

Table des matières

Glossaire	ix
Introduction Générale	1
1 Méthodes et outils de simulation	15
1.1 Équation de Schrödinger et approximation de Born-Oppenheimer . . .	15
1.1.1 Équation de Schrödinger et Hamiltonien moléculaire	15
1.1.2 Approximation de Born-Oppenheimer	17
1.2 Calcul de la structure électronique	19
1.2.1 Fonctions monoélectroniques, polyélectroniques et déterminant de Slater	20
1.2.2 Méthodes basées sur la fonction d'onde	21
1.2.3 Méthodes basées sur la densité électronique	25
1.2.4 Mécanique moléculaire et champ de force	36
1.2.5 Méthodes hybrides QM/MM	38
1.3 Dynamique Moléculaire	39
1.3.1 Notion de surface d'énergie potentielle	39
1.3.2 Dynamique moléculaire classique	41
1.3.3 Dynamique moléculaire basée sur les intégrales de chemin . . .	50
1.3.4 Dynamique moléculaire <i>Parallel-Tempering</i>	60
1.3.5 Extraction de propriétés issues de dynamiques moléculaires . .	64
2 Implémentation de la méthode Parallel-Tempering Molecular Dy- namics dans deMon2k et application à l'étude d'agrégats d'eau	69
2.1 Introduction	69
2.2 Division des communicateurs MPI dans deMon2k	74
2.2.1 Structure MPI de deMon2k avant la création de la seconde couche MPI	75

TABLE DES MATIÈRES

2.2.2	Première version de la structure MPI en 2 couches	77
2.2.3	Seconde version de la structure MPI en 2 couches	82
2.3	Implémentation de l'algorithme PTMD dans deMon2k v. 6.0.8	85
2.3.1	Préparation des communicateurs MPI et initialisation de l'algorithme PTMD	85
2.3.2	Routine en charge des échanges de répliques	89
2.3.3	Échanges asynchrones	95
2.4	Application à l'étude du changement de phase dans l'hexamère d'eau (H ₂ O) ₆	99
2.4.1	Méthodes employées et paramètres de calcul	100
2.4.2	Résultats et discussion	103
2.5	Application à l'étude du changement de phase dans l'octamère d'eau (H ₂ O) ₈	111
2.5.1	Méthodes employées et paramètres de calcul	113
2.5.2	Résultats et discussion	115
2.6	Conclusions et perspectives	118
3	Impact des effets quantiques nucléaires sur l'écrantage isotrope	
	¹³C de carboxylates linéaires	121
3.1	Introduction	121
3.2	Implémentation de l'interface entre deMon2k et i-PI	125
3.2.1	Fonctionnement du programme i-PI	125
3.2.2	Implémentation de l'interface entre i-PI et deMon2k	129
3.3	Méthodes employées et paramètres de calcul	131
3.4	Validation des méthodes	134
3.4.1	Validation de la méthode QM/MM	135
3.4.2	Convergence des NQEs	136
3.4.3	Convergence statistique	137
3.5	Résultats et discussion	139
3.5.1	Discussion sur les valeurs absolues de ¹³ C σ_{iso}	139
3.5.2	Discussion sur les valeurs relatives de ¹³ C σ_{iso}	144
3.6	Conclusions et perspectives	146
4	Impact des effets quantiques nucléaires sur la distribution d'isomères dans les agrégats d'eau (H₂O)₆ et (H₂O)₈	149

TABLE DES MATIÈRES

4.1	Introduction	149
4.2	Méthodes employées et paramètres de calcul	153
4.3	Résultats et discussion	155
4.4	Conclusions et perspectives	159
5	Impact des effets quantiques nucléaires sur les spectres d'absorption infrarouge de systèmes $(\text{H}_2\text{O})_n/\text{Ar}$ et $(\text{PAH})(\text{H}_2\text{O})/\text{Ar}$	163
5.1	Introduction	163
5.2	Méthodes employées et paramètres de calcul	167
5.2.1	Configurations initiales	167
5.2.2	Calcul de la structure électronique	168
5.2.3	Exploration de la surface d'énergie potentielle	170
5.2.4	Calcul des spectres d'absorption IR	171
5.3	Résultats et discussion	171
5.3.1	Comparaison des spectres IR de $(\text{H}_2\text{O})_n/\text{Ar}$ en MD et TRPMD-GLE	171
5.3.2	Comparaison des spectres IR de $\text{C}_{24}\text{H}_{12}/\text{Ar}$ et $(\text{C}_{24}\text{H}_{12})(\text{H}_2\text{O})_\sigma/\text{Ar}$ en MD et TRPMD-GLE	175
5.3.3	Comparaison des spectres IR de $\text{C}_{20}\text{H}_{10}/\text{Ar}$ et $(\text{C}_{20}\text{H}_{10})(\text{H}_2\text{O})/\text{Ar}$ en MD et TRPMD-GLE	181
5.4	Conclusions et perspectives	186
	Conclusion générale	189
	Annexes	195
	Bibliographie	199

Glossaire

ADFT	Auxiliary Density Functional Theory
ADFT-GIAO	Auxiliary Density Functional Theory - Gauge-Including Atomic Orbital
AIBs	Aromatic Infrared Bands
B88	Fonctionnelle d'échange GGA
BLYP	Fonctionnelle d'échange GGA B88 et fonctionnelle de corrélation LYP
BOMD	Born-Oppenheimer Molecular Dynamics
CASSCF	Complete Active Space Self-Consistent Field
CASPT2	Complete Active Space Perturbation Theory at the 2 nd order
CBS	Complete Basis Set
CC	Coupled Cluster
CCS	Coupled Cluster à excitations simples
CCSD	Coupled Cluster à excitations simples et doubles
CCSD(T)	Coupled Cluster à excitations simples et doubles traitées variationnellement et triples traitées comme des perturbations
CCSD(TQ)	Coupled Cluster à excitation simples et doubles traitées variationnellement et triples quadruples traitées comme des perturbations
CI	Configuration Interaction
CMD	Centroid Molecular Dynamics
CPMD	Car-Parrinello Molecular Dynamics
deMon2k	Programme de chimie théorique basé sur la méthode ADFT
deMonNano	Programme de chimie théorique basé sur la méthode DFTB
DFT	Density Functional Theory
DFTB	Density Functional based Tight-Binding

GLOSSAIRE

DZVP	Base orbitalaire double zêta à valence séparée
ERIs	Electronic Repulsion Integrals
FF	Force Field
GBSA	Generalized Born Surface Area
GEN-A2	
GEN-A2+	Bases auxiliaires dans deMon2k
GEN-A2*	
GGA	Generalized Gradient Approximation
GIAO	Gauge-Including Atomic Orbital
GLE	Generalized Langevin Equation
HBB2-pol	Potentiel de type champ de force pour l'eau
HF	Hartree-Fock
i-PI	Programme permettant de réaliser des simulations PIMD
INET	Type de socket pour communication inter-processus entre différents nœuds de calcul
IR	Infrarouge
ISM	Institut des Sciences Moléculaires
LCPQ	Laboratoire de Chimie et Physique Quantiques
LDA	Local Density Approximation
LYP	Fonctionnelle de corrélation
MAD	Modélisation, Agrégats, Dynamique
MC	Monte-Carlo
MCMC	Markov Chain Monte-Carlo
MCTDH	Multi-configuration time-dependent Hartree
MD	Molecular Dynamics
meta-GGA	Fonctionnelle d'échange-corrélation prenant en compte le laplacien de la densité en plus du gradient
MIS	Milieu interstellaire
MM	Molecular Mechanics
MP	Møller-Plesset
MPI	Message Passing Interface
NEVPT2	N-Electron Valence State Perturbation Theory at the 2 nd order
NHC	Nosé-Hoover chain
NQEs	Nuclear Quantum Effects

NVE	Ensemble microcanonique
NVT	Ensemble canonique
OPLS	Optimized Potential for Liquid Simulations
OPLS-AA	Ensemble de paramètres pour le champ de force OPLS.
PAHs	Polycyclic Aromatic Hydrocarbon
PBE0	Fonctionnelle hybride GGA
PES	Potential Energy Surface
PI+GLE	Thermostat GLE pour simulations PIMD
PIGLET	Thermostat GLE pour simulations PIMD
PILE	Path-Integral Langevin Equation
PILE-G	Thermostat Path-Integral Langevin Equation global
PILE-L	Thermostat Path-Integral Langevin Equation local
PIMD	Path-Integral Molecular Dynamics
PT	Parallel-Tempering
PTMC	Parallel-Tempering Monte-Carlo
PTMD	Parallel-Tempering Molecular Dynamics
PTPIMD	Parallel-Tempering Path-Integral Molecular Dynamics
q-TIP4P/F	Potentiel de type champ de force pour l'eau
QM	Quantum Mechanics
QM/MM	Quantum Mechanics / Molecular Mechanics
RDFs	Radial Distribution Functions
RMN	Résonance Magnétique Nucléaire
RP	Ring Polymer
RPMD	Ring Polymer Molecular Dynamics
SCC-DFTB	Self-Consistent-Charge Density Functional based Tight-Binding
SCF	Self-Consistent Field
SOAP	Smooth Overlap of Atomic Positions
SVR	Stochastic Velocity Rescale
TIP3P	Potentiel de type champ de force pour l'eau
TMS	Tétraméthylsilane
TRPMD	Thermostated Ring-Polymer Molecular Dynamics
TRPMD-GLE	TRPMD avec un thermostat GLE
TTM3-F	Potentiel de type champ de force pour l'eau

GLOSSAIRE

UNIX	Type de socket pour communication inter-processus au sein d'un même nœud de calcul
WHBB	Potentiel de type champ de force pour l'eau
WF	Wave-Function

Introduction Générale

Introduction Générale

La notion d'agrégat apparaît dans les années 1920 dans la théorie de la nucléation et est définie comme une phase contenant un nombre fini n d'atomes ou de molécules [1]. En 1964, dans un article portant sur l'étude théorique d'oxydes métalliques [2], Frank Albert Cotton définit le terme agrégat métallique comme un « groupe d'atomes métalliques de taille finie et dont la cohésion est assurée par des liaisons chimiques entre atomes de métal ». Depuis, qu'ils soient atomiques ou moléculaires, les agrégats n'ont cessé d'intéresser les scientifiques dans les domaines de la chimie, de la physique ou encore des sciences des matériaux. Le nombre d'études sur les agrégats s'est multiplié et la notion d'agrégat a évolué. Elle n'est plus cantonnée aux espèces métalliques et il existe aujourd'hui une profusion de familles d'agrégats qui sont toutefois classés en deux grandes familles : les **agrégats atomiques** et les **agrégats moléculaires**. Les uns comme les autres peuvent être neutres ou chargés. Les **agrégats atomiques** sont composés uniquement d'atomes et peuvent être « purs », s'ils sont composés d'un seul et même type d'atome, ou « mixtes », s'ils sont composés d'au moins deux espèces différentes. Parmi les **agrégats atomiques purs**, on compte les **agrégats métalliques** (Na_n [3], Al_n [4], Au_n [5]), les **agrégats de gaz inerte** (He_n [6], Ar_n [7, 8], Kr_n [9], Xe_n [10]), les **agrégats de semi-conducteurs** (Si_n , Ge_n) [11] ou encore les **agrégats carbonés** (C_n) [12]. Parmi les **agrégats atomiques mixtes**, on retrouve des **agrégats organométalliques** ($\text{Fe}_2(\text{CO})_9$, $\text{Mn}_2(\text{CO})_{10}$, $\text{Rh}_4(\text{CO})_{12}$, $[(\text{C}_6\text{H}_6)_4\text{Ru}_4\text{H}_4]^{2+}$ [13, 14], des **agrégats d'halogénures de métaux de transitions** (Re_3Cl_9 [15], $(\text{Mo}_6\text{Cl}_{14})^{2-}$ [16]), des **agrégats de bore** ($\text{B}_{12}\text{H}_{12}^{2-}$) [17], les **phases de Zintl** (Cs_2NaAs_7 [18], $\text{K}_{12}\text{Si}_{17}$ [19]), ou encore des **agrégats intermétalliques** ($\text{Na}_{12}\text{Cu}_{12}\text{Sn}_{21}$ [20], PbPt_{12}^{2-} [21]). Les **agrégats moléculaires** sont composés de molécules comme H_2O [22], NH_3 [23], CO_2 [24], C_6H_6 [25], ... liées entre elles par des **liaisons hydrogènes** ou par des **interactions de Van der Waals**.

INTRODUCTION GÉNÉRALE

La taille des agrégats peut aller de quelques atomes (ou molécules) à des centaines de milliers d'éléments. Pour chaque type d'agrégats, il existe une très grande diversité de propriétés en fonction de la taille et/ou de la forme de l'agrégat, ce qui engendre de nombreuses possibilités en terme d'applications. Les **agrégats métalliques** sont par exemple utilisés comme catalyseurs dans de nombreuses réactions chimiques [26, 27, 28, 29]. Des agrégats métalliques fluorescents peuvent également être utilisés en imagerie médicale ou comme vecteurs pour l'administration de médicaments [30]. Les **nanozymes** (nanoparticules présentant des propriétés similaires à celles d'une enzyme) d'oxyde de fer Fe_4O_3 ont largement été étudiées pour leur efficacité dans le traitement des tumeurs en thérapie catalytique [31]. En plus de l'intérêt « applicatif » direct que présentent de nombreux agrégats, ils peuvent également être considérés comme des « **briques** » **élémentaires de la matière condensée**. Un des exemples les plus marquants dans ce cas est celui de l'eau. En effet, la compréhension des mécanismes et des phénomènes physico-chimiques qui ont lieu dans les agrégats d'eau peut servir de support à l'étude de la phase condensée solide ou liquide [32, 33]. En 2011, Toshihiro Kaneko *et al.* [34] ont par exemple étudié la fonte des agrégats $(\text{H}_2\text{O})_{7,8,11,12}$ et ont montré que le type de transition de phase est intrinsèquement lié à la taille de l'agrégat. Une telle étude permet ainsi de mieux comprendre les mécanismes moléculaires intervenant lors des changements de phase dans l'eau liquide ou solide. **L'étude des agrégats d'eau apparaît donc comme fondamentale pour améliorer notre compréhension des phases condensées de l'eau, et c'est sur ces systèmes qu'ont porté les travaux présentés dans cette thèse.**

Depuis plus de 20 ans, l'intérêt porté aux agrégats d'eau a ainsi motivé de très nombreuses **études expérimentales**. Notons, entre autres exemples, les diverses études qui ont permis de **caractériser les isomères de plus basse énergie** de l'**hexamère d'eau** $(\text{H}_2\text{O})_6$ [35, 36, 37, 38, 39, 40]. Les agrégats d'eau sont également impliqués dans de nombreux phénomènes physico-chimiques atmosphériques [41, 42, 43, 44] ou interstellaires [45, 46, 47, 48, 43]. Par exemple, dans l'atmosphère, l'interaction entre des agrégats d'eau et des molécules de trioxide de soufre SO_3 produit de l'acide sulfurique H_2SO_4 , un composant des pluies acides [44]. Dans le milieu interstellaire, de nombreuses études montrent que les agrégats d'eau jouent un rôle important dans la formation de la glycine [46] ou du formamide [48], des molécules à l'origine de la vie.

INTRODUCTION GÉNÉRALE

Les dispositifs expérimentaux permettant l'étude d'agrégats sont en général divisés en trois parties : la **génération de l'agrégat**, la **séparation** et la **caractérisation**, comme le dispositif présenté sur la figure 1.

- **La génération des agrégats** se produit en plusieurs étapes. L'espèce chimique à étudier est d'abord évaporée dans un four si c'est un solide ou dans un ballon si c'est un liquide (a). La vapeur est ensuite entraînée par un gaz porteur (typiquement He ou Ar) (b) jusque dans une chambre d'expansion. Le mélange gazeux passe ensuite dans un canal de condensation où se forment des agrégats de différentes tailles. Les agrégats sortent finalement du canal de condensation par un diaphragme sous la forme d'un faisceau qui sera accéléré en passant par une suite d'étages de pompages différentiels. Finalement, le faisceau passe par un écorceur pour entrer dans la zone d'ionisation-accélération où les agrégats seront séparés (c).
- **La séparation des agrégats** commence par la collision du faisceau d'agrégats avec un faisceau de molécules d'une vapeur ou d'électrons (d), ce qui a pour effet de ioniser les agrégats. Les agrégats chargés sont ensuite accélérés par un champ électromagnétique avant d'entrer dans un spectromètre de masse dans lequel ils sont séparés selon leur taille par déflexion (e).
- **La caractérisation des agrégats.** Une fois séparés par taille, les agrégats peuvent être analysés par des méthodes de spectroscopie (absorption infra-rouge ou ultra-violet, émission/absorption aux rayons X, ...).

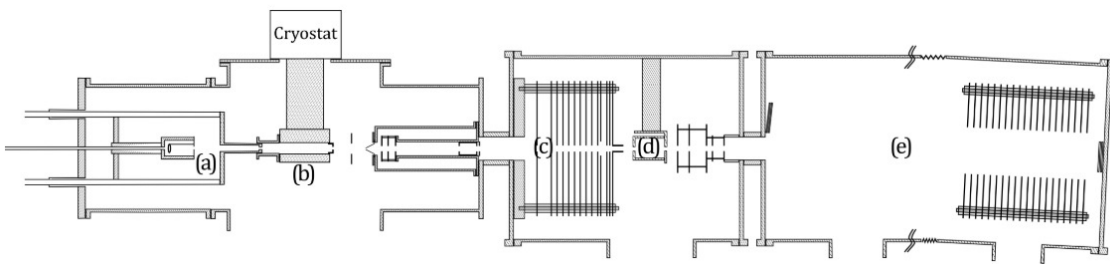


FIGURE 1 – Schéma tiré de la référence [49]. Représentation schématique du dispositif expérimental utilisé pour la création et la séparation d'agrégats.

Parmi les travaux utilisant ce type de dispositif, on peut citer ceux de l'équipe de S. Zamith du LCAR (Laboratoire Collisions Agrégats Réactivité) qui a étudié la capacité calorifique ainsi que les phénomènes d'attachement et de fragmentation

INTRODUCTION GÉNÉRALE

dans des agrégats d'eau de différentes tailles [50, 51, 52, 53, 54, 55], la thèse de G. Bruny sur la production et la caractérisation d'agrégats moléculaires protonés [56], celle de S. Legendre sur l'étude de l'ionisation et de la dissociation d'H₂O induites par collision avec des ions multichargés rapides [57] ou encore celle de R. Maisonnay sur l'étude expérimentale de la fragmentation d'agrégats d'eau induite par impact d'ions multichargés [58]. Toutefois, les dispositifs expérimentaux pour étudier les agrégats sont très coûteux et parfois compliqués à mettre en place. De plus, l'analyse des résultats peut s'avérer compliquée en raison du bruit, de la résolution des spectres mais aussi tout simplement de leur complexité. Il est donc très souvent indispensable d'avoir recours **à des outils de simulation** pour pouvoir corroborer les mesures expérimentales afin d'aider à leur interprétation. La simulation permet également d'approfondir la compréhension des mécanismes physico-chimiques mis en œuvre dans ces expériences. On retrouve de telles collaborations dans de nombreux travaux menés à Toulouse entre l'équipe de S. Zamith du LCAR et le groupe MAD (Modélisation, Agrégats, Dynamique) du LCPQ (Laboratoire de Chimie et Physique Quantiques) [59, 60, 61, 62, 63, 64, 65]. Pour ne citer qu'un seul de ces travaux, la modélisation de l'agrégat (H₂O)₂₁H⁺ a permis de reproduire la brusque rupture de pente observée expérimentalement dans la courbe de capacité calorifique (voir figure 2) de cet agrégat mais également de donner des informations sur l'évolution en température de diverses propriétés structurales et dynamiques : distribution d'angles O—O—O, nombre et taille des cycles primitifs O—O—O, fréquence de transfert de proton.

La grande variété de méthodes et d'outils que l'on peut trouver en chimie théorique permet de répondre précisément aux différents problèmes physico-chimiques mis en jeu dans les agrégats. Cependant, **la principale limitation des études théoriques réside dans l'équilibre à réaliser entre le coût de calcul, la précision des simulations, le type de propriété à simuler et la taille du système à étudier.** En effet, bien que l'augmentation constante de la puissance de calcul et l'amélioration des performances des supercalculateurs aient permis de développer de nouvelles méthodes et d'ouvrir de nouveaux champs d'applications [66, 67, 68, 69, 70, 71], une simulation sans approximation n'est pas envisageable. Chaque méthode présente ses propres approximations et permet de traiter une certaine gamme de propriétés physico-chimiques sur une gamme de tailles donnée et avec une précision donnée. Il faut donc être capable de déterminer celle qui

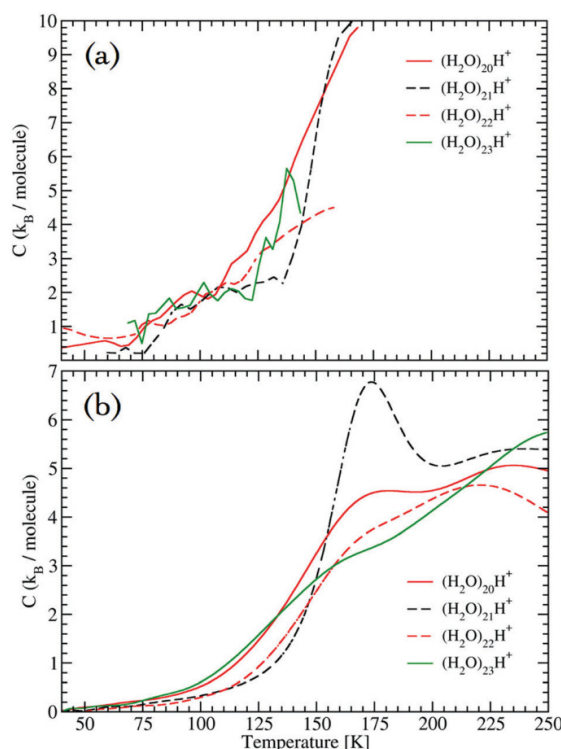


FIGURE 2 – Figure tirée de la référence [59]. Courbes (a) expérimentales et (b) théoriques de la capacité calorifique des agrégats $(\text{H}_2\text{O})_{20-23}\text{H}^+$ en fonction de la température.

sera la plus adaptée au problème à étudier. Lorsque l'on souhaite modéliser des agrégats aqueux, le choix de la méthode porte sur deux axes principaux : le **calcul de la surface d'énergie potentielle** (PES pour *Potential Energy Surface* en anglais) et l'**exploration de la surface d'énergie potentielle**.

Concernant le premier point, le **calcul de la surface d'énergie potentielle**, il existe de nombreuses méthodes permettant de calculer l'énergie d'un système chimique qui peuvent être divisées en **trois grandes familles** et dont les performances sont résumées sur la figure 3 :

- Les **méthodes basées sur la fonction d'onde** (WF pour *Wave Function* en anglais). Pour la majorité d'entre elles, elles permettent une résolution précise de l'équation de Schrödinger électronique non relativiste et indépendante du temps. La plus routinière de ces méthodes, appelée **méthode Hartree-Fock (HF)**, n'est généralement pas suffisante pour décrire correctement la physico-

INTRODUCTION GÉNÉRALE

chimie de la plupart des systèmes chimiques car elle ne tient pas compte de la **corrélacion électronique**. Pour prendre en compte cette dernière, il faut recourir aux méthodes de type **post-Hartree-Fock**. Parmi elles, la méthode dite **Full Configuration Interaction** (**full-CI**) permet de fournir une solution exacte de l'équation de Schrödinger électronique non relativiste et indépendante du temps. Ces méthodes sont très coûteuses en terme de temps de calcul et ne sont applicables que sur des petits systèmes, typiquement $<10^3$ atomes, lorsque l'on souhaite calculer l'énergie d'une structure donnée. Lorsque l'on souhaite mettre en œuvre des simulations de dynamique moléculaire avec de telles approches, les temps de simulation accessibles sont généralement inférieurs à 100 fs.

- Les **méthodes basées sur la densité électronique**. La méthode **DFT** (pour *Density Functional Theory* en anglais) est la méthode actuellement la plus répandue en chimie théorique. Contrairement aux méthodes WF, elle permet de calculer l'énergie d'un système en se basant sur sa densité électronique plutôt que sur sa fonction d'onde. L'énergie provenant de l'interaction d'échange et de la corrélation électronique est prise en compte au travers d'une **fonctionnelle d'échange-corrélation**. Il existe une grande variété de fonctionnelles permettant de traiter de nombreux systèmes atomiques et moléculaires, et de répondre à de nombreux problèmes physico-chimiques. Chaque fonctionnelle est paramétrée pour reproduire au mieux un certain nombre de propriétés pour une gamme de systèmes donnée. En général, les valeurs de référence pour ces propriétés sont des valeurs expérimentales ou des valeurs issues de calculs réalisés avec des méthodes WF. Au sein même de cette famille de méthodes basées sur la densité électronique, on trouve différentes approches basées sur des approximations spécifiques. Par exemple, l'approche **SCC-DFTB** [72] (pour *Self-Consistent-Charge Density Functional Tight-Binding* en anglais) est une formulation approximée de la DFT basée sur un développement de Taylor au second ordre de l'énergie de Kohn-Sham permettant de réaliser des calculs beaucoup plus rapidement. Un autre exemple, l'**ADFT** [73, 74] (pour *Auxiliary Density Functional Theory* en anglais) permet également de réduire le coût de calcul d'un calcul DFT en développant le potentiel d'échange-corrélation sur une densité auxiliaire issue d'un ajustement variationnel du potentiel de Coulomb. Contrairement aux méthodes WF, les méthodes basées sur la densité électronique permettent de simuler des

systèmes plus grand pouvant aller jusqu'à $\sim 10^4$ atomes en ADFT (ou DFT en général) et $\sim 10^6$ atomes en SCC-DFTB pour des calculs simple point. Pour des simulations de dynamique moléculaire, l'ADFT et la SCC-DFTB permettent d'atteindre des temps simulés de plusieurs dizaines et centaines de nanosecondes, respectivement, pour des systèmes de quelques dizaines à une centaine d'atomes.

- Enfin, les méthodes de type **champ de force** (FF pour *Force-Field* en anglais). Ces méthodes sont issues de la **mécanique classique** et calculent l'énergie potentielle d'un système à partir de fonctions analytiques paramétrées. Les paramètres de ces fonctions sont calculés afin de reproduire des données expérimentales spécifiques. Par conséquent, ils dépendent fortement du système étudié et ne sont pas ou très peu transférables d'un système à un autre. De plus, comme ils ne traitent pas explicitement la structure électronique, les méthodes de type FF ne peuvent généralement pas être employées pour étudier des propriétés impliquant les électrons. C'est le cas des processus impliquant la création et/ou la rupture de liaisons chimiques par exemple. À noter qu'un champ de force réactif, ReaxFF [75], a été développé au début des années 2000 pour pouvoir étudier des réactions chimiques par une approche FF. Toutefois, ce champ de force souffre du même problème de transférabilité que les autres champ de force, et doit être re-paramétré pour chaque nouvel ensemble d'éléments chimiques à décrire et pour chaque nouvelle réaction chimique à étudier. De façon générale, l'avantage des méthodes de type champ de force est qu'elles permettent de simuler des systèmes allant jusqu'à plusieurs millions d'atomes et d'atteindre des temps de simulations de l'ordre de la microseconde en dynamique moléculaire.

Pour une description théorique précise des agrégats, en particulier des agrégats aqueux, en plus du calcul de l'énergie potentielle du système, se pose la question de l'**exploration de cette surface d'énergie potentielle** et de l'**échantillonnage de l'ensemble thermodynamique**. En effet, les résultats issus des expériences résultent dans la plupart des cas d'une moyenne sur un très grand nombre d'agrégats pouvant présenter différentes structures et donc différentes signatures spectroscopiques. Un calcul de propriété sur une seule structure ne permet pas dans ce cas de rendre compte de la pluralité des conformations possibles que l'on retrouve nécessairement dans les expériences. Il est donc nécessaire de prendre en compte ces

INTRODUCTION GÉNÉRALE

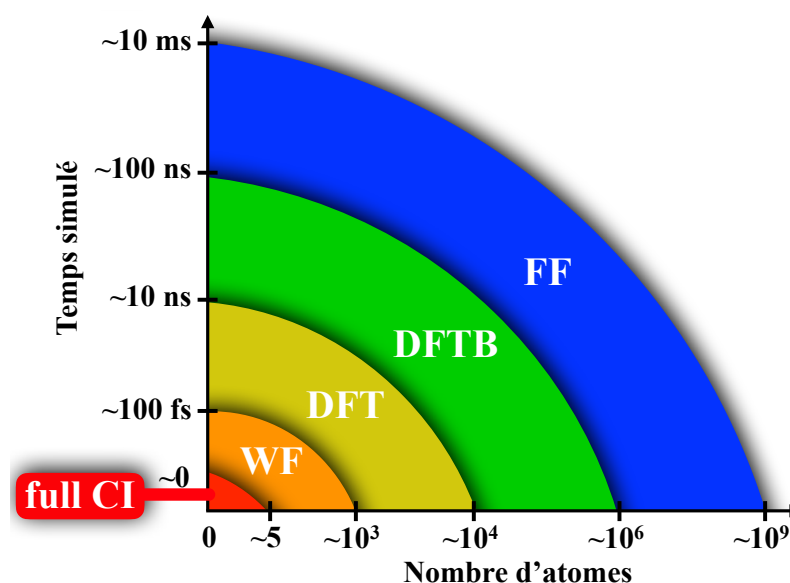


FIGURE 3 – Représentation schématique de l'efficacité des différentes méthodes de calcul d'énergie en chimie computationnelle. L'axe des abscisses indique le nombre maximum d'atomes accessible dans un calcul de type simple point pour une méthode donnée. L'axe des ordonnées indique quant à lui pour chaque méthode le temps de simulation approximatif qu'il est possible d'atteindre lors une simulation de dynamique moléculaire pour un système de taille moyenne pour la méthode considérée.

différentes conformations en réalisant une **exploration détaillée** de la **surface d'énergie potentielle**. Cette exploration détaillée peut avoir deux objectifs principaux qui peuvent nécessiter des outils méthodologiques différents :

- **Optimisation globale.** Il s'agit ici de déterminer les différents minima locaux de la surface d'énergie potentielle d'un système donné. Depuis les années 90, de nombreuses méthodes dédiées à l'optimisation globale ont été développées parmi lesquelles figurent, en autres exemples, le **recuit simulé** [76, 77], les **algorithmes génétiques** [78, 79, 80] et les **algorithmes de seuil** [81, 82]. Cependant ces méthodes permettent uniquement de trouver des minima d'énergie. Or lors de mesures expérimentales, les effets thermiques peuvent faire fluctuer les configurations autour de ces minima et il faut pouvoir prendre en compte ces fluctuations pour décrire au mieux les systèmes.
- **Échantillonnage à température finie.** Les méthodes de **dynamique moléculaire** [83] (MD pour *Molecular Dynamic* en anglais) et de **Monte-Carlo** [84](MC) prennent en compte les fluctuations thermiques

des systèmes en explorant leur **surface d'énergie potentielle** (**PES pour *Potential Energy Surface* en anglais**) à une température donnée. La MD consiste à propager en fonction du temps les positions des atomes en **intégrant les équations classiques du mouvement** à partir du calcul des gradients de l'énergie. A l'inverse, dans une simulation MC, les positions atomiques sont propagées par **déplacements aléatoires**, acceptés ou refusés, ne nécessitant donc pas le calcul des gradients de l'énergie. Les deux approches permettent de calculer des moyennes d'ensemble pour évaluer les propriétés d'un système à température finie. La MD, en faisant explicitement intervenir le temps dans la propagation des noyaux, permet également le calcul de propriétés dynamiques dépendantes du temps. Bien que ces deux méthodes permettent en théorie de réaliser une **exploration ergodique de la PES**, une telle exploration est souvent limitée par la complexité de la surface d'énergie potentielle. Une exploration ergodique de celle-ci s'avérera alors impossible pour les gammes de temps de simulation accessibles et des outils d'**échantillonnage accéléré** devront alors être mis en œuvre.

