vec{\dot{p}}_i / m_i \quad (1.54)$$

$$\vec{\dot{p}}_i = \frac{-\partial U(\mathbf{r}^N)}{\partial \vec{r}_i} - \xi \vec{p}_i \quad (1.55)$$

$$\xi = \left(\sum_i \frac{p_i^2}{m_i} \right) - \frac{g}{\beta} / Q \quad (1.56)$$

$$\frac{\dot{s}}{s} = \frac{d \ln s}{dt} = \xi \quad (1.57)$$

où Q est la mesure du couplage entre le système et le bain thermique, $\mathbf{r}$ la position, $\mathbf{p}$ la quantité de mouvement et g est le nombre de degré de liberté.

1.8 Variation d'énergie libre

Le calcul de la variation de l'énergie libre d'un système occupera une grande partie de cette thèse. Dans cette section, nous définissons les notions de bases du calcul de l'énergie libre pour ensuite dans les chapitres 5 et 6, aborder le calcul de l'énergie le long du chemin réactionnel et l'utilisation de l'échantillonnage non-Boltzmann¹² pour l'évaluation de l'énergie libre.

Pour connaître la variation d'énergie libre entre l'état A et l'état B d'un système moléculaire, on définit l'Hamiltonien pour un système de N atomes en fonction des coordonnées $\mathbf{r} \equiv (r_1, r_2, \dots, r_N)$ et de la quantité de mouvement $\mathbf{p} \equiv (p_1, p_2, \dots, p_N)$ par :

$$H(\vec{p}, \vec{r}) = \sum_{i=1}^N \frac{\vec{p}_i^2}{2m_i} + V(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N) \quad (1. 58)$$

où m_i est la masse de l'atome i et $V(\mathbf{r})$ l'énergie potentielle du système.

L'énergie libre est la mesure de la probabilité de trouver le système dans un état donné. En mécanique classique, nous pouvons, exprimer l'énergie libre d'Helmholtz, F , d'un système de N particules dans un volume V à une température T , en terme de la fonction de partition canonique Z , comme étant :

$$F = -k_B T \ln Z \quad (1. 59)$$

où k_B est la constante de Boltzmann, T la température et Z est la fonction de partition du système. La fonction de partition représente la somme de tous les états accessibles au système et se définit comme :

$$Z = N_{\text{norm}}^{-1} \iint e^{-H(\vec{p}, \vec{r}; \lambda)/k_B T} d\vec{p} d\vec{r} \quad (1. 60).$$

La variation de l'énergie entre l'état A et B sera alors :

$$\Delta F = F_B - F_A = -k_B T \ln \frac{Z_B}{Z_A} \quad (1. 61)$$

$$\Delta F = -k_B T \ln \frac{\iint e^{-H_B(\vec{p}, \vec{r})/k_B T} d\vec{p} d\vec{r}}{\iint e^{-H_A(\vec{p}, \vec{r})/k_B T} d\vec{p} d\vec{r}} \quad (1.62)$$

En multipliant le numérateur par $\exp[-H_A(\mathbf{p}, \mathbf{r})/k_B T] \exp[+H_A(\mathbf{p}, \mathbf{r})/k_B T] \equiv 1$ et en introduisant $\Delta H(\mathbf{p}, \mathbf{r}) = H_B(\mathbf{p}, \mathbf{r}) - H_A(\mathbf{p}, \mathbf{r})$, on obtient:

$$\Delta F = -k_B T \ln \frac{\iint e^{-H_A(\vec{p}, \vec{r})/k_B T} e^{-\Delta H(\vec{p}, \vec{r})/k_B T} d\vec{p} d\vec{r}}{\iint e^{-H_A(\vec{p}, \vec{r})/k_B T} d\vec{p} d\vec{r}} \quad (1.63)$$

$$\Delta F = -k_B T \ln \left\langle e^{-\Delta H(\vec{p}, \vec{r})/k_B T} \right\rangle_A \quad (1.64)$$

Dans cette expression, $\langle \dots \rangle_A$ est la moyenne sur l'ensemble statistique de l'état initial du système. Cette méthode n'est cependant efficace que si la différence entre les deux états est petite. Lorsque l'énergie libre est évaluée en faisant varier les coordonnées le long du chemin réactionnel, la variation d'énergie libre est trop grande pour être considérée comme une perturbation. Pour effectuer le calcul de la variation de l'énergie libre, on utilisera la méthode des perturbations thermodynamiques¹⁰ qui introduit un facteur de couplage λ . Ainsi, la variation d'énergie libre devient une fonction continue entre l'état A et l'état B. L'équation 1.64 peut s'écrire maintenant :

$$\Delta F = \sum_{\lambda=\lambda_i}^{\lambda_i+\Delta\lambda} -k_B T \ln \left[\left\langle \exp \left(\frac{[-H(\lambda+\Delta\lambda) - H(\lambda)]}{k_B T} \right) \right\rangle_{\lambda} \right] \quad (1.65)$$

L'équation 1.65 est une approximation pour toute perturbation. Pour des grandes perturbations, cette formulation souffre des problèmes de convergence similaire à ceux de l'équation 1.64. Pour obtenir convergence, un recouvrement significatif des régions de basses énergies des deux ensembles est essentiel. L'effet de la perturbation sur l'espace de phase échantillonné demeure très faible. Pour cette raison, la formule de perturbation thermodynamique est généralement exprimée comme étant une somme de petits intervalles comme nous le montre la Figure 1. Pour chaque fenêtre ΔF , il s'agit d'équilibrer le système à une valeur arbitraire λ_i . Par la suite, l'échantillonnage du système nous permet d'estimer les énergies $H(\lambda_i)$ et $H(\lambda_i \pm \Delta\lambda_i)$.

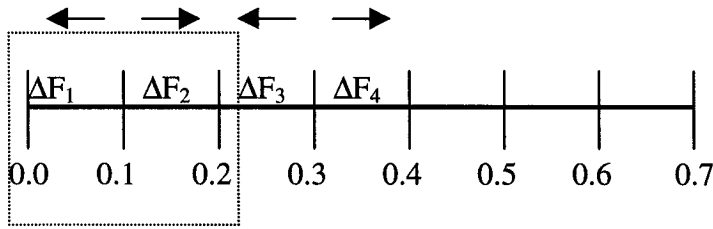


Figure 1. Définition de la direction avant et arrière pour une fenêtre.

Il est également possible d'utiliser la méthode d'intégration thermodynamique¹¹.

Pour cette méthode, la variation d'énergie libre s'écrit :

$$\Delta F = \int_{\lambda_A}^{\lambda_B} \frac{\partial F(\lambda)}{\partial \lambda} d\lambda = \int_{\lambda_A}^{\lambda_B} \left\langle \frac{\partial H(\lambda)}{\partial \lambda} \right\rangle_{\lambda} d\lambda \quad (1.66).$$

Dans cette méthode, l'intégrant est évalué par une série de points discrets, et l'intégrale est approximé en utilisant une méthode d'intégration numérique. Par exemple, la méthode des trapèzes est très souvent utilisée. Lorsque nous appliquons cette méthode à l'équation 1.66, nous obtenons :

$$\Delta F \approx \sum_0^N \left(\left\langle \frac{\partial H(\lambda)}{\partial \lambda} \right\rangle_{\lambda(i)} + \left\langle \frac{\partial H(\lambda)}{\partial \lambda} \right\rangle_{\lambda(i+1)} \right) \times \left(\frac{\lambda(i+1) - \lambda(i)}{2} \right) \quad (1.67)$$

La méthode d'intégration thermodynamique et celle de perturbation thermodynamique sont en principe équivalentes. Les deux dépendent de l'évaluation de la moyenne de l'ensemble statistique. La méthode de perturbation thermodynamique est souvent décrite comme étant exacte tandis que la méthode d'intégration thermodynamique est définie comme étant une approximation puisqu'il est nécessaire d'utiliser une intégration numérique. Si avec ces deux méthodes de calcul de l'énergie libre, on ne peut échantillonner l'espace des configurations du système moléculaire efficacement, il est possible d'utiliser un échantillonnage préférentiel (Umbrella Sampling)¹². Dans cette méthode, on ajoute un potentiel U au potentiel régulier afin de mieux échantillonner une région de l'espace de phase.

Alors, la distribution de probabilité P^* du potentiel modifié est relié à la distribution de probabilité P du potentiel régulier par :

$$P(x) = \frac{P^*(x) \exp\left[\frac{U}{k_B T}\right]}{\left\langle \exp\left[\frac{U}{k_B T}\right] \right\rangle} \quad (1.68)$$

Grâce à cette distribution de probabilité, il est possible d'évaluer l'énergie libre par :

$$\Delta F = -k_B T \ln P(x) \quad (1.69).$$

La dépendance de l'Hamiltonien avec le paramètre de couplage λ , définit le chemin de l'état initial à l'état final. Le choix du chemin n'affecte pas le changement de l'énergie libre mais affecte fortement l'efficacité informatique. Un chemin direct n'est pas nécessairement le plus efficace. Le chemin le plus efficace est un dont la période de relaxation du système en ce qui concerne le changement de l'Hamiltonien et le temps requis pour prélever un ensemble représentatif sont tous les deux réduits au minimum. Le paramètre de couplage, λ , utilisé dans cette thèse, sera défini par les coordonnées du meilleur chemin réactionnel (IRC) ou par un lien entre deux atomes. Le IRC est défini comme le chemin minimum d'énergie reliant les réactifs aux produits par l'intermédiaire de l'état de transition. Dans la pratique, l'IRC est trouvé en identifiant d'abord l'état de transition pour la réaction. Par la suite, le IRC est accompli par l'intégration numérique des équations de Fukui⁴⁵.

$$\frac{d\bar{x}}{ds} = \frac{-\vec{g}}{|\vec{g}|} = \vec{v} \quad (1.70)$$

où x est la coordonnée, s est le pas d'intégration et v (négatif) le gradient normal.

Chapitre 2 : fonctionnalité QM/MM

2.1 Concept de base

Le schéma que nous utilisons dans notre fonctionnalité est le schéma qui est le plus souvent utilisé dans la méthodologie QM/MM pour l'étude de biomolécule où les champs de force AMBER^{7,46,47} ou CHARMM^{5,48,49} sont utilisés. Les principaux problèmes avec ce schéma surgissent parce qu'il est difficile de calculer le terme $E_{QM/MM}$ exactement en présence d'un atome connecté⁵⁰ en particulier si les interactions électrostatiques sont incluses.

Dans la méthodologie QM/MM le système que nous étudions est coupé en deux sous-systèmes. Le premier sous-système contient seulement un petit nombre d'atomes généralement entre 10 et 50 atomes. Les énergies de ces atomes sont évaluées par la mécanique quantique. Le deuxième sous-système représente le reste du système et l'énergie des atomes est évaluée par un champ de force. Les méthodes QM/MM ont démontré leurs avantages dans les études de systèmes complexes. Par contre, le principal problème qui est relié à cette méthodologie est dans le traitement de l'interface entre les parties mécanique quantique et mécanique moléculaire lorsqu'il faut briser un lien covalent.

Dans l'approche QM/MM, l'énergie totale, E, du système est décrit par trois termes,

$$E = E_{QM} + E_{MM} + E_{QM/MM} \quad (2. 1)$$

où E_{QM} correspond à l'énergie de la mécanique quantique (fonctionnelle de la densité pour notre version), E_{MM} est l'énergie pour la partie classique qui décrit par la mécanique moléculaire et $E_{QM/MM}$ est le terme qui tient compte de l'interaction entre la partie classique et quantique. Les termes pour les énergies quantiques et classiques s'écrivent sous la forme suivante :

$$E_{QM}(\vec{R}_n) = \frac{-1}{2} \int d\vec{r} \sum_i \psi_i(\vec{r}) \nabla^2 \psi_i(\vec{r}) + \int d\vec{r} \sum_n \frac{Z_n}{|\vec{R}_n - \vec{r}|} \rho(\vec{r}) + \frac{1}{2} \iint d\vec{r} d\vec{r}' \frac{\rho(\vec{r}) \rho(\vec{r}')}{|\vec{r} - \vec{r}'|} + E_{xc}[\rho(\vec{r})] + \sum_n \sum_{n' > n} \frac{Z_n Z_{n'}}{r_{nn'}} \quad (2. 2)$$

$$E_{MM} = \sum_{\text{liens}} K_r (r - r_o)^2 + \sum_{\text{angles}} K_\theta (\theta - \theta_o)^2 + \sum_{\text{dièdres}} \frac{V_n}{2} [1 + \cos(n\phi - \omega)] + \sum_{i < j} \left[\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^6} + \frac{q_i q_j}{R_{ij}} \right] \quad (2. 3)$$

L'interaction QM/MM entre le sous-système quantique et le sous-système représenté par la mécanique moléculaire est divisée en trois termes : la contribution électrostatique, la contribution de van der Waals et l'interaction due à la liaison entre la partie quantique et classique. Le terme d'énergie QM/MM s'écrit comme suit :

$$E_{QM/MM}(\vec{R}_n, \vec{R}_s) = \int d\vec{r} \sum_s \frac{q_s}{|\vec{R}_s - \vec{r}|} \rho(\vec{r}) + \sum_{s,n} \frac{q_s Z_n}{r_{sn}} + \sum_{s,n} 4\epsilon_{sn} \left[\left(\frac{\sigma_{sn}}{r_{sn}} \right)^{12} - \left(\frac{\sigma_{sn}}{r_{sn}} \right)^6 \right] + E_{MM-lié} \quad (2.4)$$

Le premier terme correspond à l'interaction entre les électrons dans la région QM et les charges ponctuelles de la région MM. Ainsi, les charges ponctuelles de la région classique ont la capacité de polariser le nuage électronique dans la région quantique. Par contre, cette description de la polarisation du sous-système quantique par l'environnement classique ne peut tenir en compte des transferts de charges entre les régions QM et MM. La polarisation de la partie classique par la partie quantique est traitée par les interactions non liées. Par contre, pour avoir un meilleur traitement de ce terme, il faut inclure l'effet de la polarisation dans l'équation 2.1. Ce terme s'écrit comme suit³⁸:

$$E_{MM}^{pol} = -\frac{1}{2} \sum_M \sum_{\alpha} \frac{\mu_M^{ind}}{r_{\alpha M}^3} \vec{r}_{\alpha M} + \frac{1}{2} \sum_M \sum_{\alpha} \frac{Z_M \mu_M^{ind}}{R_{\alpha M}^3} \vec{R}_{M\alpha} \quad (2.5)$$

où μ_M est le moment dipolaire de l'atome classique.

Avec cette équation, le moment dipolaire est directement couplé à la fonction d'onde du système quantique et a besoin d'être résolu d'une façon auto-cohérente. Par contre, ce type de calcul peut être fait que sur de petit système car le coût en calcul serait beaucoup trop grand pour un système plus complexe.

Le deuxième terme est l'interaction entre les noyaux de la région quantique et les charges ponctuelles de la région classique. Le troisième terme correspond à l'énergie de van der Waals entre les atomes du sous-système QM et le sous-système MM. Le paramètre en r^{-6} du potentiel de van der Waals tient en compte de la dispersion attractive tandis que le terme en r^{-12} explique la répulsion d'échange. Ce paramètre est important à cause de l'incapacité de superposer deux nuages d'électrons car les atomes classiques ne possèdent aucun électron explicite. Le dernier terme de l'équation 2.4 correspond aux termes traités par la mécanique moléculaire pour regarder la frontière entre le sous-système quantique et classique.

Ce terme suit les deux règles suivantes³⁹:

1. Tous les liens impliquant un atome quantique et un atome classique sont calculés par le champ de force pour maintenir le lien de l'interface.
2. Les angles de liaison et les angles diédral avec un ou deux atomes classiques sont incluent dans le calcul traité par le champ de force.

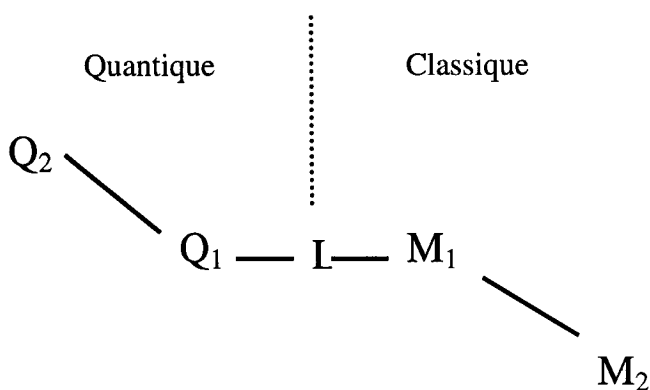


Figure 2. Représentation de l'interface entre la partie quantique et classique.

Pour un système comme celui représenté par la Figure 2 et en utilisant les règles mentionnées ci haut, l'énergie $E_{\text{MM-lié}}$ s'écrit alors comme suit :

$$E_{\text{MM-lié}} = E_{\text{lien}}(Q_1M_1) + E_{\text{angle}}(Q_2Q_1M_1) + E_{\text{angle}}(Q_1M_1M_2) + E_{\text{torsion}}(Q_2Q_1M_1M_2) \quad (2.6)$$

Pour notre implémentation de la fonctionnalité QM/MM, nous avons utilisé l'atome d'hydrogène comme atome connecté⁶. Donc, lorsqu'un lien σ C-C est coupé, un atome d'hydrogène est ajouté pour compléter la valence du sous-système quantique. Ainsi, pour une liaison C-C, l'atome connecté est placé à une distance de 1.1 Å de l'atome C QM et seulement à 0.4 Å de C MM. À cause de cette distance interatomique irréaliste, les interactions électrostatiques entre l'atome connecté et l'atome de C MM sont exclus dans la majorité des implémentations QM/MM^{5,6}. Une étude récente³⁹ a montré que les résultats obtenus par la méthode QM/MM dépendent

fortement de l'endroit où la coupure entre la partie quantique et classique est faite et du nombre d'atome MM qui sont exclus dans le calcul de l'interaction électrostatique entre la partie classique et la partie quantique. Une autre approche a été développée récemment pour améliorer le traitement des interactions électrostatiques⁵¹. Dans cette méthode, toutes les interactions électrostatiques entre les atomes QM et les atomes MM (incluant les atomes connectés) sont calculés. Par contre, les distributions de charges des atomes classiques adjacents aux atomes qui forment la frontière sont représentées par une fonction gaussienne. Les interactions de van der Waals entre l'atome connecté et tous les atomes MM sont également annulées. Les paramètres de van der Waals utilisés pour les atomes quantiques sont les mêmes que pour les atomes classiques dans notre interface QM/MM. Par contre, Gao⁵² a montré que si on optimisait les paramètres de van der Waals spécifiquement pour les atomes QM, on pouvait obtenir de meilleurs résultats.

Il ne faut pas oublier que l'atome connecté ne fait pas partie du système réel. Ainsi, lors de l'évaluation des gradients, nous devons en tenir compte. Dans la section suivante, nous allons montrer comment nous avons approché ce problème.

2.2 Traitements des atomes connectés dans Roar/DeFT

Le traitement des atomes connectés peut être divisé en trois catégories⁵³ :

1. Dans cette méthode, le lien entre l'atome quantique à la frontière et l'atome connecté (Q-L) est fixe comme pour le lien entre l'atome quantique et l'atome classique (Q-M). Dans cette méthode, aucune transformation des gradients de l'atome connecté n'est nécessaire. Cette méthodologie est celle qui est utilisée dans la version originale de IMOMM⁵⁴.
2. Dans ce traitement, le lien Q-L est flexible et le lien Q-M est également flexible. Dans cette option, le changement de la liaison Q-L est proportionnel à la longueur du lien Q-M en utilisant un facteur de correction. Cette méthodologie est utilisée dans la nouvelle version de IMOMO⁵⁵.
3. Avec ce traitement, la liaison Q-L est fixe et la liaison Q-M est flexible. Ce traitement pour les atomes connectés a été utilisé dans notre fonctionnalité QM/MM. Cette méthodologie a été implémentée dans un premier temps dans le programme Solids_EMBED⁵⁶.

L'énergie totale du système est une fonction des coordonnées des atomes réels qui font partie autant du sous-système quantique que du sous-système MM (X) et des coordonnées des atomes connectés (H).

$$E = E(\vec{r}_X, \vec{r}_H) \quad (2. 7)$$

La position de l'atome connecté est une fonction des coordonnées des atomes réels ainsi, l'énergie du système dépend seulement des positions des atomes réels.

$$E = E(\vec{r}_X, \vec{r}_H(\vec{r}_X)) = E(\vec{r}_X) \quad (2. 8)$$

En utilisant la règle de la dérivée en chaîne, on trouve que la dérivée de l'énergie en fonction des coordonnées des atomes réels s'écrit comme suit :

$$\frac{dE}{dr_{X,j}} = \frac{\partial E}{\partial r_{X,j}} + \sum_{i=1}^3 \frac{\partial E}{\partial r_{L,i}} \frac{\partial r_{L,i}}{\partial r_{X,j}} \quad (2. 9)$$

En considérant que l'atome connecté est placé le long du lien entre l'atome quantique (Q) et l'atome classique (M), le vecteur $\mathbf{r}_L$ peut s'écrire comme suit :

$$\vec{r}_L = \vec{r}_Q + g(\vec{r}_M - \vec{r}_Q) \quad (2. 10)$$

$$\text{où } g = \frac{r_{QL}}{r_{QM}} \quad (2. 11).$$

En considérant la définition de $\mathbf{r}_L$, les dérivées de $\mathbf{r}_L$ par rapport aux atomes réels du système s'écrivent :

$$\frac{\partial r_{L,i}}{\partial r_{M,j}} = -gn_i n_j + g\delta_{i,j} \quad (2.12)$$

$$\frac{\partial r_{L,i}}{\partial r_{Q,j}} = \delta_{i,j} + gn_i n_j - g\delta_{i,j} \quad (2.13)$$

où (i,j) représente les coordonnées (x,y,z), $\delta_{i,j}$ le delta de Kronecker et n est le vecteur unitaire parallèle à l'axe Q-M et pointant vers M. En mettant les équations 2.12 et 2.13 dans l'équation 2.9, il est donc possible de corriger l'effet de l'atome connecté sur les gradients des deux atomes qui forment la frontière entre la partie quantique et de la partie classique. Cette méthodologie utilisée dans notre version diffère de celle qui est utilisé dans la version QM/MM du module Roar. Dans la version originale, les coordonnées de l'atome connecté sont ajoutées à la liste d'atomes utilisés dans l'optimisation de la géométrie ou la trajectoire de dynamique moléculaire. Ainsi, les coordonnées des atomes sont libres de se déplacer sans contraintes⁶.

Chapitre 3: Validation de la fonctionnalité QM/MM

Dans cette section, nous présentons les différents tests effectués pour valider notre interface entre le module Roar⁵⁷ et le logiciel DeFT⁵⁸. Cette section nous permettra d'évaluer la méthode utilisée pour traiter les atomes connectés. L'utilisation d'atomes connectés apporte certains problèmes. L'utilisation d'un atome connecté affecte la polarité de la liaison car l'atome qui sert d'atome connecté n'a pas la même électronégativité que celui qu'il remplace. L'atome connecté ne permet pas le transfert de charge entre les atomes qui forment la frontière entre la partie quantique et classique. Le traitement des interactions électrostatiques à la frontière entre la partie quantique et classique joue un rôle majeur sur la qualité des résultats obtenus. Comme décrit dans le chapitre 2, nous utilisons les charges ponctuelles provenant du champ de force de AMBER pour calculer les interactions électrostatiques. Pour les atomes loin de la région QM, l'utilisation des charges ponctuelles est concevable, mais les charges ponctuelles peuvent modifier les interactions électrostatiques entre la région QM et les atomes MM près des atomes QM. Pour éliminer ce problème, nous avons annulé les charges sur les atomes MM entourant immédiatement la région QM. Une autre façon d'approcher ce problème est d'utiliser une distribution de charges gaussiennes sur les atomes classiques qui sont proches de la frontière^{59,60}.

Dans un premier temps, nous avons calculé le potentiel et la barrière de rotation interne de la molécule d'éthane. Dans un deuxième temps, nous avons regardé l'influence de l'atome connecté sur la liaison C-C et nous avons également calculé des affinités de proton pour une série de molécules. Pour terminer, nous avons regardé l'influence de l'endroit où le lien est coupé sur différentes conformations de la dipeptide alanine à laquelle nous ajoutons une hélice de 20 glycines.

3.1 Éthane

Nous commençons avec l'éthane car ce système est le plus simple que nous puissions étudier. Pour ce système, la coupure est le long de la liaison C-C, où chaque groupement méthyle est représenté soit par la mécanique quantique, soit par le champs de force de AMBER. Ce système est représenté à la Figure 3. Cette figure montre que, dans l'éthane, les six atomes d'hydrogène sont indiscernables.

Avec des charges ponctuelles différentes de zéro sur les atomes de la partie classique, les interactions électrostatiques entre la partie classique et la partie quantique va être beaucoup trop grandes. Pour cette raison, nous avons modifié les charges ponctuelles (à zéro) sur les atomes classiques qui forment le groupement directement lié avec le sous-système quantique.

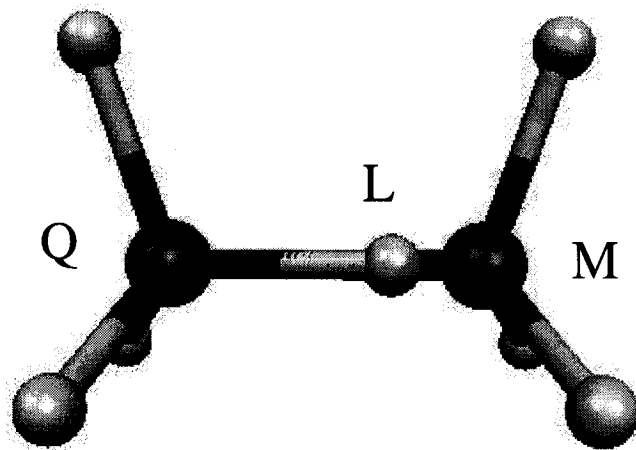


Figure 3. Modèle pour la molécule d'éthane où l'atome L représente l'atome connecté.

Pour le calcul du potentiel d'étirement du lien C-C et le potentiel de l'angle de torsion H-C-C-H, nous avons d'abord optimisé les géométries en utilisant la fonctionnelle BP86 et la base 6-31G*. Les calculs QM/MM ont été par la suite effectués sur les géométries optimisées.

La Figure 4 nous montre le potentiel d'étirement du lien σ C-C dans l'éthane. Cette figure indique que le potentiel hybride représente bien le lien σ C-C autour de la région d'équilibre. Par contre, lorsque le lien est compressé la déviation entre le potentiel hybride et le potentiel complètement quantique est plus grand. La Figure 5 représente la barrière d'énergie de torsion dans l'éthane. La barrière d'énergie de torsion est la différence d'énergie entre la conformation éclipsée (0°) et la conformation décalée (60°). Dans l'éthane, la rotation autour de la liaison σ C-C est facile donc les deux atomes de carbone tournent librement. Ainsi, les conformations décalées et éclipsées représentent deux positions particulières : les conformations éclipsées représentent des maxima tandis que les conformations décalées représentent des minima. Entre les deux formes décalées, il existe une barrière d'énergie qui est d'environ 3 kcal/mol. Cette barrière étant faible, il est possible de trouver cette énergie seulement dans un choc inter-moléculaire. La courbe de conformation nous montre que le potentiel hybride surestime légèrement la barrière de potentiel. Nous obtenons une barrière de 3.4 kcal/mol pour le calcul QM/MM et de 2.5 kcal/mol pour le calcul BP86/6-31G*.

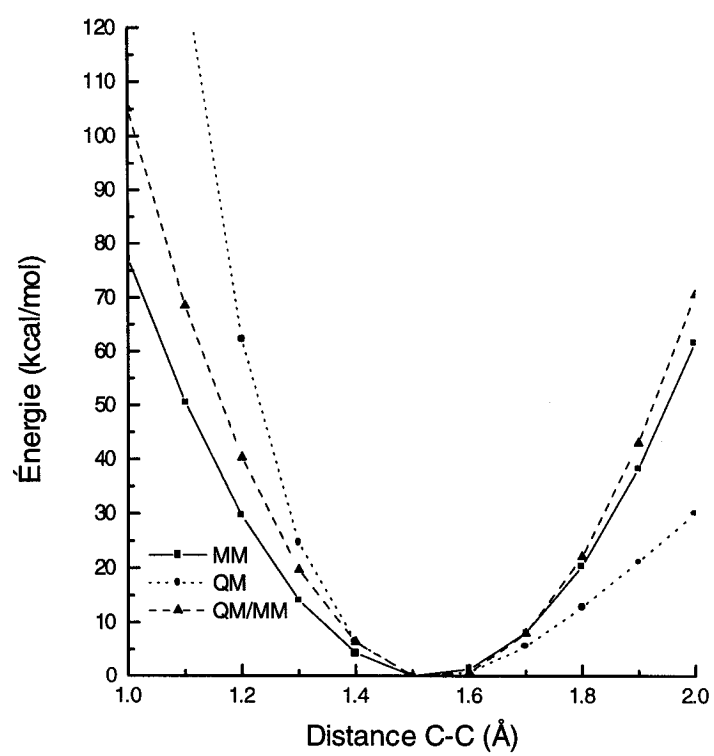


Figure 4. Potentiel d'étirement du lien C-C dans l'éthane.

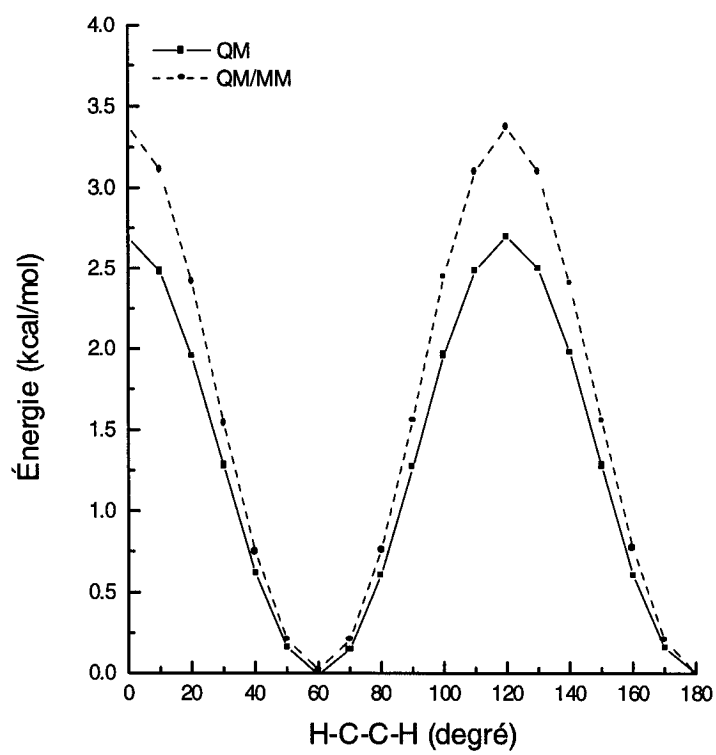


Figure 5. Potentiel de l'angle de torsion H-C-C-H dans l'éthane.

3.2 Influences de l'atome connecté sur les liaisons σ C-C

Pour observer l'effet des atomes connectés sur la liaison σ C-C, nous avons utilisé six alcools aliphatiques ($\text{HOCH}_2\text{-R}$) et trois molécules avec un groupe électronégatif dans la région classique ($\text{HOCH}_2\text{-CH}_2\text{CH}_2\text{OH}$, $\text{H}_2\text{NCH}_2\text{-CH}_2\text{CH}_2\text{OH}$, $\text{H}_2\text{NCH}_2\text{-CH}_2\text{CH}_2\text{OH}$). Cette série de molécules a également été utilisée par Antes et Thiel¹⁶ pour la vérification de leur programme (AM1/MM) où la connexion entre la partie quantique et classique est traitée par la méthode *Adjusted Connection Atoms*. Avec cette méthode, l'électronégativité de l'atome qui sert de connexion est ajustée pour être équivalent à l'atome qu'il remplace. Les auteurs ont obtenu une déviation de seulement 0.0052 Å. Néanmoins, avec l'utilisation d'un atome connecté semblable à nous, ils obtiennent une déviation moyenne de 0.015 Å.

Cette série de molécules a d'abord été optimisé au niveau BP86/6-31G* pour ensuite effectuer un calcul avec le potentiel hybride QM/MM. Les résultats sont présentés au Tableau I. Ce tableau montre que le potentiel hybride représente bien le lien σ C-C avec une erreur moyenne d'environ 0.020 Å par rapport aux résultats complètement quantiques. L'utilisation de l'atome connecté apporte une déviation plus importante que la méthode développée par le groupe de Thiel. Par contre, nos résultats se comparent très bien lorsque les auteurs utilisent une méthodologie similaire à celle utilisée dans cette thèse.

Tableau I : Résultats pour les distances C-C pour une série de molécules.

MOLECULES ^A		QM	QM/MM	DIFFÉRENCE
QM	MM	$C_{QM}-C_{QM}$	$C_{QM}-C_{MM}$	(Å)
HOCH ₂	CH ₃	1.5113	1.5284	0.0171
HOCH ₂	CH ₂ CH ₃	1.5076	1.5288	0.0212
HOCH ₂	CH ₂ CH ₂ CH ₃	1.5073	1.5303	0.0230
HOCH ₂	CH(CH ₃) ₂	1.5127	1.5300	0.0173
HOCH ₂	CH(CH ₂ CH ₃) ₂	1.5149	1.5434	0.0285
HOCH ₂	C(CH ₂ CH ₃) ₃	1.5182	1.5472	0.0290
HOCH ₂	CH ₂ CH ₂ OH	1.5068	1.5486	0.0400
H ₂ NCH ₂	CH ₂ CH ₂ OH	1.5123	1.5291	0.0168
H ₂ NCH ₂	CH ₂ CH ₂ NH ₂	1.5122	1.5307	0.0185

A) calculs quantiques avec BP86/6-31G*

3.3 Affinité de proton

Dans cette section, nous avons calculé des affinités de proton pour une série de bases organiques dans un premier temps et avec une charge ponctuelle externe dans un deuxième temps. Le calcul des affinités de proton des bases organiques fournit un excellent test pour l'exactitude de la méthode QM/MM en comparaison du modèle quantique correspondant. Nous avons déterminé les affinités de proton pour les amines primaires et leurs anions en utilisant la même série de molécules que celles utilisées par Amara et al.⁶¹.

L'affinité de proton est la différence d'énergie entre la géométrie optimisée de la forme protonée et de la forme non protonée. Pour chaque molécule, celles-ci sont coupées à différents endroits pour regarder l'effet du nombre d'atomes qui font partie du sous-système quantique. Dans la phase gazeuse, l'affinité de proton est calculée avec la formule suivante :

$$\text{Affinité de proton} = E(B) - E(HB^+) \quad (3.1)$$

où B est la base et HB^+ est l'acide conjugué. Les résultats sont présentés au Tableau II. Dans ce tableau, les groupements à droite de l'astérisque font partie du sous-système quantique et les groupements à gauche du sous-système classique. Ces résultats nous montrent qu'il est important d'inclure plus qu'un groupe CH_2 après le

groupement fonctionnel pour obtenir de bon résultat⁴⁸. Ces résultats sont en accord avec d'autres résultats qui démontraient qu'il est nécessaire de ne pas couper la molécule proche du groupement fonctionnel.

Pour compléter les calculs d'affinités de proton, nous avons regardé l'effet d'une charge externe sur l'affinité de proton en utilisant le même système que Reuter et al.⁶². Dans ce système, la charge externe est le sodium et celui-ci est traité par le champ de force. Cette charge réagit avec la molécule de propanol. Les différentes configurations étudiées sont présentées à la Figure 6. La géométrie du propanol et de sa forme acide conjuguée ($C_3H_7OH_2^+$) ont été d'abord optimisée par une méthode quantique pour ensuite ajouter l'atome de sodium pour calculer les énergies quantiques et QM/MM. Pour les calculs QM/MM, l'atome de sodium fait partie de la section MM.

Les résultats du calcul d'affinité de proton avec et sans la présence du sodium sont présentés au Tableau III. Ce tableau nous indique que la présence de l'ion Na^+ réduit l'affinité de proton du propanol. Ces résultats comme ceux présentés précédemment nous montrent qu'il est important d'avoir plus qu'un groupement CH_2 après le groupement fonctionnel dans le sous-système quantique. Avec deux groupements CH_2 , nous avons une erreur moyenne par rapport au calcul complètement quantique d'environ 7 kcal/mol. Cette valeur se compare à celle obtenue par Reuter et al.⁵¹ lorsqu'il utilise la méthode des atomes connectés.

Tableau II. Calculs d'affinités de proton sur les bases organiques

MOLÉCULES	QM ^A (KCAL/MOL)	QM/MM ^A (KCAL/MOL)
CH ₃ *-NH ⁻	425.3	441.7
CH ₃ -CH ₂ *-NH ⁻	421.7	473.5
CH ₃ *-CH ₂ -NH ⁻	421.7	428.8
CH ₃ -CH ₂ *-CH ₂ -NH ⁻	420.5	447.6
CH ₃ *-CH ₂ -CH ₂ -NH ⁻	420.5	423.4
CH ₃ *-NH ₂	226.1	215.3
CH ₃ -CH ₂ *-NH ₂	229.7	207.5
CH ₃ *-CH ₂ -NH ₂	229.7	227.0
CH ₃ -CH ₂ *-CH ₂ -NH ₂	231.0	225.4
CH ₃ *-CH ₂ -CH ₂ -NH ₂	231.0	230.4
CH ₃ -CH ₂ -CH ₂ *-CH ₂ -NH ₂	231.7	216.5
CH ₃ -CH ₂ *-CH ₂ -CH ₂ -NH ₂	231.7	231.1

A) calculs quantiques avec VWN/6-31G*

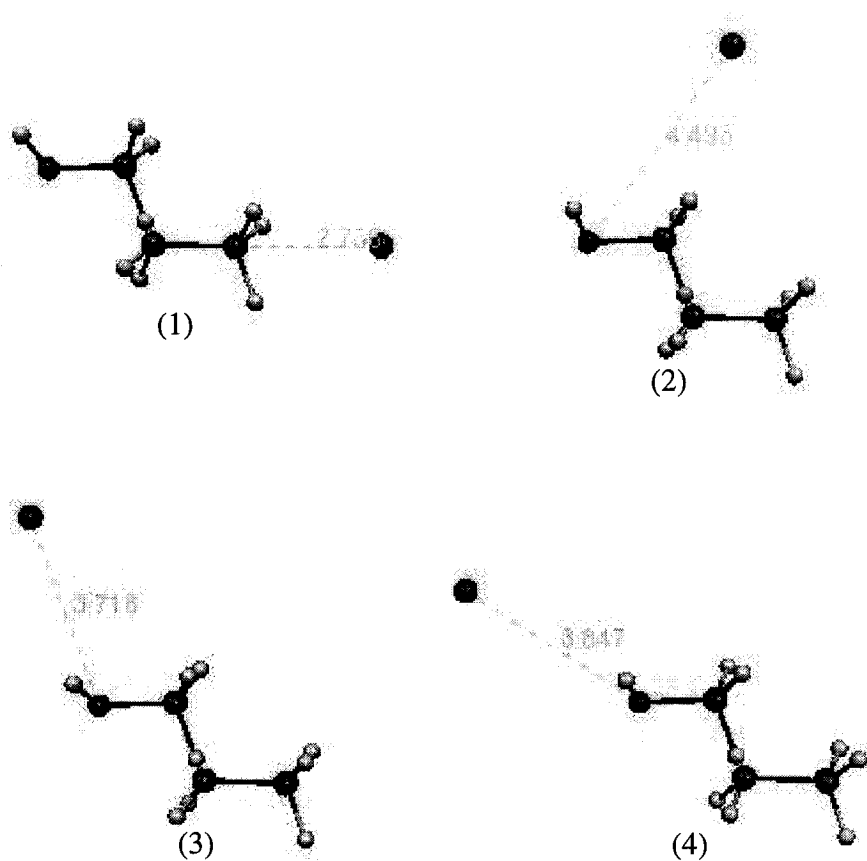


Figure 6. Les différentes structures pour le propanol + Na^+

Tableau III. Affinité de proton avec la présence d'une charge ponctuelle externe

MOLÉCULES ^A	QM	QM/MM ^B	QM/MM ^C
Propanol	198.0	174.4	195.9
Propanol + Na ⁺ (1)	134.9	99.8	122.5
Propanol + Na ⁺ (2)	121.1	85.2	112.5
Propanol + Na ⁺ (3)	103.1	72.3	100.3
Propanol + Na ⁺ (4)	109.7	77.7	105.2

A) les numéros réfèrent aux structures présentées à la Figure 6.

B) comprend un groupement CH₂ après le OH

C) comprend deux groupements CH₂ après le OH

3.3 Analyse conformationnelle de la dipeptide alanine

Pour cette analyse conformationnelle, nous avons utilisé les conformations $C7_{ax}$, $C7_{eq}$, $C5$, β_2 et α' (Figure 7) de la dipeptide alanine. Pour cette étude, la dipeptide alanine est composée d'un résidu alanine avec un groupement N-méthylamine sur l'extrémité C-terminale et une hélice alpha contenant 20 glycines sur l'extrémité N-terminale. Cette structure nous a permis de vérifier deux différents endroits où il est possible de couper la structure sans toucher à la liaison peptidique -CO-HN. La première division est entre $-CH_2-CO$ et la deuxième division est entre $-NH-CH_2$. Dans cette division, le 2^{ème} groupement fait partie du sous-système quantique et le premier groupement du sous-système classique. Les résultats sont présentés aux Tableau IV et Tableau V. Le Tableau VI présente également les résultats pour les cinq conformations de la dipeptide alanine sans l'hélice alpha sur le N-terminal mais le groupement $-COH$. Nous avons effectué ces calculs seulement pour confirmer que notre fonctionnalité QM/MM peut bien reproduire un calcul complètement quantique. Les calculs complètement quantiques ont été exécutés avec DeFT en utilisant la fonctionnalité *divide-and-conquer*⁵⁸.

Dans un premier temps, nous avons optimisé les cinq conformations dans la phase gazeuse en effectuant un calcul de niveau BP86/6-31G*. Par la suite, nous avons ajouté une hélice alpha préalablement optimisée avec la méthode semi-

empirique AM1. Un seul processus SCF a été effectué pour évaluer les énergies relatives des différentes conformations. Le résultat des énergies relatives pour les cinq conformères, sans hélice alpha, est présenté au Tableau VI. Ce tableau montre que l'ordre de la stabilité est : $C7_{eq} < C5 < C7_{ax} < \beta_2 < \alpha$. Ce résultat est identique à celui obtenu pour des calculs⁶³ de niveau B3LYP/6-31G* et MP2/6-31G*. Les ponts hydrogène internes jouent un rôle dans la stabilisation des différents conformères. Trois types de ponts hydrogène internes (N-H...O, N-H...N, C-H...O) sont présents. L'ordre de la stabilité de ces trois interactions est : N-H...O (-8.6 kcal/mol) > N-H...N (-1.6 kcal/mol) > C-H...O (-0.5 kcal/mol)^{64,65}. Pour les conformations sans hélice alpha, la stabilisation des conformères $C7_{eq}$, $C7_{ax}$ et $C5$ est due à la présence du pont hydrogène N-H...O. Les résultats QM/MM sont présentés aux Tableau IV et Tableau V. Dans ces tableaux, la première colonne correspond aux calculs effectués en utilisant un potentiel classique seulement. Le potentiel utilisé est le champ de force de AMBER. La deuxième colonne présente les résultats obtenus pour les calculs de niveau BP86/6-31G*. Toutes les autres colonnes contiennent les résultats pour les calculs QM/MM. Le ratio inscrit en haut des colonnes indique que le calcul QM/MM a été fait avec 15 atomes quantiques sur les 156 atomes du système si ratio est 15/156. Ces deux tableaux montrent que l'ordre de stabilité avec l'hélice alpha devient : $\beta_2 < C7_{ax} < C7_{eq} < C5 < \alpha$. Ces tableaux montrent également qu'il est plus avantageux de couper une liaison C-C qu'une liaison C-N. Nous remarquons que lorsque la coupure est entre -CH₂-CO, nous avons besoins de seulement 22 atomes

dans la section quantique comparativement à 67 atomes pour la coupure entre -NH-CH_2 . Nous croyons que la division entre C-N est moins efficace car celui-ci prend un peu le caractère de la liaison peptidique.

Tableau IV. Résultats : coupure est entre CH₂-CO.

	MM	QM	QM/MM							
			15/156	22/156	29/156	36/156	43/156	50/156	57/156	64/156
B ₂ -gly ₂₀	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
C _{7ax} -gly ₂₀	1.77	2.30	5.11	2.18	2.20	2.09	2.04	2.11	2.13	2.13
C _{7eq} -gly ₂₀	8.03	8.14	16.00	8.05	7.92	7.87	7.92	8.04	8.04	8.06
C ₅ -gly ₂₀	8.37	11.70	18.73	11.90	11.53	11.39	11.39	11.45	11.45	11.45
α-gly ₂₀	20.85	22.27	36.13	22.29	22.11	22.10	22.25	22.23	22.23	22.25

Tableau V. Résultats : coupure est entre NH-CH₂.

	MM	QM	QM/MM							
			18/156	25/156	32/156	39/156	46/156	53/156	60/156	67/156
B ₂ -gly ₂₀	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
C _{7ax} -gly ₂₀	1.77	2.30	2.34	2.10	2.13	2.15	2.00	2.11	2.10	2.15
C _{7eq} -gly ₂₀	8.03	8.14	7.75	7.70	7.74	7.86	7.77	7.92	7.97	8.04
C ₅ -gly ₂₀	8.37	11.70	9.74	10.46	10.65	10.85	10.71	10.93	11.06	11.16
α-gly ₂₀	20.85	22.27	22.23	21.77	21.90	22.02	21.95	22.04	22.11	22.16

Tableau VI. Résultats pour les cinq conformations de la dipeptide alanine.

CONFORMATION	ÉNERGIE (KCAL/MOL)
$C7_{eq}$	0.00
$C5$	1.71
$C7_{ax}$	2.37
β_2	3.14
α'	6.79

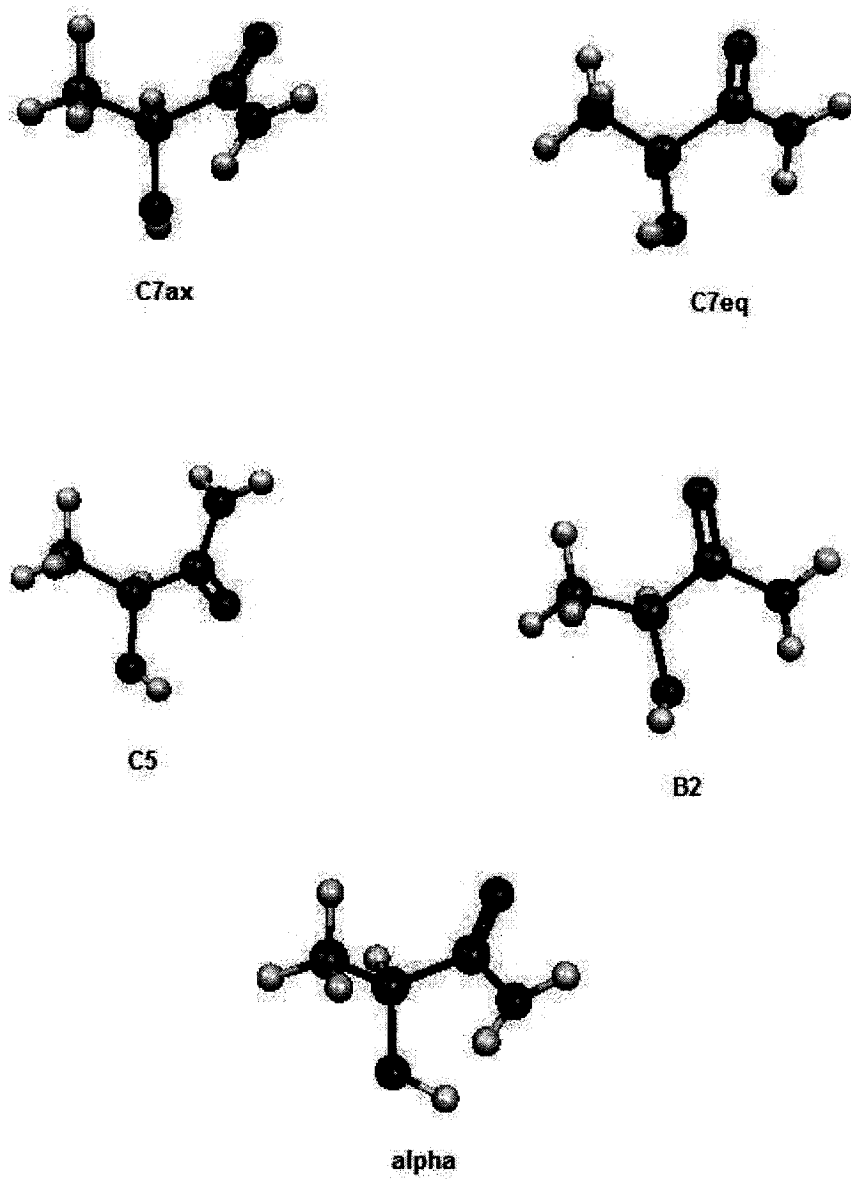


Figure 7 : Conformations de la dipeptide alanine.