Un dernier point important est à prendre en compte pour fournir une description précise des agrégats aqueux. En effet, en MD les noyaux des atomes sont généralement traités comme des **particules classiques** (masse ponctuelle chargée positivement). Dans la majorité des cas, cette approximation est justifiée et acceptable car la prise en compte des fluctuations thermiques est suffisante pour décrire correctement les propriétés du système étudié. Cependant, **pour des atomes légers**, typiquement l'hydrogène, les **effets quantiques nucléaires** (NQE pour *Nuclear Quantum Effects* en anglais) ne sont pas négligeables. Ces **effets quantiques sont d'autant plus importants que la température considérée est basse** et donc que les fluctuations thermiques sont faibles. **L'eau constitue précisément un système moléculaire pour lequel les NQE ne sont pas négligeables** et cela même à température ambiante. Il a par exemple été montré que la prise en compte des NQE jouait un rôle important pour la description des fonctions de distribution radiale (RDFs pour *Radial Distribution Functions* en anglais) O–H et H–H de l'eau liquide [85] (voir figure 4 (a)), mais également pour le calcul de sa capacité calorifique [86, 87, 88] (voir figure 4 (b)) ou encore pour décrire le phénomène d'autoprotolyse de l'eau [89] (voir figure 5). La prise en compte de ces effets est donc nécessaire pour décrire précisément les agrégats aqueux. Différentes

INTRODUCTION GÉNÉRALE

approches existent pour cela, en particulier des approches qui traitent les degrés de liberté nucléaires par les outils de la mécanique quantique, c'est le cas de l'approche MCTDH (pour *Multi-configuration time-dependent Hartree* en anglais) par exemple [90]. D'autres approches existent, en particulier la méthode **PIMD** (pour *Path-Integral Molecular Dynamic* en anglais), qui est basée sur les **intégrales de chemin** de Feynmann [91], et qui permet d'introduire les NQEs dans des simulations de MD tout en restant dans un formalisme classique.

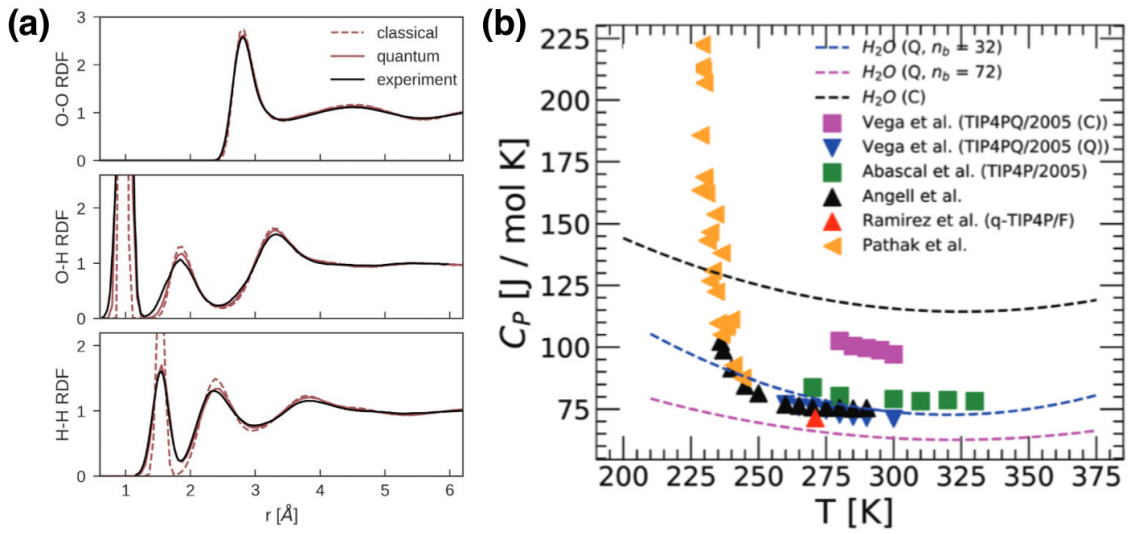


FIGURE 4 – (a) Figure tirée de la référence [85]. De haut en bas, fonctions de distribution radiale O–O, O–H et H–H expérimentales (trait plein noir), simulées en MD (trait discontinu rouge) et PIMD (trait plein rouge) dans l'ensemble NVT à 300 K. La courbe expérimentale O–O est issue de la référence [92] tandis que les courbes O–H et H–H sont issues de la référence [93]. (b) Figure tirée de la référence [86]. Capacité calorifique $C_p(T)$ de H_2O en fonction de la température calculée dans la référence [86] à partir de simulation MD (trait discontinu noir), et PIMD avec 32 répliques (trait discontinu bleu) et 72 répliques (trait discontinu violet). Les auteurs de cet article ont également inclus les valeurs de $C_p(T)$ des références [87] calculées en PIMD avec un potentiel q-TIP4P/F (triangle rouge vers le haut), [88] calculées avec un potentiel TIP4PQ/2005 en MD (carré violet) et PIMD (triangle bleu vers le bas), [94] calculées en MD avec un potentiel TIP4P/2005 (carré vert) et les valeurs expérimentales issues des références [95] et [96] sont respectivement représentées par des triangles orange vers la gauche et des triangles noirs vers le haut.

L'activité principale de l'équipe MAD du LCPQ, dans laquelle j'ai réalisé cette thèse, porte sur l'étude **d'agrégats atomiques et moléculaires par des outils de modélisation**. Une partie importante des travaux réalisés est dédiée au

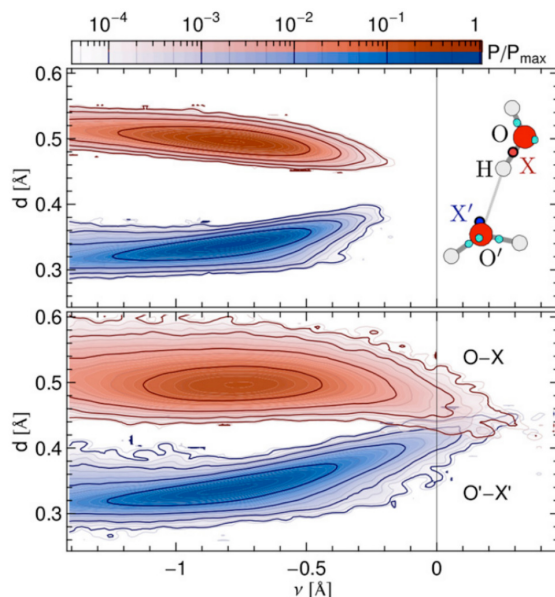


FIGURE 5 – Figure tirée de la référence [89]. Ces courbes sont issues de simulations NVT d’eau liquide à 300 K en MD (panneau du haut) et PIMD (panneau du bas). L’axe des abscisse représente le transfert de proton $\nu = d(O-H) - (O'-H)$. L’axe des ordonnées représente la distance entre l’atome O’ et le centre de Wannier X’ (en bleu) et la distance entre l’atome O et le centre de Wannier X (en rouge). La densité d’état en fonction de ces deux paramètres est proportionnelle à l’opacité de la couleur (rouge ou bleu).

développement et à l’implémentation de nouvelles méthodes dans des programmes de chimie théorique notamment dans le programme de SCC-DFTB **deMonNano** [97] qui est développé dans cette équipe. Les améliorations régulières apportées au code et l’implémentation de nouvelles fonctionnalités permettent d’aborder un grand nombre de problèmes physico-chimiques, offrant un soutien théorique aux résultats expérimentaux, favorisant ainsi les collaborations avec des expérimentateurs. Ainsi, comme mentionné précédemment, de nombreux travaux ont été réalisés en collaboration avec l’équipe Agrégats du LCAR mais également avec l’équipe de Joëlle Mascetti à l’ISM (Institut des Science Moléculaires) de Bordeaux. Cette thèse s’inscrit dans la continuité des activités de recherche du groupe MAD avec pour objectif de réaliser **différentes implémentations dans les codes deMonNano et deMon2k** [98] pour ensuite permettre d’aborder des questions nouvelles portant sur divers agrégats aqueux. Ainsi, l’étude de la **transition de phase dans les agrégats d’eau $(H_2O)_6$ et $(H_2O)_8$** a été réalisée grâce à l’implémentation de l’approche PTMD (pour *Parallel-Tempering Molecular Dynamics* en anglais) dans

INTRODUCTION GÉNÉRALE

le code deMon2k. C'est la première implémentation de cette approche dans un code basé sur le formalisme DFT. Par la suite, les codes **deMonNano** et **deMon2k** ont été interfacés avec le programme **i-PI** [99] pour pouvoir réaliser des simulations **PIMD** dans les formalismes SCC-DFTB et ADFT. À l'aide de l'interface avec deMon2k nous avons étudié l'**impact des NQEs** sur les **déplacements chimiques** ^{13}C de **carboxylates linéaires** explicitement solvatés par de l'eau. En parallèle, à l'aide de l'interface entre i-PI et deMonNano nous avons étudié l'**impact des NQEs** sur deux problèmes physico-chimiques différents : (i) la **distribution d'isomères en fonction de la température dans les agrégats d'eau** $(\text{H}_2\text{O})_6$ et $(\text{H}_2\text{O})_8$; (ii) les **spectres d'absorption infrarouge (IR) d'agrégat** $(\text{H}_2\text{O})_{1-4,6}$, $\text{C}_{24}\text{H}_{12}$, $(\text{C}_{24}\text{H}_{12})(\text{H}_2\text{O})$, $\text{C}_{20}\text{H}_{10}$ et $(\text{C}_{20}\text{H}_{10})(\text{H}_2\text{O})$ piégés dans des matrices d'argon.

Les implémentations et les travaux de recherche réalisés au cours de ces trois années de thèse seront présentés en six chapitres qui sont organisés comme suit :

- Le **premier chapitre** sera dédié à la présentation des bases théoriques, des méthodes et des outils employés tout au long de ce travail de thèse.
- Le **deuxième chapitre** présentera l'**implémentation de la méthode PTMD** dans le code deMon2k et son application à l'étude de propriétés structurales et thermodynamiques des agrégats d'eau $(\text{H}_2\text{O})_6$ et $(\text{H}_2\text{O})_8$. Concernant l'implémentation, une attention particulière sera portée au **développement de la structure parallèle MPI** (pour *Message Passing Interface* en anglais) en deux couches.
- Le **troisième chapitre** portera sur l'étude de l'**impact des NQEs sur le déplacement chimique** ^{13}C de carboxylates linéaires explicitement solvatés par de l'eau. L'**implémentation et le fonctionnement de l'interface** entre le code de *Path-Integral* **i-PI** et **deMon2k** sera introduite en début de chapitre.
- Les **deux chapitres suivant** portent sur **deux projets** qui ont eu pour objectif de déterminer comment les effets quantiques nucléaires pouvaient impacter diverses propriétés physico-chimiques qui n'avaient été jusqu'alors abordées que d'un point de vue classique.
 - * Le **chapitre quatre** porte sur l'impact des NQEs sur la distribution d'isomères en fonction de la température dans les agrégats $(\text{H}_2\text{O})_6$ et

INTRODUCTION GÉNÉRALE

(H₂O)₈. Pour répondre à ce problème, des simulations **PTPIMD** (pour *Parallel-Tempering Path-Integral Molecular Dynamics* en anglais) ont été réalisées et les résultats comparés à ceux obtenus dans la littérature.

- * Le **chapitre cinq** concerne l'étude des **spectres d'absorption infrarouge** de différents agrégats moléculaires : (H₂O)_{1-4,6}, C₂₄H₁₂, (C₂₄H₁₂)(H₂O), C₂₀H₁₀ et (C₂₀H₁₀)(H₂O), piégés dans des matrices d'argon en incluant les NQEs. Les expériences sur ce type de système sont réalisées à des températures très basses (~ 10 K). On s'attend donc à ce que la prise en compte des NQEs joue un rôle important dans la modélisation de tels systèmes. Pour évaluer l'impact des NQEs, nous avons comparé les spectres d'absorption infrarouge (IR) de ces différents systèmes issus de simulations MD et TRPMD-GLE.

Finalement, les **conclusions** et les **perspectives** de ces travaux seront présentés à la fin de ce manuscrit.

Méthodes et outils de simulation

Chapitre 1

Méthodes et outils de simulation

1.1 Équation de Schrödinger et approximation de Born-Oppenheimer

1.1.1 Équation de Schrödinger et Hamiltonien moléculaire

Équation de Schrödinger et Hamiltonien moléculaire

En mécanique quantique, l'état d'un système est décrit par sa fonction d'onde $\Psi(\mathbf{r}, t)$. La connaissance de la fonction d'onde et son évolution dans le temps sont déterminées par l'équation de Schrödinger dépendante du temps [100] :

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

Pour un système atomique ou moléculaire, cette équation peut se réécrire :

$$i\hbar \frac{\partial \Psi(\mathbf{R}, \mathbf{r}, t)}{\partial t} = \hat{\mathcal{H}} \Psi(\mathbf{R}, \mathbf{r}, t) \quad (1.2)$$

où \mathbf{R} et \mathbf{r} désignent respectivement les coordonnées des noyaux atomiques et des électrons. $\hat{\mathcal{H}}$ est l'opérateur Hamiltonien, correspondant à l'opérateur énergie totale du système. Dans le cas où l'Hamiltonien ne dépendrait pas du temps, il est possible

1. MÉTHODES ET OUTILS DE SIMULATION

de séparer les variables spatiales et temporelles dans l'équation de Schrödinger. On obtient alors l'équation de Schrödinger dite « indépendante du temps » :

$$\hat{\mathcal{H}}\Psi(\mathbf{R}, \mathbf{r}) = \mathcal{E}\Psi(\mathbf{R}, \mathbf{r}) \quad (1.3)$$

dans laquelle \mathcal{E} désigne l'énergie totale du système. L'évolution de la fonction d'onde prend alors la forme :

$$\Psi(\mathbf{R}, \mathbf{r}, t) = e^{-iEt/\hbar}\Psi(\mathbf{R}, \mathbf{r}) \quad (1.4)$$

Dans un système composé de M noyaux et N électrons, l'opérateur Hamiltonien non-relativiste s'écrit :

$$\begin{aligned} \hat{\mathcal{H}} &= \hat{T}_n + \hat{T}_e + \hat{V}_{nn} + \hat{V}_{ne} + \hat{V}_{ee} \\ &= -\frac{\hbar^2}{2} \sum_{\alpha=1}^M \frac{1}{M_{\alpha}} \nabla_{\alpha}^2 - \frac{\hbar^2}{2m_e} \sum_{i=1}^N \nabla_i^2 + \sum_{\alpha=1}^M \sum_{\beta>\alpha}^M \frac{Z_{\alpha}Z_{\beta}e^2}{4\pi\epsilon_0 R_{\alpha\beta}} \\ &\quad - \sum_{\alpha=1}^M \sum_{i=1}^N \frac{Z_{\alpha}e^2}{4\pi\epsilon_0 R_{\alpha i}} + \sum_{j=1}^N \sum_{i>j}^N \frac{e^2}{4\pi\epsilon_0 r_{ij}} \end{aligned} \quad (1.5)$$

où \hbar est la constante de Planck réduite $\hbar = h/2\pi$, M_{α} la masse du noyau α , m_e la masse d'un électron, Z_{α} le numéro atomique du noyau α , e la charge élémentaire, ϵ_0 la permittivité diélectrique du vide, $R_{\alpha\beta}$ la distance entre les noyaux α et β , $R_{\alpha i}$ la distance entre le noyau α et l'électron i et r_{ij} la distance entre les électrons i et j . \hat{T}_n et \hat{T}_e sont les opérateurs d'énergie cinétique des noyaux et des électrons. \hat{V}_{nn} , \hat{V}_{ne} , \hat{V}_{ee} sont respectivement l'opérateur d'énergie potentielle de répulsion entre les noyaux, l'opérateur d'énergie potentielle d'interaction entre les noyaux et les électrons et l'opérateur d'énergie potentielle de répulsion entre les électrons. En utilisant le système d'unité atomique (noté u.a.) défini de la manière suivante :

$$\hbar = m_e = \frac{1}{4\pi\epsilon_0} = e = 1 \quad (1.6)$$

l'équation (1.5) s'écrit :

$$\hat{\mathcal{H}} = -\frac{1}{2} \sum_{\alpha=1}^M \frac{1}{M_{\alpha}} \nabla_{\alpha}^2 - \frac{1}{2} \sum_{i=1}^N \nabla_i^2 + \sum_{\alpha=1}^M \sum_{\beta>\alpha}^M \frac{Z_{\alpha}Z_{\beta}}{R_{\alpha\beta}} - \sum_{\alpha=1}^M \sum_{i=1}^N \frac{Z_{\alpha}}{R_{\alpha i}} + \sum_{j=1}^N \sum_{i>j}^N \frac{1}{r_{ij}} \quad (1.7)$$

1.1 Équation de Schrödinger et approximation de Born-Oppenheimer

L'énergie s'exprime alors en Hartree et les distances en Bohr. Malheureusement, les **termes d'énergie potentielle incluant les électrons** (\hat{V}_{ne} et \hat{V}_{ee}) dans l'Hamiltonien donné dans l'équation (1.5) **rendent impossible la résolution analytique exacte de l'équation de Schrödinger indépendante du temps**. Différentes approximations ont donc été développées afin de pouvoir donner une solution approximée de l'équation de Schrödinger indépendante du temps.

1.1.2 Approximation de Born-Oppenheimer

L'**approximation de Born-Oppenheimer** [101] permet de contourner le problème lié au traitement du terme \hat{V}_{ne} en **découplant le mouvement des noyaux de celui des électrons**. En effet, les nucléons étant ~ 1836 fois plus lourds que les électrons, le mouvement des noyaux est beaucoup plus lent que celui des électrons. On peut considérer alors que **ces derniers s'adaptent de manière instantanément et adiabatique aux déplacements des noyaux**. En d'autres termes, les électrons se déplacent dans le champ de potentiel fixe créé par les noyaux et ne dépendent des positions atomiques \mathbf{R} que de façon paramétrique. Pour des noyaux fixes, on peut écrire un Hamiltonien électronique, c'est à dire dont les seules variables sont les degrés de liberté électroniques, comme suit :

$$\hat{H}_e = \hat{T}_e + \hat{V}_{ne} + \hat{V}_{ee} + \hat{V}_{nn} \quad (1.8)$$

L'équation 1.5 peut alors se réécrire :

$$\hat{\mathcal{H}} = \hat{H}_e + \hat{T}_n \quad (1.9)$$

En considérant l'équation de Schrödinger électronique pour l'état électronique k :

$$\hat{H}_e \Phi_R^k(\mathbf{r}) = E_R^k \Phi_R^k(\mathbf{r}) \quad (1.10)$$

où E_R^k est l'énergie de l'état électronique k , on peut exprimer la **fonction d'onde totale** $\Psi(\mathbf{R}, \mathbf{r})$ comme la somme sur les états électroniques du **produit des fonctions**

1. MÉTHODES ET OUTILS DE SIMULATION

d'ondes nucléaires $\chi^k(\mathbf{R})$ et des fonctions propres $\Phi_R^k(\mathbf{r})$ de l'équation (1.10) :

$$\Psi(\mathbf{R}, \mathbf{r}) = \sum_k \Phi_R^k(\mathbf{r}) \chi^k(\mathbf{R}) \quad (1.11)$$

A noter que les positions nucléaires \mathbf{R} sont maintenant en indice pour préciser que ce ne sont plus des variables mais bien des paramètres. En considérant que l'Hamiltonien électronique est diagonal :

$$\left\langle \Phi_R^{k'}(\mathbf{r}) \left| \hat{H}_e \right| \Phi_R^k(\mathbf{r}) \right\rangle = \delta_{kk'} E_R^k \quad (1.12)$$

et que les fonctions d'ondes électroniques sont orthonormales :

$$\left\langle \Phi_R^{k'}(\mathbf{r}) \left| \Phi_R^k(\mathbf{r}) \right\rangle = \delta_{kk'} \quad (1.13)$$

alors si l'on insère les équation (1.9) et (1.11) dans l'équation (1.3), que l'on multiplie à gauche par $\Phi_R^{k'*}(\mathbf{r})$ et que l'on intègre sur les coordonnées électronique \mathbf{r} on obtient :

$$\left[E_{k'_R} + \hat{T}_n - \mathcal{E} \right] \chi^{k'}(\mathbf{R}) = \sum_k \hat{\Lambda}^{kk'} \chi^k(\mathbf{R}) \quad (1.14)$$

dans laquelle le couplage non adiabatique $\hat{\Lambda}^{kk'}$ décrit l'interaction dynamique entre le mouvement des noyaux et celui des électrons.

$$\hat{\Lambda}^{kk'} = \left\langle \Phi_R^{k'}(\mathbf{r}) \left| \hat{T}_n \right| \Phi_R^k(\mathbf{r}) \right\rangle - \hat{T}_n \delta_{kk'} \quad (1.15)$$

Si l'on néglige ce terme de couplage, on obtient l'équation de Schrödinger dans le cadre de l'approximation de Born-Oppenheimer qui décrit le mouvement des noyaux dans un état électronique dit « adiabatique » donné.

$$\left[\hat{T}_n + E_{k'_R} \right] \chi^{k'}(\mathbf{R}) = \mathcal{E} \chi^{k'}(\mathbf{R}) \quad (1.16)$$

L'approximation de Born-Oppenheimer est valable tant que les états électroniques sont énergétiquement bien séparés. En revanche, elle n'est plus valable lorsque des états électroniques sont proches ou dégénérés. C'est souvent le cas dans des molécules radicalaires ou dans des molécules dans des états excités (effet Jahn-Teller, intersections coniques). Dans de tels cas, il faut employer des méthodes plus sophistiquées qui utilisent généralement une combinaison linéaire de bases adia-

1.2 Calcul de la structure électronique

batiques. Dans le cadre de cette thèse, nous considérerons que les électrons s'adaptent de manière instantanément et adiabatique au mouvement des noyaux de sorte que notre système soit toujours dans son état électronique fondamental ($k' = 0$). Nous devons donc résoudre l'équation :

$$\left[\hat{T}_n + E_{0_R} \right] \chi^0(\mathbf{R}) = E \chi^0(\mathbf{R}) \quad (1.17)$$

Les deux étapes suivantes consistent à :

- **traiter le mouvement des noyaux.** Pour cela il est possible de faire l'approximation qu'ils se comportent comme des particules classiques, c'est à dire que leur mouvement est régi par les équations de Newton. Cette approximation fera l'objet de la section 1.3.
- **calculer l'énergie électronique E_{0_R} .** Rappelons que le terme d'énergie potentielle de répulsion entre les électrons \hat{V}_{ee} rend impossible la résolution analytique de l'équation (1.10). Pour pouvoir calculer ce terme, plusieurs grandes familles d'approximations ont été développées. La présentation de ces approximations fait l'objet de la section suivante.

1.2 Calcul de la structure électronique

Dans cette section, nous allons présenter **trois grandes familles d'approximations** permettant de calculer l'énergie électronique E_{0_R} définie dans la section précédente :

- Les méthodes basées sur le **calcul explicite de la fonction d'onde.**
- Les méthodes basées sur le **calcul de la densité électronique.**
- Les méthodes dites de « **champ de force** ».

Chaque méthode permet de répondre à un type de problème différent. Si l'on souhaite étudier une propriété dépendant explicitement de la fonction d'onde alors une méthode de type fonction d'onde sera nécessaire tandis que pour un échantillonnage de la surface d'énergie potentielle, une méthode moins coûteuse (champ de force par exemple) sera privilégiée. Au final, chacune de ces familles présente ses

1. MÉTHODES ET OUTILS DE SIMULATION

avantages et ses inconvénients et le choix de la méthode employée devra être fait en considérant un juste milieu entre temps de calcul, précision et taille du système étudié.

1.2.1 Fonctions monoélectroniques, polyélectroniques et déterminant de Slater

Approximation orbitalaire

Avant toute chose, on suppose que chaque électron i du système peut être représenté par une fonction d'onde monoélectronique $\phi_i(\mathbf{x}_i)$. C'est l'approximation **des électrons indépendants** ou **approximation orbitalaire**. Ces **fonctions monoélectroniques** sont appelées des *spinorbitales* et s'écrivent comme suit :

$$\phi_i(\mathbf{x}_i) = \begin{cases} \varphi_i(\mathbf{r}_i)\alpha(\sigma_i) \\ \varphi_i(\mathbf{r}_i)\beta(\sigma_i) \end{cases} \quad (1.18)$$

où $\varphi_i(\mathbf{r}_i)$ est une fonction d'onde spatiale (ou orbitale) dépendant des coordonnées spatiales de l'électron i (x_i, y_i, z_i dans le cas d'un système en coordonnées cartésiennes) et α et β sont des fonctions de spin dépendant de la variable de spin σ_i . La fonction α correspond à un spin *up* et β à un spin *down*. La variable \mathbf{x}_i désigne quant à elle l'ensemble des variables d'espace et de spin d'un électron donné.

Fonction d'onde polyélectronique et déterminant de Slater

Les électrons étant des fermions, la fonction d'onde d'un système polyélectronique doit respecter le principe d'exclusion de Pauli qui stipule que deux particules fermioniques ne peuvent exister dans le même état quantique. Dans un système atomique, cela revient à dire que deux électrons ne peuvent pas présenter le même ensemble de nombres quantiques n, l, m_l et m_s . Une autre formulation de ce principe est que **la fonction d'onde d'un système polyélectronique est antisymétrique**. Pour conférer à la fonction d'onde polyélectronique ce caractère antisymétrique **on l'écrira comme un produit antisymétrisé de fonctions monoélectroniques**

1.2 Calcul de la structure électronique

qui peut se mettre sous la forme d'un déterminant que l'on appelle **déterminant de Slater**. Pour un système à N électrons occupant N spinorbitales, ce déterminant s'écrit :

$$\Phi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) = (N!)^{-1/2} \begin{vmatrix} \phi_1(\mathbf{x}_1) & \phi_2(\mathbf{x}_1) & \dots & \phi_N(\mathbf{x}_1) \\ \phi_1(\mathbf{x}_2) & \phi_2(\mathbf{x}_2) & \dots & \phi_N(\mathbf{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_1(\mathbf{x}_N) & \phi_2(\mathbf{x}_N) & \dots & \phi_N(\mathbf{x}_N) \end{vmatrix} \quad (1.19)$$

où $(N!)^{-1/2}$ est un facteur de normalisation. Sur une même ligne les électrons sont indicées de 1 à N et sur une même colonne les spinorbitales sont indicées de 1 jusqu'à N . Si l'on échange deux électrons, (ce qui revient à échanger deux lignes du déterminant) le signe du déterminant change également, en accord avec le caractère antisymétrique de la fonction d'onde.

1.2.2 Méthodes basées sur la fonction d'onde

Méthode Hartree-Fock

La méthode Hartree-Fock a joué un rôle important dans la résolution d'un grand nombre de problèmes en chimie et elle constitue la méthode de base des approches fondées sur la fonction d'onde ainsi que de la plupart des méthodes basées sur la densité électronique. En 1928, Douglas R. Hartree [102] propose d'**approximer la fonction d'onde polyélectronique comme un produit de fonctions d'ondes monoélectroniques** en se basant sur la théorie du champ moyen, c'est à dire que chaque électron ressent le potentiel moyen généré par les noyaux et les autres électrons du système. La fonction d'onde polyélectronique s'écrit alors :

$$\Phi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) = \prod_{i=1}^N \phi_i(\mathbf{r}_i, \sigma_i) \quad (1.20)$$

où $\phi_i(\mathbf{r}_i, \sigma_i)$ sont les fonctions d'ondes monoélectroniques (ou spinorbitales) définies plus haut. Cette formulation ne respecte cependant pas la propriété d'antisymétrie de la fonction d'onde polyélectronique. En 1930, John C. Slater [103] et Vladimir Fock [104] proposent indépendamment de remplacer la fonction d'onde polyélectronique proposée par Hartree par un produit antisymétrique de fonctions d'ondes

1. MÉTHODES ET OUTILS DE SIMULATION

monoélectroniques, c'est à dire un déterminant de Slater (équation (1.19)). Une fois cette forme donnée, le **théorème des variations** fournit une information importante sur la marche à suivre pour déterminer l'énergie et la fonction d'onde de l'état fondamental du système. Ce principe stipule que toute fonction d'onde arbitraire ne pourra conduire qu'à une énergie supérieure à l'énergie obtenue avec la fonction d'onde exacte du système. L'application de ce principe revient à **minimiser l'énergie totale du système par rapport à la forme des orbitales monoélectroniques** $\phi_i(\mathbf{x}_i)$ avec pour contrainte qu'elles soient normalisées. **Finalement, un calcul Hartree-Fock (ou SCF pour *Self-Consistent Field* en anglais) consiste à résoudre le système d'équations monoélectroniques couplées (ou équations de Hartree-Fock) suivant :**

$$\hat{F}_i \phi(\mathbf{x}_i) = \epsilon_i \phi(\mathbf{x}_i) \quad (1.21)$$

où \hat{F}_i est l'opérateur de Fock donné par :

$$\hat{F}_i = -\frac{1}{2}\nabla_i^2 + \hat{V}_{ne}(\mathbf{r}_i) + \sum_{j=1}^N (\hat{J}_j - \hat{K}_j) \quad (1.22)$$

En plus des deux premiers termes qui vont conduire respectivement à l'énergie cinétique de l'électron i et à l'interaction électrostatique entre l'électron i et l'ensemble des noyaux, l'opérateur de Fock présente un opérateur d'échange \hat{K}_j et un opérateur d'interaction coulombienne \hat{J}_j . Le premier est un terme purement quantique qui décrit l'énergie liée à la permutation de deux électrons, il provient de la forme antisymétrisée de la fonction d'onde polyélectronique, et le second correspond à l'énergie d'interaction coulombienne entre un électron et l'ensemble des autres électrons du système. En pratique, **résoudre le système d'équations couplées (1.21) revient à déterminer la fonction d'onde totale Φ correspondant à l'énergie la plus basse possible.** En 1951, Clemens Roothan [105] suggéra de représenter les orbitales Hartree-Fock par des combinaisons linéaires de fonctions dites « fonctions de base ». Dans la plupart des codes de chimie quantique, **ces fonctions de base $\mu(\mathbf{r})$ sont des fonctions de type gaussienne centrées sur les positions atomiques.** Des orbitales de type Slater existent aussi mais elles sont moins employées car moins efficaces d'un point de vue computationnel. L'utilisation de **combinaisons linéaires de fonctions de type gaussienne pour décrire la fonction**

d'onde est une approximation efficace qui permet de reformuler les équations de Hartree-Fock sous la forme d'un problème d'algèbre linéaire parfaitement adapté aux ordinateurs actuels.