3.4 Discussion

Différents tests de validation de la nouvelle fonctionnalité QM/MM dans DeFT nous a permis de mieux connaître le niveau d'exactitude de notre implémentation. L'utilisation d'un atome connecté utilisé dans notre fonctionnalité QM/MM est la méthode la plus facile à implémenter dans un programme de mécanique quantique comme DeFT mais n'est pas le plus précis car C-H ne peut remplacer exactement le lien covalent C-C.

L'utilisation d'un atome connecté sur une série de molécule organique nous donne une déviation d'environ 0.02 Å qui est une déviation comparable à ceux obtenue dans d'autres implémentations utilisant un atome connecté. Le calcul d'affinité protonique nous permet de mieux entrevoir la principale faiblesse de l'utilisation des atomes connectés. Par contre, comme il a été démontré dans la littérature, le Tableau II et Tableau III nous montre que lorsque nous utilisons un atome connecté qu'il est très important d'avoir aux moins deux groupement CH₂ après le groupement fonctionnelle dans le sous-système quantique pour pouvoir obtenir des résultats acceptables.

Dans la dernière section de ce chapitre, nous avons démontré qu'il est plus avantageux d'effectuer la coupure entre un lien C-C qu'un lien C-N. Donc, si l'on désire effectuer un calcul plus complexe comme l'étude d'une réaction enzymatique,

il sera très important de faire la coupure entre un lien covalent C-C et d'inclure suffisamment d'atome dans la sous-section quantique pour que les résultats soient valables.

Chapitre 4 : Énergie libre le long du chemin réactionnel (IRC)

4.1 Introduction

Le calcul de l'énergie libre est une quantité très importante dans la compréhension des réactions en chimie. Celle-ci contient toute l'information pour la prévision de la stabilité des différentes conformations et des processus réactionnels les plus susceptibles. L'énergie libre est directement liée à la probabilité d'observer une configuration dans un ensemble conformationnel. Il est donc possible de calculer cette énergie libre à partir de la distribution de probabilité que l'on obtient avec la trajectoire de dynamique moléculaire.

L'énergie libre de Helmholtz est définie comme étant :

$$\Delta F = \Delta E - T\Delta S \quad (4.1)$$

où ΔE est la variation de l'énergie interne, T est la température et ΔS est la variation de l'entropie. La qualité du calcul de l'énergie libre dépendra alors du choix des coordonnées réactionnelles que l'on va suivre au cours de la trajectoire de la dynamique moléculaire. Une bonne trajectoire est la trajectoire qui sera la plus basse en énergie. Ainsi, un mauvais choix des coordonnées réactionnelles entraîne

généralement que la trajectoire de la simulation de dynamique moléculaire ne passera pas par la région de l'état de transition et ainsi amènera une surestimation de la barrière réactionnelle et un hystérésis⁶⁶ important dans le profil de l'énergie libre. L'hystérésis est défini comme étant la différence d'énergie libre entre la trajectoire vers l'avant et la trajectoire vers l'arrière. Ce phénomène est présenté à la Figure 8.

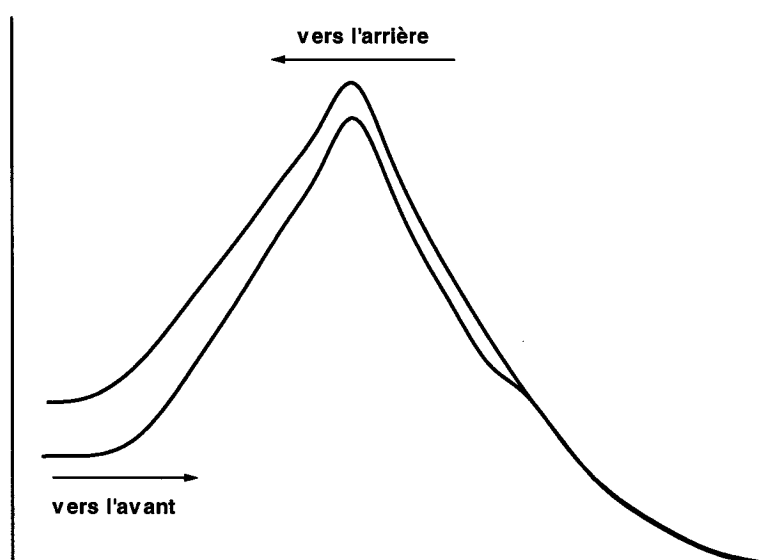


Figure 8. Figure montrant l'hystérésis entre la trajectoire vers l'avant et la trajectoire vers l'arrière.

Pour s'assurer de réduire au maximum l'hystérésis le long du profil d'énergie libre, nous utiliserons le chemin réactionnel produit par le calcul du IRC et ainsi obtenir une trajectoire micro-réversible. Une trajectoire est micro-réversible lorsque celle-ci est le parcours d'énergie minimum dans la vallée formée par la surface menant à l'état de transition puis redescendant vers le produit. Ce chemin est indépendant du sens de la réaction, alors il doit être le même pour la réaction inverse. Ainsi, en utilisant le IRC, on s'assure que la forme de la surface d'énergie potentielle dans la direction perpendiculaire ne varie presque pas pour deux points consécutifs le long du chemin réactionnel. Alors, le système oscille autour d'un minimum lorsque l'on échantillonne le système.

Il est possible de contraindre la trajectoire obtenue par la simulation de dynamique moléculaire à suivre ce chemin réactionnel. On effectue le calcul d'énergie libre en utilisant la méthode de croissance lente (slow-growth). Pour effectuer le calcul des changements d'énergie libre par la méthode de croissance lente, il est nécessaire de varier le paramètre, λ , du système à chaque étape de la trajectoire de la simulation. L'énergie libre sera obtenue en accumulant le travail effectué le long de la trajectoire. Cette méthode est simple à mettre en place et à utiliser. Ainsi, avec la méthode de croissance lente, les changements se font assez lentement pour que l'on suppose que le système est à l'équilibre à chaque étape. Il est également possible de calculer l'énergie libre en utilisant des fenêtres avec une

variation du paramètre, λ , plus grand. À chaque fenêtre, le système sera équilibré avant d'avoir l'ensemble de Boltzmann nécessaire pour évaluer l'énergie libre pour cette fenêtre.

4.2 Algorithme pour le calcul de l'énergie libre

L'algorithme utilisé pour contraindre la trajectoire à suivre le chemin réactionnel a été développé dans un premier temps par Michalak et Ziegler⁶⁷. Nous avons implanté cet algorithme dans le logiciel DeFT. Pour implanter cet algorithme, nous avons introduit dans DeFT les différents algorithmes nécessaires afin d'exécuter une trajectoire de dynamique moléculaire. Nous avons utilisé, pour intégrer les équations de mouvement, la méthode leap-frog³⁰ et pour le contrôle de la température, la méthode de Berendsen³¹. Ces méthodes sont décrites à la section 1.7.

L'algorithme pour contraindre la trajectoire est le suivant :

1. On obtient les coordonnées du chemin réactionnel en calculant le IRC;

$$X^{j,IRC} = \{X_i^j; i = 1, \dots, 3N\}, j = 1, \dots, k \quad (4.2)$$

Le IRC est défini selon Fukui³⁴ comme la trajectoire minimum d'énergie qui traverse l'état de transition. Sa direction initiale, à l'état de transition, est donnée par la

fréquence imaginaire du mode normal. Si on considère un système composé de N atomes, le vecteur de la géométrie est défini par :

$$\vec{x} = \left[\left(\frac{m_A}{\mu} \right)^{1/2} \vec{x}_A, \left(\frac{m_B}{\mu} \right)^{1/2} \vec{x}_B, \dots \right] \quad (4.3)$$

où μ est la masse réduite. Alors, le IRC sera solutionné par l'équation :

$$\frac{d\vec{x}}{ds} = \vec{v} \quad (4.4)$$

où s est la coordonnée de réaction et v est le vecteur unitaire, défini comme la valeur négative du gradient normalisé. Ce vecteur est défini par :

$$\vec{v} = \frac{-\vec{\nabla}V}{|\vec{\nabla}V|} \quad (4.5)$$

où ∇ est l'opérateur du gradient et V est l'énergie potentielle de Born-Oppenheimer. Avec l'équation 4.3, on pourra construire notre trajectoire de référence en deux temps. Premièrement, on trouvera le meilleur chemin pour passer de l'état de transition au réactif et par la suite, on produira le chemin entre l'état de transition et le produit. Pour ces deux calculs, la coordonnée de réaction Δs sera la même. Le calcul du IRC nous assure que l'état de transition relie bien le réactif et le produit que nous désirons obtenir.

2. Contraindre la trajectoire de dynamique moléculaire à suivre ce chemin réactionnel en utilisant la contrainte suivante ;

$$(X - X^{j,IRC}) \bullet G^{j,IRC} = 0 \quad (4.6)$$

où la variable G est définie comme étant :

$$G^{j,IRC} = \{N^j m_i \Delta X_i^{j,IRC}\} = \{N^j m_i (X_i^{j+1,IRC} - X_i^{j-1,IRC}), i = 1, \dots, 3N\} \quad (4.7)$$

et N^j est un facteur de normalisation.

Ainsi, la projection de la position, X, sur la direction du gradient, G, doit être constante et égale à la valeur correspondante de la géométrie provenant du chemin réactionnel. La direction du gradient sera donnée par la différence normalisée de deux géométries dans le IRC. La trajectoire de la dynamique moléculaire le long des coordonnées de la réaction est contrainte en utilisant l'équation 4.6 avec l'algorithme Shake. L'algorithme Shake est une procédure itérative qui va ajuster les positions des atomes en satisfaisant la contrainte. Cet ajustement va être calculé après avoir obtenu la nouvelle position, x_i , par l'intégration des équations du mouvement. Ce petit ajustement, δx_i , est représenté par les équations suivantes :

$$\bar{x}_{i+1} = \bar{x}_i + \delta \bar{x}_i \quad (4.8)$$

où

$$\delta x_i = -\frac{\gamma_i \Delta t^2 G_{i,IRC}}{m_i} \quad (4.9)$$

et

$$\gamma_i = \frac{\sum (x_i - x_{i,IRC}) \bullet G_{i,IRC}}{\Delta t^2 \sum \left(\frac{G_{i,IRC}^2}{m_i} \right)} \quad (4.10)$$

Ce δx_i est le déplacement qui nous assurera que la trajectoire lors de la dynamique moléculaire respectera bien la contrainte définie à l'équation 4.6. Ce déplacement est principalement influencé par les forces du système.

On pourra alors évaluer l'énergie libre d'Helmholtz en sommant la force le long de la trajectoire.

$$\Delta F = \sum_j^{nstep} F^j \Delta s^j \quad (4.11)$$

Où Δs^j correspond au changement dans la valeur de la contrainte pour l'itération j et n_{step} représente le nombre de fenêtres pour exécuter la simulation.

$$\Delta s^j = \frac{1}{2} \left[(X^{j+1,IRC} \bullet G^{j,IRC}) - (X^{j-1,IRC} \bullet G^{j,IRC}) \right] \quad (4.12)$$

4.3 Procédure pour les simulations

Pour chaque exemple montré dans les sections suivantes, nous avons dans un premier temps trouvé l'état de transition et le IRC en utilisant la fonctionnelle BP86 et la base 6-31G* avec le logiciel Gaussian 98. La trajectoire de dynamique moléculaire a été exécutée avec DeFT en utilisant 4 processeurs sur le système HPCVL. Nous avons utilisé un temps d'intégration de 0.5 fs et une température de 300 K. Pour contraindre la trajectoire à suivre le IRC, nous avons utilisé l'algorithme Shake⁶⁸. Le calcul de l'énergie libre a été exécuté avec la méthode de croissance lente. L'exécution de la trajectoire de dynamique moléculaire dans DeFT a également été faite en utilisant la fonctionnelle BP86 et la base 6-31G*. Avant chaque calcul d'énergie libre, le système a été équilibré en utilisant entre 5000 et 10000 itérations. Pour s'assurer que le système est bien à l'équilibre chaque pas le long du chemin réactionnel est équidistant de $0.0006 \text{ amu}^{-1/2} \text{ bohr}$. Tous les points qui composent la trajectoire de référence sont interpolés à partir du IRC calculé avec Gaussian 98. Nous prenons de très petit pas pour s'assurer que le système est presque à l'équilibre car nous utilisons une seule itération à chaque fenêtre avec la méthode de croissance lente. Lors de la trajectoire de dynamique moléculaire, nous contrôlons la température et la variation moyenne de la trajectoire par rapport à la trajectoire de référence. Un exemple des suivis de ces deux paramètres pour la réaction d'ouverture du cyclobutène est présenté à la Figure 9. Cette figure montre que nous avons une

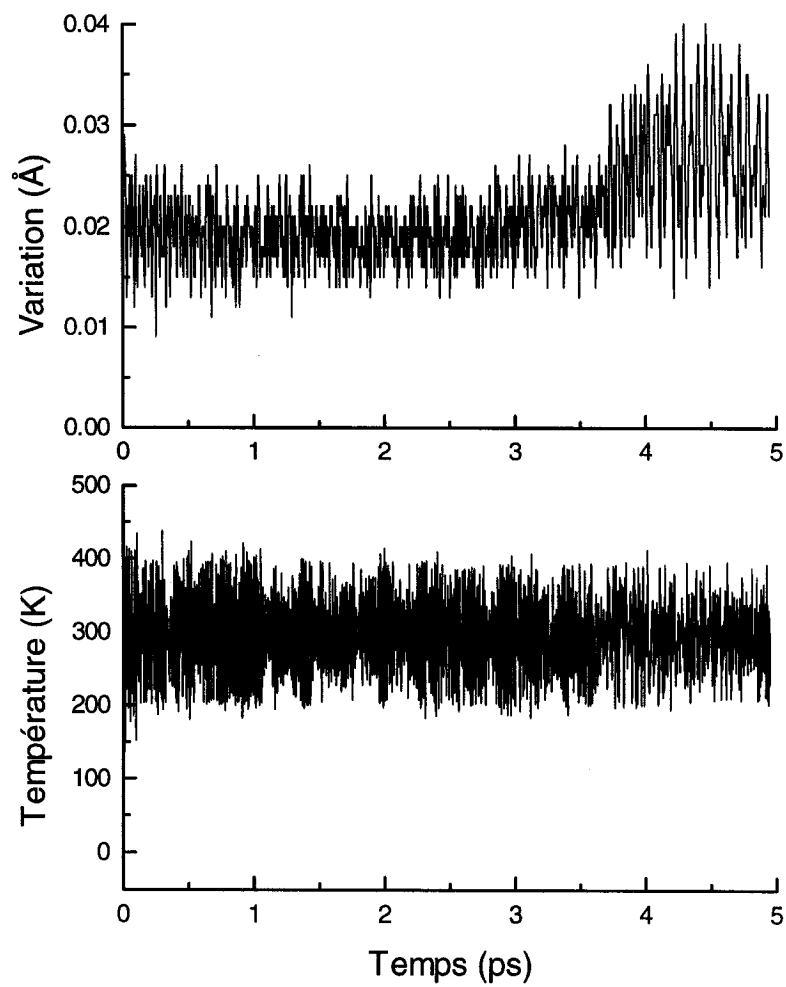


Figure 9. Paramètre de contrôle le long de la trajectoire de dynamique moléculaire

variation moyenne de 0.02 Å par rapport à la trajectoire de référence. Ce graphique indique également qu'il est plus difficile de suivre la trajectoire de référence à la fin de la trajectoire. Donc, la structure du butadiène oscille dans la simulation autour de la géométrie d'équilibre.

4.4 Résultats

Dans cette section, nous validerons cette méthode pour le calcul de l'énergie libre avec cinq réactions élémentaires. Une réaction élémentaire s'effectue en un seul acte en passant par un état de transition. Cet état de transition n'est pas un état réel et n'a pas une durée de vie finie. C'est seulement une structure de transition entre deux états stables.

Lorsqu'une réaction est exoénergétique, comme la majorité de nos exemples, le niveau d'énergie de l'état final est inférieur à celui de l'état initial. L'énergie nécessaire pour passer de l'état initial à l'état final est l'énergie d'activation. Cette énergie peut correspondre à l'énergie pour la rupture de certaines liaisons. Les trois premiers exemples sont des réactions péricycliques. Les réactions péricycliques possèdent des états transitoires cycliques. Généralement, celles-ci sont subdivisées en trois types (réactions de cycloaddition, réactions électrocycliques et réactions sigmatropiques). Nos trois exemples sont des réactions électrocycliques. Les réactions électrocycliques sont caractérisées par une transformation au cours de laquelle il y a ouverture d'un cycle (ou l'inverse) d'une molécule. Donc, la formation

et la rupture des liens se produisent par une réorganisation des électrons dans une boucle fermée des orbitales agissantes l'une sur l'autre.

Les exemples sont l'ouverture du cyclobutène pour former le 1,3-butadiène, l'ouverture du cycle de l'oxygène pour former l'acroléine et finalement la cyclisation de l'hexatriène en cyclohexadiène. Par la suite, nous montrons la migration 1,3 oxygène-oxygène pour le méthyle formate et le trifluorométhyle formate et la dernière réaction est l'isomérisation du cyanure d'hydrogène en isocyanure d'hydrogène.

4.4.1 Ouverture du cycle pour le cyclobutène en 1,3-butadiène

Dans l'étude des réactions électrocycliques, l'ouverture du cycle du cyclobutène est un exemple classique d'un mécanisme péricyclique dont les résultats sont prévus par les règles de Woodward-Hoffmann⁶⁹. Les règles de Woodward-Hoffmann nous disent que dans n'importe quel processus concerté, les orbitales occupées du réactif initial et les orbitales de la molécule issue de la réaction doivent avoir la même symétrie. Pour un système à $4n$ électrons- π comme le cyclobutène, la symétrie de l'orbital moléculaire occupée la plus haute du système est telle que l'interaction de liaison entre les atomes terminaux doivent impliquer le chevauchement entre les enveloppes des orbitales sur les atomes opposés de la

surface moléculaire. Donc, ceci peut être réalisé que dans un mouvement conrotationnel des groupements terminaux.

L'ouverture du cycle du cyclobutène est un processus spontané dès que l'on chauffe le système réactionnel. Ce processus est connu pour passer à travers un chemin conrotatoire pour une excitation thermique et un chemin disrotatoire pour une excitation photochimique. Ces deux mouvements sont représentés à la Figure 10 :

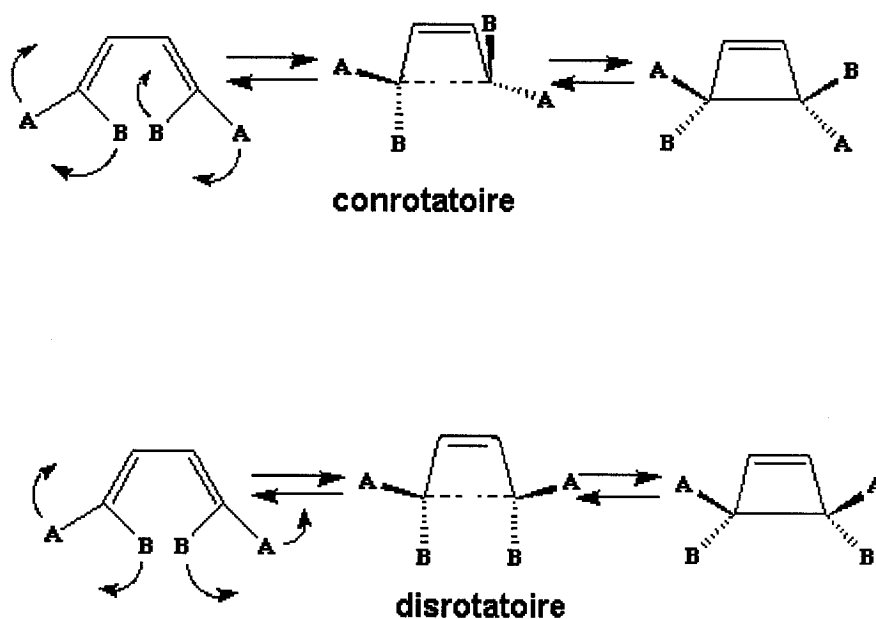


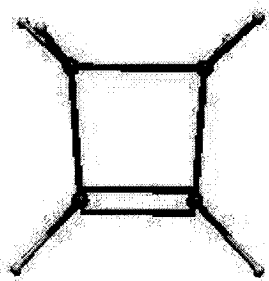
Figure 10 : Schéma du mouvement conrotatoire et disrotatoire

Pour cette ouverture du cycle, la liaison entre les carbones C₃ et C₄ va se rompre et ces carbones devront se rehybrider de sp³ à sp². Le processus étant conrotatoire, les atomes de carbone pivotent dans la même direction, soit le sens horaire ou anti-horaire. À la Figure 11, on présente le schéma de la réaction d'ouverture du cycle pour le cyclobutène. Le Tableau VII montre les différents paramètres pour les géométries du cyclobutène (1) et du 1,3-butadiène (3) obtenus lors du calcul du IRC. Ce tableau nous donne également les paramètres pour l'état de transition. Les résultats obtenus pour l'ouverture du cycle sont présentés à la Figure 12. Nous obtenons une énergie d'activation $\Delta E^\ddagger = 32.5$ kcal/mol et une variation d'énergie $\Delta E = -5.6$ kcal/mol. Pour le calcul de l'énergie libre à 300 K, nous avons un $\Delta F^\ddagger = 34.1$ kcal/mol et une énergie libre de la réaction $\Delta F = -4.5$ kcal/mol. Les résultats expérimentaux montrent que l'énergie d'activation est de $32.7 \pm 0.2^{70,71}$ kcal/mol et que l'énergie de la réaction est de -10.9 ± 0.5^{72} kcal/mol. Pour cette même réaction, Michalak et Ziegler⁵⁶ ont obtenu, en utilisant une simulation Car-Parrinello et la fonctionnelle de densité locale de Perdew et Zunger¹⁹, une valeur d'énergie libre d'activation de 31.0 kcal/mol et une énergie libre de réaction de -8.0 kcal/mol. Nous obtenons un excellent accord pour la barrière réactionnelle mais une pauvre corrélation pour l'énergie de la réaction. Cette différence peut venir de la fonctionnelle utilisée pour ce calcul (BP86). Il faut également signaler qu'il n'y a aucune correction pour l'estimation du zéro énergie. La distance C₃-C₄ passe de

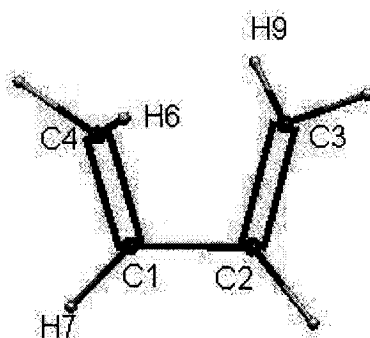
Tableau VII. Paramètre pour les géométries de cyclobutène, 1,3-butadiène et de l'état de transition

Paramètre ^a (Å,deg.)	Cyclobutène	État de transition	1,3 butadiène
R(C ₁ -C ₂)	1.351	1.386	1.469
R(C ₁ -C ₄)	1.525	1.433	1.349
R(C ₄ -C ₃)	1.572	2.156	3.013
R(C ₃ -C ₂)	1.525	1.433	1.349
R(C ₁ -H ₇)	1.096	1.098	1.098
R(C ₃ -H ₉)	1.107	1.105	1.094
∠(C ₃ -C ₂ -C ₁)	94.0	104.4	121.9
∠(C ₁ -C ₄ -H ₆)	114.6	118.8	122.6
∠(C ₂ -C ₁ -H ₇)	136.3	129.4	114.9