La principale limitation de la méthode Hartree-Fock est qu'**elle ne peut conduire qu'à une borne supérieure de l'énergie exacte** d'un système électronique dû fait de l'approximation orbitale. La **différence entre l'énergie exacte E_{exact} et l'énergie Hartree-Fock E_{HF} est communément appelée énergie de corrélation $E_{corr} = E_{exact} - E_{HF}$** . En réalité, les électrons sont déjà partiellement corrélés dans la méthode Hartree-Fock en raison de l'antisymétrie de la fonction d'onde décrite par le déterminant de Slater (1.19). **Pour accéder à cette énergie de corrélation et s'approcher de l'énergie exacte des systèmes chimiques, de nombreuses approches basées sur la méthode Hartree-Fock ont été développées.** Parmi elles, on compte la **DFT** que nous présenterons dans la section 1.2.3 et les méthodes **post Hartree-Fock** que nous allons brièvement présenter dans la section suivante.

Méthodes *post* Hartree-Fock

Les méthodes *post* Hartree-Fock ont pour objectif d'améliorer la méthode Hartree-Fock en y ré-intégrant tout ou partie de la corrélation électronique. Ces méthodes partent généralement de la fonction d'onde issue d'un calcul Hartree-Fock pour y appliquer un ensemble d'opérations supplémentaires permettant d'affiner la détermination de l'énergie de corrélation. La différence entre les différentes méthodes repose sur la manière de calculer cette énergie de corrélation. Nous allons ici présenter succinctement quelques-unes de ces méthodes.

- **La méthode Møller-Plesset (MP)** [106] calcule la corrélation électronique par une approche perturbative au second ordre (MP2) [107, 108] ou plus (MP3, [109, 110] MP4 [110, 111, 112], MP5 [113, 114, 115], MP6 [116, 117, 118, 119, 120]) d'un hamiltonien non perturbé, la plus utilisée étant MP2.
- **La méthode Coupled Cluster (CC)** [121, 122, 123, 124] multiplie le déterminant de Slater de référence par un opérateur exponentiel $e^{\hat{T}}$. L'opérateur \hat{T} est une somme des excitations simples (\hat{T}_1), doubles (\hat{T}_2) et ainsi de

1. MÉTHODES ET OUTILS DE SIMULATION

suite jusqu'à N . Dans la pratique, en raison du coût de calcul, cette méthode est appliquée sur de petits systèmes et seules les excitations simples (CCS) et doubles (CCSD) sont calculées variationnellement. Les excitations triples ou quadruples sont généralement traitées par une méthode perturbative (CCSD(T), CCSD(TQ)). Les calculs de précision CCSD(T) ou supérieure sont souvent utilisés comme calcul de référence pour d'autres méthodes.

- **Les méthodes basées sur l'interaction de configuration (CI pour *Configuration Interaction* en anglais).** Ces méthodes consistent à exprimer la fonction d'onde comme une combinaison linéaire de déterminants de Slater. Elle existe sous des **versions tronquées** où seuls les déterminants correspondant à des excitations simples, doubles, ... sont pris en compte ou dans une **version complète (full-CI)**. La méthode full-CI donne une solution exacte à l'équation de Schrödinger non-relativiste indépendante du temps dans la base utilisée. Il existe aussi la méthode **CASSCF** [125] (pour Complete Active Space Self-Consistent Field). Cette méthode consiste à réaliser un calcul full-CI dans un espace constitué par un nombre limité d'orbitales. Ces orbitales doivent être choisies à la main ce qui demande donc une connaissance chimique du système à étudier. L'énergie d'un calcul CASSCF peut ensuite être améliorée par un traitement perturbatif comme le font les méthodes CASPT2 [126, 127] ou NEVPT2 [128, 129, 130]. Toutes ces approches basées sur la l'interaction de configurations deviennent très coûteuses quand le nombre d'électrons N devient grand. Par exemple, le temps d'un calcul full-CI augmente en $N!$.

Les **méthodes basées sur la fonction d'onde** permettent d'obtenir des résultats très précis quant aux propriétés structurales et énergétiques de systèmes physico-chimiques. Elles sont d'ailleurs souvent utilisées comme méthodes de référence pour la validation d'autres méthodes faisant appel à plus d'approximations. Toutefois, leur coût de calcul élevé restreint ces méthodes à l'étude de petits systèmes et les rend totalement **inapplicables pour réaliser des simulations de dynamique moléculaire**. Ainsi, d'autres méthodes de calcul de la structure électronique, moins coûteuses en termes de temps de calcul, doivent être employées pour mettre en œuvre des simulations de dynamique moléculaire. C'est le cas par exemple de la **Théorie de la Fonctionnelle de la Densité (DFT)**.

1.2.3 Méthodes basées sur la densité électronique

Contrairement aux méthodes décrites précédemment, **la DFT se base sur la connaissance de la densité électronique** plutôt que sur la connaissance de la fonction d'onde. Cette méthode basée sur des concepts nés à la fin des années 1920, laissée de côté puis réapparue dans les années 1960 est aujourd'hui la méthode la plus utilisée en chimie théorique. Dans cette partie, nous allons présenter les développements historiques de la DFT depuis le **modèle de Thomas-Fermi** [131, 132] en 1927 en passant par les **théorèmes de Hohenberg et Kohn** [133] en 1964 jusqu'au **formalisme de Kohn-Sham** [134] sur lequel repose la grande majorité des méthodes de DFT employées par les chimistes théoriciens. Ensuite nous présenterons deux déclinaisons de la DFT que nous avons employées dans ces travaux de thèse : l'**ADFT** et la **SCC-DFTB**.

Théorie de la fonctionnelle de la densité

Modèle de Thomas-Fermi

En 1927, en se basant sur leurs travaux portant sur l'énergie cinétique d'un gaz uniforme d'électrons non interagissant, Llewellyn H. Thomas [131] et Enrico Fermi [132] proposent d'**exprimer l'énergie cinétique des électrons dans un système atomique sous la forme d'une fonctionnelle de la densité** :

$$T_{TF}[\rho(\mathbf{r})] = \frac{3}{10}(3\pi^2)^{2/3} \int \rho^{5/3}(\mathbf{r}) d\mathbf{r} \quad (1.23)$$

où $\rho(\mathbf{r})$ représente la densité électronique qui obéit à $\int \rho(\mathbf{r}) d\mathbf{r} = N$, où N est le nombre total d'électrons. En ajoutant les potentiels classiques de répulsion entre les électrons et d'attraction entre les noyaux et les électrons au terme de l'équation (1.23), on peut alors exprimer l'énergie totale d'un système moléculaire comme une fonctionnelle de la densité :

$$E_{TF}[\rho(\mathbf{r})] = T_{TF}[\rho(\mathbf{r})] - \int \rho(\mathbf{r}) v(\mathbf{r}) d\mathbf{r} + \frac{1}{2} \iint \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}' \quad (1.24)$$

où $v(\mathbf{r}) = -\sum_{\alpha} Z_{\alpha}/|\mathbf{r} - \mathbf{R}_{\alpha}|$ et Z_{α} désigne la charge atomique. Deux approximations

1. MÉTHODES ET OUTILS DE SIMULATION

majeures sont faites dans ce modèle. Premièrement, l'énergie cinétique des électrons est approximée comme étant celle d'un gaz uniforme d'électrons non interagissant. Deuxièmement, les énergies d'échange et de corrélation définies dans le modèle de Hartree-Fock sont tout simplement omises. Bien que Paul A. M. Dirac [135] tenta d'ajouter un terme d'énergie d'échange à l'expression proposée par Thomas et Fermi, la description de l'énergie cinétique des électrons est trop éloignée de la réalité. Les erreurs sur l'énergie d'échange et la non prise en compte de la corrélation électronique font que ce modèle ne donne pas de résultat satisfaisant pour la plupart des systèmes.

Théorèmes de Hohenberg et Kohn

Le **premier théorème d'Hohenberg et Kohn** stipule que pour un système de particules en interaction, la connaissance de la densité de l'état fondamental permet à elle seule de déterminer n'importe quelle propriété de ce système et de déduire la fonction d'onde de son état fondamental ainsi que de tous ses états excités. Ceci implique qu'il existe une unique fonctionnelle $F[\rho(\mathbf{r})]$ de la densité, indépendante du système considéré, telle que :

$$E[\rho(\mathbf{r})] = T[\rho(\mathbf{r})] + V_{ee}[\rho(\mathbf{r})] + \int V_{ext}(\mathbf{r})\rho(\mathbf{r})d\mathbf{r} = F[\rho(\mathbf{r})] + \int V_{ext}(\mathbf{r})\rho(\mathbf{r})d\mathbf{r} \quad (1.25)$$

où $E[\rho(\mathbf{r})]$ est l'énergie totale du système qui est donc une fonctionnelle de la densité, $T[\rho(\mathbf{r})]$ correspond à l'énergie cinétique des électrons, $V_{ee}[\rho(\mathbf{r})]$ exprime le potentiel d'interaction entre les électrons et $\int V_{ext}(\mathbf{r})\rho(\mathbf{r})d\mathbf{r}$ représente le potentiel d'interaction coulombien entre les électrons et les noyaux.

Le **second théorème** stipule quant à lui que toute densité électronique approchée d'un état fondamental non dégénéré donne lieu à une borne supérieure de l'énergie exacte de ce système. Par conséquent, le principe des variations peut être appliqué relativement à $\rho(\mathbf{r})$ pour déterminer l'énergie de l'état fondamental du système. On notera respectivement $\rho_0(\mathbf{r})$ et E_0 la densité et l'énergie de l'état fondamental. En appliquant le principe variationnel à l'équation (1.25) on peut trouver la densité électronique de l'état fondamental $\rho_0(\mathbf{r})$ à partir de laquelle on a vu que n'importe quelle propriété était calculable. Encore une fois, **le problème réside dans l'absence d'une forme analytique des termes $T[\rho(\mathbf{r})]$ et $V_{ee}[\rho(\mathbf{r})]$ pour un système de particules en interaction.**

Formalisme de Kohn-Sham

En 1965, Walter Kohn et Lu Jeu Sham supposent que la densité de l'état fondamental $\rho_0(\mathbf{r})$ d'un système de particules interagissant entre elles et soumises à un potentiel extérieur $V_{ext}(\mathbf{r})$ peut être reproduite dans un système de particules n'interagissant pas entre elles et soumises à un potentiel fictif $V_{KS}(\mathbf{r})$. La densité d'un tel système s'écrit :

$$\rho(\mathbf{r}) = \sum_i^n |\phi_i^{KS}(\mathbf{r})|^2 \quad (1.26)$$

où les $\phi_i^{KS}(\mathbf{r})$ sont des fonctions d'ondes monoélectroniques appelées orbitales de Kohn-Sham. L'énergie totale du système s'écrit alors :

$$E[\rho(\mathbf{r})] = T_{KS}[\rho(\mathbf{r})] + \overbrace{\int (V_{ext}(\mathbf{r})\rho(\mathbf{r})d\mathbf{r}) + E_H[\rho(\mathbf{r})] + E_{xc}[\rho(\mathbf{r})]}^{E_{KS}[\rho(\mathbf{r})]} \quad (1.27)$$

où $T_{KS}[\rho(\mathbf{r})]$ est l'énergie cinétique des électrons non-interagissant :

$$T_{KS}[\rho(\mathbf{r})] = \sum_i^n \langle \phi_i^{KS}(\mathbf{r}) | -\frac{1}{2}\nabla^2 | \phi_i^{KS}(\mathbf{r}) \rangle \quad (1.28)$$

$V_{ext}(\mathbf{r})$ le potentiel extérieur, et $E_H[\rho(\mathbf{r})]$ le potentiel de Hartree :

$$E_H[\rho(\mathbf{r})] = \frac{1}{2} \iint \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}d\mathbf{r}' \quad (1.29)$$

Enfin, $E_{xc}[\rho(\mathbf{r})]$ correspond à la fonctionnelle d'échange-corrélation. Ce terme fait intervenir la différence des énergies cinétiques des électrons entre le système réel et le système non interagissant et s'écrit :

$$E_{xc}[\rho(\mathbf{r})] = T[\rho(\mathbf{r})] - T_{KS}[\rho(\mathbf{r})] + V_{ee}[\rho(\mathbf{r})] - E_H[\rho(\mathbf{r})] \quad (1.30)$$

Minimiser l'énergie $E[\rho(\mathbf{r})]$ par rapport aux orbitales de Kohn-Sham revient à résoudre **les équations de Kohn-Sham** :

$$\left(-\frac{1}{2}\nabla^2 + V_{KS}[\rho(\mathbf{r})] \right) \phi_i^{KS}(\mathbf{r}) = \epsilon_i \phi_i^{KS}(\mathbf{r}) \quad (1.31)$$

où ϵ_i sont les énergies des orbitales de Kohn-Sham $\phi_i^{KS}(\mathbf{r})$. Le potentiel de Kohn-

1. MÉTHODES ET OUTILS DE SIMULATION

Sham contient le potentiel extérieur, le potentiel de Hartree et le potentiel d'échange-corrélation :

$$\begin{aligned} V_{KS}[\rho(\mathbf{r})] &= V_{ext}(\mathbf{r}) + V_H[\rho(\mathbf{r})] + V_{xc}[\rho(\mathbf{r})] \\ &= V_{ext}(\mathbf{r}) + \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' + \frac{\delta E_{xc}[\rho(\mathbf{r})]}{\delta \rho(\mathbf{r})} \\ &= \frac{\delta E_{KS}[\rho(\mathbf{r})]}{\delta \rho(\mathbf{r})} \end{aligned} \quad (1.32)$$

La densité $\rho(\mathbf{r})$ minimisant l'énergie peut être trouvée par une procédure auto-cohérente (*Self-Consistent* en anglais) comme pour la méthode Hartree-Fock. Si la fonctionnelle d'échange-corrélation exacte était connue, la théorie de la fonctionnelle de la densité (dans le cadre de l'approximation de Born-Oppenheimer) serait exacte également. Cependant, la forme mathématique de cette fonctionnelle demeure jusqu'à présent inconnue et **il est donc nécessaire d'utiliser une fonctionnelle approchée pour réaliser un calcul de DFT. Ces fonctionnelles peuvent être classées en trois grandes familles** que nous allons présenter dans ce qui suit.

Fonctionnelles d'échange-corrélation

Depuis que le formalisme de Kohn-Sham a été établi en 1965, de nombreuses fonctionnelles d'échange-corrélation ont été développées pour pouvoir répondre à toute sorte de problèmes physico-chimiques. Chaque fonctionnelle présente ses avantages et ses défauts et **on les groupe généralement en trois grandes familles** :

- **les fonctionnelles LDA** (pour *Local Density Approximation* en anglais). Ces fonctionnelles sont les premières à avoir été proposées. Elles se basent sur le gaz uniforme d'électron pour lequel on connaît la forme analytique de l'énergie d'échange. L'énergie de corrélation quant à elle n'est pas connue analytiquement mais elle peut être approximée par des méthodes numériques de chimie quantique. En dépit du caractère local de ce type de fonctionnelles, elles ont donné lieu à des résultats satisfaisants, notamment sur des systèmes dans lesquels la distribution électronique est relativement homogène comme les métaux par exemple.
- **les fonctionnelles GGA** (pour *Generalized Gradient Approximation* en anglais). Dans cette approximation, on introduit le gradient de densité $\nabla \rho(\mathbf{r})$ afin de prendre en compte les variations locales de la densité électronique. Ces

fonctionnelles donnent de meilleurs résultats que les fonctionnelles LDA pour les systèmes présentant de fortes variations de densité, comme les molécules par exemple. Parmi les nombreuses combinaisons de fonctionnelles GGA que l'on peut trouver, **nous avons utilisé la fonctionnelle BLYP** (fonctionnelle d'échange GGA B88 [136] et fonctionnelle de corrélation LYP [137, 138, 139]) **dans ces travaux**. Il existe aussi les fonctionnelles meta-GGA qui prennent en compte le laplacien de la densité en plus du gradient.

- **les fonctionnelles hybrides**. Ces fonctionnelles introduisent une partie ou la totalité de l'énergie d'échange Hartree-Fock dans l'expression de l'énergie d'échange-corrélation. **Dans ces travaux, nous avons utilisé la fonctionnelle PBE0** [140, 141], qui contrairement à la grande majorité des fonctionnelles hybrides, ne fait apparaître aucun paramètre empirique dans sa définition.

Théorie de la fonctionnelle de la densité auxiliaire

L'ADFT est une méthode dérivée de la DFT utilisant une densité **approximée**, la densité auxiliaire, afin de réduire le temps d'un calcul DFT. Dans cette section, nous allons présenter brièvement les différentes approximations qui permettent d'accélérer considérablement le calcul de l'énergie DFT sans pour autant diminuer la précision des résultats. Nous nous limiterons aux principales approximations existantes dans le programme deMon2k, programme que nous avons utilisé pour réaliser les calculs de DFT dans nos travaux.

Formalisme ADFT dans le code deMon2k

Dans la plupart des méthodes de chimie théorique, le goulot d'étranglement d'un calcul *ab-initio* réside dans le calcul des intégrales électroniques à quatre centres que l'on retrouve dans l'interaction coulombienne, l'interaction d'échange et l'interaction de corrélation. Nous avons vu qu'en DFT, l'interaction d'échange et l'interaction de corrélation sont prises en compte dans la fonctionnelle d'échange-corrélation et ne génèrent donc pas de calcul d'intégrales électroniques à quatre centres, elles n'interviennent que dans l'interaction coulombienne. La méthode d'ajustement de densité (DF pour *Density fitting* en anglais) permet de réduire les intégrales à quatre centres dans le potentiel de Coulomb (aussi appelée intégrales de répulsion électronique ou ERIs pour *Electronic Repulsion Integrals* en anglais) en intégrales à trois centres.

1. MÉTHODES ET OUTILS DE SIMULATION

Cette méthode consiste à décomposer la densité orbitale $\rho(\mathbf{r})$ sur des fonctions auxiliaires en minimisant l'énergie :

$$\varepsilon = \frac{1}{2} \iint \frac{[\rho(\mathbf{r}) - \tilde{\rho}(\mathbf{r})][\rho(\mathbf{r}') - \tilde{\rho}(\mathbf{r}')] }{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}' \quad (1.33)$$

où $\tilde{\rho}(\mathbf{r})$ est appelée la densité auxiliaire. Elle s'exprime comme une combinaison linéaire de fonctions auxiliaires $F_i(\mathbf{r})$:

$$\tilde{\rho}(\mathbf{r}) = \sum_i f_i F_i(\mathbf{r}) \quad (1.34)$$

où les f_i sont les coefficients de décomposition. La densité orbitale $\rho(\mathbf{r})$ s'exprime quant à elle comme une combinaison linéaire de produits d'orbitales :

$$\rho(\mathbf{r}) = \sum_{\mu, \nu} P_{\mu\nu} \mu(\mathbf{r}) \nu(\mathbf{r}) \quad (1.35)$$

où $\mu(\mathbf{r})$ et $\nu(\mathbf{r})$ sont des orbitales atomiques de type gaussienne contractée et $P_{\mu\nu}$ un élément de la matrice densité. Dans le code deMon2k, **les fonctions auxiliaires sont des fonctions d'Hermite-Gauss centrées sur les atomes**. L'intérêt de telles fonctions est qu'elles peuvent être calculées en utilisant les relations de récurrence des polynômes d'Hermite. On peut donc écrire une version approximée de l'équation (1.27) faisant intervenir la densité orbitale ainsi que la densité auxiliaire :

$$\begin{aligned} E[\rho(\mathbf{r}), \tilde{\rho}(\mathbf{r})] = & \sum_{\mu, \nu} P_{\mu\nu} H_{\mu\nu} + \sum_{\mu, \nu} \sum_i P_{\mu\nu} \frac{\langle \mu(\mathbf{r}) \nu(\mathbf{r}) | F_i(\mathbf{r}') \rangle f_i}{|\mathbf{r} - \mathbf{r}'|} \\ & - \frac{1}{2} \sum_{i, j} f_i f_j \frac{\langle F_i(\mathbf{r}) | F_j(\mathbf{r}) \rangle}{|\mathbf{r} - \mathbf{r}'|} + E_{xc}[\tilde{\rho}(\mathbf{r})] \end{aligned} \quad (1.36)$$

Cette expression est valide pour les calculs avec des fonctionnelles LDA et GGA mais un terme additionnel est nécessaire pour prendre en compte l'énergie d'échange Hartree-Fock lors des calculs utilisant une fonctionnelle hybride.

Le **premier terme** de cette équation contient les contributions énergétiques monoélectroniques (l'énergie cinétique et l'énergie d'attraction électrons noyaux). Les **deuxième et troisième termes** représentent l'interaction de Coulomb répul-

1.2 Calcul de la structure électronique

sive entre les électrons telle qu'elle est décrite dans l'approche ADFT. On voit que ce terme fait intervenir une intégrale à trois centres et une intégrale à deux centres, réduisant ainsi de N^4 à N^2M la dépendance du coût de calcul au nombre de fonctions de base, N étant le nombre de fonctions de base orbitales et M le nombre de fonctions de base auxiliaires. Enfin, le **dernier terme** de l'équation (1.36) correspond à l'énergie d'échange-corrélation calculée à partir de la densité auxiliaire $\tilde{\rho}(\mathbf{r})$. **La dépendance en $\tilde{\rho}(\mathbf{r})$ de l'énergie d'échange-corrélation présente un intérêt majeur.** En effet, le temps de calcul des intégrales de répulsion électroniques étant dorénavant proportionnel à N^2M , le goulot d'étranglement d'un calcul DFT réside alors dans la décomposition du potentiel électronique sur les bases orbitales. Hideo Sambe et Ronald H. Felton [74] ont proposé une solution à ce problème en ajustant les coefficients de décomposition sur une grille par la méthode des moindres carrés. Cependant, cette méthode ne permet pas d'avoir une formule analytique pour les gradients mais seulement des gradients approximatifs. De tels gradients peuvent engendrer des problèmes lors de calculs d'optimisation de géométrie et ne permettent pas d'obtenir des dérivées d'ordre supérieur. En revanche, il est possible de trouver une formulation variationnelle de l'énergie d'échange-corrélation à partir de la densité auxiliaire issue de l'ajustement variationnel du potentiel de Coulomb. Il est donc possible de dériver une expression analytique du gradient de l'énergie d'échange-corrélation dépendant de la densité auxiliaire. Ainsi, le goulot d'étranglement qu'était le calcul de l'énergie d'échange-corrélation n'en est plus un car le potentiel d'échange-corrélation est à présent décomposé sur les bases auxiliaires. **C'est cette formulation de l'approche ADFT que nous avons employée dans nos travaux.**

La DFT permet d'obtenir des résultats précis pour de nombreuses propriétés électroniques et structurales sur des systèmes pouvant aller jusqu'à plusieurs milliers d'atomes et des propriétés thermodynamiques pour des systèmes allant jusqu'à la centaine d'atomes. Les méthodes basées sur la fonction d'onde quant à elles, ne peuvent pas dépasser quelques centaines d'atomes et sont trop coûteuses pour pouvoir calculer des propriétés thermodynamiques qui nécessitent un échantillonnage statistique suffisamment grand. Cependant, **pour étudier des systèmes plus grands encore ou pour réaliser des trajectoires plus longues**, il est nécessaire de réduire d'avantage le coût du calcul de la structure électronique avec d'autres approximations. C'est ce que propose la méthode SCC-

1. MÉTHODES ET OUTILS DE SIMULATION

DFTB que nous allons présenter dans la prochaine section.

Théorie « *Self-Consistent-Charge Density-Functional based Tight-Binding* »

Nous avons vu dans la section précédente que les parties coûteuses en temps de calcul résident dans le calcul d'intégrales multiples. Dans cette partie, nous allons voir comment la méthode DFTB permet de simplifier le calcul de la structure électronique en réduisant drastiquement le nombre d'intégrales à calculer. Comme pour l'ADFT, nous nous limiterons aux principales équations et approximations présentes dans l'implémentation de la méthode SCC-DFTB dans le programme deMonNano que nous avons utilisé dans nos travaux.

L'approche SCC-DFTB peut être dérivée de la DFT à partir de quatre approximations :

- seuls les électrons de valence sont traités explicitement.
- les orbitales moléculaires (dénotées dans ce qui suit $\psi_i(\mathbf{r})$) sont développées sur une base atomique minimale, la base des orbitales atomiques de valence (dénotées dans ce qui suit $\phi_\nu(\mathbf{r})$).
- un développement de Taylor à l'ordre 2 (pour la SCC-DFTB) de l'énergie est réalisé autour d'une densité de référence.
- les intégrales faisant intervenir plus de deux centres ne sont pas calculées.

Si l'on reprend l'équation (1.27), que l'on exprime la densité de l'état fondamental $\rho_f(\mathbf{r})$ comme la somme d'une densité de référence $\rho_0(\mathbf{r})$ et de sa variation $\delta\rho(\mathbf{r})$, un développement de Taylor au second ordre conduit à l'expression de l'énergie d'un système atomique suivante :

$$\begin{aligned}
 E[\rho_0(\mathbf{r}) + \delta\rho(\mathbf{r})] = & \sum_i^{occ} \langle \psi_i(\mathbf{r}) | \hat{H}^{DFTB} | \psi_i(\mathbf{r}) \rangle \Big\} \mathbf{E}_{\text{band}} \\
 & + \frac{1}{2} \iint \left(\frac{1}{|\mathbf{r} - \mathbf{r}'|} + \frac{\delta^2 E_{xc}}{\delta\rho(\mathbf{r})\delta\rho(\mathbf{r}')} \Big|_{\rho_0(\mathbf{r})} \right) \delta\rho(\mathbf{r})\delta\rho(\mathbf{r}') d\mathbf{r}d\mathbf{r}' \Big\} \mathbf{E}_{\text{2nd}} \\
 & - \frac{1}{2} \iint \frac{\rho_0(\mathbf{r}) + \delta\rho_0(\mathbf{r})}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}d\mathbf{r}' - \int V_{xc}[\rho_0(\mathbf{r})]\rho_0(\mathbf{r}) + E_{xc}[\rho_0(\mathbf{r})] \Big\} \mathbf{E}_{\text{rep}}
 \end{aligned} \tag{1.37}$$

1.2 Calcul de la structure électronique

Énergie des liaisons fortes E_{band}

La première ligne de la partie droite de l'équation (1.37) correspond à ce que l'on appelle l'énergie des liaisons fortes. Elle fait intervenir les orbitales moléculaires $\psi_i(\mathbf{r})$ et l'hamiltonien \hat{H}^{DFTB} . Dans le formalisme SCC-DFTB, les orbitales moléculaires sont exprimées comme une combinaison linéaire d'orbitales atomiques $\phi_\nu(\mathbf{r})$:

$$\psi_i(\mathbf{r}) = \sum_{\nu}^{occ} c_{\nu i} \phi_\nu(\mathbf{r} - \mathbf{R}_\alpha) \quad (1.38)$$

où les $c_{\nu i}$ sont les coefficients de décomposition et $\mathbf{r} - \mathbf{R}_\alpha$ représente la distance au noyau α . Rappelons que seules les orbitales de valence sont prises en compte.

Les éléments de la matrice hamiltonienne \hat{H}^{DFTB} sont les suivant :

$$\hat{H}_{\mu\nu}^{DFTB} = \begin{cases} \epsilon_\mu^{\text{atome libre}} & \text{si } \mu = \nu \\ \langle \phi_\mu^\alpha | \hat{T} + \hat{V}_\alpha + \hat{V}_\beta | \phi_\nu^\beta \rangle & \text{si } \alpha \neq \beta, \mu \in \alpha, \nu \in \beta \\ 0 & \text{autrement} \end{cases} \quad (1.39)$$

Les éléments diagonaux $\epsilon_\mu^{\text{atome libre}}$ sont les valeurs propres de l'atome libre. Les éléments hors diagonaux sont calculés uniquement pour des paires d'atomes. \hat{V}_α est égal à la somme du potentiel de Hartree $V_H[\rho_0(\mathbf{r})]$, du potentiel appliqué par les noyaux $V_{ext}(\mathbf{r})$ et du potentiel d'échange-corrélation $V_{xc}[\rho_0(\mathbf{r})]$ exprimés à la densité de référence $\rho_0(\mathbf{r})$. Tous les éléments de la matrice hamiltonienne $\hat{H}_{\mu\nu}^{DFTB}$ sont pré-calculés pour différentes paires d'atomes et sont compilés dans des tables appelées tables de Slater-Koster. Il existe différentes tables selon le type de système étudié (*mio* pour molécules organiques, *mat* pour solides/matériaux). Dans les travaux de cette thèse nous avons uniquement utilisé les tables *mio*.

Énergie du 2nd ordre E_{2nd}

La deuxième ligne de la partie droite de l'équation (1.37) correspond à ce que l'on appelle l'énergie du second ordre. Cette énergie peut s'écrire :

$$E_{2nd} = \frac{1}{2} \sum_{\alpha, \beta} \Delta q_\alpha \Delta q_\beta \gamma_{\alpha\beta} \quad (1.40)$$

où $\Delta q_\alpha = q_\alpha - q_\alpha^0$ est la différence entre la charge sur l'atome α lorsqu'il est dans la

1. MÉTHODES ET OUTILS DE SIMULATION

molécule et lorsqu'il est isolé. Les paramètres $\gamma_{\alpha\beta}$ correspondent aux paramètres de Hubbard lorsque $\alpha = \beta$ et sont calculés par une approche de type Coulomb écranté lorsqu'ils sont sur des sites différents ($\alpha \neq \beta$). Encore une fois, ces paramètres sont pré-calculés et écrits dans des fichiers au sein desquels le programme va chercher les valeurs. Dans la formulation originale, les charges q_α étaient estimées par des charges de Mulliken mais elles ont ensuite été remplacées par des charges *Class IV / Charge Model 3* (CM3) [142]. Ces dernières ont l'avantage de mieux décrire la polarisation des liaisons interatomiques et donnent de meilleurs résultats quant à la description du dipôle électrique, du potentiel électrostatique, des charges atomiques partielles ou encore du potentiel Coulombien intermoléculaire dans des agrégats de PAHs (pour *Polycyclic Aromatic Hydrocarbon* en anglais).

Énergie de répulsion E_{rep}

La dernière ligne de la partie droite de l'équation (1.37) correspond à l'énergie de répulsion dans le système. Elle est approximée comme étant la somme de l'énergie répulsive entre chaque paire d'atomes :

$$E_{rep} = \frac{1}{2} \sum_{\alpha, \beta} V_{rep}(\mathbf{R}_\alpha - \mathbf{R}_\beta) \quad (1.41)$$

En pratique, $V_{rep}(\mathbf{R}_\alpha - \mathbf{R}_\beta)$ est calculée comme étant la différence entre l'énergie du système calculée en DFT et les énergies des liaisons fortes E_{band} et du second ordre E_{2nd} pour la paire d'atomes α, β donnée :

$$V_{rep}(\mathbf{R}_\alpha - \mathbf{R}_\beta) = E_{DFT}(\mathbf{R}_\alpha - \mathbf{R}_\beta) - E_{band}(\mathbf{R}_\alpha - \mathbf{R}_\beta) - E_{2nd}(\mathbf{R}_\alpha - \mathbf{R}_\beta) \quad (1.42)$$

Les valeurs de $V_{rep}(\mathbf{R}_\alpha - \mathbf{R}_\beta)$ sont également compilées dans les tables de Slater-Koster pour chaque paire d'atomes.

Équations séculaires

On peut à présent reformuler l'énergie totale SCC-DFTB sous une forme ne dépendant pas explicitement de la densité :

$$E_{SCC-DFTB} = \sum_i^{occ} \langle \psi_i(\mathbf{r}) | \hat{H}^{DFTB} | \psi_i(\mathbf{r}) \rangle + \frac{1}{2} \sum_{\alpha\beta} \Delta q_\alpha \Delta q_\beta \gamma_{\alpha\beta} + \frac{1}{2} \sum_{\alpha\beta} V_{rep}(\mathbf{R}_\alpha - \mathbf{R}_\beta) \quad (1.43)$$

1.2 Calcul de la structure électronique

Minimiser cette énergie revient à résoudre les équations séculaires :

$$\sum_{\nu} c_{\nu i} (H_{\mu\nu}^{SCC-DFTB} - \epsilon_i S_{\mu\nu}) = 0 \quad (1.44)$$

avec :

$$H_{\mu\nu}^{SCC-DFTB} = \overbrace{\langle \phi_{\mu}(\mathbf{r}) | \hat{H}^{DFTB} | \phi_{\nu}(\mathbf{r}) \rangle}^{H_{\mu\nu}^{DFTB}} + \overbrace{\frac{1}{2} S_{\mu\nu} \sum_{\xi} (\gamma_{\alpha\xi} + \gamma_{\beta\xi} \Delta q_{\xi})}^{H_{\mu\nu}^1} \quad (1.45)$$

où \hat{H}^{DFTB} est l'hamiltonien DFTB présenté dans l'équation (1.39) et $S_{\mu\nu}$ est la matrice des recouvrements.

Énergie de dispersion E_{disp}

L'énergie de dispersion pour prendre en compte les interactions de Van der Waals s'écrit :

$$E_{disp}(\mathbf{R}_{\alpha\beta}) = - \sum_{\alpha} \sum_{\beta \neq \alpha} f_{damp}(\mathbf{R}_{\alpha\beta}) \frac{C_{\alpha\beta}^6}{\mathbf{R}_{\alpha\beta}} \quad (1.46)$$

où $f_{damp}(\mathbf{R}_{\alpha\beta})$ est une fonction de coupure pour éviter les divergences à trop courte distance et $C_{\alpha\beta}$ est un coefficient empirique pré-calculé pour chaque paire d'atomes.

Calcul des forces

Le calcul des forces ou gradients est nécessaire pour réaliser des optimisations de géométries et pour intégrer les équations du mouvement en dynamique moléculaire. Les forces s'expriment comme le gradient de l'énergie SCC-DFTB par rapport aux coordonnées atomiques :

$$\mathbf{F}_{\alpha} = - \sum_i^{occ} \rho_i(\mathbf{r}) \sum_{\mu\nu} c_{\mu i} c_{\nu i} \left[\frac{\partial H_{\mu\nu}^0}{\partial \mathbf{R}_{\alpha}} - \left(\epsilon_i - \frac{H_{\mu\nu}^1}{S_{\mu\nu}} \right) \frac{\partial S_{\mu\nu}}{\partial \mathbf{R}_{\alpha}} \right] - \Delta q_{\alpha} \sum_{\xi}^N \frac{\partial \gamma_{\alpha\xi}}{\partial \mathbf{R}_{\alpha}} \Delta q_{\xi} - \frac{\partial E_{rep}}{\partial \mathbf{R}_{\alpha}} \quad (1.47)$$

1. MÉTHODES ET OUTILS DE SIMULATION

1.2.4 Mécanique moléculaire et champ de force

La mécanique moléculaire (MM) permet de calculer l'énergie potentielle d'un système atomique en décrivant les interactions interatomiques par la mécanique classique. Cette méthode est conceptuellement différente des méthodes que l'on vient de décrire car elle ne traite pas les électrons de manière explicite et ne nécessite donc pas de résoudre l'équation de Schrödinger. Par conséquent, **son coût de calcul est nettement plus faible que les méthodes basées sur la fonction d'onde ou sur la densité électronique**. Elle peut donc être utilisée pour traiter des systèmes atomiques beaucoup plus gros pouvant aller jusqu'à des milliards d'atomes ou pour réaliser des dynamiques moléculaires de l'ordre de la milliseconde. Dans ce qui suit, nous allons présenter les équations de la MM telle qu'elle est implémentée dans le code deMon2k. L'énergie totale d'un système atomique s'exprime comme suit :

$$E_{total} = \overbrace{E_{liaison} + E_{angle} + E_{Urey} + E_{tors} + E_{itors}}^{E_{covalentes}} + \overbrace{E_{VW} + E_{Coulomb}}^{E_{non-covalentes}} \quad (1.48)$$

Termes d'interactions covalentes

La première partie de l'équation (1.48) contient les termes d'interactions covalentes. Le **premier terme** $E_{liaison}$ décrit l'énergie d'élongation interatomique. Il peut être exprimé par un potentiel harmonique :

$$E_{liaison,Harm} = \frac{1}{2} \sum_{liaisons} k_l (R - R_{eq})^2 \quad (1.49)$$

où k est la constante de raideur harmonique ou constante de force de la liaison et R_{eq} la distance d'équilibre de la liaison. Ces paramètres dépendent des paires d'atomes considérées. Dans deMon2k, l'énergie d'élongation interatomique est décrite par un tel potentiel. À titre informatif, il existe d'autres formes de potentiel pour décrire l'énergie d'élongation comme le potentiel de Morse :

$$E_{liaison,Morse} = \sum_{liaisons} D_{eq} (1 - e^{-a(R-R_{eq})}) \quad (1.50)$$

où D_{eq} est la profondeur du puit à l'équilibre et $a = \sqrt{\frac{k_{eq}}{2D_{eq}}}$ définit la largeur du puit. Ce potentiel est plus proche de la réalité que le potentiel harmonique car il

1.2 Calcul de la structure électronique

tend asymptotiquement vers l'infini quand R tend vers 0 et vers 0 quand R tend vers $+\infty$, mais il est également plus coûteux en terme de temps de calcul en raison de l'exponentielle. Le **deuxième terme** E_{angle} représente l'énergie associée aux déformations angulaires. Il fait intervenir trois atomes et peut également s'exprimer comme un potentiel harmonique :

$$E_{angle} = \frac{1}{2} \sum_{angles} k_{ang} (\theta - \theta_{eq})^2 \quad (1.51)$$

ou k_{ang} est la constante de raideur. Le **troisième terme** E_{Urey} est un potentiel harmonique interatomique comme dans (1.49) mais il fait intervenir deux atomes situés aux extrémités d'un angle :

$$E_{Urey} = \frac{1}{2} \sum_{liaisons} k_{Urey} (R - R_{eq})^2 \quad (1.52)$$

Les **quatrième et cinquième termes** E_{tors} et E_{itors} correspondent à l'énergie associée aux variations d'angle dièdre et font intervenir quatre atomes. Les deux termes ont la même expression :

$$E_{tors} = \frac{1}{2} \sum_{diedre} \sum_n V_n (1 + \cos(n\phi - \phi_n)) \quad (1.53)$$

La différence vient de la connectivité des atomes sur lesquels le potentiel est calculé. Dans le premier cas, les atomes se suivent formant une chaîne linéaire et le nombre de coefficients V_n et de facteurs de phase ϕ_n est de quatre. Dans le second cas, trois atomes sont connectés à un atome central et le nombre de coefficients V_n et de facteurs de phase ϕ_n est de trois.

Termes d'interactions non-covalentes

La deuxième partie de l'équation (1.48) contient les termes d'interactions non-covalentes ou interactions à longue distance. Ces termes décrivent les interactions électrostatiques entre atomes qui ne sont pas nécessairement liés entre eux. Le **premier terme** est le terme d'interaction de Van der Waals qui est modélisé par un

1. MÉTHODES ET OUTILS DE SIMULATION

potentiel de Lennard-Jones :

$$E_{VW} = \sum_{VW} 4\epsilon_0 \left(\left(\frac{\sigma}{R} \right)^{12} - \left(\frac{\sigma}{R} \right)^6 \right) \quad (1.54)$$

où les paramètres diatomiques ϵ_0 et σ correspondent respectivement à la profondeur du puit du potentiel et à la distance à laquelle le potentiel s'annule. Quand ils font intervenir deux types d'atomes différents, ils sont respectivement calculés comme la moyenne arithmétique et géométrique des valeurs homonucléaires ϵ_α et σ_α qui elles, sont connues :

$$\begin{aligned} \epsilon_0 &= \frac{1}{2}(\epsilon_\alpha + \epsilon_\beta) \\ \sigma &= (\sigma_\alpha \sigma_\beta)^{1/2} \end{aligned} \quad (1.55)$$

Enfin le **dernier terme** de l'équation (1.48) $E_{Coulomb}$ représente l'interaction Coulombienne :

$$E_{Coulomb} = \sum_{\alpha, \beta} \frac{q_\alpha q_\beta}{R} \quad (1.56)$$

Nous avons à présent tous les ingrédients pour pouvoir calculer l'énergie d'un système en mécanique moléculaire. Tous les paramètres (constantes de raideurs, profondeurs de puit, largeurs de puit, ...) sont prédéfinis et tabulés dans des fichiers. Différents jeux de paramètres existent. C'est ce que l'on appelle les champs de force. Dans les travaux réalisés au cours de cette thèse, nous avons utilisé les champs de force OPLS-AA [143, 144] et TIP3P [145] qui ont été respectivement paramétrés pour étudier des molécules organiques et l'eau liquide.

1.2.5 Méthodes hybrides QM/MM

L'approche QM/MM (pour *Quantum Mechanics / Molecular Mechanics* en anglais) [146, 147, 148] consiste à traiter certains atomes avec une méthode quantique (QM) par exemple l'ADFT ou la SCC-DFTB, et d'autres en mécanique classique (MM). Ainsi, cette méthode offre la possibilité de modéliser des systèmes trop grands pour être simulés exclusivement en QM tout en traitant explicitement la structure électronique de la partie qui nous intéresse. Par exemple si l'on étudie une molécule solvatée, on peut traiter la molécule solvatée en QM et le solvant en MM. Nous nous

limiterons ici à la présentation de la méthode QM/MM tel qu'elle est implémentée dans deMon2k. L'énergie QM/MM s'écrit :

$$E_{QM/MM} = E_{QM} + E_{MM} + E_{QMMM} + E_{NN} \quad (1.57)$$

où E_{QM} , E_{MM} , E_{QMMM} et E_{NN} correspondent à l'énergie de la partie QM, l'énergie de la partie MM, l'énergie d'interaction QMMM et la répulsion des noyaux entre les parties QM et MM. L'énergie d'interaction QMMM a la même forme que l'équation (1.54) à la différence que cette fois-ci un des atomes appartient à la partie QM :

$$E_{QMMM} = \sum_{\alpha}^{QM} \sum_{\beta}^{MM} 4\epsilon_0 \left(\left(\frac{\sigma}{R} \right)^{12} - \left(\frac{\sigma}{R} \right)^6 \right) \quad (1.58)$$

Les paramètres ϵ_0 et σ sont calculés comme dans l'équation (1.55). Le terme de répulsion entre les noyaux s'écrit :

$$E_{NN} = \sum_{\alpha > \beta}^{QM} \frac{Z_{\alpha} Z_{\beta}}{R} + \sum_{\alpha}^{QM} \sum_{\beta}^{MM} \frac{Z_{\alpha} q_{\beta}}{R} \quad (1.59)$$

1.3 Dynamique Moléculaire

1.3.1 Notion de surface d'énergie potentielle

Un système de taille finie composé de N atomes est caractérisé par $3N - 6$ degrés de liberté. Le nombre 3 correspond aux projections sur les axes \mathbf{x} , \mathbf{y} , \mathbf{z} des coordonnées de chaque atome et le nombre 6 traduit les degrés de rotations et de translations. Pour décrire le paysage énergétique d'un système polyatomique, on utilise la notion de **surface d'énergie potentielle** (PES pour *Potential Energy Surface* en anglais). La PES est une fonction donnant l'énergie d'un système en fonction de l'ensemble des coordonnées atomiques de ce système, c'est donc une fonction **hautement multidimensionnelle** dans la plupart des cas. Dans le cas d'un système diatomique, elle peut être représentée par une courbe à deux dimensions $E = f(R)$ où R est la distance interatomique (voir figure 1.1 (a)). Pour des systèmes contenant plus de deux atomes, il n'existe pas de repré-

1. MÉTHODES ET OUTILS DE SIMULATION

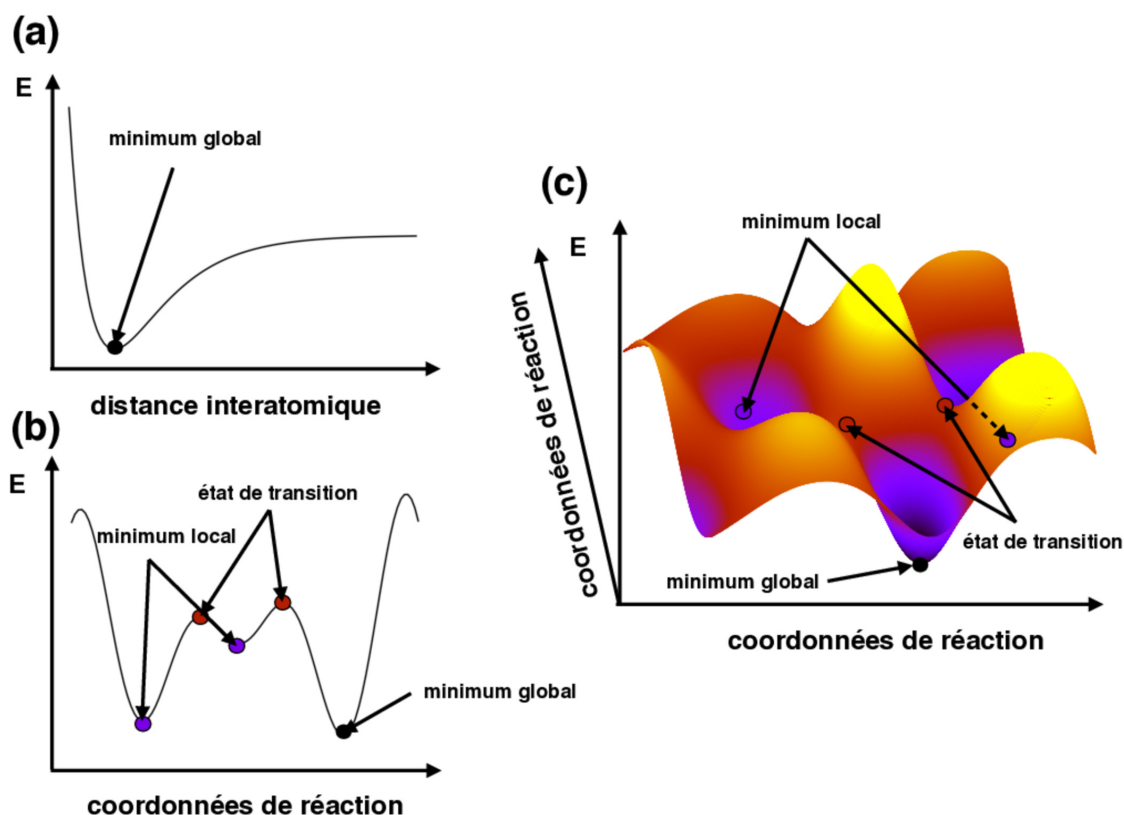


FIGURE 1.1 – Représentation schématique d’une surface d’énergie potentielle (a) en deux dimensions pour un système diatomique, (b) en deux dimensions pour un système avec plus de deux atomes, (c) en trois dimensions pour un système avec plus de deux atomes.

sentation simple de la PES. Elle peut être schématisée par une courbe en deux dimensions $E = f(x)$ où x représente une **coordonnée de réaction**, appelée aussi **variable collective** (voir figure 1.1 (b)) ou comme une isosurface dépendant de deux coordonnées de réaction (voir figure 1.1 (c)). Il est important de noter que les coordonnées de réaction peuvent être **des quantités chimiques concrètes** : distances interatomiques, angles, nombre de coordination ou autre, ou **des quantités beaucoup plus abstraites** telles que la variable *sketch-map* par exemple [149]. Ces dernières sont souvent utilisées pour permettre une représentation en basse dimension d’une PES complexe car l’utilisation d’une seule ou même de deux coordonnées de réaction ne permet de représenter qu’une tranche à 2 ou 3 dimensions de la PES.

Pour mieux visualiser le concept de surface d’énergie potentielle, on peut faire le

parallèle avec un récif montagneux. Les vallées du récif montagneux correspondent à des minima d'énergie locaux sur la PES, les cols correspondent à des états de transition entre ces minima locaux, ils sont également appelés point-selle, et les sommets correspondent à des maxima d'énergie. **Les points de la PES qui nous intéressent** d'un point de vue chimique sont **les minima, locaux et global**, car ils correspondent à des configurations stables, et **les états de transitions** car ils définissent les chemins de plus basse énergie entre deux configurations stables. Mathématiquement, ils correspondent tous à des **points critiques de la PES**, c'est à dire que leurs gradients par rapport aux coordonnées atomiques sont nuls. Cependant, la matrice hessienne des minima ne présente que des valeurs propre positive, alors que celle des états de transitions présente une valeur propre négative. **La PES est une notion clé pour comprendre les propriétés physico-chimiques d'un système.**

Dans la partie précédente, nous avons vu comment calculer l'énergie d'un système atomique pour un jeu de coordonnées atomiques, c'est à dire un point de la PES. Dans cette partie, nous allons présenter les outils permettant d'explorer l'ensemble des configurations chimiquement intéressantes de la PES. Dans une première partie, nous allons présenter les principes et concepts à la base de la **dynamique moléculaires** et les outils développés pour pouvoir échantillonner les ensembles thermodynamiques **microcanonique (NVE)** et **canonique (NVT)**. Dans une deuxième partie, nous expliquerons comment traiter les **effets quantiques des noyaux (NQEs)** pour *Nuclear Quantum Effects* en anglais) à l'aide de la **dynamique moléculaire basée sur les intégrales de chemin (PIMD)** pour *Path-Integral Molecular Dynamics* en anglais). Dans une troisième partie, nous présenterons comment accélérer l'exploration de la PES par l'approche **PTMD** (pour *Parallel-Tempering Molecular Dynamics* en anglais). Enfin, dans une quatrième et dernière partie, nous présenterons les méthodes employées dans cette thèse pour extraire des propriétés à partir des trajectoires de dynamique moléculaire.

1.3.2 Dynamique moléculaire classique

La **dynamique moléculaire (MD)** pour *Molecular Dynamics* en anglais) est une méthode qui permet de modéliser l'évolution temporelle d'un système atomique.

1. MÉTHODES ET OUTILS DE SIMULATION

En MD, les noyaux peuvent être traités comme des particules classiques, c'est à dire comme des masses ponctuelles chargées positivement. On parle dans ce cas de **dynamique moléculaire classique** par opposition à la dynamique quantique. **L'évolution du système est alors régie par les équations de Newton.** La dynamique quantique quant à elle consiste à résoudre l'équation de Schrödinger dépendante du temps (équation (1.1)). Toutes les méthodes présentées dans ce qui suit sont des méthodes de dynamique classique. Le premier travail mentionnant la MD a été réalisé en 1957 par Berni Julian Alder et Thomas Everette Wainwright [150] qui ont étudié la transition de phase d'un système de sphères solides. Depuis, la MD a permis de résoudre un grand nombre de problèmes dans les domaines de la chimie, biochimie, physique, biophysique ou de l'ingénierie. La MD permet de calculer des propriétés dynamiques telles que la diffusion [151, 152, 153] ou la viscosité [152, 154, 155, 156] mais aussi spectroscopiques [157, 158, 159, 160, 161], thermodynamiques [162, 163, 164, 165, 166] ou structurales [167, 168, 169, 170].

Pour pouvoir propager les positions atomiques selon les équations de Newton, il faut pouvoir calculer l'énergie du système en fonction de ces positions. Cela peut être réalisé à l'aide d'une des méthodes présentées dans la section 1.2. Dans le cas où l'énergie est calculée à partir d'une méthode de type champ de force (voir section 1.2.4), les noyaux et les électrons ne sont pas distinguables et constituent une seule et même particule. Comme nous l'avons dit précédemment, cette méthode a l'avantage d'être rapide et permet donc d'atteindre des temps de simulation longs pour des systèmes contenant de nombreux atomes. En revanche, elle ne permet pas de traiter des problèmes électroniques comme par exemple des réactions chimiques au sein desquelles des créations et/ou ruptures de liaison ont lieu. La MD basée sur les champ de force est par exemple très utilisée en biochimie pour étudier le repliement des protéines [171, 172, 173, 174]. Lorsque le calcul de l'énergie se fait par la résolution de l'équation de Schrödinger électronique, on parle alors de **dynamique moléculaire *ab-initio***. La **dynamique moléculaire Born-Oppenheimer (BOMD)** est une méthode de MD *ab-initio* dans laquelle la propagation des degrés de liberté des noyaux et des électrons sont découplés. Il existe d'autres méthodes, telle que la **dynamique moléculaire de Car-Parrinello** (CPMD pour *Car-Parrinello Molecular Dynamics* en anglais) dans laquelle les degrés de liberté électroniques et les degrés de liberté nucléaires sont propagés de manière couplée [175].

Intégration des équations du mouvement

Dans un système atomique, chaque atome interagit avec les autres selon un potentiel défini par les positions atomiques. La MD consiste à faire évoluer ces positions à partir des forces selon la deuxième loi de Newton :

$$\mathbf{F}_\alpha = m_\alpha \mathbf{a}_\alpha = m_\alpha \frac{d^2 \mathbf{R}_\alpha}{dt^2} \quad (1.60)$$

où m_α , \mathbf{a}_α et \mathbf{F}_α sont respectivement la masse, l'accélération et la force s'appliquant sur l'atome α . Cette dernière est définie comme le gradient de l'énergie potentielle en fonction des coordonnées \mathbf{R}_α de l'atome α :

$$\mathbf{F}_\alpha = - \frac{\partial V}{\partial \mathbf{R}_\alpha} \quad (1.61)$$

Pour un système de N atomes, cela conduit à un système de $3N$ équations différentielles couplées. A partir des positions au temps t , il est possible de calculer le potentiel à l'aide de l'une des méthodes qui a été présentées dans la section 1.2. À partir des positions et des forces à cet instant t , il est alors possible de les propager jusqu'au temps $t + \delta t$ en utilisant **un algorithme d'intégration**. Le temps est alors discrétisé en une infinité de petits intervalles de temps appelés **pas de temps** δt . De nombreux algorithmes d'intégration ont été développés parmi lesquels on compte l'algorithme d'Euler [176, 177], la méthode de Runge-Kutta [178, 179, 180], les méthodes de type prédictor-correcteur [181, 182], les algorithmes Verlet simple [183], Leapfrog Verlet [184] et Velocity Verlet [185]. Aujourd'hui, la grande majorité des programmes de MD classique utilisent l'algorithme Velocity Verlet pour intégrer les équations du mouvement. C'est le cas des codes deMonNano et deMon2K que nous avons utilisés au cours de cette thèse. Nous nous limiterons donc à la description de cet algorithme. En appliquant un développement limité pour exprimer les positions et les vitesses à l'instant $t + \delta t$, les équations suivantes sont obtenus :

$$\mathbf{R}_\alpha(t + \delta t) = \mathbf{R}_\alpha(t) + \mathbf{v}_\alpha(t) \delta t + \frac{1}{2} \mathbf{a}_\alpha(t) \delta t^2 \quad (1.62)$$

$$\mathbf{v}_\alpha(t + \delta t) = \mathbf{v}_\alpha(t) + \frac{\mathbf{a}_\alpha(t) + \mathbf{a}_\alpha(t + \delta t)}{2} \delta t \quad (1.63)$$

En pratique, l'implémentation de ces équations se fait généralement en quatre étapes :

1. MÉTHODES ET OUTILS DE SIMULATION

1. Les vitesses à l'instant $t + \frac{1}{2} \delta t$ sont calculées :

$$\mathbf{v}_\alpha \left(t + \frac{1}{2} \delta t \right) = \mathbf{v}_\alpha(t) + \frac{1}{2} \mathbf{a}_\alpha(t) \delta t \quad (1.64)$$

2. À partir de ces nouvelles vitesses, les positions à l'instant $t + \delta t$ sont calculées :

$$\mathbf{R}_\alpha(t + \delta t) = \mathbf{R}_\alpha(t) + \mathbf{v}_\alpha \left(t + \frac{1}{2} \delta t \right) \delta t \quad (1.65)$$

3. Le potentiel V associé à ces nouvelles coordonnées est calculé et dérivé pour trouver l'accélération à l'instant $t + \delta t$:

$$\mathbf{a}_\alpha(t + \delta t) = -\frac{\nabla V}{m_\alpha} \quad (1.66)$$

4. Les vitesses à l'instant $t + \delta t$ sont finalement calculées :

$$\mathbf{v}_\alpha(t + \delta t) = \mathbf{v}_\alpha \left(t + \frac{1}{2} \delta t \right) + \frac{1}{2} \mathbf{a}_\alpha(t + \delta t) \delta t \quad (1.67)$$

Cet algorithme est numériquement stable et **il respecte la réversibilité temporelle**. Cela implique que la trajectoire est la même quelle que soit la direction du temps (t ou $-t$). **Il respecte aussi la conservation de l'hamiltonien à condition que le pas de temps soit suffisamment petit**. Ainsi pour réaliser une simulation de dynamique moléculaire, il faut choisir un pas de temps qui soit suffisamment grand pour pouvoir accéder au temps de simulation désiré et suffisamment petit pour que l'hamiltonien soit conservé tout au long de la simulation. En pratique, la vérification de la conservation de l'hamiltonien se fait en examinant la conservation de l'énergie totale du système. L'algorithme Velocity Verlet permet d'accéder aux positions \mathbf{R}_α , aux vitesses \mathbf{v}_α et aux forces \mathbf{F}_α à la fin de chaque pas d'intégration. Cela présente un grand intérêt si l'on doit modifier les vitesses au cours de la simulation. Or c'est le cas lors de simulations avec des contraintes ou lors de simulations réalisées dans l'ensemble NVT. Dans la section suivante, nous allons justement voir les différents ensembles qu'il est possible d'explorer en MD et les outils qui permettent cette exploration.

Thermostats

En raison de la conservation de l'énergie dans une simulation de MD, l'ensemble thermodynamique qui est échantillonné est l'ensemble microcanonique ou NVE. Dans cet ensemble, le nombre de particules N , le volume V et l'énergie totale E sont constantes tout au long de la simulation. Or, en général, les expériences sont réalisées dans l'ensemble canonique NVT pour lequel ce n'est plus l'énergie qui est constante mais la température. En MD, il est possible d'explorer cet ensemble en **couplant le système étudié à un thermostat**. Ce thermostat va agir sur les vitesses des noyaux atomiques de sorte à ce que la valeur moyenne de l'énergie cinétique corresponde à une énergie cinétique cible E_0 . L'énergie cinétique moyenne est calculée à partir des énergies cinétiques instantanées $E_{cin}(t)$ visitées le long de la simulation. Ces dernières sont reliées à la température instantanée du système par l'équation :

$$T_{cin}(t) = \frac{2E_{cin}(t)}{k_b 3N} \quad (1.68)$$

où k_b représente la constante de Boltzmann, $3N$ le nombre de degrés de liberté ($3N - 6$ si l'on ne tient pas compte des mouvements de translation et de rotation du système) et $E_{cin}(t)$ est l'énergie cinétique instantanée calculée au temps t comme suit :

$$E_{cin}(t) = \sum_{\alpha} \frac{1}{2} m_{\alpha} v_{\alpha}^2(t) \quad (1.69)$$

En pratique, le thermostat va augmenter les vitesses des noyaux si celles-ci sont trop lentes ou au contraire les diminuer si elles sont trop rapides. Pour explorer correctement l'ensemble canonique, un thermostat doit être ergodique, c'est-à-dire qu'il doit permettre au système étudié d'échantillonner correctement la PES accessible à une température donnée. La quête de cette ergodicité a donné lieu au développement de nombreux type de thermostats au fil des années.

Le premier thermostat imaginé et le plus simple conceptuellement est le ***velocity rescale***. Il consiste simplement à renormaliser les vitesses à chaque pas de simulation en les multipliant par un coefficient α pour que celles-ci conduisent à l'énergie cinétique cible E_0 correspondant à la température cible T_0 . Ce coefficient est calculé selon l'équation :

$$\alpha = \sqrt{\frac{E_0}{E_{cin}(t)}} \quad (1.70)$$

1. MÉTHODES ET OUTILS DE SIMULATION

où E_0 est l'énergie cinétique cible correspondant à la température T_0 . Le problème de cette approche est qu'elle impose de brusques impulsions « non physiques » aux vitesses des noyaux à chaque pas de temps et qu'elle ne permet pas à l'énergie cinétique instantanée $E_{cin}(t)$ de fluctuer autour de l'énergie cinétique cible E_0 . En 1984, **Herman J. C. Berendsen a proposé d'introduire une constante de temps τ** dans la renormalisation des vitesses pour que celle-ci soit moins brusque [186]. Le coefficient de renormalisation s'écrit alors :

$$\alpha = \sqrt{1 + \frac{\delta t}{\tau} \left(\frac{E_0}{E_{cin}(t)} - 1 \right)} \quad (1.71)$$

Dans le cas où la constante de temps τ est égale au pas de temps δt , on retrouve l'équation du *velocity rescale*. Au contraire, pour des valeurs de τ très grandes devant δt , on explorera l'ensemble microcanonique (NVE). Tout comme le *velocity rescale*, le thermostat de Berendsen ne reproduit par les fluctuations de l'énergie cinétique de l'ensemble canonique.