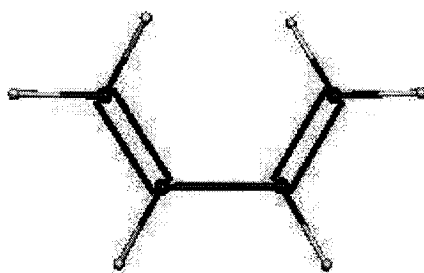
a) optimisé avec la fonctionnelle BP86 et la base 6-31G*



(1)



(2)



(3)

Figure 11. Schéma pour la réaction de l'ouverture du cycle du cyclobutène (1)

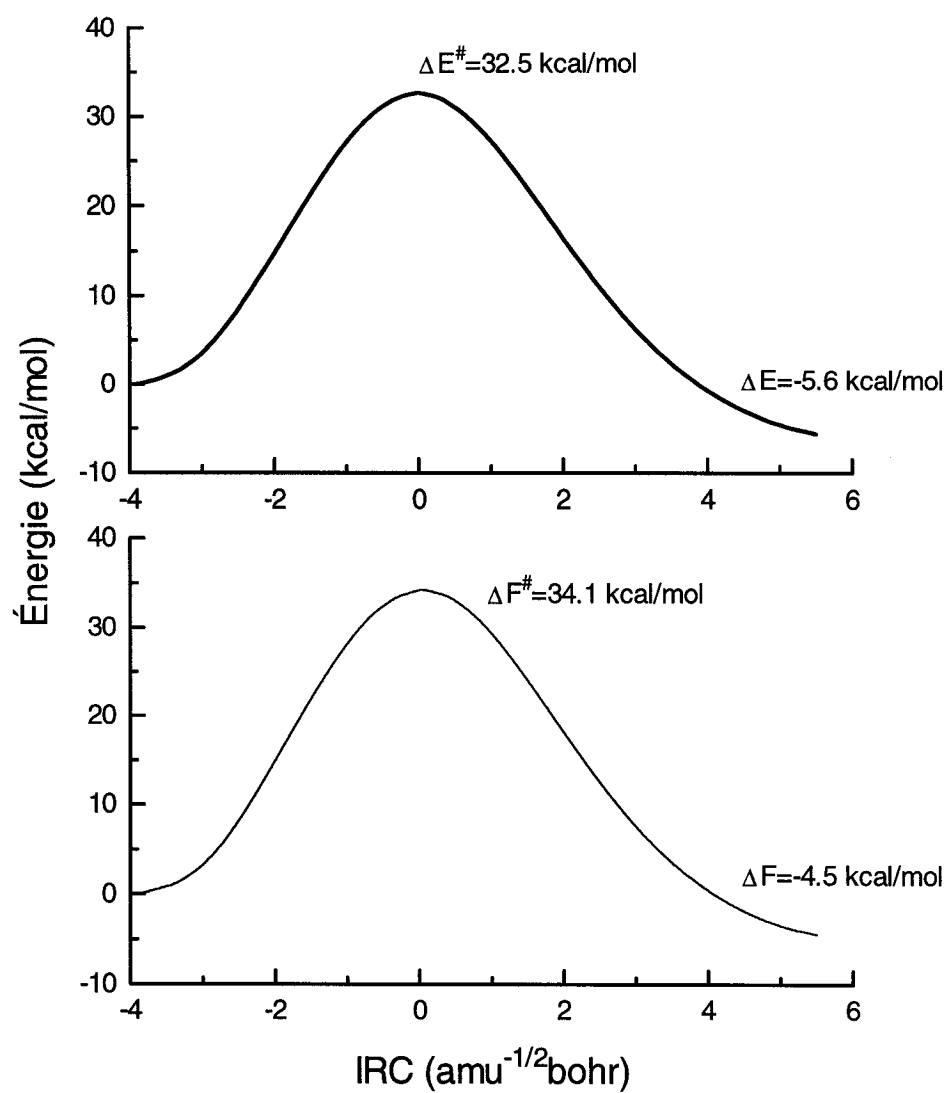


Figure 12. Profil d'énergie et d'énergie libre pour la réaction d'ouverture du cycle du cyclobutène le long du chemin réactionnel.

1.572 Å pour le cyclobutène à 2.156 Å pour l'état de transition. Cette distance C-C se compare aux autres distances (2.144 – 2.292 Å) qui ont été calculées pour d'autres états de transition de réaction péricyclique.

4.4.2 Réaction de l'ouverture du cycle de l'oxetène

Dans ce deuxième exemple, nous regardons l'effet de la substitution d'un groupement méthyle par un atome d'oxygène dans l'ouverture d'un cycle. En générale, les composés hétérocycliques contenant l'oxygène qui passe de l'hybridation sp^3 à sp^2 , sont plus exothermiques que les cycles analogues avec seulement des atomes de carbone⁷³. Le schéma de l'ouverture du cycle de l'oxetène en acrolein est présenté à la Figure 13. Le Tableau VIII nous montre les paramètres pour les géométries d'oxytène, d'acrolein et de l'état de transition.

Le profil d'énergie libre obtenu par dynamique moléculaire à 300 K est présenté à la Figure 14. Cette figure nous montre que l'énergie d'activation est $\Delta E^\ddagger = 22.9$ kcal/mol et l'énergie de la réaction est $\Delta E = -23.7$ kcal/mol. Nous obtenons une énergie libre d'activation de $\Delta F^\ddagger = 24.2$ kcal/mol et une énergie libre pour l'ouverture du cycle de $\Delta F = -23.2$ kcal/mol. La valeur expérimentale pour l'énergie d'activation est de 24.1 ± 1.5 kcal/mol^{74,75}. Par contre, il n'y a pas de valeur expérimentale pour l'énergie de la réaction. Les valeurs disponibles dans la littérature

Tableau VIII. Paramètre des géométries d'oxetène, d'acrolein et de l'état de transition

Paramètre ^a (Å,deg.)	Oxetène	État de transition	Acrolein
R(O ₂ -C ₇)	1.499	1.982	2.749
R(O ₂ -C ₃)	1.398	1.319	1.228
R(C ₃ -C ₄)	1.351	1.382	1.484
R(C ₄ -C ₇)	1.513	1.443	1.345
R(C ₃ -H ₅)	1.096	1.103	1.126
R(C ₇ -H ₈)	1.107	1.106	1.085
∠(C ₃ -O ₂ -C ₇)	85.4	76.5	61.1
∠(O ₂ -C ₃ -C ₄)	99.5	108.0	121.5
∠(C ₃ -C ₄ -C ₇)	86.6	96.6	116.7
∠(O ₂ -C ₃ -H ₅)	120.1	119.8	121.1
∠(O ₂ -C ₇ -H ₈)	110.5	127.9	173.4
∠(C ₇ -C ₄ -H ₆)	135.7	129.2	123.2

a) optimisé avec la fonctionnelle BP86 et la base 6-31G*

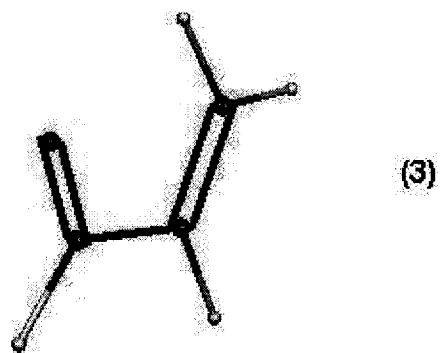
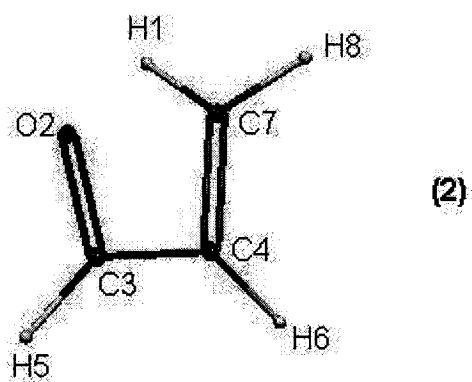
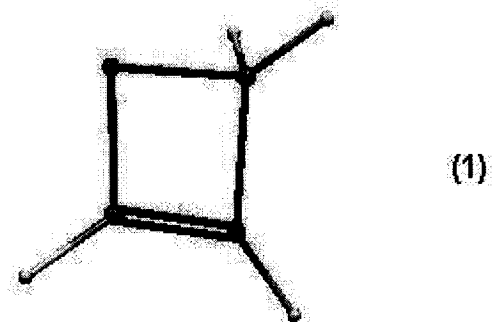


Figure 13. Schéma pour la réaction d'ouverture du cycle de l'oxetène (1)

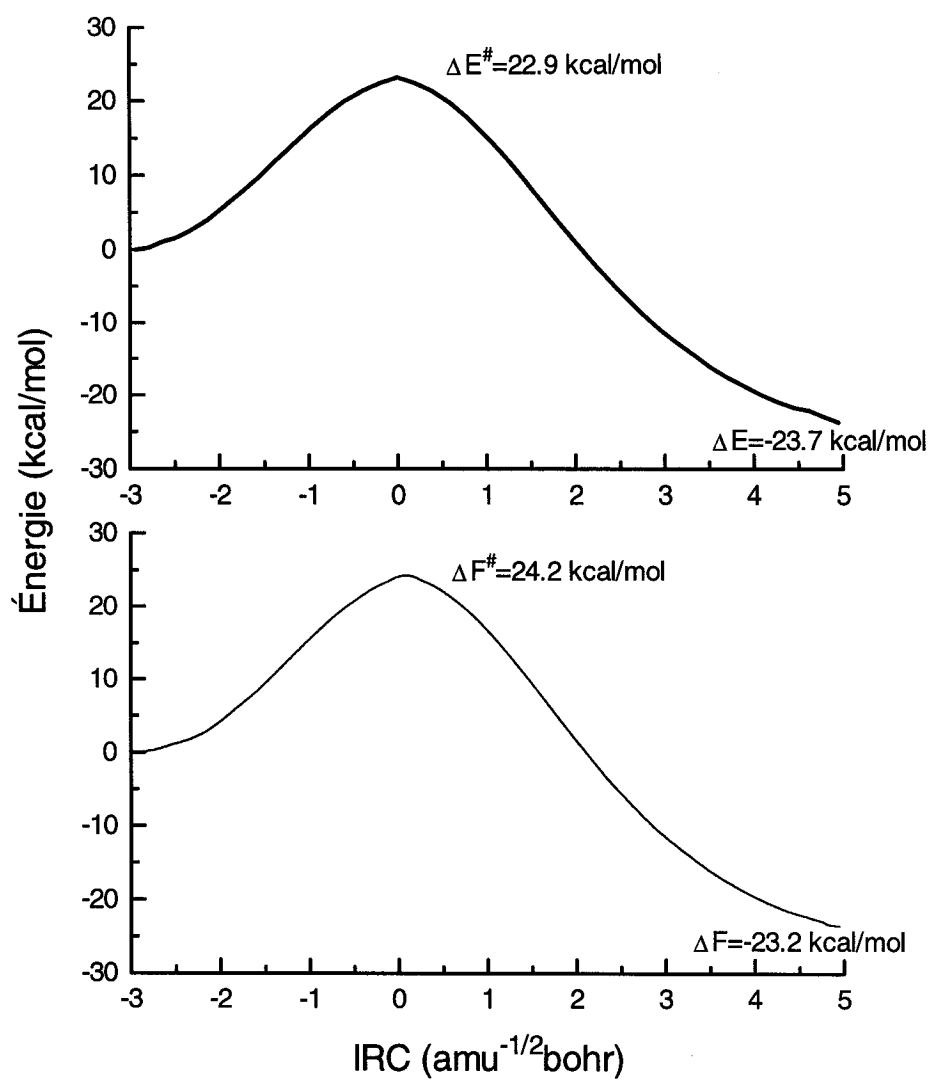


Figure 14. Profil d'énergie et d'énergie libre pour la réaction d'ouverture du cycle de l'oxetène le long du chemin réactionnel.

sont des calculs⁶⁴ HF/6-31g* et CISD/6-31g* qui sont respectivement -35.12 kcal/mol et -27.5 kcal/mol. Nos résultats nous montrent très bien que l'ouverture du cycle de l'oxetène est beaucoup plus exothermique (une différence de 18.7 kcal/mol) que l'ouverture du cyclobutène comme on le prévoyait. Si on regarde la distance O₂-C₇, on voit que celle-ci passe de 1.499 Å pour l'oxetène à 1.982 Å à l'état de transition. La présence de l'oxygène à la place d'un carbone fait en sorte de diminuer cette distance à l'état de transition par rapport à système avec seulement des atomes de carbones.

4.4.3 Réaction de cyclisation de l'hexatriène

La cyclisation de l'hexatriène en cyclohexadiène est considérée comme une réaction péricyclique de base. Il est connu que cette réaction passe à travers un chemin disrotatoire lorsqu'elle est soumise à une excitation thermique. Dans un système ouvert contenant $4n+2$ électrons- π comme l'hexatriène, l'interaction de la liaison de la liaison terminale, C₁-C₂, dans l'état fondamental de la molécule exige le chevauchement des enveloppes des orbitales sur les atomes du même côté du système. Ce chevauchement est seulement possible pour des déplacements disrotatoire. Dans un processus disrotatoire, les 2 carbones pivotent dans des sens opposés.

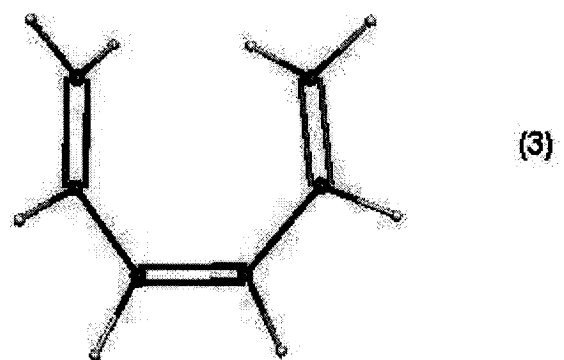
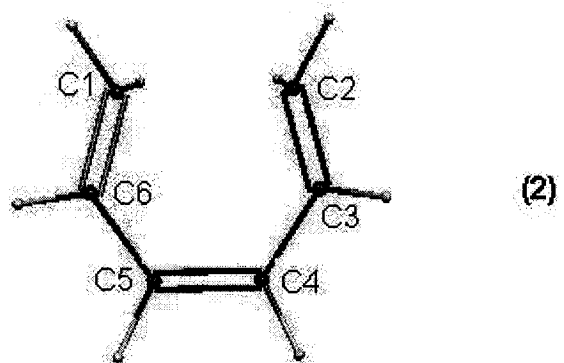
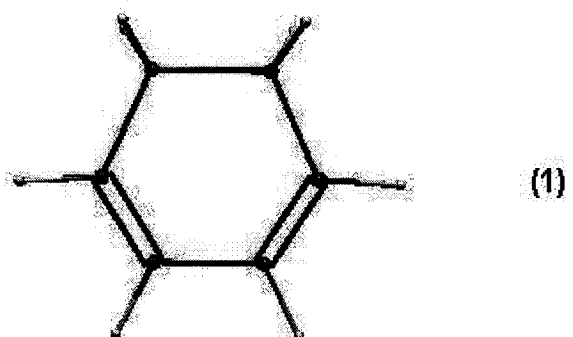


Figure 15. Schéma pour la réaction de cyclisation du hexatriène (3) en cyclohexadiène (1)

La cinétique unimoléculaire de la fermeture du cycle a été déterminée expérimentalement par Lewis et Steiner⁷⁶ dans la phase gazeuse. Ils ont trouvé une énergie d'activation de 29.9 ± 0.5 kcal/mol pour la barrière entre la forme 1,3 cis-5 hexatriène (C_{2v}) et l'état de transition. La Figure 15 présente le schéma pour la réaction de cyclisation entre le précurseur pour l'électrocyclisation (3) et le cyclohexadiène (1). Nous avons arrêté au précurseur car la surface d'énergie potentielle entre ce précurseur et la forme la plus stable de l'hexatriène est très plate et il est difficile d'obtenir un bon IRC dans cette zone. Les paramètres pour les géométries sont présentés au Tableau IX. Pour cette réaction, nous avons utilisé la fonctionnel VWN et la base 6-31G*.

Le profil d'énergie libre entre la forme précurseur de l'hexatriène et le cyclohexadiène est présenté à la Figure 16. Cette figure indique que l'énergie d'activation $\Delta E^\ddagger = 11.5$ kcal/mol et que l'énergie de la réaction $\Delta E = -24.6$ kcal/mol. Nous obtenons une valeur de 9.6 kcal/mol pour l'énergie libre d'activation et une énergie libre de réaction de -32.3 kcal/mol. En utilisant un calcul de niveau B3LYP/6-31G**, Rodriguez-Otero⁷⁷ a obtenu une énergie d'activation de 20.34 kcal/mol pour la réaction entre le précurseur et le cyclohexatriène et une énergie d'activation de 30.11 kcal/mol pour la réaction complète. Dans une autre étude, Schleyer⁷⁸ a obtenu une énergie d'activation entre la forme précurseur et le cyclohexatriène de 16 kcal/mol pour un calcul MP2/6-31G*. Pour la réaction

Tableau IX. Paramètres pour les géométries de cyclohexadiène, d'hexatriène et de l'état de transition

Paramètre ^a (Å,deg)	Cyclohexadiène	État de transition	hexatriène
R(C ₆ -H ₁₀)	1.102	1.102	1.103
R(C ₁ -H ₁₂)	1.107	1.099	1.096
R(C ₆ -C ₅)	1.352	1.408	1.459
R(C ₆ -C ₁)	1.514	1.403	1.353
R(C ₁ -C ₂)	1.574	2.313	2.964
R(C ₄ -C ₅)	1.471	1.412	1.368
∠(C ₅ -C ₆ -C ₁)	121.3	125.6	129.3
∠(C ₁ -C ₂ -C ₃)	115.5	103.7	92.5
∠(C ₅ -C ₆ -H ₁₀)	120.6	115.6	113.7
∠(C ₆ -C ₁ -H ₁₂)	110.3	118.4	120.6

a) optimisé avec la fonctionnelle BP86 et la base 6-31G*

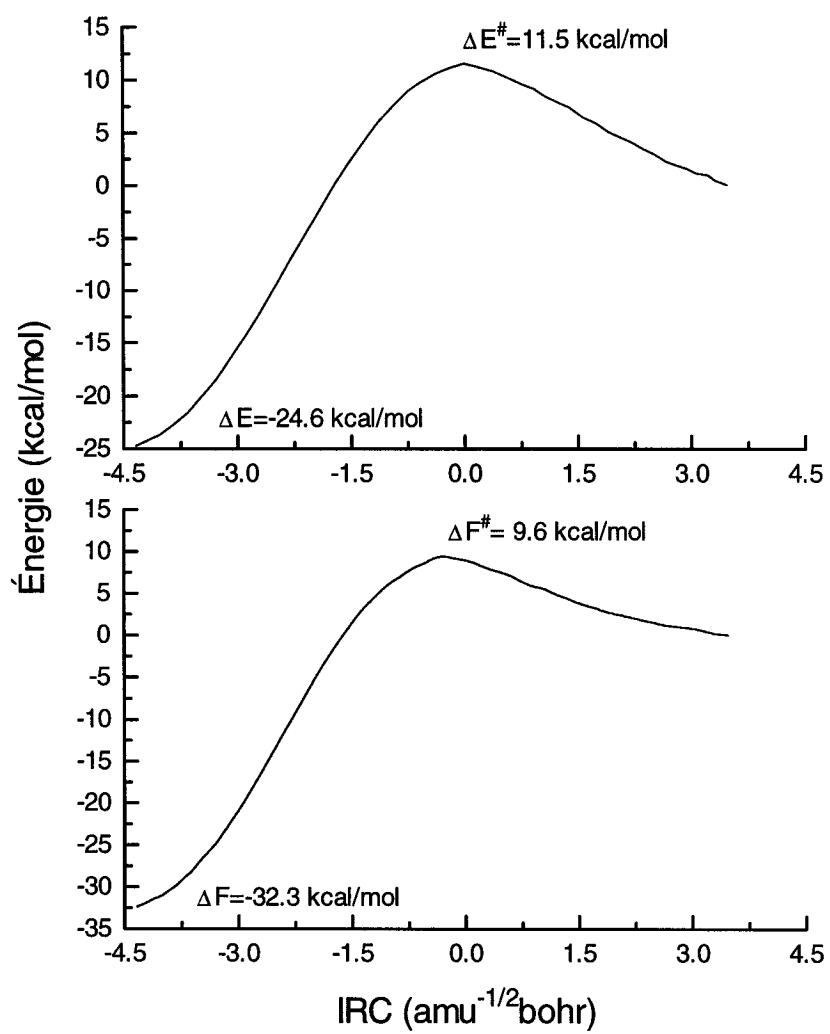


Figure 16. Profil d'énergie et d'énergie libre pour la réaction de cyclisation de l'hexatriène le long du chemin réactionnel.

complète, il obtient une énergie d'activation de 26.8 kcal/mol et une énergie de la réaction de -21.2 kcal/mol. Notre résultat pour l'énergie d'activation avec VWN/6-31G* se rapproche de la valeur obtenue avec la méthode MP2 et la même base. On remarque que le résultat de l'énergie libre de la réaction dévie énormément de l'énergie de la réaction. On a une différence de 7 kcal/mol entre les deux calculs. Le résultat de l'énergie libre pourrait paraître erroné mais celui-ci a été confirmé avec un calcul de l'énergie libre en utilisant l'approximation harmonique. Ce résultat est présenté au Tableau XII.

4.4.4 Migration 1,3 oxygène-oxygène du méthyle formate et du trifluorométhyle formate

Dans cette section, nous regardons l'énergie libre de la barrière d'activation pour la migration entre l'oxygène 1 et 3 dans deux esters différents. Nous utilisons le méthyle et le trifluorométhyle de formate. Pour exécuter ces calculs, nous utilisons la conformation Z car celle-ci est beaucoup plus stable dans la majorité des esters⁷⁹. La Figure 17 représente le schéma de la réaction de migration entre l'oxygène 1 et 3 pour le méthyle formate. Les paramètres pour l'état de transition pour le méthyle et le trifluorométhyle sont présentés au Tableau X. Pour effectuer ce calcul, nous avons utilisé le chemin réactionnel complet (le groupement méthyle ou trifluorométhyle passait de l'oxygène 1 à l'oxygène 3). Pour ces deux exemples, nous avons utilisé

Tableau X. Paramètres des états de transition pour la migration entre l'oxygène 1 et 3 pour le méthyle et le trifluorométhyle méthyle de formate.

Paramètre ^a (Å,deg.)	Méthyle formate	Trifluorométhyle formate
R(H ₁ -C ₂)	1.109	1.109
R(C ₂ -O ₃)	1.262	1.262
R(O ₃ -C ₅)	2.114	2.094
R(C ₅ -H ₇)	1.091	
R(C ₅ -F ₇)		1.332
∠(H ₁ -C ₂ -O ₃)	118.6	119.2
∠(O ₄ -C ₂ -O ₃)	122.8	121.6
∠(C ₂ -O ₃ -C ₅)	86.9	87.4
∠(H ₆ -C ₅ -H ₈)	115.4	
∠(F ₆ -C ₅ -F ₈)		112.4
∠(C ₂ -O ₃ -C ₅ -H ₈)	92.9	
∠(C ₂ -O ₃ -C ₅ -F ₈)		93.3

a) optimisé avec la fonctionnelle BP86 et la base 6-31G*

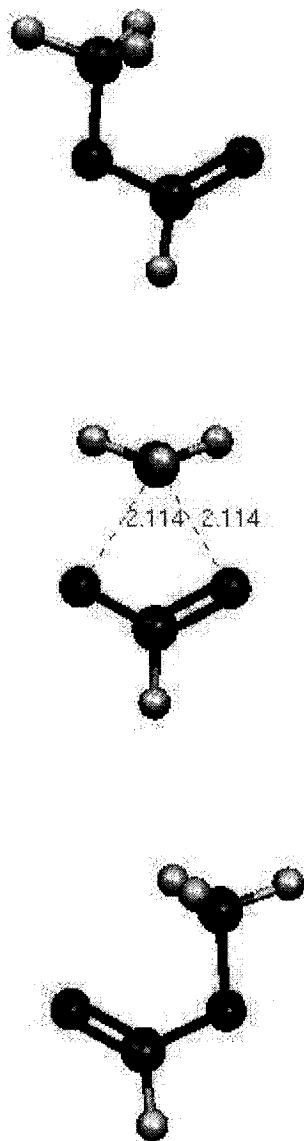


Figure 17. Schéma représentant la réaction de migration du méthyle formate entre l'oxygène 1 et 3.

des fenêtres au lieu de la technique de la croissance lente. Pour la simulation avec le méthyle formate, nous avons utilisé 56 fenêtres et 33 fenêtres pour la réaction avec le trifluorométhyle formate. Nous avons accumulé entre 500 et 1000 géométries dans notre ensemble pour évaluer la variation d'énergie libre d'une fenêtre.

Les Figure 18 et Figure 19 montrent les résultats pour le calcul de l'énergie libre à 300 K. Pour ces réactions, nous obtenons une valeur identique pour la barrière de réaction et la barrière d'énergie libre de la réaction. Pour le méthyle formate, nous obtenons une barrière de 46.0 kcal/mol et pour le trifluorométhyle formate, une barrière de 49.4 kcal/mol. Nous n'avons aucune valeur expérimentale pour la comparaison mais un calcul⁸⁰ MP2/6-31+g*. Avec ce niveau de calcul, la barrière d'activation est de 55.30 kcal/mol pour le méthyle formate et de 62.98 kcal/mol pour le trifluorométhyle formate.

Le réactif et le produit étant identique, nous savons que l'énergie libre de la réaction devrait être zéro si nous n'avons aucune hystérésis dans notre système. Nous obtenons une énergie libre de la réaction de 0.47 kcal/mol ce qui démontre que cette technique apporte un hystérésis très faible.

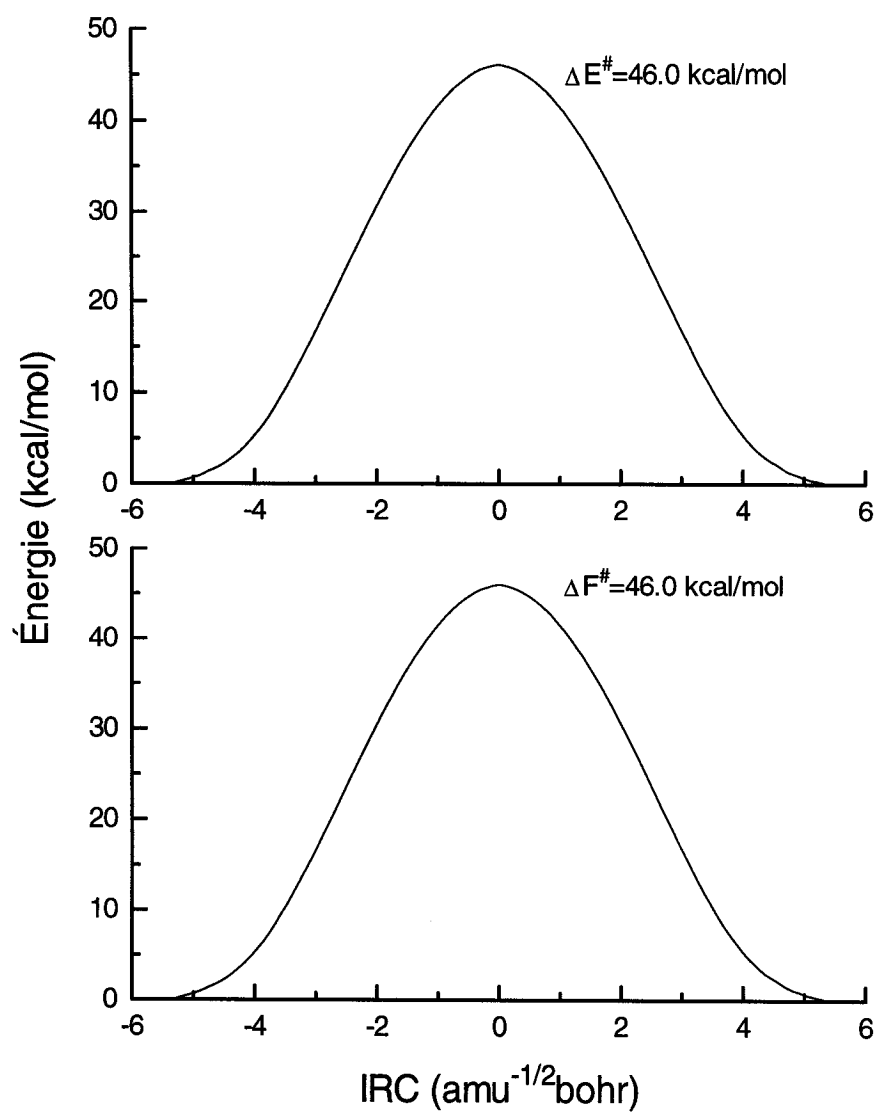


Figure 18. Profil d'énergie libre pour la migration 1,3 oxygène-oxygène du méthyle formate le long du chemin réactionnel.

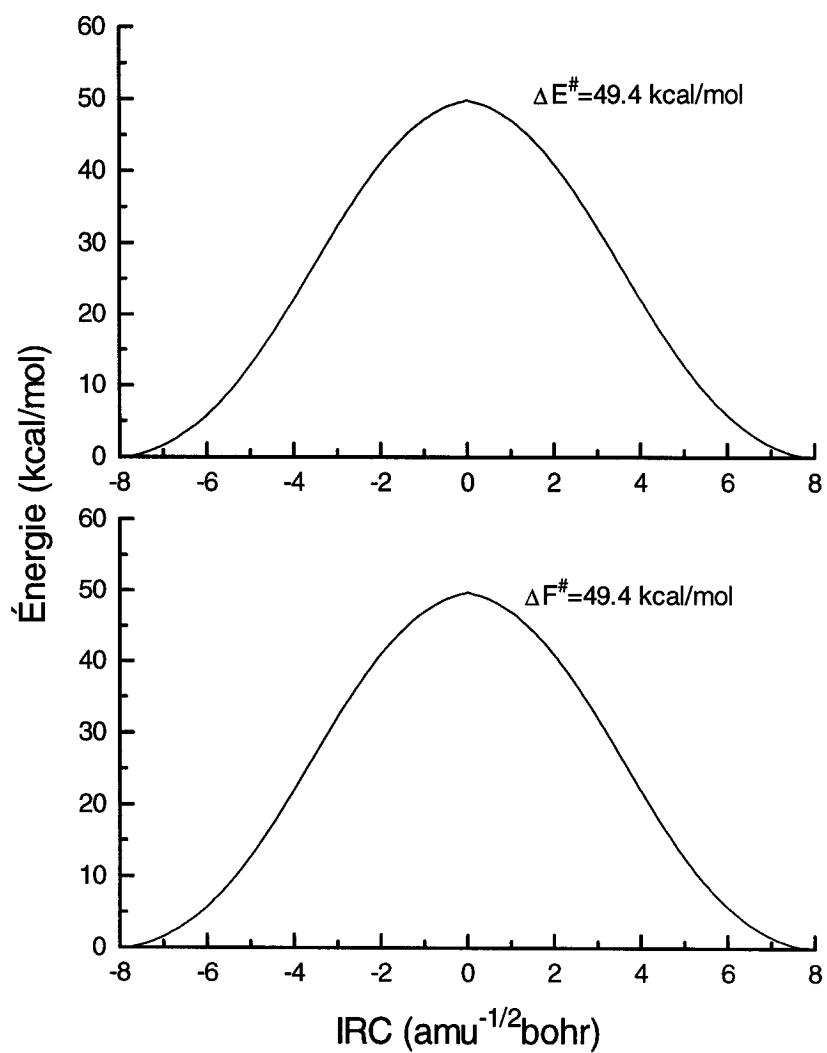


Figure 19. Profil d'énergie libre pour la migration 1,3 oxygène-oxygène pour le trifluorométhyle formate le long du chemin réactionnel.

4.4.5 Isomérisation du cyanure d'hydrogène

L'isomérisation du cyanure d'hydrogène en isocyanure d'hydrogène est un exemple simple très souvent utilisé par les théoriciens^{81,82}. Les paramètres des structures des molécules de HCN, CNH et l'état de transition sont présentés au Tableau XI.

Le profil d'énergie du IRC et le profil d'énergie libre le long du IRC sont présentés à la Figure 20. Le profil d'énergie nous montre que la réaction d'isomérisation est endothermique avec un $\Delta E = 15.0$ kcal/mol pour la réaction. Le profil d'énergie libre à 300 K nous montre également que l'énergie libre de la réaction est $\Delta F = 15.0$ kcal/mol et l'énergie libre de la barrière de réaction $\Delta F^\ddagger = 44.0$ kcal/mol. Nous n'avons effectué aucune correction pour l'énergie du point zéro. L'énergie libre de la réaction se compare avec le résultat expérimental⁸³ qui est de 14.8 ± 2 kcal/mol. Les plus récents résultats de profil d'énergie libre obtenu par Michalak et Ziegler⁵⁶ montre une énergie libre de 14.1 kcal/mol et une énergie libre de la barrière d'activation de 44.6 kcal/mol.

Tableau XI. Paramètre pour l'isomérisation du cyanure d'hydrogène

Paramètres ^a	HCN	CNH	État de transition
R(C-N)	1.207	1.222	1.276
R(N-H)	2.304	1.048	1.356
R(C-H)	1.097	2.270	1.299
∠(CNH)	180	180	63.6

a) optimisé avec la fonctionnelle BP86 et la base 6-31G*

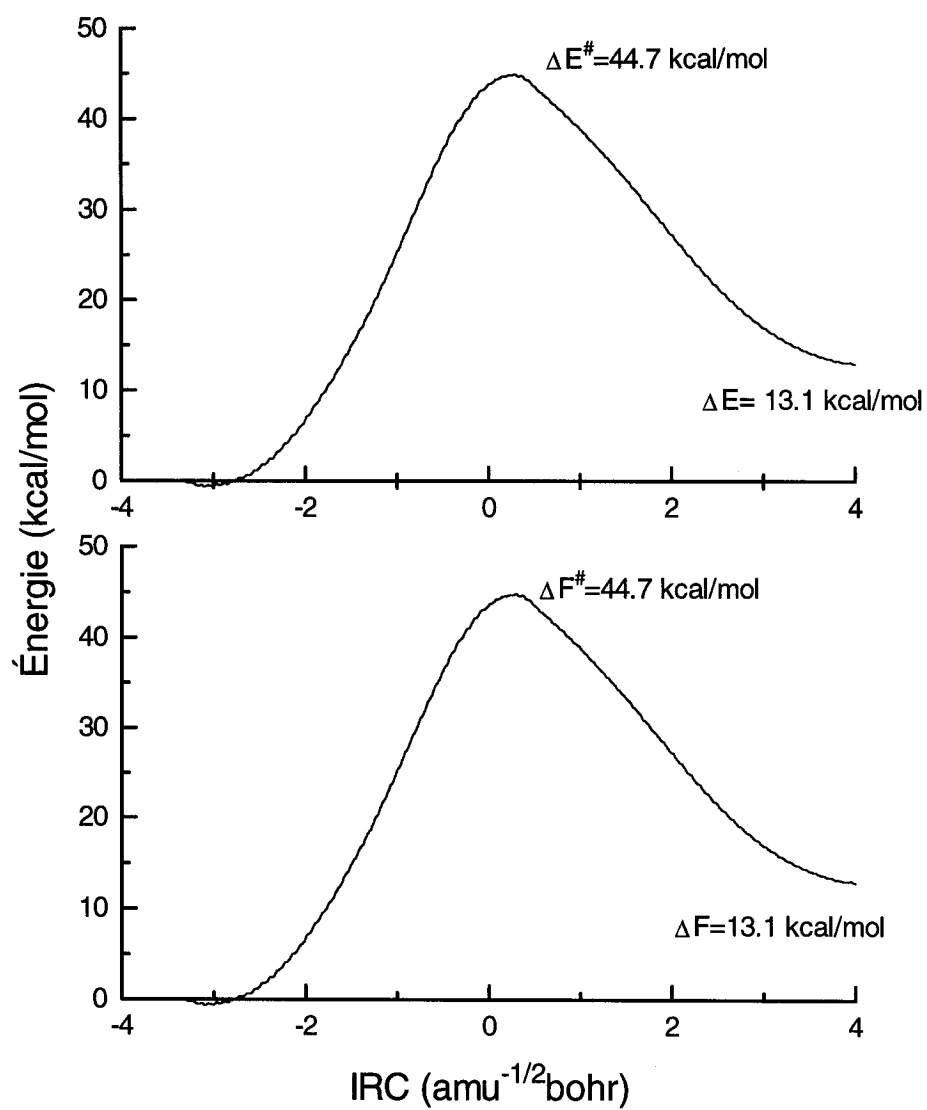


Figure 20. Profil d'énergie libre pour la réaction d'isomérisation du cyanure d'hydrogène le long du chemin réactionnel.

4.5 Discussion

Les différents exemples montrent très bien qu'il est possible d'obtenir des profils d'énergie libre très lisses en suivant le IRC. Lorsque l'on regarde ce qu'il se passe durant une simulation, on voit que tous les paramètres des molécules (liens, angles,...) oscillent régulièrement autour du chemin réactionnel. A notre connaissance, c'est la première fois que des simulations de dynamiques moléculaires Born-Oppenheimer (BOMD) le long du IRC sont exécutées avec un programme utilisant la théorie de la fonctionnelle de la densité avec des orbitales gaussiennes. La seule implémentation décrite dans la littérature utilise la dynamique moléculaire Car-Parrinello (CPMD) avec des fonctions d'onde plane. La dynamique moléculaire Car-Parrinello est beaucoup plus rapide car la fonction d'onde électronique ne doit pas être convergée à chaque itération comme pour la dynamique moléculaire Born-Oppenheimer mais elle est seulement propagée. Par contre, pour avoir une trajectoire de dynamique moléculaire stable, il est nécessaire d'utiliser un temps d'intégration petit. En convergeant la fonction d'onde à chaque itération, on pourrait croire que ce type de simulation est très coûteux en temps de calculs. Cependant, parce que les forces sont réelles, on peut utiliser un temps d'intégration d'environ 100 fois supérieure à celui utilisé avec la méthode CPMD⁸⁴. Généralement dans les

simulations Car-Parrinello pour augmenter le temps d'intégration, les masses sont changées⁵⁶ comme par exemple pour le C,N et O passe de 12, 14 et 16 amu à 2 amu.

L'évaluation de l'énergie libre en utilisant une trajectoire de dynamique moléculaire est beaucoup plus dispendieuse en temps de calculs que l'utilisation de calculs statiques en utilisant l'approximation harmonique. L'approximation harmonique est une excellente approximation près d'un minimum de la surface d'énergie potentielle. Ainsi, il est possible d'évaluer l'énergie libre en construisant la fonction de partition à partir des fréquences de vibrations de la molécule qui nous intéresse. Pour des petites molécules, il est relativement facile d'évaluer les fréquences du système mais lorsque dans le processus les forces intermoléculaires faibles dominant⁸⁵, l'approximation harmonique est moins efficace et l'analyse vibrationnelle devient très coûteuse en temps de calcul et en espace mémoire nécessaire pour exécuter ce calcul. De plus, l'approximation harmonique est moins valable pour de tels systèmes car ils ont un grand nombre de minimums sur la surface d'énergie potentielle. Également, l'approximation harmonique échoue lorsque nous avons un rotateur libre dans une molécule. Alors, il n'est plus possible de calculer exactement les propriétés thermodynamiques avec le calcul des modes normaux. Même si l'approximation harmonique n'est pas valide dans certaine condition, il est toujours possible d'utiliser de la dynamique moléculaire (BOMD) pour le calcul de l'énergie libre car celle-ci ne dépend pas de l'approximation harmonique. Pour des

systèmes plus complexes comme des réactions en solution ou des réactions enzymatiques, il serait possible de coupler cette méthode pour évaluer l'énergie libre avec la fonctionnalité QM/MM décrit au chapitre 2. Pour l'instant le calcul de l'énergie libre le long du IRC n'est pas encore inclus dans notre version QM/MM.

Les valeurs d'énergie libre obtenues avec les exemples de réaction péricyclique dans ce chapitre pour la validation de la méthode de calcul de l'énergie libre le long du chemin réactionnel auraient pu être également obtenues en utilisant l'approximation harmonique car ces systèmes sont très simples. Pour évaluer les énergies libres avec l'approximation harmonique, nous avons utilisé les géométries obtenues avec le calcul du IRC, c'est-à-dire, la géométrie de l'état de transition utilisée comme point de départ pour le calcul du IRC et les structures pour le réactif et le produit obtenu avec le calcul du chemin réactionnel. Les résultats pour ces calculs sont présentés au Tableau XII. Ce tableau montre qu'il existe un bon accord entre les résultats obtenus par dynamique moléculaire et par les calculs statiques. On obtient une déviation moyenne d'environ 2.5 kcal/mol entre les deux méthodes. Ces résultats montrent que l'évaluation de l'énergie libre avec des simulations de dynamique moléculaire est un excellent complément à la méthode statique.

Tableau XII. Comparaison des résultats d'énergie libre entre la dynamique moléculaire et l'analyse vibrationnelle.

Énergie libre d'activation (kcal/mol)		
réaction	Dynamique moléculaire	Analyse vibrationnelle
Cyclobutène	34.1	30.8
Oxytène	24.2	21.6
Hexatriène	9.6	9.6
Énergie libre de la réaction (kcal/mol)		
Cyclobutène	-4.5	-7.1
Oxytène	-23.2	-24.7
Hexatriène	-32.3	-28.7

4.6 Conclusion

Le calcul de l'énergie libre le long du chemin réactionnel en utilisant une contrainte linéaire a été implémenté dans DeFT. Cette fonctionnalité nous permet d'étudier des réactions chimiques en utilisant la théorie de la fonctionnelle de la densité pour explorer la surface d'énergie potentielle. Les exemples utilisés dans ce chapitre démontrent très bien que l'utilisation de la méthode de croissance lente donne des profils d'énergie libre très lisse et cela même si l'énergie libre dévie de la trajectoire de référence. L'usage du IRC comme chemin de référence permet de diminuer l'hystérésis en s'assurant que l'on passe réellement dans la région de l'état de transition. Par contre, la procédure de croissance lente demande un très grand nombre de fenêtres et coûte très cher en temps CPU.

Chapitre 5 : Développement d'une méthode hybride AM1/DFT pour le calcul de l'énergie libre

5.1 Introduction

L'évaluation de l'énergie libre représente un défi majeur dans l'étude de réaction chimique. Cette évaluation à partir de simulations numériques est donc très importante et souvent très coûteuse en temps de calculs. Lors du calcul de l'énergie libre, l'échantillonnage de l'espace de phase est l'étape qui est la plus coûteuse en temps de calculs. Dans le chapitre précédent, nous avons montré qu'il était possible d'évaluer l'énergie libre le long du chemin réactionnel en utilisant uniquement un calcul de haut niveau (la fonctionnelle de la densité). Cette méthode nous donne un profil d'énergie libre très lisse mais a besoin de ressources informatiques importantes. Comme nous l'avons vu au chapitre 1, l'énergie libre peut être obtenue avec l'équation suivante :

$$\Delta F_{BA} = -k_B T \ln \left\langle e^{\frac{H(\lambda_B) - H(\lambda_A)}{k_B T}} \right\rangle_{\lambda_A} \quad (5.1).$$

En utilisant cette équation, il est possible d'évaluer l'énergie libre à partir de la probabilité relative d'observer une configuration. Dans l'équation 5.1, les configurations, λ_B , sont prélevées dans l'ensemble λ_A . Donc, il faut que

l'Hamiltonien de l'ensemble λ_B puissent décrire correctement l'ensemble λ_A . Pour obtenir la convergence de notre ensemble, le chevauchement des régions de basses énergies des deux ensembles doit être significatives. Le choix du chemin n'aura pas d'effet sur l'énergie libre, car l'énergie libre est une fonction d'état, mais aura un énorme effet sur le temps de calculs. Ainsi, lors de l'échantillonnage de l'ensemble de Boltzmann par la trajectoire de dynamique moléculaire, le temps d'intégration devra être suffisamment petit pour bien décrire les vibrations et pour bien balayer la surface d'énergie potentielle. Si ce temps d'intégration est trop grand, nous aurons une mauvaise description des vibrations de notre système. Une mauvaise description nous apportera une mauvaise trajectoire.

Dans ce chapitre, nous regarderons une méthode basée sur l'utilisation d'un échantillonnage non-Boltzmann pour l'évaluation du profil de l'énergie libre avec la méthode de perturbation thermodynamique. Dans cette méthode, nous utiliserons un potentiel de référence moins coûteux en temps CPU pour effectuer l'échantillonnage. Le potentiel utilisé doit préserver l'endroit des minimums et son influence sur l'énergie libre est corrigée par une procédure comme celle utilisée dans la méthode développée par Torrie et Valleau⁸⁶.

Cette méthode a été utilisée dans un premier temps par le groupe de Brünger⁸⁷ pour évaluer l'énergie libre de la mutation du butane en propanol en utilisant des potentiels classiques seulement. Avec l'utilisation de potentiel classique, il est

impossible d'observer la formation ou la rupture de liaison. Donc, l'utilisation de potentiel classique n'est pas très attrayante pour l'étude de réaction chimique. Par la suite, VandeVondel et Rothlisberger^{88,89,90} ont utilisé un potentiel polarisé (bias potential) classique pour augmenter la vitesse du calcul pour l'évaluation du profil de l'énergie libre de l'angle de torsion H-C-C-H dans l'éthane. Dans cette technique, le potentiel polarisé est ajouté au potentiel régulier pour permettre un échantillonnage préférentiel ce qui permet de mieux représenter l'ensemble de Boltzmann. Par la suite, les groupes de Rothlisberger et Woo⁹¹ ont développé un potentiel polarisé utilisant l'information chimique codée dans la structure électronique, c'est-à-dire les énergies des orbitales (HOMO/LUMO). Ils ont utilisé ce potentiel pour étudier une série de réactions chimiques le long du chemin réactionnel avec des simulations de dynamique moléculaire ab-initio. Dernièrement, le groupe de Warshel⁹² ont utilisé l'approche QM_(ai)/MM pour calculer l'énergie libre de l'autodissociation de l'eau en solution. Dans cette approche, les paramètres EVB sont choisis pour bien reproduire les résultats ab-initio en phase gazeuse d'un modèle pour la réaction le long du chemin sélectionné. Alors, des trajectoires de dynamique moléculaire sont exécutées avec le potentiel EVB pour échantillonner le système le long du chemin de la réaction. Les énergies potentielles QM_(ai)/MM sont calculées pour les conformations prélevées. Les variations d'énergie libre le long du chemin réactionnel sont évaluées avec la différence d'énergie potentielle entre les énergies EVB et QM/MM. Dans cette thèse, nous utilisons l'échantillonnage non-Boltzmann pour reproduire le profil

d'énergie libre, calculée complètement par la fonctionnelle de la densité, d'une réaction en utilisant comme potentiel de référence la méthode semi-empirique AM1. L'utilisation d'un potentiel semi-empirique nous permet d'étudier des réactions où la rupture ou la création d'une liaison est présente.

5.2 Méthode des perturbations thermodynamiques non-Boltzmann^{73,93}

Comme on l'a vu dans la section sur le calcul de la variation d'énergie libre, on écrit la variation d'énergie libre d'Helmholtz entre l'état A et B comme étant :

$$\Delta F = -k_B T \ln \frac{Z_B}{Z_A} \quad (5.2)$$

où la fonction de partition Z s'écrit comme :

$$Z = \iint e^{-\beta H(\vec{p}, \vec{r})} d\vec{p} d\vec{r} \quad (5.3)$$

$$\frac{Z_B}{Z_A} = \frac{\iint e^{-\beta H_B(\vec{p}, \vec{r})} d\vec{p} d\vec{r}}{\iint e^{-\beta H_A(\vec{p}, \vec{r})} d\vec{p} d\vec{r}} \quad (5.4).$$

Si on introduit au numérateur l'identité $1 = \exp (+\beta H_{A,r}(\mathbf{p},\mathbf{r})) \exp (-\beta H_{A,r}(\mathbf{p},\mathbf{r}))$.