Le **thermostat de Nosé-Hoover** [187, 188] consiste à remplacer le système initial par un système « étendu » dans lequel l'évolution temporelle du système étudié est couplée au mouvement d'une variable artificielle ξ de masse Q . Les équations du mouvement à intégrer dans l'algorithme de Verlet deviennent alors :

$$m_\alpha \mathbf{a}_\alpha = -\mathbf{F}_\alpha - m_\alpha \frac{\xi}{Q} \mathbf{v}_\alpha \quad (1.72)$$

$$\dot{\xi} = \sum_{\alpha} m_\alpha \mathbf{v}_\alpha^2 - D k_b T \quad (1.73)$$

où D est le nombre de degrés de liberté dans le système et $\dot{\xi}$ correspond à un coefficient de friction. Le système et le thermostat peuvent échanger de l'énergie au travers de la variable ξ . L'énergie totale du système n'est donc plus conservée. **La somme de l'énergie totale du système et de l'énergie du thermostat quant à elle est bien conservée.** Dans cette approche, l'énergie cinétique peut fluctuer autour de l'énergie cinétique cible permettant ainsi une exploration ergodique de l'ensemble canonique. Toutefois, il a été montré que le thermostat de Nosé-Hoover manque de chaoticité et n'est pas ergodique dans certains cas comme par exemple

pour l'oscillateur harmonique [189].

Une solution à ce problème a été proposée par Glenn J. Martyna *et al.* [190] qui suggèrent de **remplacer le thermostat de Nosé-Hoover par une chaîne de thermostats** (NHC pour *Nosé-Hoover chain* en anglais). Dans ce cas, la particule fictive couplée à une particule du système est elle-même attachée à une autre particule qui est elle-même attachée à une autre particule et ainsi de suite. Cette approche a montré de très bons résultats dans de nombreuses études [191, 192, 193, 194, 195, 196, 197, 198] sans induire un surplus de temps de calcul. En pratique, lorsque l'on utilise le NHC dans une simulation de MD, on définit la longueur de la chaîne, c'est à dire le nombre de particules fictives qu'elle contient ainsi que la fréquence de couplage entre le thermostat et le système étudié. Les masses des particules fictives du NHC sont calculées à partir de cette fréquence et de la température cible T_0 . Le thermostat NHC existe également sous une forme locale appelée « *massive Nosé-Hoover chain* » [199] dans laquelle une chaîne de thermostats est couplée à chaque degré de liberté du système. Cette approche permet d'accélérer grandement la phase d'équilibration pour des systèmes possédant de nombreux degrés de liberté (typiquement des protéines [190] ou de l'eau liquide [200]) et s'avère efficace pour échantillonner localement la PES, c'est à dire le fond d'un puits. En revanche, le caractère local de ce thermostat amorti les fluctuations thermiques ce qui a pour effet de défavoriser le passage de barrières énergétiques et peut donc engendrer des problèmes d'ergodicité. De façon générale, on qualifie de « global » un thermostat qui agit sur l'énergie cinétique totale du système et de « local » un thermostat qui agit sur l'énergie cinétique de chaque particule.

Une autre famille de thermostats se base sur une approche stochastique. Nous présenterons trois thermostats stochastiques dits de « Langevin » : le **thermostat de Langevin** [201], le **stochastic velocity rescale (SVR)** [202] et le **thermostat basé sur l'équation de Langevin généralisée (GLE pour *Generalized Langevin equation* en anglais)** [203, 204, 205]. Les deux premiers thermostats permettent de réaliser une exploration ergodique de l'ensemble canonique, comme le fait le NHC, à la différence qu'ils sont conceptuellement plus simple que ce dernier et plus proches de la « réalité physique ». L'implémentation de ces thermostats s'en trouve donc allégée. Le thermostat GLE consiste à attacher un thermostat de Langevin local à chaque degrés de liberté mais contrairement au thermostat de Langevin

1. MÉTHODES ET OUTILS DE SIMULATION

où le coefficient de friction est le même pour tous les degrés de liberté, il y a un coefficient de friction différent pour chacun d'entre eux.

Thermostat de Langevin

Pour la clarté des explications qui suivent, il convient de reformuler les équations de Newton sous forme hamiltonienne :

$$\begin{aligned}\dot{\mathbf{R}}_\alpha &= \mathbf{p}_\alpha/m_\alpha \\ \dot{\mathbf{p}}_\alpha &= -\frac{\partial V}{\partial \mathbf{R}_\alpha}\end{aligned}\tag{1.74}$$

où $\mathbf{p}_\alpha = \dot{\mathbf{R}}_\alpha \times m_\alpha$ est la quantité de mouvement. Initialement pensée pour décrire le mouvement brownien, dans le cadre de la MD, l'équation de Langevin permet de décrire l'évolution d'un système atomique en contact avec un thermostat extérieur. Elle modifie les équations de Hamilton (1.74) :

$$\begin{aligned}\dot{\mathbf{R}}_\alpha &= \mathbf{p}_\alpha/m_\alpha \\ \dot{\mathbf{p}}_\alpha &= -\frac{\partial V}{\partial \mathbf{R}_\alpha} - \gamma \dot{\mathbf{R}}_\alpha + \sqrt{2m_\alpha \gamma k_B T / \xi_\alpha}\end{aligned}\tag{1.75}$$

en introduisant un terme de frottement visqueux $\gamma \dot{\mathbf{R}}_\alpha$ et un terme de bruit blanc aléatoire $\sqrt{2m_\alpha \gamma k_B T / \xi_\alpha}$. Ce dernier fait intervenir un coefficient de friction γ et un nombre aléatoire à distribution gaussienne ξ_α . La valeur de γ détermine le couplage entre le système et le thermostat. Dans le cas où $\gamma = 0$, on retrouve l'ensemble NVE. Pour une valeur de γ très grande, on se retrouve dans un régime diffusif. Dans ce cas, l'évolution du système ne dépend plus de l'accélération due au potentiel : c'est le mouvement brownien.

Stochastic velocity rescale

Le thermostat SVR est une version globale du thermostat de Langevin. Il consiste à renormaliser les vitesses par un coefficient dépendant du rapport entre l'énergie cinétique cible et l'énergie cinétique instantanée comme dans le cas du thermostat de Berendsen (eq (1.71)), à la différence que l'énergie cinétique cible est cette fois-ci issue de la distribution canonique de l'énergie cinétique à l'équilibre. On notera cette énergie $E_t(t)$. Le coefficient de renormalisation des vitesses s'écrit

donc :

$$\alpha = \sqrt{\frac{E_t(t)}{E_{cin}(t)}} \quad (1.76)$$

La constante de temps τ qui apparaît explicitement dans le thermostat de Berendsen est ici caché dans la variable E_t . En pratique, E_t est calculée selon l'équation :

$$E_t(t) = E_{cin}(t) + (1 - e^{(-\delta t/\tau)}) \left(\frac{\xi_1^2 + \sum_{k=2}^{N_{deg}} \xi_k^2}{N_{deg}} E_0 - E_{cin} \right) + 2\xi_1^2 \sqrt{E_{cin}(t) E_0 \frac{(1 - e^{-\delta t/\tau}) e^{-\delta t/\tau}}{N_{deg}}} \quad (1.77)$$

où ξ est un nombre aléatoire suivant une loi normale, et N_{deg} représente le nombre de degrés de liberté. Comme pour le thermostat de Berendsen, la constante de temps τ assure une évolution progressive de l'énergie cinétique sans introduire de brusques impulsions dans les vitesses. Dans la limite $\tau = 0$ alors les vitesses sont renormalisées à chaque pas de temps et l'on retrouve un équivalent stochastique du velocity rescale (eq (1.70)) ce qui n'est pas souhaitable car l'exploration de la PES ne serait pas ergodique. Au contraire, pour des valeurs de τ très grandes, on retrouve une fois de plus l'ensemble NVE. Il a été montré que ce thermostat explore l'ensemble canonique plus rapidement que le thermostat de Langevin [206].

Generalized Langevin Equation Thermostat (GLE)

Un problème inhérent au thermostat de Langevin est qu'il impose le même coefficient de friction à tous les degrés de liberté, alors que l'échantillonnage optimal de chaque degré de liberté nécessiterait une valeur de γ différente pour chacun. C'est ce constat qui a inspiré le développement de thermostats basés sur l'équation de Langevin généralisée utilisant un bruit coloré au lieu d'un bruit blanc. L'équation de Langevin généralisée s'écrit :

$$\begin{aligned} \dot{\mathbf{R}}_\alpha &= \frac{\mathbf{p}_\alpha}{m_\alpha} \\ \dot{\mathbf{p}}_\alpha &= -\frac{\partial V}{\partial \mathbf{R}_\alpha} - \int_{-\infty}^t dt' \mathcal{K}(t-t') \mathbf{p}_\alpha(t') + \xi_\alpha(t) \end{aligned} \quad (1.78)$$

1. MÉTHODES ET OUTILS DE SIMULATION

où $\mathcal{K}(t)$ est une fonction décrivant la friction et dépend de l'historique des états passés. $\xi_\alpha(t)$ est un vecteur de nombres aléatoires normaux indépendants. En écrivant la matrice mémoire sous la forme :

$$\mathcal{K}(t) = \text{Re} \left(\sum_k^{N_{deg}} c_k e^{-t(\gamma_k + i\omega_k)} \right) \quad (1.79)$$

avec $\gamma_k > 0$, il est possible de réécrire l'équation (1.78) :

$$\begin{aligned} \dot{\mathbf{R}}_\alpha &= \frac{\mathbf{P}_\alpha}{m_\alpha} \\ \begin{pmatrix} \dot{\mathbf{p}}_\alpha \\ \dot{\mathbf{s}}_\alpha \end{pmatrix} &= \begin{pmatrix} -\frac{\partial V}{\partial \mathbf{R}_\alpha} \\ 0 \end{pmatrix} - \begin{pmatrix} a_{pp} & \mathbf{a}_p^T \\ \bar{\mathbf{a}}_p & \mathbf{A} \end{pmatrix} \begin{pmatrix} p_\alpha \\ \mathbf{s}_\alpha \end{pmatrix} + \begin{pmatrix} b_{pp} & \mathbf{b}_p^T \\ \bar{\mathbf{b}}_p & \mathbf{B} \end{pmatrix} (\xi) \end{aligned} \quad (1.80)$$

où $\mathbf{s}_\alpha = (p_\alpha, s_{\alpha 1}, \dots, s_{\alpha N_{deg}})^T$ est un vecteur composé de la quantité de mouvement p_α et de N_{deg} degrés de liberté fictifs $s_{\alpha i}$, et ν_i est un vecteur composé de nombres aléatoires suivant une loi normale. Les matrices \mathbf{A} et \mathbf{B} sont des matrices réelles qui déterminent l'évolution temporelle des moments p_i et qui dépendent de $\mathcal{K}(t)$. De façon générale, l'intérêt des thermostats basés sur l'équation de Langevin généralisée réside dans la flexibilité de la conception de $\mathcal{K}(t)$ qui offre un champ très vaste d'applications. On peut trouver les paramètres pour un certain nombre des ces thermostats sur le site web [207].

1.3.3 Dynamique moléculaire basée sur les intégrales de chemin

La dynamique moléculaire classique telle que nous l'avons présenté précédemment traite indépendamment les degrés de liberté nucléaires et les degrés de liberté électroniques. Si les seconds peuvent être traités explicitement par des méthodes ab-initio, les noyaux sont toujours considérés comme des particules classiques (c'est à dire une masse ponctuelle chargée positivement). Il n'est pas possible d'avoir accès aux propriétés quantiques des noyaux par cette méthode. Or **si l'approximation classique des noyaux est légitime pour de atomes lourds ou pour des simulations dans lesquelles l'énergie thermique est grande devant l'énergie quantique $k_b T \gg h\nu$, les NQEs deviennent importants pour de noyaux**

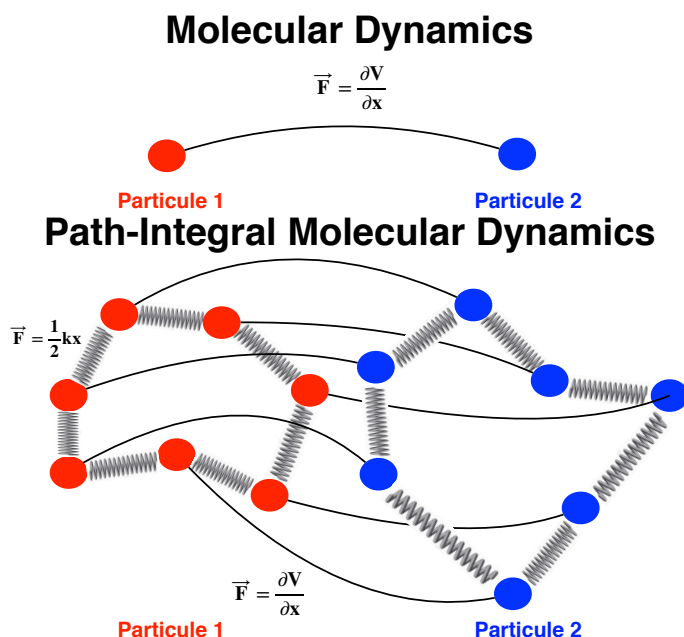


FIGURE 1.2 – Illustration schématique d’un système composé de deux particules dans les formalismes MD et PIMD. En MD, le système est composé uniquement des deux particules qui interagissent entre elles via un potentiel V . En PIMD chaque particule est répliquée n fois (ici 7). Chaque particule au sein d’une réplique est alors soumise au potentiel V et à un potentiel harmonique la reliant aux répliques voisines.

légers et ce d’autant plus que la température est faible. Par exemple, il a été démontré que les NQEs ont un effet non négligeable sur les propriétés de l’eau même à température ambiante [85, 86, 87, 88, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217, 218]. **L’approche PIMD permet de prendre en compte les NQEs tout en restant dans un formalisme de dynamique moléculaire classique.** Dans cette partie, nous expliquerons le fonctionnement de l’approche PIMD. Dans un premier temps, nous présenterons brièvement son formalisme. Ensuite nous verrons comment **intégrer les équations du mouvement en PIMD**, et enfin comment **contrôler la température dans une telle simulation** à l’aide des thermostats stochastiques présentés dans la section 1.3.2.

1. MÉTHODES ET OUTILS DE SIMULATION

Formalisme du PIMD

La méthode PIMD repose sur l'isomorphisme entre la fonction de partition d'un système quantique et la fonction de partition classique d'un système dans l'espace des phases étendu d'un polymère cyclique (RP pour *Ring Polymer* en anglais) [219]. Par polymère cyclique, nous nous référons ici à un système cyclique composé de plusieurs répliques d'un même système atomique reliées entre elles par une force harmonique. Voyons à présent d'où vient cet isomorphisme. L'hamiltonien d'un système de N particules s'écrit :

$$\begin{aligned} H &= \sum_{\alpha}^N \frac{\mathbf{p}_{\alpha}^2}{2m_{\alpha}} + V(\mathbf{R}_1, \dots, \mathbf{R}_N) \\ &= T + V \end{aligned} \quad (1.81)$$

En mécanique quantique, la fonction de partition canonique d'un tel système s'écrit :

$$Z = \text{Tr}(e^{-\beta H}) \quad (1.82)$$

où $\beta = \frac{1}{k_b T}$ est la température inverse. En utilisant une factorisation de Suzuki-Trotter symétrique :

$$e^{-\beta H} = \lim_{n \rightarrow +\infty} [e^{\frac{-\beta V}{2n}} e^{\frac{-\beta T}{n}} e^{\frac{-\beta V}{2n}}] \quad (1.83)$$

il est alors possible d'approximer Z par :

$$Z_n \approx \left(\frac{1}{(2\pi\hbar)^{Nn}} \right) \int d^{Nn} \mathbf{p} \int d^{Nn} \mathbf{R} e^{-\beta_n H_n(\mathbf{p}, \mathbf{R})} \quad (1.84)$$

où $\beta_n = \beta/n$ et $H_n(\mathbf{p}, \mathbf{R})$ est l'hamiltonien classique d'un polymère cyclique composé de n répliques du système de N particules. Les particules identiques au sein de deux répliques voisines sont reliées par une force harmonique. Les fonctions de partition quantiques Z et classique du polymère cyclique Z_n sont égales dans la limite où n tend vers l'infini. Il est donc théoriquement possible de décrire l'état d'un système quantique par un système classique étendu. Écrivons à présent l'hamiltonien classique du polymère cyclique :

$$H_n(\mathbf{p}, \mathbf{R}) = H_n^0(\mathbf{p}, \mathbf{R}) + V_n(\mathbf{R}) \quad (1.85)$$

où $H_n^0(\mathbf{p}, \mathbf{R})$ est l'hamiltonien du polymère libre :

$$H_n^0(\mathbf{p}, \mathbf{R}) = \sum_{\alpha=1}^N \sum_{j=1}^n \left(\frac{[\mathbf{p}_{\alpha}^{(j)}]^2}{2m_{\alpha}} + \frac{1}{2} m_{\alpha} \omega_n^2 (\mathbf{R}_{\alpha}^{(j)} - \mathbf{R}_{\alpha}^{(j-1)}) \right) \quad (1.86)$$

et $V_n(\mathbf{R})$ le potentiel physique créé par l'interaction des particules au sein d'une même réplique.

$$V_n(\mathbf{R}) = \sum_{j=1}^n V(\mathbf{R}_1^{(j)}, \dots, \mathbf{R}_N^{(j)}) \quad (1.87)$$

Comme précédemment, les atomes sont indexés par les indices α . Les indices j servent ici à indexer les répliques au sein d'un polymère cyclique. L'hamiltonien du polymère libre décrit le système des particules non interagissant entre elles au sein d'une même réplique. Il fait intervenir les positions $\mathbf{R}_{\alpha}^{(j)}$ et vitesses $\mathbf{p}_{\alpha}^{(j)}$ des particules dans toutes les répliques ainsi que sa pulsation $\omega_n = 1/\beta_n \hbar$. La condition cyclique est imposée par $\mathbf{R}_{\alpha}^0 \equiv \mathbf{R}_{\alpha}^n$.

L'évolution temporelle des coordonnées $\mathbf{R}_{\alpha}^{(j)}$ est régie par les équations du mouvement :

$$\begin{aligned} \dot{\mathbf{p}}_{\alpha}^{(j)} &= - \frac{\partial H_n(\mathbf{p}, \mathbf{R})}{\partial \mathbf{R}_{\alpha}^{(j)}} \\ \dot{\mathbf{R}}_{\alpha}^{(j)} &= + \frac{\partial H_n(\mathbf{p}, \mathbf{R})}{\partial \mathbf{p}_{\alpha}^{(j)}} \end{aligned} \quad (1.88)$$

En utilisant l'isomorphisme $\langle A \rangle = \lim_{n \rightarrow +\infty} \langle A \rangle_n(\mathbf{R})$, la moyenne thermique d'un opérateur dépendant des positions :

$$\langle A \rangle = \frac{1}{Z} \text{Tr} \left(e^{-\beta \hat{H}} \hat{A} \right) \quad (1.89)$$

peut être calculée comme :

$$\langle A \rangle_n(\mathbf{R}) = \frac{1}{(2\pi\hbar)^n Z_n} \int d\mathbf{p} \int d\mathbf{R} e^{-\beta_n H_n(\mathbf{p}, \mathbf{R})} A_n(\mathbf{R}) \quad (1.90)$$

1. MÉTHODES ET OUTILS DE SIMULATION

où la valeur de $A_n(\mathbf{R})$ est moyennée sur les répliques du polymère cyclique :

$$A_n(\mathbf{R}) = \frac{1}{n} \sum_{j=1}^n A(\mathbf{R}_1^{(j)}, \dots, \mathbf{R}_N^{(j)}) \quad (1.91)$$

Pour la suite, il convient de réécrire l'hamiltonien du polymère libre dans la base des modes normaux de ce dernier :

$$H_n^0(\tilde{\mathbf{p}}, \tilde{\mathbf{R}}) = \sum_{\alpha=1}^N \sum_{k=-l}^{l+1} \left(\frac{[\tilde{\mathbf{p}}_{\alpha}^{(k)}]^2}{2m_{\alpha}} + \frac{1}{2} m_{\alpha} \omega_k^2 [\tilde{\mathbf{R}}_{\alpha}^{(k)}]^2 \right) \quad (1.92)$$

avec :

$$\omega_k = 2\omega_n \sin(k\pi/n) \quad (1.93)$$

et

$$\begin{aligned} \tilde{\mathbf{p}}_{\alpha}^{(k)} &= \sum_{j=1}^n \mathbf{p}_{\alpha}^{(j)} C_{jk} \\ \tilde{\mathbf{R}}_{\alpha}^{(k)} &= \sum_{j=1}^n \mathbf{R}_{\alpha}^{(j)} C_{jk} \end{aligned} \quad (1.94)$$

où C_{jk} sont les éléments de la matrice de transformation qui pour n pair s'écrit :

$$C_{jk} = \begin{cases} \sqrt{1/n} & \text{si } k = 0 \\ \sqrt{2/n} \cos(2\pi jk/n) & \text{si } 1 \leq k \leq n/2 - 1 \\ \sqrt{1/n} (-1)^j & \text{si } k = n/2 \\ \sqrt{2/n} \sin(2\pi jk/n) & \text{si } n/2 + 1 \leq k \leq n - 1 \end{cases} \quad (1.95)$$

Dans l'équation (1.92) **il est possible d'attribuer une masse fictive aux différents modes normaux** afin de modifier leur fréquence propre. Comme ces masses n'apparaissent pas dans les équations à venir nous ne les faisons pas apparaître explicitement. Toutefois, il faut savoir que le choix de ces masses fictives est à la base des méthodes **PIMD**, **RPMD** (pour *Ring Polymer Molecular Dynamics* en anglais) et **CMD** (pour *Centroid Molecular Dynamics* en anglais). Nous ne présenterons pas en détails ces trois méthodes mais voyons brièvement les différents choix de masses qui les différencient. Dans l'approche PIMD, **les masses fictives**

associées aux modes normaux sont choisies de sorte à ce que tous les modes normaux vibrent à la même fréquences. Cela permet d'optimiser l'efficacité de l'échantillonnage de l'espace des phase du polymère cyclique. De plus, pour assurer un échantillonnage ergodique, il convient d'attacher un thermostat local à chaque degrés de liberté. Ainsi, **la méthode PIMD permet d'échantillonner efficacement la PES**. En revanche, elle ne permet pas de calculer des fonctions de corrélation à temps réel et donc de calculer des propriétés dépendantes du temps. **Cette méthode s'avère cependant très efficace pour obtenir un échantillonnage des configurations atomiques incluant les NQEs**. Dans la méthode RPMD [220], **les masses attribuées aux modes normaux correspondent aux masses physiques** (somme des masses atomiques au sein d'une réplique). La dynamique des modes normaux a donc un sens physique. On peut donc calculer des fonctions de corrélation à temps réel à partir de simulations RPMD à condition de ne pas employer de thermostat, mais ce choix induit des problèmes d'ergodicité. Une solution à ce problème consiste à moyenner les fonctions de corrélation sur différentes trajectoires RPMD, dont les conditions initiales ($\mathbf{p}_0; \mathbf{R}_0$) sont issues d'un équilibre thermodynamique calculé par une simulation PIMD thermostatée. Un autre problème de l'utilisation des masses physiques dans la RPMD réside dans les modes normaux du polymère libre. Ces modes vibrent à des fréquences se situant généralement dans la même gamme que celles du système physique, ce qui peut engendrer un couplage non physique entre les modes du RP et les modes du système physique. Ce couplage entraîne des complications pour l'étude de propriétés vibrationnelles comme par exemple les spectres d'absorption infrarouge (IR). Enfin, dans la méthode CMD [221], **le mode propre du centroïde $w_{k=0}$ se voit associer une masse physique** tandis que **les masses associées aux autres modes sont choisies de sorte à optimiser l'échantillonnage**.

Intégration des équations du mouvement

Les équations du mouvement sont intégrées selon un algorithme similaire au Velocity Verlet. Comme l'hamiltonien totale $H_n(\mathbf{p}, \mathbf{R})$ est la somme de l'hamiltonien du polymère cyclique libre $H_n^0(\mathbf{p}, \mathbf{R})$ et du potentiel physique $V_n(\mathbf{R})$, on peut faire évoluer les quantités de mouvement et les positions atomiques sous l'effet de $H_n^0(\mathbf{p}, \mathbf{R})$ et $V_n(\mathbf{R})$ séparément. **L'algorithme de propagation est composé**

1. MÉTHODES ET OUTILS DE SIMULATION

de six étapes :

1. Les quantités de mouvements sont propagées de $\frac{1}{2} \delta t$ sous l'effet de $V_n(\mathbf{R})$:

$$\mathbf{p}_\alpha^{(j)} \leftarrow \mathbf{p}_\alpha^{(j)} - \frac{\delta t}{2} \frac{\partial V(R_1^{(j)}, \dots, R_N^{(j)})}{\partial R_\alpha^{(j)}} \quad (1.96)$$

2. Les quantités de mouvement et les positions sont transformées dans la base des modes normaux :

$$\begin{aligned} \tilde{\mathbf{p}}_\alpha^{(k)}(t) &\leftarrow \sum_{j=1}^n \mathbf{p}_\alpha^{(j)}(t) C_{jk} \\ \tilde{\mathbf{R}}_\alpha^{(k)}(t) &\leftarrow \sum_{j=1}^n \mathbf{R}_\alpha^{(j)}(t) C_{jk} \end{aligned} \quad (1.97)$$

3. Les quantités de mouvement et les positions sont propagées de δt sous l'effet de $H_n^0(\mathbf{p}, \mathbf{R})$:

$$\begin{pmatrix} \tilde{\mathbf{p}}_\alpha^{(k)}(t + \delta t) \\ \tilde{\mathbf{R}}_\alpha^{(k)}(t + \delta t) \end{pmatrix} \leftarrow \begin{pmatrix} \cos(\omega_k \delta t) & -m_\alpha \omega_k \sin(\omega_k \delta t) \\ [1/m_\alpha \omega_k] \delta t & \cos(\omega_k \delta t) \end{pmatrix} \begin{pmatrix} \tilde{\mathbf{p}}_\alpha^{(k)}(t) \\ \tilde{\mathbf{R}}_\alpha^{(k)}(t) \end{pmatrix} \quad (1.98)$$

4. Les quantités de mouvement et les positions sont transformées dans la base des répliques :

$$\begin{aligned} \mathbf{p}_\alpha^{(j)}(t + \delta t) &\leftarrow \sum_{k=0}^{n-1} C_{jk} \tilde{\mathbf{p}}_\alpha^{(k)}(t + \delta t) \\ \mathbf{R}_\alpha^{(j)}(t + \delta t) &\leftarrow \sum_{k=0}^{n-1} C_{jk} \tilde{\mathbf{R}}_\alpha^{(k)}(t + \delta t) \end{aligned} \quad (1.99)$$

5. Le potentiel $V(\mathbf{R}_1^{(j)}, \dots, \mathbf{R}_N^{(j)})$ et les forces associées aux nouvelles coordonnées sont calculés pour chaque réplique.
6. Les quantités de mouvement sont à nouveau propagées sur $\frac{1}{2} \delta t$ sous l'effet de $V_n(\mathbf{R})$:

$$\mathbf{p}_\alpha^{(j)} \leftarrow \mathbf{p}_\alpha^{(j)} - \frac{\delta t}{2} \frac{\partial V(R_1^{(j)}, \dots, R_N^{(j)})}{\partial R_\alpha^{(j)}} \quad (1.100)$$

Maintenant que nous avons vu comment intégrer les équations du mouvement dans le formalisme PIMD, nous allons voir comment les thermostats stochastiques présentés dans la section (1.3.2) permettent d'optimiser l'exploration de la PES en PIMD, et comment dans certains cas ils peuvent même résoudre des problèmes inhérents à la méthode PIMD tel que le couplage entre les modes physiques et ceux du polymère cyclique dans l'approche RPMD.

Thermostats stochastique pour le PIMD

Bien qu'il soit possible d'utiliser un thermostat NHC dans des simulations PIMD [222], nous ne discuterons ici que des thermostats stochastiques basés sur l'équation de Langevin. Les trois thermostats que nous présenterons sont des adaptations au formalisme PIMD des thermostats de Langevin présentés dans la section (1.3.2).

Thermostats PILE-L et PILE-G

En PIMD, les thermostats sont appliqués dans la base des modes normaux avant les équations (1.96) et (1.100) dans l'algorithme d'intégration des équations du mouvement. **Il faut donc appliquer les trois étapes suivantes avant les équations (1.96) et (1.100) :**

1. Les quantités de mouvement sont transformées dans la base des modes normaux :

$$\tilde{\mathbf{p}}_{\alpha}^{(k)} \leftarrow \sum_{j=1}^n \mathbf{p}_{\alpha}^{(j)} C_{jk} \quad (1.101)$$

2. Les quantités de mouvement sont renormalisées dans la base des modes normaux selon l'équation de Langevin :

$$\tilde{\mathbf{p}}_{\alpha}^{(k)} \leftarrow c_1^{(k)} \tilde{\mathbf{p}}_{\alpha}^{(k)} + \sqrt{\frac{m_{\alpha}}{\beta_n}} c_2^{(k)} \xi_i^{(k)} + \sqrt{\frac{m_{\alpha}}{\beta_n} (1 - e^{-(\delta t/2)\gamma^{(k)}})} \xi_i^{(k)} \quad (1.102)$$

3. Les quantités de mouvement sont transformées dans la base des répliques :

$$\mathbf{p}_{\alpha}^{(j)} \leftarrow \sum_{k=0}^{n-1} C_{jk} \tilde{\mathbf{p}}_{\alpha}^{(k)} \quad (1.103)$$

1. MÉTHODES ET OUTILS DE SIMULATION

Les coefficients $c_1^{(k)}$ et $c_2^{(k)}$ de l'équation (1.102) s'écrivent :

$$c_1^{(k)} = e^{-(\delta t/2)\gamma^{(k)}} \quad (1.104)$$

et

$$c_2^{(k)} = \sqrt{\frac{m_\alpha}{\beta_n} (1 - e^{-(\delta t/2)\gamma^{(k)}})} \quad (1.105)$$

L'intérêt d'appliquer le thermostat aux quantités de mouvement dans la base des modes normaux est le suivant : la valeur du coefficient de friction dans la base des modes normaux $\gamma^{(k)}$ permettant l'exploration la plus rapide de l'espace des phases des modes excités ($k > 0$) du RP libre peut être calculée analytiquement et vaut $\gamma^{(k)} = 2\omega_k$. Pour le mode du centroïde ($k = 0$), $\omega_k = 0$. **Il faut donc définir une constante de temps τ_0 pour le centroïde comme lorsque l'on utilise le thermostat de Langevin en MD classique.** Le coefficient de friction pour le centroïde vaut $\gamma^{(0)} = 1/\tau_0$.