Alors, on peut écrire le rapport Z_B/Z_A :

$$\frac{Z_B}{Z_A} = \frac{\iint e^{-\beta H_B(\vec{p},\vec{r})} e^{+\beta H_{A,r}(\vec{p},\vec{r})} e^{-\beta H_{A,r}(\vec{p},\vec{r})} d\vec{p} d\vec{r}}{\iint e^{-\beta H_A(\vec{p},\vec{r})} d\vec{p} d\vec{r}} \quad (5.5).$$

Lorsque l'on désire effectuer l'échantillonnage dans l'état A avec un potentiel différent que le potentiel original, on peut écrire :

$$H_{A,r}(\vec{p},\vec{r}) = H_A(\vec{p},\vec{r}) - \Delta H(\vec{p},\vec{r}) \quad (5.6)$$

où $H_{A,r}$ est le potentiel de référence (semi-empirique), H_A est le potentiel original qui nous intéresse (fonctionnelle de la densité dans notre cas) et ΔH est la différence entre le potentiel original et le potentiel de référence. Alors, si on introduit la définition du potentiel de référence (équation 5.5) dans le rapport Z_B/Z_A , l'équation 5.4 devient :

$$\frac{Z_B}{Z_A} = \frac{\iint e^{-\beta H_B(\vec{p},\vec{r})} e^{+\beta H_{A,r}(\vec{p},\vec{r})} e^{-\beta H_{A,r}(\vec{p},\vec{r})} d\vec{p} d\vec{r}}{\iint e^{-\beta H_{A,r}(\vec{p},\vec{r})} e^{-\beta \Delta H(\vec{p},\vec{r})} d\vec{p} d\vec{r}} \quad (5.7).$$

Sachant que la densité de probabilité $P_{A,r}(\mathbf{p},\mathbf{r})$ de trouver une géométrie dans le système avec le potentiel de référence peu s'exprimer :

$$P_{A,r}(\vec{p},\vec{r}) = \frac{e^{-\beta H_{A,r}(\vec{p},\vec{r})}}{Z_{A,r}(\vec{p},\vec{r})} \quad (5.8)$$

Il est alors possible si on divise le numérateur et le dénumérateur par $Z_{A,r}$, d'obtenir :

$$\frac{Z_B}{Z_A} = \frac{\iint e^{-\beta H_B(\vec{p},\vec{r})} e^{+\beta H_{A,r}(\vec{p},\vec{r})} e^{-\beta H_{A,r}(\vec{p},\vec{r})} d\vec{p}d\vec{r} / Z_{A,r}}{\iint e^{-\beta H_{A,r}(\vec{p},\vec{r})} e^{-\beta \Delta H(\vec{p},\vec{r})} d\vec{p}d\vec{r} / Z_{A,r}} \quad (5.9)$$

$$\frac{Z_B}{Z_A} = \frac{\iint e^{-\beta H_B(\vec{p},\vec{r})} e^{+\beta H_{A,r}(\vec{p},\vec{r})} P_{A,r}(\vec{p},\vec{r}) d\vec{p}d\vec{r}}{\iint e^{-\beta \Delta H(\vec{p},\vec{r})} P_{A,r}(\vec{p},\vec{r}) d\vec{p}d\vec{r}} \quad (5.10)$$

$$\frac{Z_B}{Z_A} = \frac{\iint e^{-\beta(H_B - H_{A,r} + \Delta H)} P_{A,r}(\vec{p},\vec{r}) d\vec{p}d\vec{r}}{\iint e^{-\beta \Delta H(\vec{p},\vec{r})} P_{A,r}(\vec{p},\vec{r}) d\vec{p}d\vec{r}} \quad (5.11)$$

$$\frac{Z_B}{Z_A} = \frac{\langle e^{-\beta(H_B - H_A + \Delta H)} \rangle_{A,r}}{\langle e^{-\beta \Delta H} \rangle_{A,r}} \quad (5.12)$$

Ainsi, la variation d'énergie libre d'Helmholtz est :

$$\Delta F = -k_B T \ln \frac{\langle e^{-\beta(H_B - H_A + \Delta H)} \rangle_{A,r}}{\langle e^{-\beta \Delta H} \rangle_{A,r}} \quad (5.13)$$

où H_A et H_B est le potentiel que l'on désire (fonctionnel de la densité) dans l'état A et B.

5.3 Protocole pour les simulations

Dans cette section, nous avons utilisé la version de Roar interfacée avec DeFT comme décrit au chapitre 2. Toutes les simulations de dynamique moléculaire ont été exécutées dans l'ensemble canonique NVT à une température de 300 K. Le contrôle de la température est effectué avec la méthode de Nosé-Hoover. Le temps d'intégration est de 1 fs et nous avons utilisé 2000 itérations pour équilibrer le système et 2000 itérations pour l'échantillonnage. Pour la réaction en solution, nous avons utilisé une goutte d'eau contenant 202 molécules qui sont traitée par la mécanique moléculaire en utilisant le modèle TIP3P⁹⁴. Le calcul de l'énergie a été

exécuté en utilisant la méthode de perturbation thermodynamique avec un échantillonnage non-Boltzmann. La procédure utilisée est la suivante :

- I. Détermination du meilleur chemin réactionnel pour la réaction étudiée. Chaque géométrie obtenue devient ainsi la géométrie de départ pour les différentes fenêtres dans le calcul de l'énergie libre.
- II. Calcul de l'énergie libre pour une fenêtre en utilisant le potentiel de référence. Nous avons utilisé comme potentiel de référence la méthode semi-empirique AM1.
- III. Dans l'ensemble de référence obtenu en II, on retire chaque 20^{ième} géométrie. Doren et al.⁹³ ont montré, lors de calculs de la fonction $g(r)$ à partir d'un potentiel de référence classique, qu'il est nécessaire d'avoir un ensemble contenant au moins 200 géométries quantiques pour obtenir de bon résultat. Donc, pour un ensemble contenant 2000 géométries, on va retirer 200 géométries.
- IV. Pour ces géométries, nous recalculons l'énergie potentielle de la géométrie de référence avec un potentiel de plus haut niveau (fonctionnelle de la densité dans notre cas) et nous évaluons également les énergies potentielles pour les géométries perturbées ($\lambda \pm \Delta\lambda$) avec le potentiel de plus haut niveau.

- V. On calcule la variation d'énergie libre pour cette fenêtre avec la formule 5.12 développée dans la section précédente.

5.4 Test pour diminuer l'erreur dans le calcul de l'énergie libre

Nous avons évalué trois tests permettant d'estimer l'erreur dans le calcul de l'énergie libre⁹⁵. Nous avons utilisé la réaction d'ouverture du cyclobutène pour ces tests. Pour effectuer l'équilibration et l'échantillonnage, nous avons utilisé le potentiel AM1.

1. On regarde la différence entre les valeurs « vers l'avant » et « vers l'arrière ».
Pour nos simulations, cette différence est généralement d'environ 0.3 kcal/mol en utilisant 2000 itérations pour l'équilibration et l'échantillonnage à chaque fenêtre.
2. Pour s'assurer que l'utilisation de 2000 itérations, pour l'équilibration et l'acquisition des données, nous permet de s'assurer que la convergence a été atteinte, nous avons exécuté le test suivant :
 - On exécute une série de simulations qui diffèrent dans les quantités respectives de prélèvements. Les résultats pour ces séries de simulations sont présentés au Tableau XIII. Ces résultats nous montrent que, si nous avons obtenu la convergence lorsque nous utilisons 1000 itérations pour

l'équilibration et l'échantillonnage à chaque fenêtre, la valeur obtenue avec 2000 itérations devrait être très près de la première simulation. Le Tableau XIII, montre que 1000 itérations ne sont pas assez long pour obtenir la convergence mais que 2000 itérations pour l'équilibration et l'échantillonnage donnent des résultats acceptables. Cette méthode fournit une bonne assurance que la convergence est atteinte mais n'est pas une garantie⁸¹.

3. Pour regarder l'importance de l'hystérésis dans nos profils d'énergie libre lorsqu'ils sont produit avec le potentiel de référence (AM1), nous avons fait deux simulations. Premièrement, nous avons exécuté une simulation où λ passe de 0 $\rightarrow$ 1 et une seconde simulation où le λ passe de 1 $\rightarrow$ 0. Les résultats pour ces simulations sont présentés à la Figure 21. Ces deux simulations devraient montrer la même valeur pour l'énergie libre mais on obtient plutôt une différence d'environ 5 kcal/mol. Cette valeur correspond à l'hystérésis que nous avons obtenu pour cette réaction en utilisant la méthode semi-empirique AM1. Cette hystérésis provient du fait que l'on ne passe pas directement dans la région de l'état de transition. La Figure 21 indique également que la déviation arrive seulement dans la deuxième partie de la courbe et commence dans la région de l'état de transition. Cette hystérésis nous indique que nous devrions diminuer le $\Delta\lambda$ pour diminuer cette hystérésis.

Tableau XIII. Test pour le nombre d'itération pour équilibrer et échantillonner le système.

ÉQUILIBRER	ÉCHANTILLONNER	VERS L'AVANT	VERS L'ARRIÈRE
1000	1000	-2.7178	2.7478
1000	2000	-2.7292	2.7637
1000	3000	-2.7552	2.7838
1000	4000	-2.7678	2.7988
1000	5000	-2.7734	2.7947
1000	6000	-2.7716	2.8009
1000	7000	-2.7689	2.8086
2000	2000	-2.7701	2.8225
2000	3000	-2.7715	2.8253
2000	4000	-2.7730	2.8087
2000	5000	-2.7832	2.8133
2000	6000	-2.7867	2.8181
2000	7000	-2.7933	2.8238
2000	8000	-2.7949	2.8304

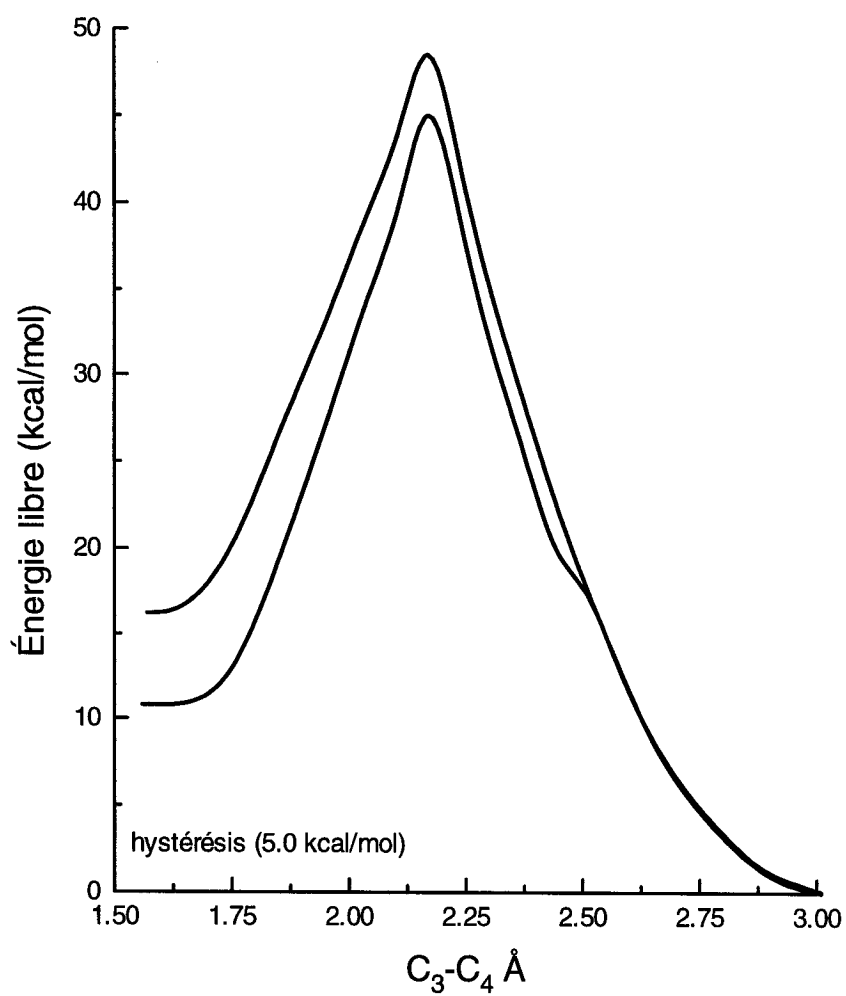


Figure 21. Simulation où λ passe de $0 \rightarrow 1$ et de $1 \rightarrow 0$ pour la réaction d'ouverture du cyclobutène en utilisant le potentiel AM1.

5.5 Choix du potentiel de référence

Il a déjà été démontré dans la littérature que l'utilisation d'un potentiel polarisé^{74,75,76} (bias potential) peut accélérer la convergence dans le calcul des valeurs thermodynamiques moyennes dans une simulation de dynamique moléculaire ou une simulation Monté Carlo d'une façon importante. Par contre, l'efficacité de cette méthode dépendra du choix de ce potentiel.

Pour connaître les critères de sélection du potentiel de référence, nous avons effectué les tests suivants :

- I. Avec un potentiel de référence qui possède une énergie d'activation inférieure au potentiel que l'on désire reproduire;
- II. Avec un potentiel de référence qui possède une énergie d'activation supérieure au potentiel que l'on désire reproduire;
- III. Nous avons évalué si le signe de la différence d'énergie pour une fenêtre donnée influence le résultat (le signe est différent entre le potentiel de référence et le potentiel que l'on désire reproduire).

Pour effectuer ces différents tests, nous avons utilisé la réaction entre le NH_3 et le H_2CO . Nous avons d'abord trouvé le IRC pour cette réaction en utilisant un calcul de niveau BP86/6-31G*. Par la suite, nous avons fait un calcul d'énergie sur ces géométries en utilisant les méthodes semi-empiriques AM1, PM3 et MNDO. La variation d'énergie le long du chemin réactionnel est montrée à la Figure 22 pour les différents niveaux de calculs. Cette figure nous indique qu'il est possible pour les tests I et II d'utiliser les potentiels PM3 et MNDO, car il existe une grande différence dans la barrière d'énergie d'activation entre ces deux méthodes. Alors, pour le test I, nous avons utilisé comme potentiel de référence la méthode PM3 pour reproduire les résultats de la méthode MNDO, tandis que pour le test II c'est la méthode MNDO qui a servi de potentiel de référence pour reproduire la courbe PM3. Les résultats sont présentés à la Figure 23. Cette figure nous montre que la hauteur de la barrière d'énergie d'activation du potentiel de référence n'influence pas la capacité de ce potentiel de bien reproduire un autre potentiel. La Figure 23 nous montre une légère déviation à la fin du profil de l'énergie libre qui est due à l'hystérésis. Cette hystérésis sera plus ou moins importante selon que l'on passe ou non correctement dans la région de l'état de transition. Lorsqu'on regarde la déviation pour chaque fenêtre entre la valeur obtenue pour la simulation PM3//MNDO et PM3, on obtient une déviation moyenne ≈ 0.23 kcal/mol. Par contre, le profil s'obtient par la sommation de chaque fenêtre. Ainsi, on obtient une déviation finale d'environ 3.0 kcal/mol pour l'énergie libre.

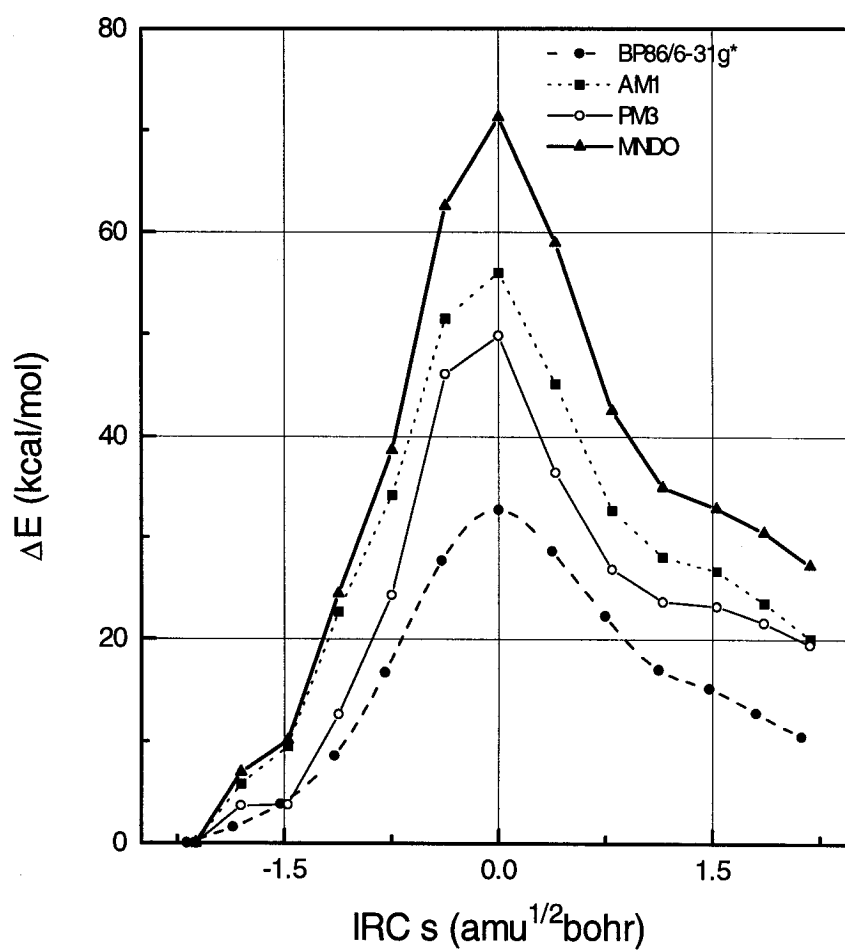


Figure 22. Variation de l'énergie le long du chemin réactionnel pour la réaction entre le NH_3 et H_2CO .

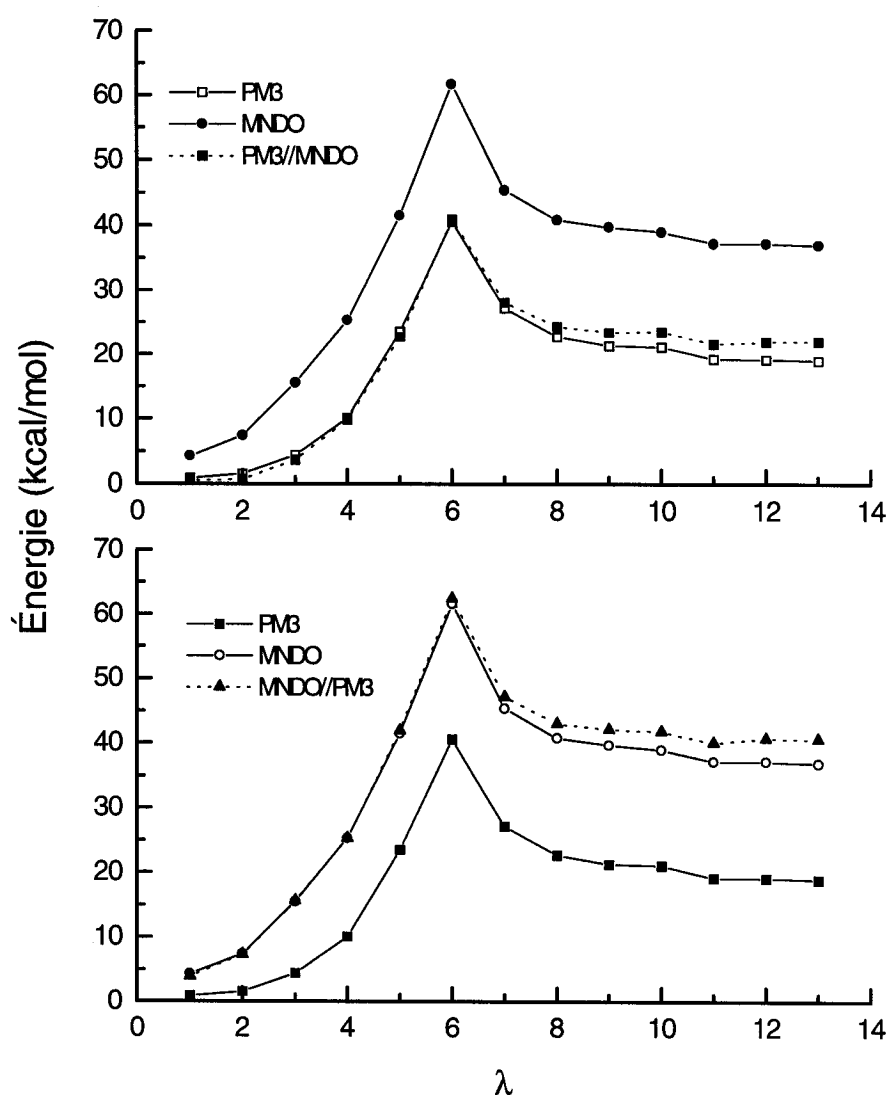


Figure 23. Résultats pour les tests I et II.

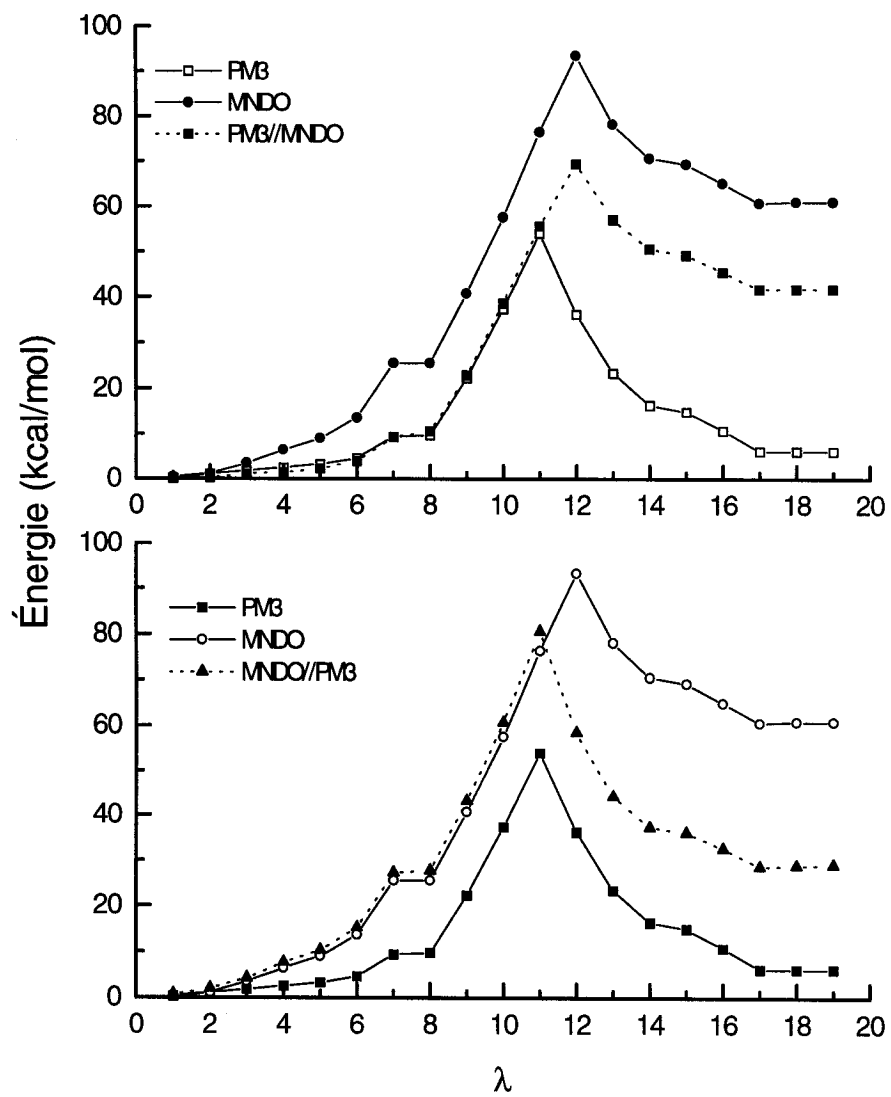


Figure 24. Résultats pour le test III.

Pour le test III, nous avons fait varier la largeur de $\Delta\lambda$ pour obtenir une fenêtre où le signe du potentiel de référence est différent du signe du potentiel que nous désirons reproduire. La Figure 24 démontre qu'il est très important pour bien représenter un potentiel que le signe de la variation d'énergie libre pour chaque fenêtre entre le potentiel de référence et celui que l'on désire reproduire soit identique.