Les modes de vibration du RP d'un système moléculaire ne sont pas exactement les mêmes que ceux du RP libre. Cependant pour les modes excités de haute fréquence, les fréquences propres sont généralement proches dans les deux cas et découplées des fréquences propres du système étudié. Ces modes normaux excités de haute fréquence devraient donc être correctement thermostatés sans pour autant induire un couplage fictif et non physique avec le système moléculaire. La version du thermostat PILE (pour *Path-Integral Langevin Equation* en anglais) que l'on vient de présenter est **une version locale** dans le sens où chaque mode normal est thermostaté indépendamment des autres. Comme nous l'avons expliqué plus haut, il peut-être parfois intéressant d'employer un thermostat moins agressif. **Une version globale** a également été développée pour répondre à ce problème. Dans la version globale, les modes normaux excités sont thermostatés de la même façon que dans la version locale. Le mode du centroïde ($k = 0$) quant à lui est thermostaté avec le stochastic velocity rescale, qui va agir de façon globale sur toutes les quantités de mouvement en les renormalisant par un coefficient α identique à celui calculé dans l'équation (1.76) :

$$\tilde{\mathbf{p}}_\alpha^{(k)} \leftarrow \alpha \tilde{\mathbf{p}}_\alpha^{(k)} \quad (1.106)$$

Les versions locale et globale de ce thermostat sont respectivement nommées **PILE-L** et **PILE-G**.

Thermostats PI+GLE et dérivés

Le thermostat basé sur l'équation de Langevin généralisée présenté dans la section (1.3.2) peut aussi être employé pour thermostatier efficacement une simulation PIMD. Ce thermostat peut couvrir une large gamme de fréquences couvrant celles du polymère cyclique libre d'une part et celles du système physique d'autre part. Dans ce cas, la renormalisation des quantités de mouvement se fait dans l'espace des répliques. Les équations (1.101), (1.102) et (1.103) sont simplement remplacées par :

$$\mathbf{p}_\alpha^{(j)} \leftarrow \mathbf{C}_1 \mathbf{p}_\alpha^{(j)} + \sqrt{\frac{m_\alpha}{\beta_n}} \mathbf{C}_2 \xi_\alpha^j \quad (1.107)$$

où \mathbf{C}_1 et \mathbf{C}_2 sont équivalents aux équations (1.104) et (1.105) sous forme matricielle. Ces matrices sont calculées à partir de la matrice A . **Ce thermostat permet de réduire drastiquement le nombre de répliques nécessaires à la convergence des effets quantiques nucléaires dans une simulation PIMD.**

L'équation de Langevin généralisée peut également servir à thermostatier des simulations RPMD. Comme nous l'avons vu précédemment, la RPMD souffre d'un manque d'ergodicité mais elle ne peut pas être thermostatée sans quoi on ne pourrait plus calculer de fonction de corrélation temporelle et donc calculer des propriétés dépendantes du temps. L'autre problème que l'on rencontre en RPMD est **le couplage non physique entre les fréquences de vibration du système physique et celles du polymère cyclique**. L'équation de Langevin généralisée peut résoudre ce deuxième problème. Elle peut être appliquée aux modes normaux du polymère pour en atténuer les fréquences de vibrations qui se trouvent dans la gamme de fréquences d'intérêt du système physique pour éviter un couplage fictif. Cette méthode est particulièrement intéressante pour simuler les spectres infrarouges. Là où la méthode CMD permettait d'étudier les fréquences hautes et la RPMD les fréquences basses, **la TRPMD (pour *Thermostated Ring-Polymer Molecular Dynamics*) [223, 224] peut a priori traiter l'intégralité du spectre de fréquences.**

1. MÉTHODES ET OUTILS DE SIMULATION

1.3.4 Dynamique moléculaire *Parallel-Tempering*

Pendant le temps d'une mesure expérimentale, le système étudié se trouve statistiquement dans toutes les configurations disponibles à la température d'étude. En dynamique moléculaire, les temps de simulation accessibles sont beaucoup plus courts (voir figure 3), même pour les méthodes les moins coûteuses en temps de calcul. Or les surfaces d'énergie potentielle sont souvent très complexes et présentent de nombreux minima locaux et cela d'autant plus que le nombre de degrés de liberté du système est élevé. Parfois, ces minima sont séparés par des barrières énergétiques élevées qui ne peuvent pas être franchies à basse température ou alors sur des temps de simulation beaucoup trop longs pour être accessibles. Cela engendre des simulations non ergodiques à partir desquelles il n'est pas possible d'extraire de résultat statistique. En général, les barrières franchissables à une température donnée sont de l'ordre de $k_b T$. La figure 1.3 permet de visualiser ce concept. Sur cette figure sont représentées en (a) les zones d'une PES accessibles selon la température de simulation et en (b) la densité d'état pour différentes températures. Admettons que le système se trouve initialement au fond du puits de potentiel le plus profond. À basse température (bleu), le système ne peut que visiter le fond de ce puits. À température intermédiaire (violet), le système pourrait se trouver dans les deux puits de potentiels. Cependant il ne possède pas assez d'énergie pour pouvoir franchir la barrière d'énergie potentielle séparant les deux puits. À haute température (rouge), le système peut facilement passer la barrière d'énergie potentielle et visiter les deux puits de potentiel.

La méthode *Parallel Tempering* (PT) consiste à simuler le même système à différentes températures et à échanger périodiquement les répliques entre les différents ensembles visités. Originellement nommée *Replica Exchange* ou *Markov Chain*, elle a été appliquée pour la première fois par Robert H. Swendsen et Jian Sheng Wan en 1986 afin d'obtenir les différents minima d'un système de verre de spin [225]. Dans cette étude, l'échange de répliques n'était pas combiné à la MD mais à la méthode Monte-Carlo. On parle alors de Markov Chain Monte-Carlo (MCMC). La méthode MCMCs a ensuite été massivement étudiée et améliorée dans les années 90 [226, 227, 228, 229]. La première application de cette méthode à un système moléculaire a été réalisée en 1999 par Marco Falconi et Michael W. Deem pour l'étude structurale de zéolithes [230]. La même

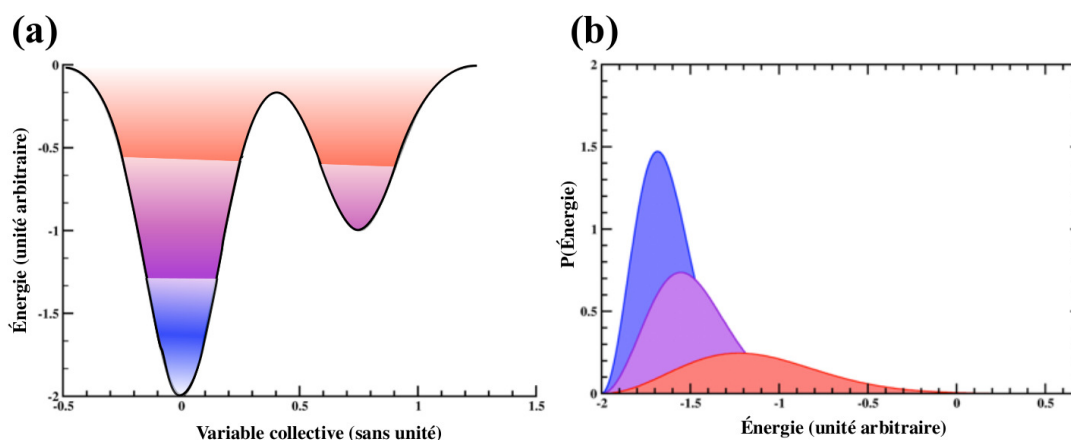


FIGURE 1.3 – (a) Représentation schématique des zones de la PES accessibles en MD en fonction de la température de la simulation. À basse température (bleu), le système a accès aux conformations au fond du puits le plus profond. A température intermédiaire (violet), le système a accès à une partie plus grande du puits le plus profond et à une partie de l'autre puits. En revanche, à cette température, il n'y a pas assez d'énergie pour que le système puisse franchir la barrière séparant les deux puits. Dans une simulation de MD, le système restera bloqué dans le puits où il se trouve initialement. À haute température (rouge), le système a accès à l'intégralité des deux puits de potentiel et il possède assez d'énergie pour pouvoir passer la barrière les séparant. (b) Probabilité de trouver le système à une certaine énergie à basse (bleu), moyenne (violet) et haute (rouge) température.

année, Yuji Sugita et Yuko Okamoto proposent de combiner les chaînes des Markov et la MD [174]. **C'est l'apparition de l'approche PTMD.** Dans cet article, les auteurs utilisent la méthode PTMD pour améliorer l'échantillonnage des configurations dans une étude portant sur le repliement des protéines. À partir de cette date, la méthode PTMD sera intégrée dans un nombre croissant de programmes de MD classique basés sur des potentiels de types champ de force.

Le fonctionnement de la méthode PTMD est représenté sur la figure 1.4. Le même système moléculaire est répliqué N fois et chaque réplique évolue en MD dans l'ensemble NVT à une température différente. Les répliques à basse température vont efficacement explorer le fond des puits de la PES tandis que les répliques à haute température pourront aisément franchir les barrières d'énergie les plus hautes. Tous les X pas de MD, les répliques sont échangées avec une probabilité d'acceptation de

1. MÉTHODES ET OUTILS DE SIMULATION

l'échange définie par le critère de Metropolis-Hasting [231] :

$$p = \min \left(1, e^{(E_i - E_j) \left(\frac{1}{k_b T_i} - \frac{1}{k_b T_j} \right)} \right) \quad (1.108)$$

où E_i et E_j sont les énergies potentielles des répliques évoluant respectivement aux températures T_i et T_j avec $T_j > T_i$. D'après ce critère, si $E_i > E_j$, alors l'échange est accepté. En effet, dans ce cas, l'argument de l'exponentielle est positif et la probabilité d'accepter l'échange est alors supérieure à 1. Au contraire, si $E_i < E_j$, alors l'exponentielle est comparée à un nombre tiré dans une distribution aléatoire uniforme entre 0 et 1. Si l'exponentielle est plus grande que ce nombre alors l'échange est accepté. Sinon, il est refusé. **La qualité d'une simulation PTMD dépend de plusieurs paramètres.** La température la plus basse T_1 doit être choisie de sorte à explorer efficacement les fonds de puits de potentiel tandis que la température la plus haute T_N doit permettre de franchir rapidement les barrières d'énergie séparant deux puits voisins. Enfin, **le nombre de températures entre T_1 et T_N ainsi que leur distribution doit permettre le recouvrement optimal des densités d'états entre températures voisines.** Si le recouvrement est trop faible, les échanges seront refusés trop souvent et les répliques mettront trop de temps à diffuser dans l'espace des températures. Au contraire, si le recouvrement est trop important, les répliques à des température voisines échantillonneront pratiquement le même ensemble canonique. On aura alors des redondances entre les répliques et un gaspillage des ressources de calcul.

La méthode PTMD est une **méthode d'optimisation de l'échantillonnage**, ou **méthode d'échantillonnage accéléré**, non biaisée. Cela implique qu'aucune force non physique n'est appliquée au système. Parmi d'autres méthodes non biaisées on retrouve : **l'échantillonnage de chemins de transition** [233, 234, 235], **la dynamique accélérée en température** [236, 237, 238], ou **le recuit simulé** [77, 239, 240]. Il existe aussi des méthodes biaisées telles que **la metadynamics** [241, 242, 243, 244, 245, 246], ou **l'échantillonnage en parapluie** [247, 248, 249]. N'utilisant pas de potentiel extérieur, les méthodes non biaisées échantillonnent le bon ensemble canonique. Il est donc possible d'extraire directement des propriétés thermodynamiques de telles simulations.

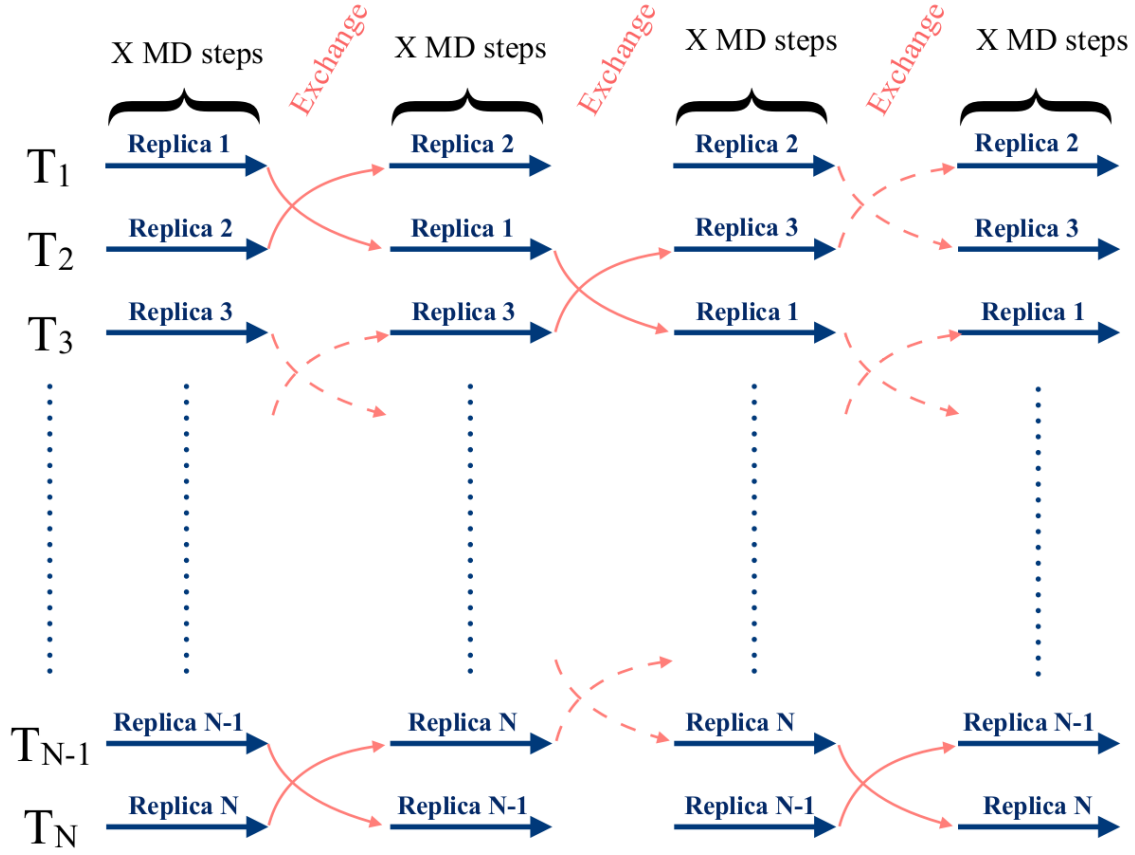


FIGURE 1.4 – Tirée de [232]. Représentation schématique d’une simulation PTMD. N répliques du même système moléculaire sont simulées à différentes températures allant de 1 à N . Après X pas de MD (flèche bleu rectiligne), les répliques sont échangées entre températures voisines selon le critère de Metropolis-Hasting (flèches rouges incurvées). Les échanges acceptés sont représentés par des flèches pleines tandis que les échanges refusés sont représentés par des flèches en pointillés. Une fois que les répliques sont échangées, la simulation reprend pour N pas de MD et ainsi de suite.

Dans le chapitre 2, nous présenterons l’implémentation de la méthode PTMD synchrone et asynchrone dans le programme de DFT deMon2k ainsi que son application pour l’étude de la transition de phase dans les agrégats d’eau $(\text{H}_2\text{O})_6$ et $(\text{H}_2\text{O})_8$. Dans le chapitre 4, nous présenterons l’utilisation de la méthode PTMD couplée à l’approche PIMD pour étudier l’impact des NQEs dans les distributions d’isomères de ces mêmes agrégats d’eau.

1. MÉTHODES ET OUTILS DE SIMULATION

1.3.5 Extraction de propriétés issues de dynamiques moléculaires

Comme nous venons de le voir, la MD permet d'échantillonner un ensemble thermodynamique. Il est donc possible de calculer des moyennes d'ensemble à partir des états échantillonnés au cours de la dynamique. Dans cette dernière section, nous allons voir comment calculer une distribution d'isomères, la capacité calorifique et les spectres d'absorption IR à partir de simulations de MD.

Distribution d'isomères

Pour faire le lien entre les changements de conformation et les changements de pente dans les courbes de capacité calorifique lors de transitions de phase dans les agrégats d'eau, nous avons analysé les distributions des isomères de plus basse énergie en fonction de la température. Pour cela, il faut disposer d'un ensemble de configurations géométriques à différentes températures. Ces conformations peuvent être obtenues à l'aide d'une simulation PTMD. Les structures de plus basse énergies sont obtenues en optimisant un grand nombre de géométries issues de trajectoires réalisées à différentes températures. En ordonnant les structures ainsi obtenues par énergie croissante, il est possible de discriminer différents isomères. Concernant les agrégats d'eau que nous avons étudiés ($(\text{H}_2\text{O})_6$ et $(\text{H}_2\text{O})_8$), nous avons décidé de caractériser les isomères par le nombre de cycles O—O et leur taille. Par taille nous entendons ici le nombre d'atomes d'oxygène par cycle. Pour définir les cycles, nous avons utilisé l'algorithme de Xianglong Yuan et Alistair N. Cormak [250]. Cet algorithme est basé sur la méthode des cycles primitifs. Il permet de faire une analyse topologique de systèmes cycliques dans lesquels les atomes sont traités comme des nœuds et les liaisons chimiques comme des liens entre ces nœuds. Pour cela, il faut définir la taille maximum d'un lien. Dans toutes les distributions d'isomères que nous avons étudiées, nous avons choisi une distance O—O limite de 3,25 Å.

Il existe d'autres méthodes telles que la méthode des arbres binaires enracinés qui a été employée pour décrire les réseaux de liaisons hydrogènes dans des agrégats d'eau protonés [251, 252, 253, 254] ou la méthode SOAP (pour *Smooth Overlap of Atomic Positions* en anglais) [255].

Capacité calorifique

L'étude de la capacité calorifique C en fonction de la température permet de caractériser les transitions de phase. En effet, lors d'une transition de phase (ou d'une isomérisation dans le cas d'agrégats moléculaires) la pente de la courbe de capacité calorifique change. Dans cette thèse, les capacités calorifiques ont été calculées à l'aide de la méthode des histogrammes multiples développée par Pierre Labastie et Robert L. Whetten [256, 257]. Cette méthode consiste à évaluer la capacité calorifique à une température donnée en fonction de la probabilité pour que le système étudié se trouve à une certaine énergie. Cette probabilité est calculée à partir de l'échantillonnage de la PES à différentes températures. Cela peut être fait en réalisant des simulations de MD à différentes températures ou à l'aide de la méthode PTMD. La distribution d'énergie se présente sous la forme d'un histogramme comme ceux de la figure 1.3 (b). À partir de cette probabilité, on peut calculer l'entropie puis la fonction de partition du système. La fonction de partition donne accès à l'énergie interne U . La capacité calorifique C est définie comme la dérivée de l'énergie interne par rapport à la température et se calcule par :

$$C(T) = \frac{\partial U}{\partial T} = \frac{N_{deg} k_B}{2} + \frac{\langle E^2 \rangle - \langle E \rangle^2}{k_B T^2} \quad (1.109)$$

où E est l'énergie potentielle du système.

Spectre d'absorption infrarouge

Les spectres d'absorption IR peuvent être calculés à partir de la transformée de Fourier de la fonction d'autocorrélation temporelle du moment dipolaire :

$$\alpha(\omega) \propto \omega^2 \int_0^{+\infty} \langle \boldsymbol{\mu}(0) \cdot \boldsymbol{\mu}(t) \rangle e^{-i\omega t} dt \quad (1.110)$$

où ω est le nombre d'onde et $\boldsymbol{\mu}$ le moment dipolaire.

1. MÉTHODES ET OUTILS DE SIMULATION

Calcul de tenseurs d'écrantage magnétique par la méthode ADFT-GIAO

La spectroscopie par **résonance magnétique nucléaire** (RMN) compte parmi les méthodes les plus couramment employées par les expérimentateurs pour obtenir une description de la structure des molécules. Les paramètres d'écrantage obtenus en RMN donnent une information précise sur la distribution de la densité électronique aux abords des noyaux étudiés à partir desquelles il est possible d'extraire des propriétés structurales et électroniques. Le calcul théorique de ces paramètres d'écrantage est fondamental pour pouvoir comparer des résultats obtenus en simulations aux résultats expérimentaux. Dans cette section nous expliquerons les principales équations de la méthode ADFT-GIAO telle qu'elle est implémentée dans le code deMon2k. La méthode **GIAO** (**pour *gauge-including atomic orbital en anglais***) a initialement été développée par Fritz London en 1937 pour étudier les relations de courant annulaire dans des composés aromatiques [258], puis a ensuite été développée par d'autres personnes [259, 260, 261, 262, 263, 264, 265]. La première combinaison de la méthode GIAO avec la DFT a été réalisée en 1990 par Karlheinz Friedrich *et al.* [266] en se basant sur la formulation GIAO de Robert Ditchfield [267] puis a été amélioré par Georg Schreckenbach et Tom Ziegler [268].

La méthode **ADFT-GIAO** repose sur l'utilisation d'orbitales atomiques indépendantes de jauge (**GIAO** pour *Gauge-Invariant Atomic Orbitals* en anglais). L'utilisation de ce type d'orbitales rend les résultats indépendants de l'origine de jauge. Ainsi, les orbitales atomiques $\mu(\mathbf{r})$ et $\nu(\mathbf{r})$ dans l'équation (1.36) sont remplacées par des orbitales GIAO qui s'écrivent comme une combinaison linéaire d'orbitales atomiques de type gaussienne contractée :

$$\phi_a(\mathbf{r}, \mathcal{H}) = a(\mathbf{r})e^{-\frac{i}{2}\mathcal{H}\cdot(\mathbf{A}-\mathbf{G})\times\mathbf{r}} \quad (1.111)$$

où \mathcal{H} représente le champ magnétique externe, $a(\mathbf{r})$ une orbitale atomique qui ne dépend pas du champ magnétique externe, et \mathbf{A} et \mathbf{G} représentent respectivement les vecteurs positions de l'atome A sur lequel l'orbitale $a(\mathbf{r})$ est centrée et de l'origine de jauge.

Les orbitales moléculaires s'écrivent :

$$\psi_j(\mathbf{r}, \mathcal{H}) = \sum_{\mu} c_{\mu j}(\mathcal{H}) \phi_{\mu}(\mathbf{r}, \mathcal{H}) \quad (1.112)$$

où ψ_j est la j^{me} orbitale moléculaire et $c_{\mu j}$ son coefficient de décomposition dépendant du champ magnétique \mathcal{H} .

L'interaction entre le champ magnétique et la structure électronique est introduite en modifiant l'opérateur impulsion dans le terme $H_{\mu\nu}$ de l'équation (1.36) :

$$\mathbf{p} \rightarrow \pi = -i\nabla + \mathcal{A} \quad (1.113)$$

où $-i\nabla$ est l'opérateur quantique d'impulsion indépendant champ magnétique et \mathcal{A} le potentiel vecteur à la position de l'électron.

En traitant le champ magnétique externe, le champ magnétique local et le moment magnétique de spin comme des perturbations de l'hamiltonien indépendant du champ, il est possible de calculer le tenseur d'écrantage pour un noyau i comme :

$$\begin{aligned} \sigma_{xy,i} &= \left(\frac{\partial^2 E}{\partial \mathcal{M}_y(i) \partial \mathcal{H}_x} \right)_{\mathcal{H}=0} \\ &= \sum_{\mu,\nu} P_{\mu\nu} H_{\mu\nu}^{(xy)}(i) + \sum_{\mu,\nu} P_{\mu\nu}^{(x)} H_{\mu\nu}^{(y)}(i) \end{aligned} \quad (1.114)$$

où $\mathcal{M}_y(i)$ est la composante y du moment magnétique nucléaire intrinsèque de i et les éléments de matrice $H_{\mu\nu}^{(y)}$ et $H_{\mu\nu}^{(xy)}$ représentent les dérivées première et seconde des éléments de l'hamiltonien de cœur perturbé par rapport au champ magnétique et aux composantes du moment nucléaire. $P_{\mu\nu}$ et $P_{\mu\nu}^{(x)}$ représentent respectivement les matrices de densité et de densité perturbée. Les éléments de l'hamiltonien perturbé qui ne s'annulent pas dans les dérivées première et seconde s'écrivent :

$$\begin{aligned} H_{ab} &= \langle \phi_a^* | \hat{H}^{(0)} | \phi_b \rangle + \vec{\mathcal{H}}^T \langle \phi_a^* | \hat{\xi}^p | \phi_b \rangle \\ &\quad + \sum_i \vec{\mathcal{M}}^T(i) \langle \phi_a^* | \vec{\sigma}^p(i) | \phi_b \rangle \\ &\quad + \sum_i \vec{\mathcal{H}}^T \langle \phi_a^* | \hat{\sigma}^d(i) | \phi_b \rangle \vec{\mathcal{M}}(i). \end{aligned} \quad (1.115)$$

1. MÉTHODES ET OUTILS DE SIMULATION

Le premier terme de l'équation (1.115) correspond à la valeur moyenne de l'hamiltonien non perturbé indépendant de l'origine de jauge sur les GIAOs. Les deuxième et troisième termes correspondent à la contribution paramagnétique de la magnétisabilité et de l'écrantage respectivement. Enfin, le dernier terme correspond à la contribution diamagnétique de l'écrantage.

Dans les travaux réalisés au cours de cette thèse nous avons étudié les paramètres RMN d'atomes de ^{13}C dans des ions carboxylates solvatés par un agrégat d'eau (voir chapitre 3). Pour pouvoir comparer nos résultats aux données expérimentales nous avons calculé les déplacements chimiques isotropes des ^{13}C au cours de simulations de dynamique moléculaire. Le déplacement chimique isotrope d'un atome de carbone i $\delta_{iso,i}$ se calcule selon l'expression :

$$\delta_{iso,i} = \sigma_{iso,TMS} - \sigma_{iso,i} \quad (1.116)$$

où $\sigma_{iso,TMS}$ et $\sigma_{iso,i}$ correspondent respectivement à l'écrantage isotrope des ^{13}C du tétraméthylsilane (TMS) et celui d'un atome de carbone i . L'écrantage isotrope $\sigma_{iso,i}$ s'exprime comme le tiers de la trace du tenseur d'écrantage :

$$\sigma_{iso,i} = \frac{1}{3} [\sigma_{xx,i} + \sigma_{yy,i} + \sigma_{zz,i}] \quad (1.117)$$

Implémentation de la méthode
Parallel-Tempering Molecular
Dynamics dans deMon2k et
application à l'étude d'agrégats d'eau

Chapitre 2

Implémentation de la méthode Parallel-Tempering Molecular Dynamics dans deMon2k et application à l'étude d'agrégats d'eau

2.1 Introduction

Nous avons vu dans la section 1.3.4 que l'algorithme d'échantillonnage accéléré *Parallel-Tempering Molecular Dynamics* (PTMD) permet d'**accélérer considérablement l'exploration de surfaces d'énergie potentielle**. Bien que la méthode dite « *Parallel-Tempering* » puisse s'appliquer aux simulations de dynamique moléculaire (MD) comme aux simulations Monte-Carlo (MC), **nous ne traiterons ici que de son application en dynamique moléculaire**. Nous avons également vu que cet algorithme étant intrinsèquement parallèle, **son implémentation s'adapte parfaitement à l'architecture parallèle des supercalculateurs modernes**. En effet, en assignant un nœud de calcul par température, chaque réplique du système peut évoluer dans son propre ensemble thermodynamique indépendamment des autres répliques et une communication entre les différents nœuds est nécessaire uniquement lors des étapes d'échange.

A ce stade, soulignons ici que **deux types d'implémentation peuvent être**

2. IMPLÉMENTATION DE LA MÉTHODE PTMD DANS DEMON2K ET APPLICATION À L'ÉTUDE D'AGRÉGATS D'EAU

envisagées. Dans la **première**, ce sont les **températures** qui **sont échangées**, un nœud ne traite alors que la dynamique d'une seule réplique tout au long de la simulation. Ce sont uniquement les températures des thermostats qui sont échangées lors de l'étape d'échange. Cette approche est plus efficace car elle entraîne moins de données à échanger lors des communications entre nœuds. En revanche, elle nécessite un effort supplémentaire de post-traitement pour réordonner les fichiers de sortie afin de générer un fichier par température plutôt qu'un fichier par réplique. Dans la **seconde**, ce sont les **répliques** (coordonnées atomiques et vitesses) qui **sont échangées** entre nœuds lors de l'étape d'échange. Dans ce cas, la quantité de données échangées est plus grande mais les fichiers de sortie sont directement utilisables pour analyse sans étape de post-traitement. Lorsque l'on considère des simulations longues, ou sur de grands systèmes, ces fichiers peuvent atteindre des tailles conséquentes, plusieurs dizaines de Go par fichier. Dans ce cas, cette étape de post-traitement n'est pas anodine et engendre de nombreuses phases d'écriture/suppression sur les disques de stockage. Dans ce qui suit, **c'est cette seconde approche qui sera considérée**, principalement pour éviter l'effort de post-traitement supplémentaire lors de l'utilisation intensive de l'algorithme.