Nous avons exécuté ces tests sans utiliser la fonctionnelle de la densité car des calculs complets de profil d'énergie libre en utilisant la fonctionnelle de la densité demande un temps de calculs très important. Nous sommes conscients qu'il peut avoir une différence entre utilisé un potentiel semi-empirique pour reproduire un calcul de niveau de la fonctionnelle de la densité que pour deux potentiels semi-empiriques. Néanmoins, nous considérons que les résultats obtenus avec les potentiels semi-empiriques sont directement transférables à un potentiel de plus haut niveau.

Donc, pour obtenir un bon potentiel de référence, nous devons suivre ces trois règles :

1. Le signe de la variation d'énergie pour chaque fenêtre entre le potentiel de référence et le potentiel que l'on désire reproduire doit être identique. Pour ce

critère, nous n'avons aucun contrôle mais il est toujours possible de reparamétriser les paramètres AM1 ou PM3 pour mieux satisfaire ce critère.

2. Sélectionner les coordonnées de la réaction qui représentent le mieux la réaction afin de diminuer au maximum l'hystérésis dans le profil d'énergie libre.
3. Le calcul de l'énergie et des forces correspondantes pour le potentiel de référence doit être informatiquement très peu coûteux comparativement au potentiel que l'on désire reproduire. Les méthodes semi-empiriques sont énormément moins coûteuses en temps de calculs que la méthode de la fonctionnelle de la densité.

5.6 Exemples d'utilisation de la méthode de perturbation thermodynamique non-Boltzman

Dans cette section, nous avons également utilisé deux réactions électrocycliques déjà étudiées dans la section précédente pour pouvoir effectuer une bonne comparaison des méthodes. Nous avons regardé l'ouverture du cyclobutène, l'ouverture de l'oxetène en utilisant la méthode de perturbation thermodynamique. Nous avons également utilisé cette méthode pour la réaction du transfert de proton dans glycine dans un milieu aqueux en utilisant notre version QM/MM de Roar/DeFT.

5.6.1 Ouverture du cycle pour le cyclobutène

Comme indiqué dans la section 5.3, nous avons utilisé les géométries du IRC comme structure de départ pour chaque fenêtre. Pour exécuter la simulation, nous avons utilisé 21 trajectoires de dynamique moléculaire. Les trajectoires ont été exécutées à 300 K pour une durée de 4 ps, 2 ps pour le temps d'équilibration et 2 ps pour le temps d'échantillonnage, avec un temps d'intégration de 1 fs. Cette procédure a été utilisée avec le potentiel de référence AM1 et pour la fonctionnelle de la densité. Les résultats pour la réaction complète sont présentés à la Figure 25. Cette figure nous montre les trois profils d'énergies. On remarque que le profil obtenu avec la méthode mixte se compare très bien avec le profil d'énergie libre calculée en utilisant seulement la fonctionnelle de la densité. Pour la méthode mixte, nous avons utilisé comme potentiel de référence un potentiel semi-empirique (AM1) et son profil d'énergie libre est également présenté sur cette figure. Avec la méthode mixte, on obtient une énergie libre d'activation de 36.9 kcal/mol et une énergie libre de réaction de -3.0 kcal/mol. Les résultats expérimentaux pour cette réaction sont de 34.5 kcal/mol pour la barrière d'activation et l'énergie de la réaction est de -10.9 kcal/mol. Nous obtenons ainsi des résultats très proches des résultats expérimentaux tous en augmentant la vitesse de calcul par un facteur 10 entre les calculs complètement quantiques et les calculs mixtes.

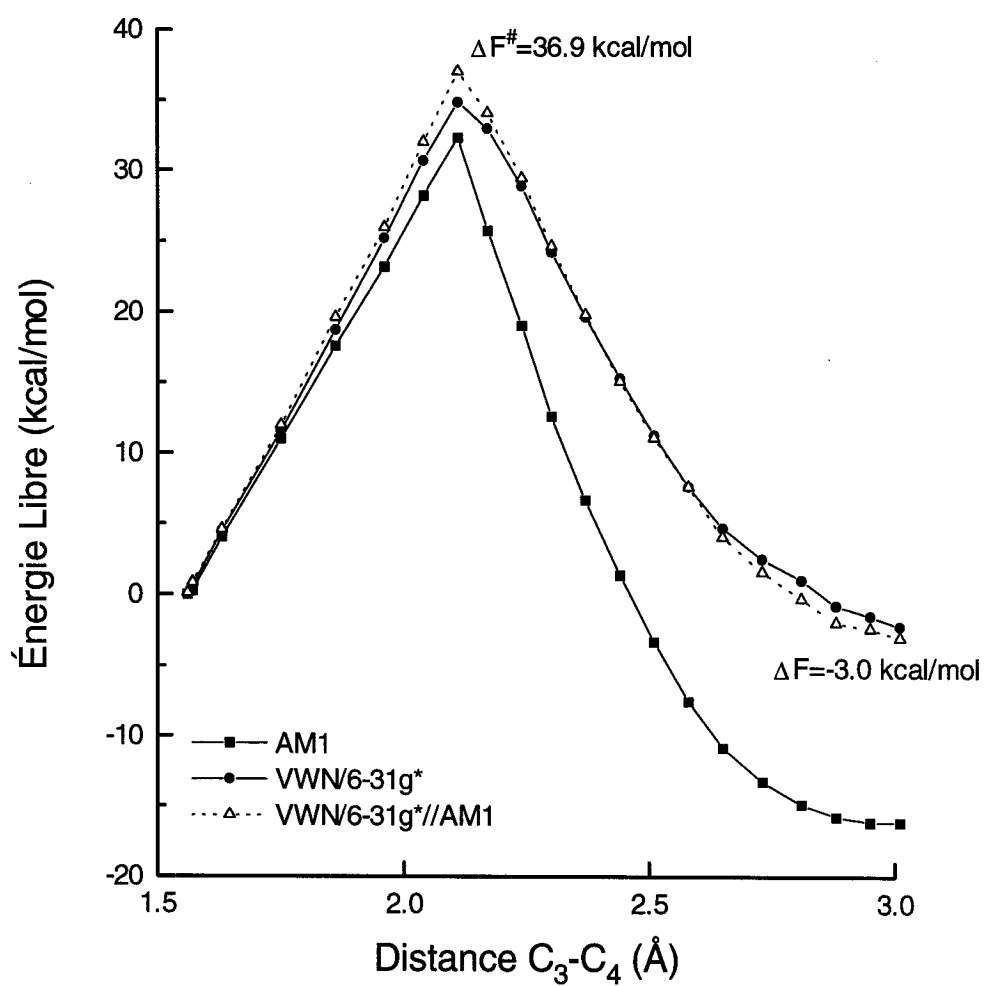


Figure 25. Profil d'énergie libre pour la réaction d'ouverture du cyclobutène en utilisant la méthode de perturbation thermodynamique.

5.6.2 Ouverture du cycle pour l'oxetène

Pour cette réaction, nous avons utilisé 22 trajectoires de dynamique moléculaire. Les trajectoires ont été exécutées à 300 K pour une durée de 4 ps, 2 ps pour le temps d'équilibration et 2 ps pour le temps d'échantillonnage, avec un temps d'intégration de 1 fs. Les résultats pour l'ouverture de l'oxetène sont présentés à la Figure 26. Cette figure montre qu'il y a un léger décalage entre le profil mixte et le profil DFT. Pour cette réaction, on obtient une énergie libre d'activation de 28.4 kcal/mol. La valeur expérimentale est de 24.1 kcal/mol. Comme nous l'avons montré à la section 5.4, il existe une hystérésis important dans le profil de référence lorsque nous utilisons la méthode de perturbation thermodynamique. Cette hystérésis nous occasionne de la difficulté à bien représenter la région de l'état de transition. La Figure 26, montre que le début du décalage débute dans la région de l'état de transition. Ce décalage entraîne une déviation pour le reste du profil par le fait que l'énergie libre est la somme d'une série de fenêtres. Le Tableau XIV présente le pourcentage d'erreur absolue entre la fenêtre du calcul mixte (VWN/6-31G*//AM1) et la fenêtre complètement quantique (VWN/6-31G*). On remarque rapidement que la fenêtre où l'erreur est très importante correspond à la région de l'état de transition.

Tableau XIV. Erreur à chaque fenêtre entre les résultats de la méthode mixte et la méthode complètement DFT.

Fenêtre	% d'erreur	Fenêtre	% d'erreur	fenêtre	% d'erreur	fenêtre	% d'erreur
1	2	7	1	13	7	19	2
2	7	8	5	14	3	20	22
3	13	9	> 100	15	14	21	31
4	5	10	4	16	0	22	28
5	2	11	10	17	13	Le % erreur moyen sans la fenêtre 9 est de 10 %	
6	5	12	13	18	21		

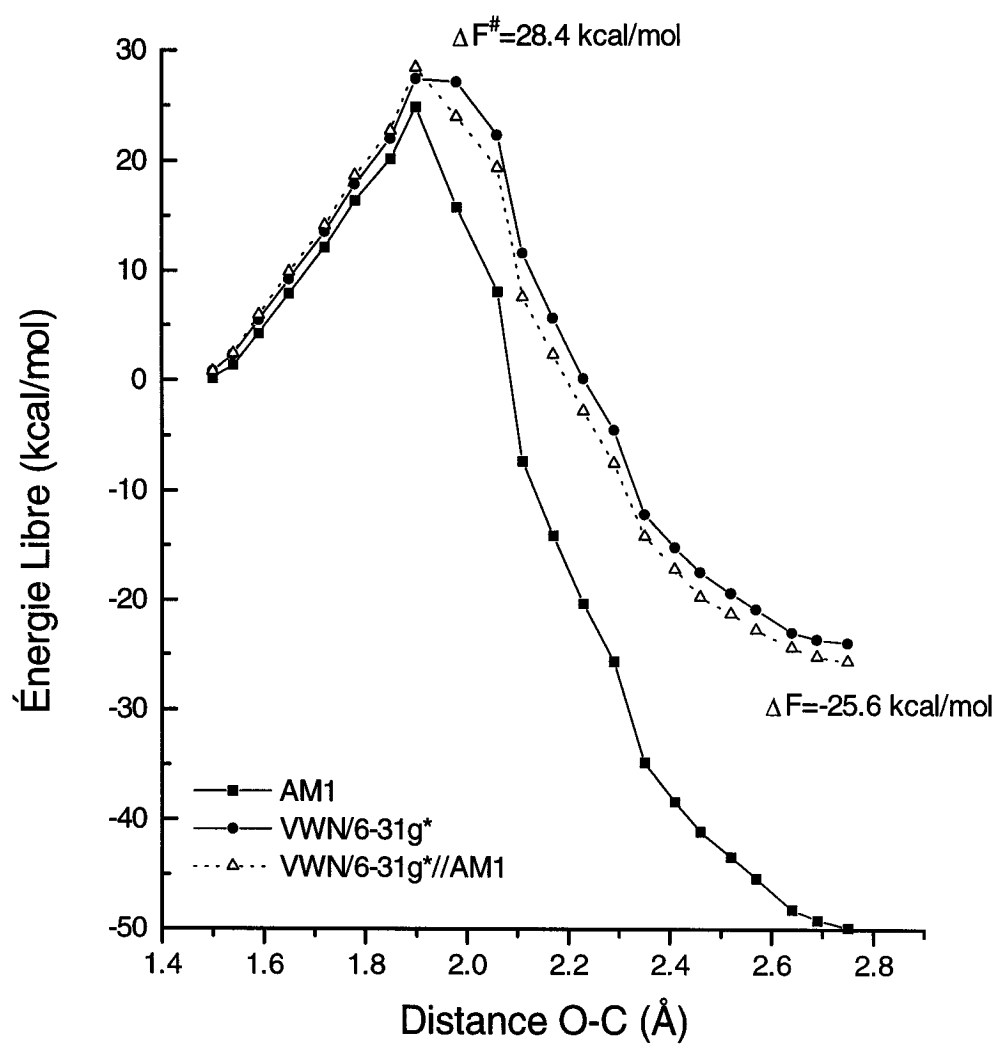


Figure 26. Profil d'énergie libre pour la réaction d'ouverture de l'oxetène en utilisant la méthode de perturbation thermodynamique.

Si nous regardons l'erreur moyenne pour chaque fenêtre à l'exception de la fenêtre 9, nous trouvons une erreur moyenne de 10% par rapport au calcul complètement quantique. Pour ces calculs d'énergie libre pour la mutation du butane en propanol, Brüger⁶⁴ a trouvé une erreur entre 3% à 10% entre les calculs utilisant la méthode classique pour le calcul de l'énergie libre et la méthode où il utilise l'échantillonnage non-Boltzmann. Pour ce système, Brüger a utilisé deux potentiels classiques pour faire les calculs. Malgré ce faible décalage, la méthode de perturbation thermodynamique non-Boltzmann permet de très bien représenter le calcul DFT en économisant du temps de calculs.

5.6.3 Transfert de proton dans la Glycine

Dans de nombreux tissus vivants, les protéines sont les principaux constituants. Ainsi, les protéines jouent un rôle primordial dans divers processus vitaux. Les protéines sont constituées par une série d'acides aminés. L'acide aminé le plus simple qui peut servir de prototype pour l'étude de plus grands systèmes est la glycine. Les transferts de proton sont impliqués dans un grand nombre de mécanismes biologiques. Ainsi, le transfert de proton dans la glycine a été énormément étudié par les théoriciens^{96,97,98,99}.

Les résultats expérimentaux ont montré que la barrière d'activation pour passer de la forme de l'ion dipolaire à la forme neutre est de 14.36 kcal/mol et

l'énergie de la réaction est 7.3 kcal/mol. La majorité des études théoriques pour l'étude du transfert de proton dans la glycine ont été exécutées avec un modèle de solvation. Ruiz-Lopez et al.⁹⁵ ont trouvé, en utilisant B3LYP/6-31G* + continuum, une énergie d'activation de 5.42 kcal/mol et une énergie de réaction de 4.77 kcal/mol pour le transfert direct du proton. Ces résultats indiquent qu'il y a seulement une barrière de 0.65 kcal/mol pour passer de la forme neutre à la forme de l'ion dipolaire. Cette barrière passe à 11.0 kcal/mol si on utilise la méthode Hartree-Fock et 2.4 kcal/mol avec MP2^{100,101}. Un calcul de dynamique moléculaire QM/MM a également montré que la conversion de la forme neutre à la forme de l'ion dipolaire se passe très rapidement⁹⁴.

Nous avons utilisé notre méthode mixte pour évaluer l'énergie libre pour le transfert du proton intramoléculaire dans la glycine, placée dans une bulle d'eau, par dynamique moléculaire QM/MM. Pour faire ceci, il faut que les fonctions amine et acide carboxylique soient correctement orientées pour faire un transfert direct du proton. La forme de l'ion dipolaire et neutre est présentée à la Figure 27. La forme neutre présentée sur cette figure est la forme la plus stable en solution par contre cette structure n'est pas la plus stable dans la phase gazeuse⁹⁷.

Pour exécuter ce calcul, nous avons placé la molécule de glycine dans une bulle contenant 212 molécules d'eau. Avant de commencer nos fenêtres pour le calcul de l'énergie libre, nous avons optimisé les molécules d'eau autour de la

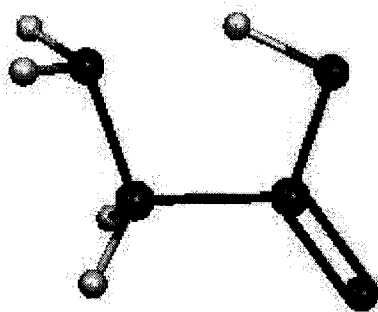
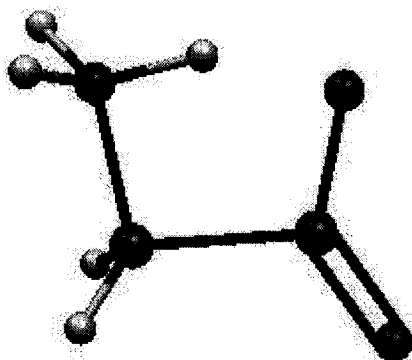


Figure 27 : Structure pour la forme neutre et la forme de l'ion dipolaire

glycine en gardant celle-ci fixe avec l'option Belly du module Roar. Avec cette option, les déplacements et les vitesses sont remis à zéro à chaque itération. Les géométries de départ ont été obtenues en optimisant la glycine en gardant la distance N-H fixe. Ces calculs ont été exécutés en utilisant la fonctionnelle BP86 et la base 6-31G* avec le programme DeFT. Comme pour les autres exemples précédents, nous avons utilisé le potentiel AM1 comme potentiel de référence.

Pour ces calculs, nous avons utilisé 16 trajectoires de dynamique moléculaire. Les trajectoires ont été exécutées à 300 K pour une durée de 4 ps, 2 ps pour le temps d'équilibration et 2 ps pour le temps d'échantillonnage, avec un temps d'intégration de 1 fs. Le profil d'énergie libre du transfert de proton en solution de la glycine est présenté à la Figure 28. Cette figure montre que nous obtenons une barrière d'activation de 15.5 kcal/mol et une énergie libre pour la réaction de 13.8 kcal/mol. Ainsi, nous avons une barrière de 1.7 kcal/mol entre la forme neutre à la forme de l'ion dipolaire. Nos résultats sont donc compatibles avec l'autre résultat QM/MM⁹⁴ qui indiquait également que cette barrière était très faible. Par contre, nous sommes incapables de reproduire les résultats expérimentaux en utilisant la fonctionnelle BP86. Pour comparer les résultats obtenus en utilisant la méthode QM/MM avec un modèle de solvation, nous avons utilisé nos géométries de départ pour le calcul de l'énergie libre et nous avons calculé le profil d'énergie avec un modèle de solvation.

Par la suite, nous avons rajouté trois molécules d'eau explicite et avons recalculé le profil d'énergie en utilisant également le modèle de solvation. Ces deux profils sont présentés à la Figure 29. Cette figure montre que lorsque nous utilisons un modèle de solvation sans molécule d'eau explicite, nous obtenons une barrière d'activation de 4.9 kcal/mol et une énergie de réaction de 2.25 kcal/mol. Par contre, avec trois molécules d'eau explicites, la barrière d'activation est de 14.7 kcal/mol et l'énergie de la réaction est de 10.4 kcal/mol. Ces résultats nous indiquent qu'il semble important d'avoir des molécules d'eau explicites pour bien représenter cette réaction. Une autre étude⁹⁶ sur le transfert du proton dans la glycine a également montré que lorsque nous ajoutons des molécules d'eau explicites avec un modèle de solvation que la barrière d'activation augmente par rapport aux résultats avec un modèle de solvation simple. Dans cette étude, la barrière passe de 0.4 kcal/mol à 7.6 kcal/mol lorsqu'il ajoute trois molécules d'eau explicites en plus du modèle de solvation. Les auteurs de cette étude ont utilisé la fonctionnelle B3LYP et la base 6-31G*.

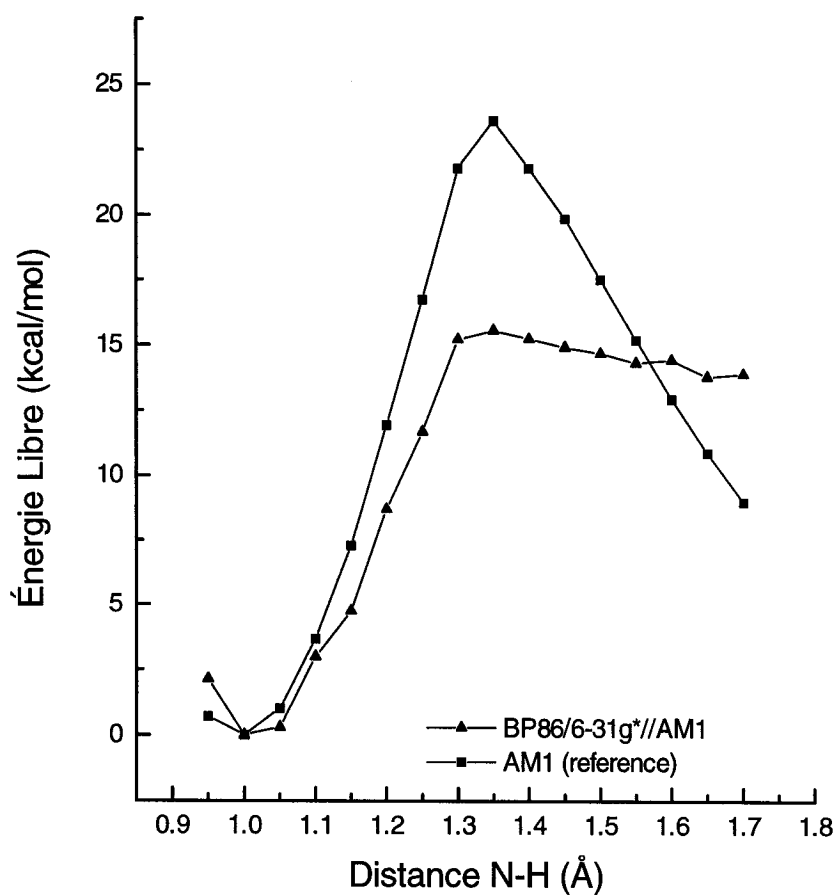


Figure 28. Profil d'énergie libre pour la réaction du transfert de proton dans la glycine en utilisant la méthode de perturbation thermodynamique.

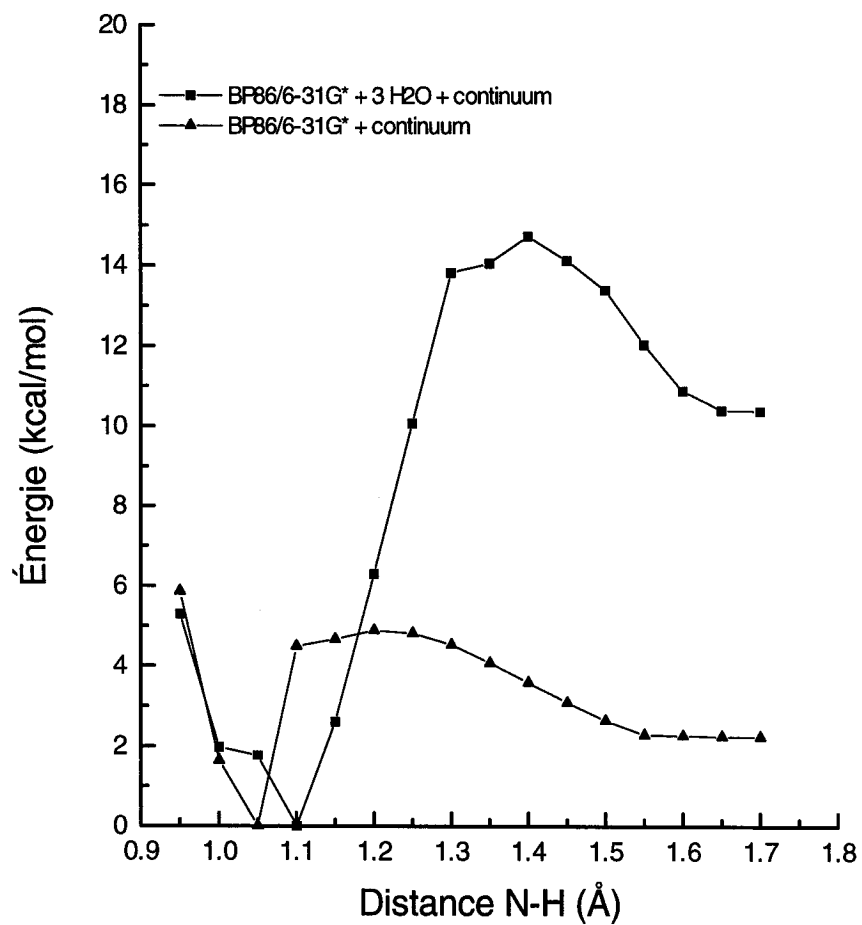


Figure 29 . Profil d'énergie pour la réaction de transfert du proton dans la glycine en utilisant un modèle de solvation.

5.6.4 Comparaison entre les méthodes pour le calcul de l'énergie libre

Dans cette section, nous regardons les résultats obtenus en utilisant la méthode de perturbation thermodynamique non-Boltzmann. Nous allons les comparer aux résultats obtenus en utilisant la méthode standard de perturbation thermodynamique avec un potentiel quantique puisque le but principal de l'utilisation de cette technique est de pouvoir reproduire correctement les profils d'énergie libre quantique tout en diminuant le temps de calculs pour chaque fenêtre.

Les résultats pour l'énergie libre d'activation et l'énergie libre de la réaction sont présentés au Tableau XV. La différence en temps CPU est également montrée dans ce tableau. Le Tableau XV démontre que l'utilisation de la méthode de perturbation thermodynamique non-Boltzmann peut être utilisée pour obtenir des énergies libres qui se comparent très bien à des résultats où un calcul de niveau supérieur a été utilisé. En plus de bien reproduire les profils quantiques, il est possible d'accélérer le temps de calcul pour chaque fenêtre par un facteur d'environ 12. Ces résultats nous laissent entrevoir qu'il serait possible d'obtenir un profil d'énergie libre équivalent à un calcul de niveau DFT pour un système plus gros avec un temps de calcul raisonnable.

Tableau XV. Comparaison entre la méthode de perturbation thermodynamique complète quantique et mixte (semi-empirique / quantique)

Énergie libre de la barrière d'activation (kcal/mol)			
Molécules	(DFT)	(AM1/DFT)	(AM1)
Cyclobutène	34.8	36.9	32.3
Oxetène	27.3	28.4	24.9
Énergie libre de la réaction (kcal/mol)			
Cyclobutène	-2.2	-3.0	-16.9
Oxetène	-25.9	-25.6	-49.8
Temps CPU (heures)			
Cyclobutène	228 ^a	19 ^a	0.08 ^a
Oxetène	205 ^a	16 ^a	0.06 ^a

a) par fenêtre

Par contre, en utilisant seulement un lien pour faire les calculs d'énergie libre, on cause généralement une hystérésis important car il est très rare qu'une réaction soit gouvernée que par la rupture ou la formation d'un lien sans que d'autres paramètres interviennent (angle, torsion). Il serait donc intéressant d'utiliser la méthode mixte, mais en obligeant la trajectoire à suivre le chemin réactionnel comme dans le chapitre 4. Présentement, notre version de Roar/DeFT ne permet pas ce genre de calculs.

Il serait possible d'utiliser des fenêtres aux lieux de la méthode *slow growth* et alors le calcul de l'énergie libre en utilisant la méthode d'échantillonnage non-Boltzmann serait donnée par l'équation suivante :

$$\Delta F = \sum_{i=1}^n \frac{\left\langle \frac{\partial E(x, \lambda)}{\partial \lambda} \exp[-\beta \Delta E(x, \lambda)] \right\rangle_{\text{référence}}}{\left\langle \exp[-\beta \Delta E(x, \lambda)] \right\rangle_{\text{référence}}} (\Delta \lambda) \quad (5.14)$$

Cette façon augmenterait seulement la qualité des profils d'énergie libre pour des réactions simples. Cette méthode a besoin du IRC de la réaction qui est malheureusement très difficile à obtenir pour des réactions complexes.

Pour les réactions électrocycliques, nous avons obtenu d'excellents résultats car les minimums et les maximums le long du lien perturbé étaient très semblables

pour le potentiel de référence et le potentiel de plus haut niveau. Par contre, si l'on désire étudier la réaction SN2 ($\text{Cl-CH}_3\text{Cl}$) les résultats ne seraient pas probants car les minimums et les maximums sont trop différents comme le montre la Figure 30. Ainsi, pour l'étude de certaines réactions, il serait nécessaire de reparamétriser le potentiel semi-empirique que l'on désire utiliser, comme potentiel de référence, pour obtenir un excellent résultat. Par contre, il serait nécessaire de reparamétriser pour obtenir seulement un bon accord pour la position des minimums et des maximums. Il n'est pas nécessaire comme on l'a démontré dans la section 5.5 d'avoir un accord pour la barrière de l'énergie d'activation.

L'utilisation d'un potentiel de référence de niveau semi-empirique permet d'obtenir d'excellent résultat. Par contre, il serait intéressant de développer dans le future une procédure qui nous permettrait de bien représenter le potentiel de toutes les réactions qui nous intéressent. Cette procédure devrait pouvoir faire varier les paramètres AM1 ou PM3 pour s'assurer que les minimums et les maximums du potentiel de référence seraient à la même position que pour le potentiel de haut niveau que l'on désire reproduire.

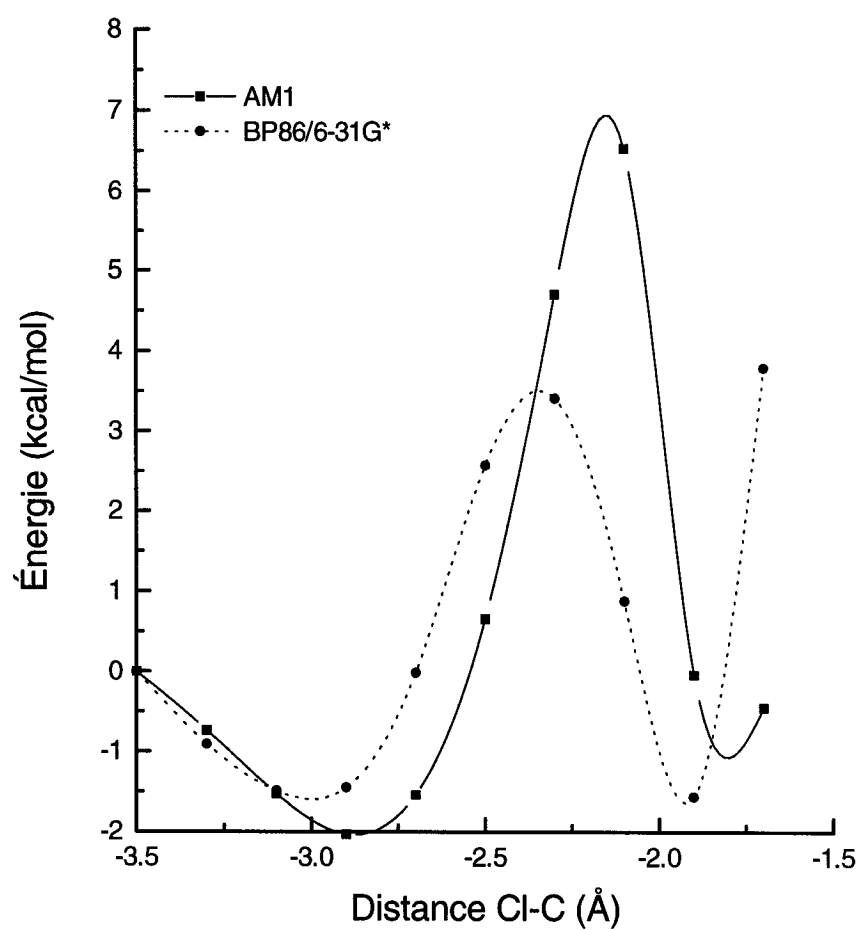


Figure 30. Profil d'énergie pour la réaction SN2 (Cl CH₃Cl).

5.7 Conclusion

La méthode de perturbation thermodynamique non-Boltzmann permet d'avoir un profil d'énergie libre de calibre DFT tout en diminuant le temps de calcul par un facteur 12. Ce profil représentera correctement le profil quantique si on utilise un potentiel de référence qui représente bien le potentiel DFT tout en diminuant le temps nécessaire pour le calcul de l'énergie et de son gradient. Les exemples ont montré que la région de l'état de transition était très difficile à bien représenter avec la méthode hybride. Pour améliorer cette situation, on pourrait augmenter la statistique que l'on utilise pour le calcul de la variation d'énergie libre à chaque fenêtre. Pour les courbes présentées dans ce chapitre, nous avons utilisé chaque 20^{ième} géométrie mais ce nombre pourrait diminuer dans la zone de l'état de transition. Il est important de noter que lorsque nous utilisons chaque 40^{ième} géométrie, nous obtenons des fluctuations importantes dans les résultats comme il a été démontré par Doren et al⁹³.

Conclusion

Dans cette thèse, nous avons montré comment nous avons incorporé le logiciel DeFT, utilisant la théorie de fonctionnelle de la densité, avec le module Roar de la suite Amber. Pour faire la connexion entre la partie quantique et la partie classique, nous avons utilisé la méthode de l'atome connecté en distribuant le gradient de cet atome sur le premier atome des sections QM et MM qui sert de séparation entre les deux parties. La validation a été faite en calculant l'affinité de proton sur une série de molécules et en regardant si on pouvait reproduire correctement des résultats complètement quantiques pour différentes conformations de la dipeptide alanine à la quelle on a ajouté une hélice alpha contenant 20 glycines. Dans cette section, nous avons démontré que nous avons une déviation d'environ 0.02 Å lorsque nous optimisons un lien C-C et qu'il est très important d'inclure au moins un groupement méthyle après le groupement fonctionnelle pour obtenir des résultats qui se comparent à des résultats qui seraient obtenus pour un calcul complètement quantique.

Dans la deuxième partie, nous avons montré comment nous avons inclus dans DeFT deux différentes façons de calculer l'énergie libre d'une réaction. Dans un premier temps, nous avons inclus la méthode de calcul d'énergie libre le long du

chemin réactionnel développée dans un premier temps par le groupe de Ziegler. Nous avons montré qu'il était possible de trouver des profils d'énergie libre très lisse avec cette méthode pour différents types de réaction. Par contre, cette méthode est très coûteuse en temps de calcul. Nous avons également optimisé une méthode hybride AM1/DFT pour le calcul de l'énergie libre. Nous avons démontré qu'il est possible de reproduire assez fidèlement un profil d'énergie libre de niveau de calcul DFT en diminuant considérablement le temps de calcul. Au Tableau XV, nous avons démontré que pour les réactions d'ouverture du cycle du cyclobutène et de l'oxytène que l'on pouvait diminuer le temps de calculs par un facteur ~ 12 .

Pour améliorer les interactions électrostatiques entre la partie quantique et la partie classique de notre version QM/MM, il serait intéressant d'introduire une distribution de charge autour de l'atome de la partie classique qui est la plus proche de l'atome connecté pour mieux représenter cette interaction. En utilisant cette méthode, Field⁵¹ a montré qu'on pouvait améliorer considérablement la qualité des calculs d'affinité de proton. Cette amélioration et la méthode hybride pour le calcul de l'énergie libre, pourrait être utilisé avec des systèmes beaucoup plus complexe comme une réaction enzymatique. Par contre, il serait nécessaire de paralléliser l'interface entre Roar et DeFT pour envisager ce genre de calculs. Cette incapacité de pouvoir faire des calculs en parallèle nous a limité dans l'utilisation de notre version QM/MM durant cette thèse.

Pour pouvoir compléter notre travail, il serait nécessaire pour pouvoir publier nos résultats du chapitre 5 de faire certains calculs pour mieux compléter nos résultats. Une analyse complète des erreurs pour toutes les trajectoires de dynamique moléculaire va nous permettre de mieux supporter nos résultats pour la comparaison entre le profile d'énergie libre qui utilise seulement la fonctionnelle de la densité et le profile d'énergie libre où la méthode hybride AM1/DFT est utilisée.

Originalité du travail

- Traitement différent de l'atome connecté par rapport à la version originale de Roar et implémentation de la possibilité d'utiliser la fonctionnelle de la densité pour la partie quantique dans les calculs QM/MM avec Roar.
- Implémentation de la possibilité de contraindre une trajectoire de dynamique moléculaire à suivre le IRC dans un programme DFT utilisant des orbitales gaussiennes.
- Utilisation pour la première fois de la méthode de perturbation thermodynamique non-Boltzmann pour reproduire un profil d'énergie libre de haut niveau avec l'utilisation d'un potentiel semi-empirique.

Conférence

E. Paradis and A. St-Amant, *Introducing the QM/MM functionality within the DeFT Software Package and its Application to Free Energy Calculations*, International Conference on Modeling Chemical Reactivity, France (2003).

Bibliographie

1. W.D. Cornell, P. Cieplak, C.L. Bayly, I.R. Gould, K.M. Merz Jr., D.M.Ferguson, Spellmeyer, T. Fox, J.W. Caldwell, P.A. Kollman, J Am Chem Soc., **117**, 5179 (1995)
2. W.R.P. Scott, P.H. Hunenberger, I.G. Tironi, A.E. Mark, S.R. Billeter, J. Fennen, A.E. Torda, T. Huber, P. Kruger, W.F. vanGunsteren, J Phys Chem A, **103**, 3596 (1999)
3. A.D. Mackell, J. Wiorkiewicz, M. Karplus, J Am Chem Soc, **117**, 11946 (1995)
4. M. Head-Gordon, J. A. Pople and M. J. Frisch, Chem Phys Lett, **153**, 503 (1988)
5. J. Cizek, Adv Chem Phys, **14**, 35 (1969)
6. W. Yang, Phys Rev Lett, **66**, 1438 (1991)
7. S. L. Dixon and K. M. Merz, J Chem Phys, **104**, 6643 (1996)
8. M. W. Wong, M. J. Frisch and K. B. Wiberg, J Am Chem Soc, **113**, 4776 (1991)
9. S. Miertus, E. Scrocco and T. Tomasi, Chem Phys, **55**, 117 (1981)
10. V. Barone, M Cossi and T. Tomasi, J Comp Chem, **19**, 104 (1998)
11. A.Warshel and M.Levitt, J Mol Biol, **103**, 227 (1976)
12. M.J. Field, P.A. Bash and M. Karplus, J Comput Chem, **11**, 700 (1990)
13. D. Bakowies and W. Thiel, J Phys Chem, **100**, 10580 (1996)

-
14. U.C. Sing, P.A. Kollman, *J Comput Chem*, **7**, 718 (1986)
 15. V. Théry, D. Rinaldi, J.L. Rivail, B. Maigret, G.G Ferenczy, *J Comput Chem*, **15**, 269 (1994)
 16. I.Antes, W.J. Thiel, *J Phys Chem A*, **103**, 9290 (1999)
 17. R. W. Zwanzig, *J Chem Phys*, **22**, 1420 (1954)
 18. J.G. Kirkwood, *J Chem Phys*, **3**, 300 (1935)
 19. G.M. Torrie and J.P. Valleau, *J Comput Phys*, **23**, 187 (1977)
 20. J.L. Rivail, *Eléments de Chimie quantique à l'Usage des Chimistes*, InterEditions/CNRS Editions, Paris (1994)
 21. M. Born and J. R. Oppenheimer, *Ann Physik*, **84**, 457 (1927)
 22. J. C. Slater, *Phys Rev*, **34**, 1293 (1929)
 23. D.R. Hartree, *Proc Cambridge Phil Soc*, **24**, 89 (1928)
 24. V. Fock, *Z Physik*, **61**, 126 (1930)
 25. C. C. J. Roothan, *Rev. Mod. Phys.*, **23**, 69 (1951)
 26. A.St-Amant, in *Reviews in Computational Chemistry*, Vol. 7, Chap. 5, K.B. Lipkowitz and D.B. Boyd eds., VCH, New York (1996)
 27. P. Hohenberg and W. Kohn, *Phys Rev B*, **136**, 864 (1964)
 28. W. Kohn and L.J. Sham, *Phys Rev A*, **140**, 1133 (1965)
 29. J. P. Perdew and A. Zunger, *Phys Rev B*, **23**, 5048 (1981)
 30. S.H. Vosko, L. Wilk and M. Nusair, *Can J Phys*, **58**, 1200 (1980)

-
31. F. Sim, A. St-Amant, I. Papai and D. R. Salahub, *J Am Chem Soc*, **114**, 4391 (1992)
32. A.D. Becke, *Phys Rev A*, **38**, 3098 (1988)
33. J. P. Perdew, *Phys. Rev. B*, **33**, 8822 (1986)
34. A.D. Becke, *J Chem Phys*, **98**, 5648 (1993)
35. J. C. Slater, *Phys Rev*, **36**, 57 (1930)
36. W.J. Hehre, L. Radom, P.v.R. Schleyer and J.A. Pople, *Ab Initio Molecular Orbital Theory*, Wiley, New York (1986)
37. M.J.S. Dewar, E.F. Zoebisch, E.F. Healy, J.J.P. Stewart, *J Am Chem Soc*, **107**, 3902 (1985)
38. J.J.P. Stewart, *J Comput Chem*, **10**, 209 (1989)
39. M.P. Allen and D.J. Tildesley, *Computer Simulations of Liquids*, Oxford University Press, New York (1994)
40. L. Verlet, *Phys Rev*, **159**, 98 (1967)
41. R.W. Hockrey and J.W. Eastwood, *Computer Simulation Using Particles*, McGraw-Hill, New-York (1981)
42. H.J.C. Berendsen, J.P.M. Postma, V.F. vanGunsteren, A. DiNola and J.R. Haak, *J Chem Phys*, **81**, 3684 (1984)
43. Shuichi Nosé, *Mol Phys*, **52**, 255 (1984)
44. A. Cheng and K.M. Merz Jr., *J Phys Chem*, **100**, 1927 (1996)

-
45. K. Fukui, *Acc Chem Res*, **14**, 363 (1981)
46. B. Waszkowycz, I. H. Hillier, N. Gensmantel and D. W. Payling, *J. Chem Soc Perkin Trans*, **2**, 2025 (1991)
47. R.V. Stanton, D. S. Hartsough and K. M. Merz, *J Chem Phys*, **97**, 11858 (1993)
48. Y. S. Lee, M. Hodoscek, B. R. Brooks, P. F. Kador, *Biophys Chem*, **70**, 203 (1998)
49. P. D. Lyne, M. Hodoscek and M. Karplus, *J Phys Chem A*, **103**, 3462 (1999)
50. K.P. Eurenium, D.C. Chatfield and B.R. Brooks, *Int J Quant Chem*, **60**, 1189 (1996)
51. P. Amara and M.J. Field, *Theor Chem Acc*, **109**, 43 (2003)
52. M. Freindorf, J. Gao, *J Comput Chem*, **17**, 386 (1996)
53. T. Kerdcharoen, U. Birkenheuer, S. Kruger, A. Woiterski, N. Rosch, *Theor Chem Acc*, **109**, 285 (2003)
54. T. Matsubara, S. Sieber, K. Morokuma, *Int J Quantum Chem*, **60**, 1101 (1996)
55. S. Dapprich, I. Komaromi, K. S. Byun, K. Morokuma, M. J. Frisch, *J Mol Struc (THEOCHEM)*, **461**, 1 (1999)
56. U. Eichler, C.M. Kolnel and J. Sauer, *J Comput Chem*, **4**, 463 (1996)
57. A. Chang, R. S. Stanton, J. J. Vincent, A. van der Vaart, K. V. Damodaran, S.L. Dixon, D. S. Hartsough, M. Mori, S. A. Best, G. Monard, M. Garcia-Viloca, L. VanZant and K. M. Merz, *ROAR 2.0*, The Pennsylvania State University, 1999

-
58. S. K. Goh, R. T. Gallant and A. St-Amant, *Int. J. Quantum Chem.*, **69**, 405 (1998)
59. P. Amara and M.J. Field, *Theor Chem Acc*, **109**, 43 (2003)
60. D. Debananda, K. P. Eurenus, E. M. Billings, P. Sherwood, D. C. Chatfield, M. Hodoscek and B. R. Brooks, *J Chem Phys*, **117**, 10534 (2002)
61. P. Amara, M.J. Field, C. Alhambra, J. Gao, *Theor Chem Acc*, **104**, 336 (2000)
62. N. Reuter, A. Dejaegere, B. Maigret, M. Karplus, *J Phys Chem A*, **104**, 1720 (2000)
63. R. Vargas, J. Garza, B. P. Hay and D. A. Dixon, *J Phys Chem A*, **106**, 3213 (2002)
64. K. Kim, R. A. Friesner, *J Am Chem Soc*, **119**, 12952 (1997)
65. R. Vargas, J. Garza, R. A. Friesner, H. Stern, B. P. Hay, D.A. Dixon, *J Phys Chem A*, **105**, 4963 (2001)
66. T.K. Woo, P.M. Margl, L. Deng, L. Cavallo, T. Ziegler, *Catalysis Today*, **50**, 479 (1999)
67. A. Michalak, T. Ziegler, *J. Phys Chem A*, **105**, 4333 (2001)
68. J.P. Ryckaert, G. Ciccotti, H.J. Berendsen, *J Comput Phys*, **23**, 327 (1977)
69. B.R. Woodward, R. Hoffmann, *J Am Chem Soc*, **87**, 395 (1965)
70. W. Cooper, W.D. Walters, *J Am Chem Soc*, **80**, 4220 (1958)
71. R.W. Carr, W.D. Walters, *J. Phys Chem*, **59**, 1073 (1965)

-
72. K.B. Wiberg, R.A. Fenoglio, J Am Chem Soc, **90**, 3395 (1968)
73. L.E. Friedrich, G.B. Schuster, J Am Chem Soc, **8**, 4602 (1971)
74. P.C. Martino, P.B. Shevlin, J Am Chem Soc, **102**, 5430 (1980)
75. H. Yu, W. T. Chan, D. Goddard, J Am Chem Soc, **112**, 7529 (1990)
76. K. E. Lewis and H. J. Steiner, J Am Chem Soc, 3080 (1964)
77. J. Rodriguez , J Org Chem, **64**, 6848 (1999)
78. H. Jiao and P. von Ragué Schleiter, J Am Chem Soc, **117**, 11529 (1995)
79. T. Uchimaru, S. Tsuzuki, M. Sugie, A. Sekiya, Chem Phys Lett, **373**, 182 (2003)
80. D. Cain, D.M. Pawar, M. Stewart, H. Billings Jr., E.A. Noe, J Org Chem, **66**, 6092 (2001)
81. P.K. Pearson, H.F. Schaefer, J Chem Phys, **62**, 350 (1975)
82. L.T. Redman, G.D. Purvis, R.J. Bartlett, J Chem Phys, **72**, 986 (1980)
83. C.F. Pan, W.J. Hehre, J Phys Chem, **86**, 321 (1982)
84. J. R. Chelikowsky, J. J. Derby, V. V. Godlevsky, M. Jain and J. Y. Raty, J Phys : Condens Matter, **13**, R817 (2001)
85. D. L. Beveridge, F. M. DiCapua, Annu Rev Biophys Chem, **18**, 431 (1989)
86. G.M. Torrie, J.P. Valleau, J Comput Phys, **23**, 187 (1977)
87. N. Ota, A.T. Brünger, Theor Chem Acc, **98**, 171 (1997)
88. J. VandeVondele, U. Rothlisberger, J Chem Phys, **113**, 4863 (2000)
89. J. VandeVondele, U. Rothlisberger, J Am Chem Soc, **124**, 8163 (2002)

-
90. J. VandeVondele, U. Rothlisberger, J. Phys Chem B, **106**, 203 (2002)
91. N.J. Mosey, A. Hu, T.K. Woo, Chem Phys Lett, **373**, 498 (2003)
92. M. Strajbl, G. Hong, A. Warshel, J Phys Chem B, **106**, 13333 (2002)
93. R.H. Wood, W. Liu, D.J. Doren, J Phys Chem A, **106**, 6689 (2002)
94. W.L. Jorgensen, J. Chandrasekhar, J. Madura, R.W. Imprey, M.L. Klein, J Chem Phys, **79**, 926 (1983)
95. D.A. Pearlman, P.A. Kollman, K Chem Phys, **94**, 4532 (1991)
96. N. Okuyama-Yoshida, M. Nagaoka, T. Yamabe, J Phys Chem A, **102**, 285 (1998)
97. I. Tuñón, E. Silla, C. Millot, M. T. C. Martins-Costa, M. F. Ruiz-López, J Phys Chem A, **102**, 8673 (1998)
98. I. Tuñón, E. Silla, M. F. Ruiz-López, Chem Phys Lett, **321**, 433 (2000)
99. E. Kassab, J. Langlet, E. Evleth, Y. Akacem, J Mol Struc (THOECHEM), **531**, 267 (2000)
100. F. R. Tortonda, J. L. Pascual-Ahuir, E. Silla, I. Tuñón, Chem Phys Lett, **260**, 21 (1996)
101. F. R. Tortonda, J. L. Pascual-Ahuir, E. Silla, I. Tuñón, F. J. Ramírez, J Chem Phys, **109**, 592 (1998)