Comme nous l'avons précisé dans la section 1.3.4, le travail d'implémentation que nous présentons dans ce chapitre a été motivé par plusieurs point, le premier d'entre eux ayant été de **mettre à la disposition de la communauté des chimistes théoriciens une implémentation native de l'algorithme PTMD dans un code de DFT**. En effet, si la méthode PTMD est d'ores et déjà implémentée dans la plus part des codes de dynamique moléculaire basés sur des champs de force comme les codes Abalone [269], Amber [270], Ascalaph Designer [269], GROMACS [71], ou LAMMPS [271], il n'existait pas, au début de ce projet, de programme de dynamique moléculaire basé sur un potentiel DFT et possédant une implémentation native de l'algorithme PTMD. À notre connaissance, la seule mention d'une telle implémentation se trouve dans la référence [272] dans laquelle les auteurs font mention d'une implémentation « *on the top of CP2K* ». Ce dernier est un code de DFT périodique utilisant une approche mixte fonctions gaussiennes et ondes planes pour la description de la structure électronique [273]. Toutefois, nous n'avons pu trouver d'autre mention de cette implémentation et elle n'apparaît pas à ce jour dans les fonctionnalités affichées de CP2K. Ainsi, **aucune implémentation de l'algorithme PTMD dans un code DFT**, et plus généralement *ab-initio*, uti-

lisant une base d'ondes planes ou une base de fonctions localisées, **n'a été reportée jusqu'à présent.**

Or, si les méthodes de type champ de force ou semi-empiriques telles que la DFTB sont rapides et permettent un grand nombre d'applications, **certains problèmes nécessitent l'utilisation d'une méthode certes plus coûteuse mais plus précises, par exemple la DFT.** Ainsi, actuellement, pour pouvoir réaliser une simulation PTMD en traitant la structure électronique au niveau DFT, une approche possible est de réaliser une interface entre un code de DFT et un logiciel externe qui traitera la dynamique moléculaire et les échanges de répliques. Par exemple, **i-PI** [99], qui est un programme initialement développé pour réaliser des simulations *Path-Integral Molecular Dynamics* (PIMD), (voir section 1.3.3 pour plus de détails), peut être employé pour réaliser des simulations PTMD. Il suffit pour cela de l'interfacer avec un programme pouvant calculer des énergies et des forces au niveau DFT. L'avantage de tels programmes externes est double. Ils peuvent être combinés avec *a priori* n'importe quel autre programme et peuvent donc être utilisés pour étudier un grand nombre de problèmes en fonction du code avec lequel ils sont interfacés. De plus, le travail d'implémentation pour réaliser l'interface est bien moindre que celui à fournir pour implémenter directement tous les algorithmes que contient le programme externe. Réaliser ce type d'interface avec des programmes externes présente toutefois plusieurs problèmes. Premièrement, compilation et utilisation du code sont rendues plus complexes ce qui peut être dissuasif pour certains utilisateurs. Ensuite, **les interfaces ne sont pas toujours optimisées aux architectures des supercalculateurs** ce qui peut limiter leur utilisation pour un certain nombre d'applications du fait d'une efficacité dégradée des programmes. C'est particulièrement le cas lorsque des communications inter-processus sont requises comme dans le cas de l'algorithme PTMD. Notons qu'en plus de cette approche par interface, une autre manière de réaliser des simulations PTMD avec un potentiel *ab-initio* est de réaliser N dynamiques moléculaires à différentes températures et d'utiliser un script externe pour traiter les échanges de répliques. C'est le cas par exemple de la suite d'outils MMTSB qui consiste en un ensemble de bibliothèques et d'utilitaires comportant la méthode PTMD [274]. Cette approche présente les mêmes inconvénients que les interfaces externes mentionnées ci-dessus.

C'est donc **cette absence de programme de DFT incluant nativement**

2. IMPLÉMENTATION DE LA MÉTHODE PTMD DANS DEMON2K ET APPLICATION À L'ÉTUDE D'AGRÉGATS D'EAU

l'algorithme PTMD, ainsi que les défauts des interfaces, qui nous ont motivé à implémenter cette approche dans le code deMon2k [98]. Comme présenté dans la section 1.2.3, deMon2k est un programme dédié au calcul de structure électronique par l'approche DFT qui utilise des combinaisons linéaires de fonctions gaussiennes pour la résolution des équations de Kohn-Sham. L'utilisation d'une base de fonctions auxiliaires qui permet d'éviter le calcul des intégrales de répulsion électronique à quatre centres **rend ce code particulièrement efficace et donc bien adapté à des simulations intensives de type PTMD**. Ce code intègre également un certain nombre de fonctionnalités parmi lesquelles l'approche *Auxiliary Density Functional Theory* (ADFT) [275], l'approche QM/MM (voir section 1.2.5) [276], une large variété de fonctionnelles LDA, GGA, meta-GGA et hybrides, des outils de dynamique moléculaire, la *time-dependent* ADFT, et **il est intégralement parallélisé à l'aide de bibliothèques MPI** (pour *Message Passing Interface* en anglais) [277] ce qui le rend particulièrement bien adapté à l'étude d'une large gamme de problème chimique pouvant aller jusqu'à l'échelle nanoscopique. À noter enfin qu'**une communauté active de développeurs s'attelle à le mettre à jour et à y implémenter de nouvelles fonctionnalités**. Comme, nous le verrons plus bas, une implémentation parallèle de l'algorithme PTMD dans un code qui possède nativement une structure parallèle pose un certain nombre de complications. En effet, **la condition première de notre travail d'implémentation était de ne modifier en aucun cas la structure parallèle MPI existante de deMon2k**. Dans le cas contraire, cela aurait engendré un travail considérable de vérification du code. Ainsi, comme décrit plus cas, notre implémentation parallèle de l'algorithme PTMD au dessus la couche de parallélisation préexistante a nécessité **le développement d'une architecture parallèle en deux couches**.

Le **principal défaut de la méthode PTMD** dans sa formulation dite « synchrone » résulte de la perte de temps de calcul, et donc de ressources de calcul, lors des phases d'échanges. En effet, s'il est nécessaire que toutes les répliques réalisent le même nombre de pas N de MD entre deux échanges, rien ne certifie qu'elles mettront le même temps à les réaliser. Ainsi, si une réplique a réalisé les N pas de MD déterminés depuis le dernier échange, la phase d'échange n'aura lieu que lorsque toutes les répliques auront parcouru ces N pas. Il apparaît donc clairement que le temps entre deux échanges consécutifs est déterminé par le temps que met la réplique la plus lente à réaliser ses N pas de dynamique moléculaire. Ainsi, entre

l'instant où une réplique, qui n'est pas la plus lente, a parcouru N pas et l'instant où l'échange a effectivement lieu, les ressources de calcul attribuées à cette réplique « en attente » sont inutilisées. Cette perte de temps et de ressources de calculs, due à la synchronisation des échanges, est multipliée par le nombre de répliques et par le nombre total d'échanges dans la simulation. **L'origine de cette désynchronisation peut être multiple.** Elle peut être due à l'**inhomogénéité des ressources de calcul** employées, c'est à dire à la différence de performance entre plusieurs nœuds, ou à la **méthode mise en œuvre pour le calcul de l'énergie**. En 2009, **Giovanni Bussi a publié une étude importante portant sur le problème de désynchronisation dans l'algorithme PTMD** [278]. Dans cette étude, Giovanni Bussi a comparé l'efficacité parallèle λ , définie comme le ratio entre le nombre de pas de MD total effectué lors d'une simulation PTMD avec échanges et une simulation équivalente sans échanges, d'une implémentation synchrone avec celle d'une implémentation asynchrone de la méthode PTMD. Une partie des résultats obtenus pour un liquide de Lennard-Jones est présentée sur la figure 2.1. On peut voir sur cette figure que **l'efficacité parallèle peut chuter jusqu'à 10% lorsque plus de 100 répliques sont utilisées et que les échanges sont réalisés de façon synchrone**. En revanche, l'efficacité parallèle de la version asynchrone reste constante à $\sim 90\%$ et ce, quelque soit le nombre de répliques utilisé. Selon l'auteur, cette différence considérable en terme de performances est due à une mise à jour de la liste des atomes les plus proches voisins plus fréquente aux hautes températures qu'aux basses températures. Cette source de désynchronisation est donc propre à un algorithme de champ de force. Dans le cas de calculs DFT, on peut imaginer que l'on rencontrera des problèmes similaires puisque les répliques évoluant à hautes températures donneront lieu à des géométries peut-être plus désordonnées, nécessitant plus de cycles pour attendre la convergence SCF. **Nous avons donc implémenté deux approches d'échange dans l'algorithme PTMD de deMon2k.** Une **approche synchrone** et une **approche asynchrone** que nous allons détailler plus bas.

Ce chapitre présentera de manière détaillée **le travail d'implémentation que nous avons réalisé de l'algorithme PTMD dans le code deMon2k**. Dans un premier temps, nous discuterons brièvement de la **structure parallèle originale du code deMon2k**. Nous expliquerons ensuite en détails la mise en place de la **2^{nde} couche de parallélisation MPI** nécessaire à l'implémentation parallèle de l'al-

2. IMPLÉMENTATION DE LA MÉTHODE PTMD DANS DEMON2K ET APPLICATION À L'ÉTUDE D'AGRÉGATS D'EAU

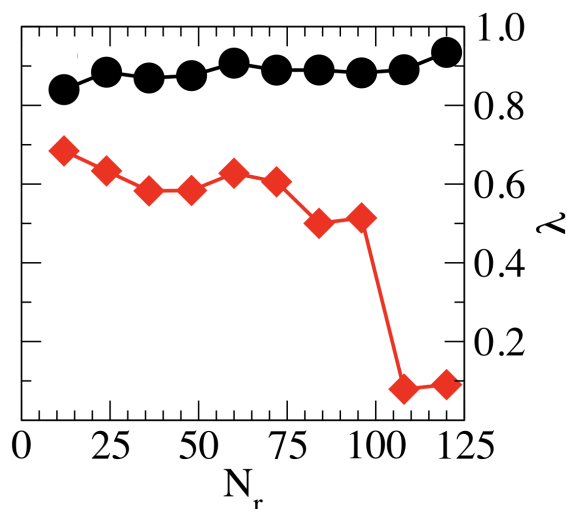


FIGURE 2.1 – Efficacité parallèle, λ , en fonction du nombre de répliques N_r pour une implémentation synchrone (losanges) et une asynchrone (cercles) de l’algorithme PTMD. λ est définie comme le ratio entre le nombre de pas de MD total effectué lors d’une simulation PTMD avec échanges et une simulation équivalente sans échanges. Figure tirée de la référence [278].

gorithme PTMD. Nous présenterons enfin les **implémentations de l’algorithme PTMD** sous ses formes synchrone et asynchrone. Cet algorithme sera finalement appliqué à l’étude des propriétés structurales et thermodynamiques des agrégats d’eau neutres $(\text{H}_2\text{O})_6$ et $(\text{H}_2\text{O})_8$, étude pour la première fois réalisée au niveau DFT. Il est à noter pour finir que ce travail a donné lieu à une publication en 2021 dans le journal *Theoretical Chemistry Accounts* [232].

2.2 Division des communicateurs MPI dans deMon2k

Dans cette section, nous allons présenter l’implémentation de **la structure parallèle MPI en deux couches** qui nous a ensuite permis d’implémenter l’algorithme PTMD. Nous avons développé en 2019 une première architecture et implémenté la méthode PTMD dans la version 6.0.2 de deMon2k. Ce travail a été présenté lors du *19th deMon developer Workshop* le 28 mai 2019 à Fréjus. Les discussions fructueuses qui ont suivi nous ont ensuite motivé à développer une nouvelle version

2.2 Division des communicateurs MPI dans deMon2k

de cette architecture parallèle en deux couches dans la version 6.0.8 de deMon2k **en collaboration avec Gerald Geudtner et Andreas M. Köster** du CINVESTAV à Mexico. Gerald Geudtner est notamment à l'origine de la parallélisation MPI de deMon2k. Pour bien rendre compte de l'intégralité de la démarche et du travail fourni, nous présenterons dans ce qui suit les deux versions bien que seule la 2^{nde} ait été retenue pour apparaître dans la version officielle de deMon2k. Dans ce qui suit, pour faciliter la lecture, **le nom des routines** est indiqué en minuscule alors que **celui des communicateurs ou variables fortran** est indiqué en majuscule. Tous les noms qui commencent par « MPI_ » désignent **les routines ou variables intrinsèques MPI** alors que ceux qui commencent par « mp_ » désignent **les routines deMon2k d'interface** vers les routines intrinsèques MPI.

2.2.1 Structure MPI de deMon2k avant la création de la seconde couche MPI

Lorsque nous avons commencé à implémenter l'algorithme PTMD dans la version 6.0.2 de deMon2k, certaines parties du programme étaient d'ores et déjà parallélisées à l'aide de bibliothèques MPI. C'est le cas des routines d'algèbre linéaire et des routines dédiées au calcul de la structure électronique. Comme mentionné plus haut, l'implémentation de l'algorithme PTMD a nécessité une modification de cette structure MPI du code. Afin de mieux expliciter ces modifications, nous allons d'abord présenter la structure MPI qui précédait celle que nous avons mis en place.

Comme on peut le voir sur la Figure 2.2, avant toute chose, **le programme initialise la structure MPI du code dans la routine mp_start**. Cette routine fait appel aux **trois principales routines nécessaires à l'initialisation d'un programme MPI**. La **routine mpi_init** initialise les processus MPI. Le nombre de processus MPI correspond au nombre de CPUs alloués lors de la soumission du calcul. La **routine mpi_comm_size** attribue quant à elle un communicateur à tous ces processus. Dans la structure MPI originale de deMon2k il s'agit du communicateur MPI_COMM_WORLD qui est le communicateur par défaut des bibliothèques MPI. Enfin, **mpi_comm_rank** attribue un rang à chaque processus au sein de ce communicateur. Voici comment sont appelées ces trois routines dans le code :

```
CALL MPI_INIT(INITERR)
```

2. IMPLÉMENTATION DE LA MÉTHODE PTMD DANS DEMON2K ET APPLICATION À L'ÉTUDE D'AGRÉGATS D'EAU

```
CALL MPI_COMM_SIZE (MPI_COMM_WORLD, NCPU, IERRS)
CALL MPI_COMM_RANK (MPI_COMM_WORLD, ME, IERRR)
```

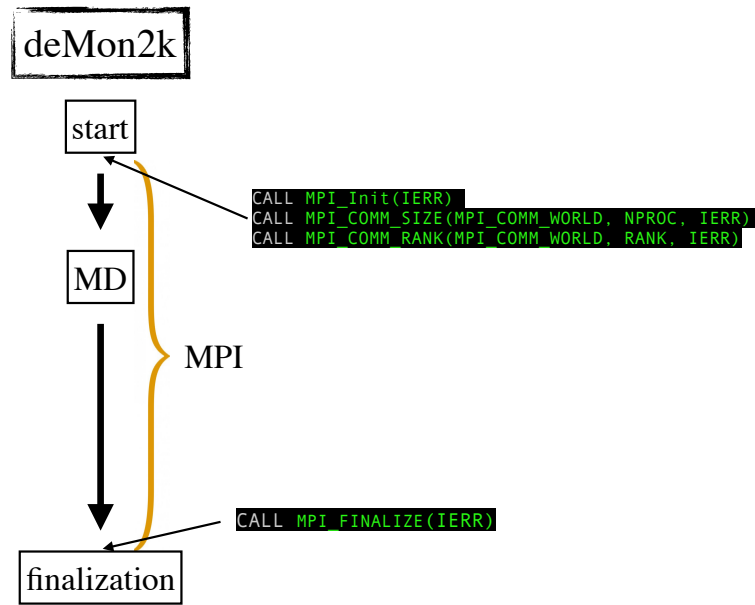


FIGURE 2.2 – Représentation schématique de la structure parallèle MPI du code deMon2k avant la création de la structure parallèle en deux couches.

A ce stade, `MPI_COMM_WORLD` est le seul communicateur employé dans tout le code. Cela signifie qu'à chaque fois qu'une tâche est parallélisée, tous les processus MPI travaillent de concert à la réalisation de cette tâche. Par exemple, si un processus appelle la commande `mpi_bcast`, tous les autres processus recevront la donnée envoyée. Si cette parallélisation fonctionne efficacement pour accélérer le calcul de l'énergie d'un système, elle ne permet pas d'avoir une implémentation parallèle du PTMD. En effet, dans ce cas, il faut que l'énergie de chacune des N répliques soit calculée en parallèle mais indépendamment des autres répliques, ce qui n'est pas possible quand tous les processus communiquent ensemble au travers d'un seul communicateur MPI. **Une implémentation parallèle de l'algorithme PTMD nécessite donc une modification de l'architecture parallèle**

du code deMon2k. Une solution élégante et efficace qui permet de ne pas avoir à modifier toute les parties parallèles du code consiste à **subdiviser le communicateur qui contient tous les processus MPI**, ici `MPI_COMM_WORLD`, **en sous-communicateurs**.

2.2.2 Première version de la structure MPI en 2 couches

L'idée principale de notre modification de l'architecture MPI consiste à créer une seconde couche de parallélisation afin de ne pas avoir à intervenir dans la structure MPI précédant notre implémentation. Pour ce faire, **il nous faut diviser le nombre total de processus MPI requis par un calcul en groupes de processus MPI indépendants les uns des autres**. Chacun de ces groupes de processus se verra ensuite assigner une des températures de l'algorithme PTMD sans modification des routines de calcul de la structure électronique. Un groupe additionnel de processus créé à partir des processus maîtres de ces premiers groupes servira pour les échanges de réplique entre les différentes températures.

Duplication de `MPI_COMM_WORLD` dans `mp_start`

On souhaite que les routines de deMon2k servant d'interface vers les bibliothèques MPI fonctionnent indifféremment si l'on réalise une simulation PTMD ou un autre type de simulation. Le soucis, c'est que ces routines d'interface ne prennent pas le communicateur MPI comme argument, elles utilisent toutes `MPI_COMM_WORLD` par défaut. Pour remédier à ce problème, à la fin de la routine `mp_start`, après l'initialisation des processus MPI, **le communicateur `MPI_COMM_WORLD` est dupliqué en `COMM_TEMPWORLD`** à l'aide de la routine `mpi_comm_dup`. Il faut ensuite dimensionner et classer ce nouveau communicateur avec `mpi_comm_size` et `mpi_comm_rank`. On remarquera que les communicateurs `COMM_TEMPWORLD` et `MPI_COMM_WORLD` ont la même dimension, décrite par l'entier `NCPU`, et que les rangs des processus, décrits par l'entier `ME`, est le même dans ces deux communicateurs. Voici les commandes impliquées :

```
CALL MPI_COMM_DUP (MPI_COMM_WORLD, COMM_TEMPWORLD, IERR)
```

2. IMPLÉMENTATION DE LA MÉTHODE PTMD DANS DEMON2K ET APPLICATION À L'ÉTUDE D'AGRÉGATS D'EAU

CALL MPI_COMM_SIZE (COMM_TEMPWORLD, NCPU, IERRS)

CALL MPI_COMM_RANK (COMM_TEMPWORLD, ME, IERRR)

En remplaçant MPI_COMM_WORLD par COMM_TEMPWORLD dans toutes les routines d'interfaces vers les routines intrinsèques MPI, le fonctionnement du code est inchangé.

Création de groupes de processus dans mp_group

Nous avons créé la routine **mp_group** afin de gérer la création de nouveaux groupes de processus et de nouveaux communicateurs MPI. Le programme appelle cette routine uniquement dans le cas d'une simulation PTMD. C'est cette routine qui est en charge de la **mise en place de l'architecture parallèle MPI en deux couches**. Son fonctionnement est décrit dans ce qui suit. A l'issue de l'appel à cette routine, **trois communicateurs MPI seront créés** :

- **COMM_ALLWORLD** au travers duquel tous les processus pourront communiquer.
- **COMM_TEMPWORLD** au travers duquel seuls les processus en charge de la dynamique à une température donnée pourront communiquer.
- **COMM_PARENTWORLD** au travers duquel les processus maîtres au sein de COMM_TEMPWORLD pourront communiquer pour réaliser les échanges de réplique.

La structure parallèle MPI en deux couches qui résulte de la création de ces sous-groupes est représentée de façon schématique sur la Figure 2.3. Expliquons maintenant son fonctionnement en détails.

Comme les rangs et rôles des processus (esclave ou maître) vont être amenés à changer dans les nouveaux communicateurs MPI, il faut commencer par sauvegarder leurs attributs dans le communicateur actuel (COMM_TEMPWORLD) de la manière suivante :

NCPU_TOT = NCPU

MPGREATMASTER = MPMMASTER

MPGREATSLAVE = MPMSLAVE

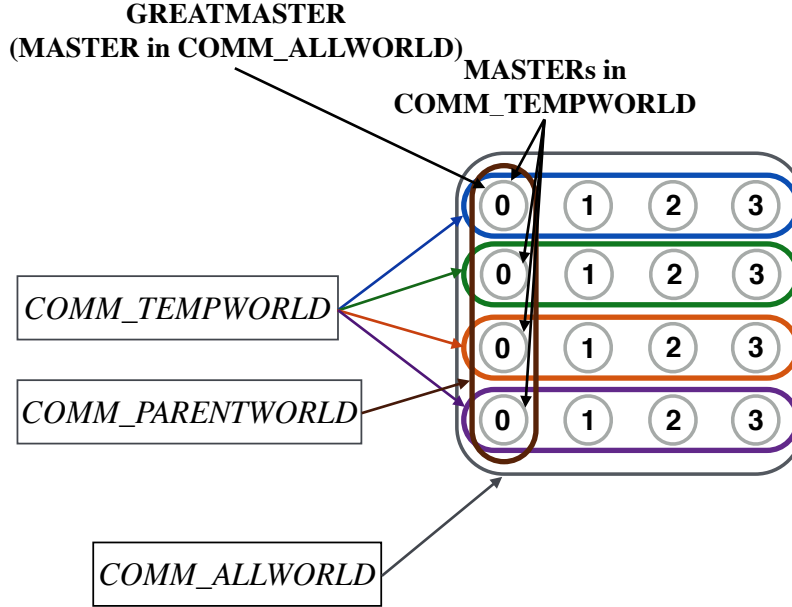


FIGURE 2.3 – Représentation schématique de l’architecture parallèle MPI en deux couches créée à l’aide de la première version de la routine `mp_group`. Le communicateur `COMM_TEMPWORLD` est associé aux sous-groupes de processus créés à partir du communicateur `COMM_ALLWORLD`. Les processus se voient attribuer un nouveau rang dans chacun de ces sous-groupes et les processus avec le rang 0 sont désignés comme étant les processus maîtres (MASTER). Un autre sous-groupe est créé à partir des processus maîtres dans `COMM_TEMPWORLD` auquel est associé le communicateur `COMM_PARENTWORLD`. Les processus se voient aussi associer un nouveau rang dans ce communicateur et le rôle de processus maître est aussi associé au processus de rang 0. Par soucis de clarté nous ne l’avons pas fait apparaître sur cette figure. Le processus maître dans `COMM_PARENTWORLD` et dans le `COMM_TEMPWORLD` de COLOR 0 est aussi le processus maître dans `COMM_ALLWORLD` (GREATMASTER).

GREATMASTER = MASTER

COMM_ALLWORLD = COMM_TEMPWORLD

La dimension du communicateur `COMM_TEMPWORLD` après la division en groupes de processus MPI est ensuite calculée. Cela correspond au nombre de processus travaillant de concert sur une température. Ce nombre est égal au nombre total de processus MPI dédié au calcul, `NCPU`, divisé par le nombre de températures (`NUMREP`) dans la simulation PTMD.

2. IMPLÉMENTATION DE LA MÉTHODE PTMD DANS DEMON2K ET APPLICATION À L'ÉTUDE D'AGRÉGATS D'EAU

`NPROCS = NCPU/NUMREP`

A noter que dans la présente implémentation, le nombre total de CPUs requis pour le calcul, NCPU, doit strictement être un multiple du nombre de températures, NUMREP, dans la simulation PTMD. Cette contrainte peut-être levée mais cela complexifie significativement la constitution des groupes de processus. Un vecteur `PARENT_GRP()` de taille NUMREP est ensuite alloué :

`ALLOCATE (PARENT_GRP (NUMREP))`

et contiendra les rangs des processus au sein du communicateur `COMM_PARENTWORLD`. Ces processus seront les processus maîtres au sein du communicateur `COMM_TEMPWORLD`.

Chaque processus calcule ensuite le rang qu'il aura dans `COMM_TEMPWORLD`, détermine le rang actuel du processus qui sera son processus maître dans `COMM_TEMPWORLD` et inscrit cette valeur dans le vecteur `PARENT_GRP()`.

```
DO I=1,NUMREP
  NLOW = (I-1)*NPROCS
  PARENT_GRP(I) = NLOW
  HIGH = I*NPROCS - 1
  IF ( (ME.GE.NLOW) .AND. (ME.LE.NHIGH) ) THEN
    MASTER = 0
    IF (ME.EQ.NLOW) THEN
      MPMaster = .TRUE.
      MPslave = .FALSE.
    END IF
    NEW_ME = ME - NLOW
    REPL_ID = I
  END IF
END DO
```

La routine `mpi_comm_split` permet maintenant de créer de nouveaux communicateurs à partir de communicateurs préexistants et des données calculées ci-dessus. Le communicateur `COMM_TEMPWORLD` va être redéfini à partir des valeurs des variables `COLOR` et `KEY` de chaque processus. Les processus partageant la même valeur de `COLOR` au sein de `COMM_TEMPWORLD` communiqueront

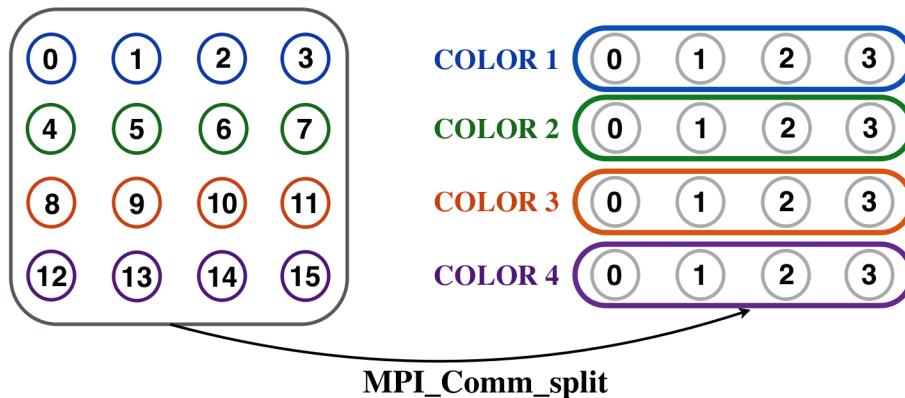


FIGURE 2.4 – Représentation schématique de la division de communicateur MPI à l’aide de la routine `mpi_comm_split`. Cette routine crée des sous-groupes de processus MPI suivant la valeur de la variable `COLOR` attribuée à chaque processus et associe un nouveau communicateur à ces sous-groupes.

ensemble et seront donc en charge de la dynamique d’une même température. Pour chaque processus, la valeur de la variable `COLOR` est égale à la partie entière du quotient de son rang, `ME`, par le nombre de processus travaillant sur une même température, `NPROCS`. La variable `KEY` détermine l’ordre des processus dans le nouveau communicateur. Parmi les processus ayant la même valeur de `COLOR`, celui qui aura la plus petite valeur de `KEY` aura le rang 0 dans le nouveau communicateur. Le suivant aura le rang 1 et ainsi de suite. Ainsi, le rang `ME` attribué à un processus après l’appel à `mpi_comm_split` sera le même que celui calculé avant, `NEW_ME`. Les commandes correspondantes sont les suivantes et le fonctionnement de la commande `mpi_comm_split` est schématisé sur la figure 2.4 :

```
COLOR = ME/NPROCS
```

```
KEY=ME
```

```
CALL MPI_COMM_SPLIT (COMM_ALLWORLD, COLOR, KEY, COMM_TEMPWORLD, IERR)
```

Finalement on attribue à `NCPU` la dimension de `COMM_TEMPWORLD` et à `ME` le rang de chaque processus au sein de `COMM_TEMPWORLD`.

```
CALL MPI_COMM_SIZE (COMM_TEMPWORLD, NCPU, IERRS)
```

```
CALL MPI_COMM_RANK (COMM_TEMPWORLD, ME, IERR)
```

A ce stade, les processus partageant la même valeur de `COLOR` peuvent communiquer au travers de `COMM_TEMPWORLD` et tous les processus peuvent

2. IMPLÉMENTATION DE LA MÉTHODE PTMD DANS DEMON2K ET APPLICATION À L'ÉTUDE D'AGRÉGATS D'EAU

communiquer ensemble au travers de COMM_ALLWORLD. Il ne reste plus qu'à créer un communicateur contenant les processus maîtres de chaque COLOR de COMM_TEMPWORLD. Pour cela, on commence par attribuer à la variable ALLWORLD le groupe associé au communicateur COMM_ALLWORLD avec la routine `mpi_comm_group`. Ensuite, avec la routine `mpi_comm_incl` on crée le groupe PARENTWORLD à partir des processus de ALLWORLD dont les rangs sont contenus dans PARENT_GRP(). Finalement, on crée le nouveau communicateur COMM_PARENTWORLD associé à ce groupe. Voici les commandes mises en jeu :

```
CALL MPI_COMM_GROUP (COMM_ALLWORLD, ALLWORLD, IERR)
CALL MPI_GROUP_INCL (ALLWORLD, NUMREP, PARENT_GRP, PARENTWORLD, IERR)
CALL MPI_COMM_CREATE (COMM_ALLWORLD, PARENTWORLD, COMM_PARENTWORLD, IERR)
```

2.2.3 Seconde version de la structure MPI en 2 couches

Comme mentionné précédemment, après avoir présenté cette implémentation au *19th deMon developer Workshop*, nous avons commencé à collaborer avec Andreas M. Köster et Gerald Geudtner pour implémenter l'algorithme PTMD dans la dernière version de deMon2k. A ce moment, Gerald Geudtner avait commencé à développer une routine permettant de créer et de manipuler des sous-groupes de communicateurs MPI similaire à notre routine `mp_group`. Après quelques échanges, nous avons décidé d'utiliser sa routine qui présente plusieurs avantages :

- Elle ne nécessite pas la copie des routines d'interface vers les routines intrinsèques MPI.
- Elle permet plus de flexibilité quant à la création de nouveaux communicateurs. Par exemple, elle permet de créer des sous-groupes qui ne présentent pas le même nombre de processus.
- Elle peut-être utilisée pour créer et manipuler des communicateurs MPI pour d'autres applications que le PTMD.

Nous allons dans ce qui suit expliquer le fonctionnement de cette seconde routine `mp_group`.

La routine est appelée avec trois arguments : GROUP, COLOR, OP-

2.2 Division des communicateurs MPI dans deMon2k

TION. L'argument **OPTION** peut prendre 2 valeurs : « SET » ou « GEN ». Si elle est appelée avec la valeur GEN, alors la routine va créer un nouveau communicateur à partir du communicateur MPI_COMM_WORLD, par un appel à `mpi_comm_split` :

```
CALL MPI_COMM_SPLIT (MPI_COMM_WORLD, COLOR, KEY, NEWGROUP, IERR)
CALL MPI_COMM_SIZE (NEWGROUP, NEWNCPU, IERR)
CALL MPI_COMM_RANK (NEWGROUP, NEWME, IERR)
```

En assignant au préalable une valeur à la variable **COLOR** pour chaque processus, à l'issu de l'appel à `mp_group` avec la valeur « GEN », tous les processus ayant la même valeur de **COLOR** communiqueront ensemble à l'intérieur du nouveau communicateur **NEWGROUP**. La **variable KEY** correspond au rang du processus dans MPI_COMM_WORLD. Ainsi, dans le nouveau groupe, les processus sont rangés dans le même ordre que dans MPI_COMM_WORLD. `mpi_comm_size` assigne à la variable **NEWNCPU** le nombre de processus dans le nouveau groupe et `mpi_comm_rank` assigne un rang à chaque processus au sein du nouveau communicateur. Toutes les variables concernant le nouveau groupe sont ensuite sauvegardées dans les vecteurs **MPGNCPU**, **MPGME**, **MPGMASTER** et **MPGCOMM** pour que le communicateur puisse par la suite être déclaré comme étant le communicateur actif à n'importe quel endroit du programme.

```
MPGNCPU (GROUP) = NEWNCPU
MPGME (GROUP) = NEWME
MPGMASTER (GROUP) = 0
IF (GROUP.NE.MPGALL) THEN
    IF (MPGME (MPGALL) .EQ.MPGMASTER (MPGALL) ) THEN
        MPGMASTER (GROUP) = NEWME
    END IF
END IF
MPGCOMM (GROUP) = NEWGROUP
```

Si la routine est appelée avec la valeur SET, alors le communicateur **GROUP** passé en argument sera désigné comme étant le communicateur actif jusqu'au prochain appel de `mp_group` avec la valeur SET et un communicateur **GROUP** différent. Dans le cas d'un appel à `mp_group` avec la valeur SET, la valeur de la variable **COLOR** n'a aucune importance puisqu'elle n'est pas utilisée pour activer un communicateur.

2. IMPLÉMENTATION DE LA MÉTHODE PTMD DANS DEMON2K ET APPLICATION À L'ÉTUDE D'AGRÉGATS D'EAU

Cette routine permet donc de créer et de manipuler simplement et avec beaucoup de flexibilité de nouveaux communicateurs MPI dans deMon2k sans avoir à modifier la structure parallèle préexistante. L'unique modification qui a dû être apportée en même temps que l'ajout de `mp_group` est la déclaration du communicateur MPGALL comme étant le communicateur contenant tous les processus (soit un équivalent à `MPI_COMM_WORLD`). Cette déclaration est faite lors de l'initialisation de MPI dans `mp_start`. Après l'appel à `mpi_init`, tous les processus appellent la routine `mp_group` avec la valeur GEN et la même valeur de COLOR (ONE) de GROUP (MPGALL) créant ainsi le communicateur MPGALL qui contient tous les processus MPI. MPGALL est ensuite désigné comme étant le communicateur actif par un second appel à `mp_group` avec la valeur SET.

```
ONE = 1
ZERO = 0

! *** Initialization of MPI ***

CALL MPI_INIT(INITERR)

! *** Generate global group including all tasks ***

CALL MP_GROUP(MPGALL, ONE, 'GEN')

! *** Set global group a the active one ***

CALL MP_GROUP(MPGALL, ZERO, 'SET')
```

Cette routine a été ajoutée officiellement à deMon2k pour donner lieu à la version 6.0.7. Maintenant que l'on a expliqué son fonctionnement en détails, nous allons à présent voir comment nous l'avons utilisée pour implémenter l'algorithme PTMD dans deMon2k, **implémentation disponible dans les versions 6.0.8 et ultérieures de deMon2k.**

2.3 Implémentation de l'algorithme PTMD dans deMon2k v. 6.0.8

Une fois la routine `mp_group` développée, l'implémentation de l'algorithme PTMD dans deMon2k nécessite des modifications légères de la routine de dynamique moléculaire et le développement d'une routine dédiée aux échanges de répliques. C'est que nous allons présenter dans ce qui suit.

2.3.1 Préparation des communicateurs MPI et initialisation de l'algorithme PTMD

Après avoir initialisé tous les processus MPI et défini le groupe MPGALL dans la routine `mp_start`, et après avoir lu le fichier input deMon.inp, les processus rentrent dans la routine `ptmdprep` que nous allons décrire dans ce qui suit et dont le fonctionnement est résumé sur la figure 2.5.

```
IF (MPMASTER) THEN

    REWIND (INP)

    SUCCESS = .FALSE.
    CALL FINDKEY('PARATEMP', SUCCESS)

END IF

CALL MP_SEND_ALL(SUCCESS, 1, MASTER, MP_LOG)

! *** Quick return if PTMD keyword was not found ***

IF (.NOT.SUCCESS) RETURN

! *** Get number of replica from PTMD options ***

IF (MPMASTER) THEN

    CALL FINDOPT('PARATEMP', OPTIONS, NOPTION)
```

2. IMPLÉMENTATION DE LA MÉTHODE PTMD DANS DEMON2K ET APPLICATION À L'ÉTUDE D'AGRÉGATS D'EAU

```
NUMREP = 0
DO IOPTION=1,NOPTION
  IF (OPTIONS(IOPTION) (1:7) .EQ. 'NUMREP=') THEN
    NUMBER = OPTIONS(IOPTION) (8:)
    NUMREP = INT(DIGIT('PARATEMP', 'NUMREP=', '> 0', NUMBER))
  END IF
END DO
END IF

CALL MP_SEND_ALL(NUMREP, 1, MASTER, MP_INT)

IF (NUMREP.LE.0) THEN
  CALL ERRMSG('PTMDPREP', 'NUMBER OF REPLICAS HAS NOT BEEN '//
    'DEFINED IN PARATEMP KEYWORD', 1)
END IF

! *** Define MPI groups needed in a PTMD calculation ***

CALL PTMDGDEF

! *** Activate 'all master' group and prepare input files ***

CALL MP_GROUP(MPGPTM, 0, 'SET')
IF (MPINGRP(MPGPTM)) THEN
  CALL MYINPUT
  CALL MYREPINP
END IF

! *** Active PTMD group ***

CALL MP_GROUP(MPGPT, 0, 'SET')
```

Le processus maître recherche tout d'abord la présence du **mot clef PARATEMP** dans le fichier input. Dans le cas où ce mot clef est présent, il lit ensuite le nombre de répliques NUMREP, *i.e.* de températures, requises par l'utilisateur pour la simulation PTMD et envoie ce nombre à tous processus MPI. Les processus rentrent ensuite dans la **routine ptmdgdef** dans laquelle sont générés les deux communicateurs nécessaires au PTMD (voir partie gauche de la figure 2.5) :

- **MPGPT** est le communicateur au sein duquel communiquent tous

les processus travaillant sur une même température du PTMD. Pour le créer, il faut définir le nombre de processus à l'intérieur du nouveau communicateur NPROCS, soit le nombre de processus par température du PTMD. Comme précédemment, il est calculé comme étant le nombre total de processus MPI divisé par le nombre de répliques, *i.e.* températures. Ensuite une réplique est attribuée à chaque processus, *i.e.* la variable COLOR pour l'appel à `mpi_comm_split`, qui est la valeur entière de leur rang dans le communicateur MPGALL divisé par le nombre de processus attribué à chaque réplique.

```
NPROCS = NCPU/NUMREP
REPID = ME/NPROCS
CALL MP_GROUP (MPGPT, REPID, 'GEN')
```

- **MPGPTM est le communicateur qui permettra de communiquer entre les différentes températures du PTMD au moment des échanges.** Il contient les processus maîtres de chacun des groupes MPGPT. Lors de la création de ce groupe, la routine `mp_group` sera appelée avec la variable MCOLOR pour l'argument COLOR. Puisque les processus maîtres dans MPGPT seront les seuls à communiquer dans MPGPTM, la même valeur de MCOLOR leur est attribuée (MCOLOR = 1). Les autres processus se voient assigner la valeur MPNOCOLOR (équivalente à `MPI_UNDEFINED`). Ainsi, ces processus feront aussi appel aux routines collectives MPI mais ils ne recevront ni n'envoieront aucune donnée lors des communications utilisant le communicateur MPGPTM. On attribue une valeur de COLOR (MCOLOR = 1) à chaque processus maître au sein du communicateur MPGPT. Aux autres processus, on leur assigne la valeur MPNOCOLOR qui correspond à `MPI_UNDEFINED`. Ainsi, ces processus feront aussi appel aux routines collectives MPI mais ils ne recevront ni n'envoieront aucune donnée lorsque le communicateur actif sera MPGPTM. On appelle ensuite la routine `mp_group` avec la valeur GEN et MCOLOR pour l'argument COLOR.

```
IF (MPGME (MPGPT) == MPGMASTER (MPGPT)) THEN
  MCOLOR = 1
ELSE
  MCOLOR = MPNOCOLOR
END IF
CALL MP_GROUP (MPGPTM, MCOLOR, 'GEN')
```

2. IMPLÉMENTATION DE LA MÉTHODE PTMD DANS DEMON2K ET APPLICATION À L'ÉTUDE D'AGRÉGATS D'EAU

Dans deMon2k, pour les calculs utilisant la parallélisation MPI, la routine myinput sert à créer une copie du fichier input deMon.inp par processus MPI. Pour cela, après la lecture de deMon.inp par le processus maître, ce dernier envoie les données du fichier input ligne par ligne aux autres processus. Dans l'implémentation du PTMD, cette procédure est répétée deux fois. Une première fois avec le communicateur MPGPTM actif, à l'issue de quoi chaque processus maître dans MPGPT a créé son fichier input. Une seconde fois avec le communicateur MPGPT actif pour que tous les processus aient leur propre fichier d'input.

Arrivé à cette étape du code, la majorité des modifications nécessaires à l'implémentation de l'algorithme PTMD ont été réalisées. En effet, il ne reste alors qu'à implémenter **l'attribution d'une température différente pour le thermostat de chaque COLOR dans le communicateur MPGPT** et à implémenter **la routine d'échanges**. Pour tout le reste, **le code fait appel aux routines de dynamique moléculaire « normales »**. La différence introduite par la double couche de parallélisation est que, au lieu de communiquer tous ensembles, les processus MPI communiqueront seulement avec ceux qui partagent la même valeur de COLOR au sein du communicateur MPGPT. **Avant le début de la boucle MD**, la **graine aléatoire** utilisée pour la **génération des nombres aléatoires** est initialisée. Ils seront nécessaires à l'algorithme d'échange basé sur le critère de Metropolis-Hasting[231]. Chaque processus calcule également **la température de chacun des thermostats** à partir de la température maximum, minimum, du nombre total de répliques et du numéro de chaque réplique. La **distribution des températures peut être linéaire ou exponentielle**. Les processus rentrent ensuite dans la boucle de dynamique moléculaire (voir partie droite de la figure 2.5) et intègrent les équations du mouvement à la température qui est attribuée à leur groupe. A la fin d'une itération de la boucle MD, si le nombre de pas de MD parcouru est un multiple du nombre de pas requis par l'utilisateur pour qu'une tentative d'échange ait lieu, ce qui correspond à une implémentation synchrone de l'algorithme, **les processus rentrent dans la routine d'échange ptmdexchange** comme on peut le voir sur la partie droite de la figure 2.5.

```
IF (PTMD) THEN
  IF (MDFLUSH(MDCYC,EXSTEP)) THEN
    CALL MDTRJ('FLUSH')
    CALL WRTREPGEO
    IF (MDCYC.NE.MDMIN+MDMAX) THEN
```

2.3 Implémentation de l'algorithme PTMD dans deMon2k v. 6.0.8

```
CALL PTMD EXCHANGE
END IF
END IF
ELSE IF (MDFLUSH(MDCYC, TRJINT)) THEN
CALL MDTRJ('FLUSH')
END IF
```

2.3.2 Routine en charge des échanges de répliques

Au début de la routine ptmdexchange, le groupe MPGPTM est déclaré comme étant le groupe actif (voir la partie droite de la figure 2.5).

```
CALL MP_GROUP(MPGPTM, 0, 'SET')
```

Chaque processus écrit ensuite l'énergie potentielle de sa réplique dans la case du vecteur ENERGIES() qui lui correspond.

```
IF (ME.GE.0) ENERGIES(ME+1) = EPOT
```

Le processus maître additionne tous les vecteurs des autres processus et redistribue ensuite son vecteur ENERGIES(), qui contient l'énergie potentielle de toutes les répliques, à tous les processus actifs dans le groupe MPGPTM. Comme mentionné au chapitre 1, il existe différentes manières de réaliser les échanges dans un algorithme PTMD. Dans notre cas, nous avons décidé qu'**un processus donné ne pourra échanger sa réplique qu'avec les processus voisins**, voisins en terme de température. Cette façon de réaliser les échanges est beaucoup plus facile à mettre en place qu'une approche de type « *all-to-all* ». De plus, elle engendre moins de communication au moment de l'échange. Le résultat diffère cependant peu entre les deux approches puisque la probabilité que deux répliques qui évoluent à des températures non voisines soient échangées est très faible. Pour éviter des échanges successifs entre deux simulations de température voisine qui n'apportent rien en terme d'échantillonnage, nous avons également fait le choix de tentatives d'échange alternées entre simulations de rang pairs et impairs. Ainsi, une fois sur deux, les simulations ayant un rang pair tenteront des échanges avec les simulations adjacentes de température supérieure. L'autre moitié des échanges seront tentées par les simulations de rang impair avec les simulations adjacentes de température supérieure comme décrit ci-dessous :

2. IMPLÉMENTATION DE LA MÉTHODE PTMD DANS DEMON2K ET APPLICATION À L'ÉTUDE D'AGRÉGATS D'EAU

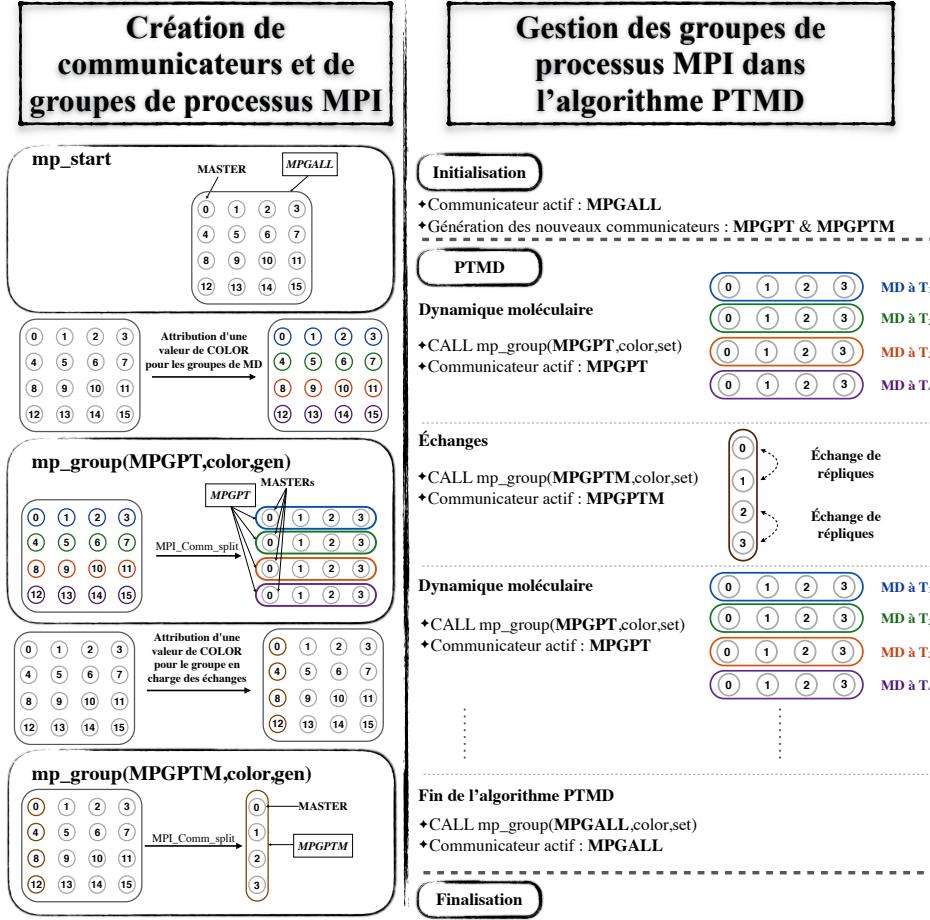


FIGURE 2.5 – (a) Représentation schématique de la création de la structure parallèle MPI en deux couches pour l'algorithme PTMD. Après l'initialisation des processus MPI, chaque processus se voit assigner une valeur de COLOR. La routine `mp_group` crée des sous-groupes de processus selon le même fonctionnement que celui présenté sur la Figure 2.4 et associe le communicateur MPGPT à chacun de ces sous-groupes. Les processus au sein du même sous-groupe seront en charge de la dynamique moléculaire d'une même réplique dans l'algorithme PTMD. A partir des processus maîtres dans ces sous-groupes, un autre sous-groupe est créé et le communicateur MPGPTM lui est associé. (b) Représentation schématique de la manipulation des sous-groupes de processus dans l'implémentation de l'algorithme PTMD dans deMon2k. Les communicateurs MPGPT et MPGPTM sont créés avant le début de l'algorithme PTMD. Au début de l'algorithme PTMD, le communicateur MPGPT est désigné comme étant le communicateur actif. Ainsi, les processus au sein d'un même sous-groupe dans le communicateur MPGPT travaillent de concert sur la dynamique moléculaire d'une même réplique à une température donnée. Au moment de l'étape d'échange, le communicateur MPGPTM est désigné comme étant le communicateur actif. A la fin de l'étape d'échange, MPGPT redevient le communicateur actif. A la fin de l'algorithme PTMD, MPGALL est désigné comme étant le communicateur actif jusqu'à la fin du programme.

2.3 Implémentation de l'algorithme PTMD dans deMon2k v. 6.0.8

```
PARACOUNT = MDCYC/EXSTEP
IF (MODULO (PARACOUNT, 2) .EQ.0) THEN
  DO IEXC=1, NUMREP-1, 2
    DELTA (COUNTER) = (1.0/REP (IEXC+1) %TEMP -
$                      1.0/REP (IEXC) %TEMP) *
$                      (ENERGIES (IEXC+1) - ENERGIES (IEXC))
    COUNTER = COUNTER + 1
  END DO
ELSE
  DO IEXC=2, NUMREP-2, 2
    DELTA (COUNTER) = (1.0/REP (IEXC+1) %TEMP -
$                      1.0/REP (IEXC) %TEMP) *
$                      (ENERGIES (IEXC+1) - ENERGIES (IEXC))
    COUNTER = COUNTER + 1
  END DO
END IF
```

Tous les processus calculent la variable DELTA, Δ , intervenant dans le critère de Metropolis entre les températures concernées par l'échange à une étape donnée. L'exponentielle de Δ/k_b est comparé au nombre aléatoire METRORNG compris entre 0 et 1. Ce nombre est le même pour toutes les répliques. Si $\exp(\Delta/k_b) > \text{METRORNG}$ alors **l'échange de répliques entre les deux températures concernées est accepté** et cette information est écrite dans le vecteur EXC(). Notons que pour une paire de répliques, si l'énergie de la réplique à la température la plus basse est plus élevée que l'énergie de la réplique à la température la plus haute, alors l'échange sera forcément accepté car dans ce cas $\Delta > 0$ et $\exp(\Delta/k_b) > 1$.

```
! *** Choose the random number for exchanges ***
```

```
METRORNG = RANDOM(PTSEED)
```

```
! *** It broadcasts this value to all the other colors ***
```

```
CALL MP_SEND_ALL (METRORNG, 1, MASTER, MP_REAL)
```

```
COUNTER = 1
```

```
! *** Number of exchanges set to 0 ***
```

2. IMPLÉMENTATION DE LA MÉTHODE PTMD DANS DEMON2K ET APPLICATION À L'ÉTUDE D'AGRÉGATS D'EAU

```
NUMEXC = 0

! *** Every color checks if wether or not it will exchange ****

IF (MODULO (PARACOUNT, 2) .EQ. 0) THEN
  DO IEXC=1, NUMREP-1, 2
    IF (EXP (DELTA (COUNTER) / BOLTZ) .GE. METRORNG) THEN
      NUMEXC = NUMEXC + 1
      EXCH (NUMEXC) = IEXC
    END IF
    COUNTER = COUNTER + 1
  END DO
ELSE
  DO IEXC=2, NUMREP-2, 2
    IF (EXP (DELTA (COUNTER) / BOLTZ) .GE. METRORNG) THEN
      NUMEXC = NUMEXC + 1
      EXCH (NUMEXC) = IEXC
    END IF
    COUNTER = COUNTER + 1
  END DO
END IF
```

On procède ensuite à l'**échange des répliques** entre processus concernés. Si par exemple l'échange de répliques entre les processus n et $n + 1$ est accepté, le processus n copie ses vecteurs de coordonnées, de vitesses et d'accélération, c'est à dire de gradients, dans des vecteurs temporaires. Ensuite, le processus $n + 1$ envoie les coordonnées, vitesses et accélérations de sa réplique au processus n à l'aide de la routine `mp_send`. Le processus n les reçoit à l'aide de la routine `mp_get` et envoie ensuite les coordonnées, vitesses et accélérations qu'il avait copié dans des tableaux temporaires au processus $n + 1$ qui les reçoit à son tour avec la commande `mp_get`. Les **vitesses sont réajustées** grâce à un facteur de « *scaling* » classique pour correspondre aux nouveaux ensembles thermodynamiques dans lesquels se trouvent les répliques. Dans le cas où les différences de température entre répliques sont faibles, ces réajustements sont faibles et ne perturbent que peu la dynamique des répliques. Enfin, les **thermostats sont réinitialisés** et le communicateur MPGPT est rétabli comme étant le communicateur actif. Les différentes commandes impliquées pour réaliser l'échange de répliques sont données ci-dessous :

```
! *** Process the exchanges of atomic coordinates, velocities ***
```

2.3 Implémentation de l'algorithme PTMD dans deMon2k v. 6.0.8

```
! *** gradients and NHC between the corresponding replicas. ***

IF (NUMEXC.GT.0) THEN
  DO IEXC=1,NUMEXC

! *** First, exchange setcoord ***

    JREP = EXCH(IEXC)
    TEMPSET = REP(JREP)%SETCOORD
    REP(JREP)%SETCOORD = REP(JREP+1)%SETCOORD
    REP(JREP+1)%SETCOORD = TEMPSET

! *** Then exchange the coordinates in four steps ***

! *** 1 - Copy the atomic coordinates of the first proc in a ***
! ***      temporary variable. ***

    IF (ME.EQ.REP(JREP)%RANK) THEN
      TEMPCOORD(:, :) = COORD(:, 1:NATOM_ALL, CCSET)
      TEMPVELOS(:, :) = VELOS(:, 1:NATOM_ALL)
      TEMPACCEL(:, :) = ACCEL(:, 1:NATOM_ALL)
      IF (MDBATH.EQ.'NOSE') THEN
        TEMPNHPOS(:) = NHCOORD(1:NNHC)
        TEMPNHVEL(:) = NHVELOS(1:NNHC)
      END IF
      TEMPEPOT = EPOT
    END IF

! *** 2 - Send atomic coordinates of the second color to the ***
! ***      first one ***

    IF (ME.EQ.REP(JREP+1)%RANK) THEN
      CALL MP_SEND(COORD(1, 1, CCSET), 3*NATOM_ALL, REP(JREP)%RANK,
                   MP_REAL)
      CALL MP_SEND(VELOS, 3*NATOM_ALL, REP(JREP)%RANK, MP_REAL)
      CALL MP_SEND(ACCEL, 3*NATOM_ALL, REP(JREP)%RANK, MP_REAL)
      IF (MDBATH.EQ.'NOSE') THEN
        CALL MP_SEND(NHCOORD, NNHC, REP(JREP)%RANK, MP_REAL)
        CALL MP_SEND(NHVELOS, NNHC, REP(JREP)%RANK, MP_REAL)
      END IF
      CALL MP_SEND(EPOT, 1, REP(JREP)%RANK, MP_REAL)
    END IF
```

2. IMPLÉMENTATION DE LA MÉTHODE PTMD DANS DEMON2K ET APPLICATION À L'ÉTUDE D'AGRÉGATS D'EAU

```

! *** 3 - The first color receives the atomic coordinates of ***
! *** the second proc ant sends its previous atomic ***
! *** coordinates (stored in TEMPCOORD) to the second color. ***

IF (ME.EQ.REP (JREP) %RANK) THEN
  CALL MP_GET (COORD (1, 1, CCSET), 3*NATOM_ALL, REP (JREP+1) %RANK,
    MP_REAL)
  CALL MP_GET (VELOS, 3*NATOM_ALL, REP (JREP+1) %RANK, MP_REAL)
  IF (MDBATH.EQ.'NOSE') THEN
    CALL MP_GET (NHCOORD, NNHC, REP (JREP+1) %RANK, MP_REAL)
    CALL MP_GET (NHVELOS, NNHC, REP (JREP+1) %RANK, MP_REAL)
  END IF
  CALL MP_GET (EPOT, 1, REP (JREP+1) %RANK, MP_REAL)
  CALL MP_SEND (TEMPCOORD, 3*NATOM_ALL, REP (JREP+1) %RANK, MP_REAL)
  CALL MP_SEND (TEMPVELOS, 3*NATOM_ALL, REP (JREP+1) %RANK, MP_REAL)
  CALL MP_SEND (TEMPACCEL, 3*NATOM_ALL, REP (JREP+1) %RANK, MP_REAL)
  IF (MDBATH.EQ.'NOSE') THEN
    CALL MP_SEND (TEMPNHPOS, NNHC, REP (JREP+1) %RANK, MP_REAL)
    CALL MP_SEND (TEMPNHVEL, NNHC, REP (JREP+1) %RANK, MP_REAL)
  END IF
  CALL MP_SEND (TEMPEPOT, 1, REP (JREP+1) %RANK, MP_REAL)

! *** Rescale the velocities of the atoms and ***
! *** of nose to the new temperature ensemble. ***

  CALL NUCKIN (NDEG)
  CALL VELOSCALE (MDTEMP, .FALSE.)
  DO INHC=1, NNHC
    NHVELOS (INHC) = SQRT (MDTEMP / REP (JREP+1) %TEMP) *
      NHVELOS (INHC)
  END DO
END IF

! *** 4 - Finally the second color receives the coordinates of ***
! *** the first one. ***

IF (ME.EQ.REP (JREP+1) %RANK) THEN
  CALL MP_GET (COORD (1, 1, CCSET), 3*NATOM_ALL, REP (JREP) %RANK,
    MP_REAL)
  CALL MP_GET (VELOS, 3*NATOM_ALL, REP (JREP) %RANK, MP_REAL)
  CALL MP_GET (ACCEL, 3*NATOM_ALL, REP (JREP) %RANK, MP_REAL)

```


2.3 Implémentation de l'algorithme PTMD dans deMon2k v. 6.0.8

```
IF (MDBATH.EQ.'NOSE') THEN
  CALL MP_GET (NHCOORD, NNHC, REP (JREP) %RANK, MP_REAL)
  CALL MP_GET (NHVELOS, NNHC, REP (JREP) %RANK, MP_REAL)
END IF
CALL MP_GET (EPOT, 1, REP (JREP) %RANK, MP_REAL)

! *** Rescale the velocities of the atoms and ***
! *** of nose to the new temperature ensemble. ***

CALL NUCKIN (NDEG)
CALL VELOSCALE (MDTEMP, .FALSE.)
DO INHC=1, NNHC
  NHVELOS (INHC) = SQRT (MDTEMP / REP (JREP) %TEMP) * NHVELOS (INHC)
END DO
END IF
END DO

*** Reinitialization of the thermostats ***

IF (MDBATH.EQ.'NOSE') THEN
  CALL NOSEDATA (MDTEMP)
  FRICTION = NHVELOS (1)
ELSE IF (MDBATH.EQ.'HOOVER') THEN
  THOOVER = SQRT (MDTEMP / TKIN)
END IF
PTEXCH = .TRUE.

END IF
```

2.3.3 Échanges asynchrones

Comme mentionné plus haut, un défaut de l'algorithme PTMD lorsqu'il est implémenté de façon synchrone provient du temps « d'attente » au moment des échanges. En effet, pour que les processus MPI rentrent dans la routine d'échange, il faut que toutes les répliques aient parcouru depuis le dernier échange les EXSTEP pas de MD requis par l'utilisateur. Or comme nous l'avons dit, les sources potentielles de désynchronisation sont multiples et il peut en résulter un temps de calcul

2. IMPLÉMENTATION DE LA MÉTHODE PTMD DANS DEMON2K ET APPLICATION À L'ÉTUDE D'AGRÉGATS D'EAU

qui rend la simulation impraticable. Nous avons donc implémenter une **version du PTMD dite « asynchrone » pour résoudre ce problème**. Dans cette version de l'algorithme PTMD, ce n'est plus le nombre de pas de MD parcouru qui est utilisé comme condition pour rentrer dans la routine d'échange PTMD mais le **temps horloge écoulé depuis le dernier échange**. La figure 2.6 présente de façon schématique l'algorithme PTMD dans ses versions synchrone et asynchrone avec le temps horloge comme condition d'échange. La variable correspondant au temps horloge écoulé entre deux échanges a été appelée EXTIME. Tout comme EXSTEP, la valeur de EXTIME est définie par l'utilisateur. Toutefois, le nombre de pas de MD parcourus pendant un temps horloge donné peut varier en fonction du *hardware* ou du nombre de processus MPI alloué à chaque température de l'algorithme PTMD. Ainsi une phase de calibration doit être réalisée avant de lancer des calculs de production pour s'assurer que la valeur de EXTIME soit optimale. Si cette valeur est trop faible, les répliques ne pourront pas explorer efficacement la PES à la température qui leur est attribuée. Au contraire, si cette valeur est trop grande, un « sur-échantillonnage » sera réalisé et l'algorithme PTMD perdra en efficacité.

Nous allons à présent examiner les modifications réalisées dans le code pour implémenter cette version asynchrone de l'algorithme PTMD. La condition de fin de simulation n'est plus un nombre maximum de pas de MD mais un temps horloge maximum, MAXTIME, que tous les processus MPI doivent avoir réalisé. Cette condition se traduit par la variable logique FINISH. Voyons dans un premier temps les différentes variables nécessaires à cette implémentation asynchrone :

- Le vecteur **REP_FINISH()** sert à **enregistrer la valeur de FINISH propre à chaque processus MPI**. La valeur de FINISH est initialisée à `.FALSE.` pour tous les processus MPI.
- La variable **REAL_MDTIME** est la variable correspondant au **temps horloge écoulé durant l'algorithme PTMD**. Elle sera **incrémentée et comparée à MAXTIME après chaque échange**. Elle est initialisée à 0,0 avant le début de la boucle de MD.
- La fonction **DETIME()** donne accès au temps CPU. La **différence entre deux valeurs de DETIME()** enregistrées à des moments différents donne accès au temps horloge écoulé pour un CPU.
- La variable **REAL_MDTIME_INIT** sert à enregistrer la valeur de **DETIME()** avant le début de la boucle MD.

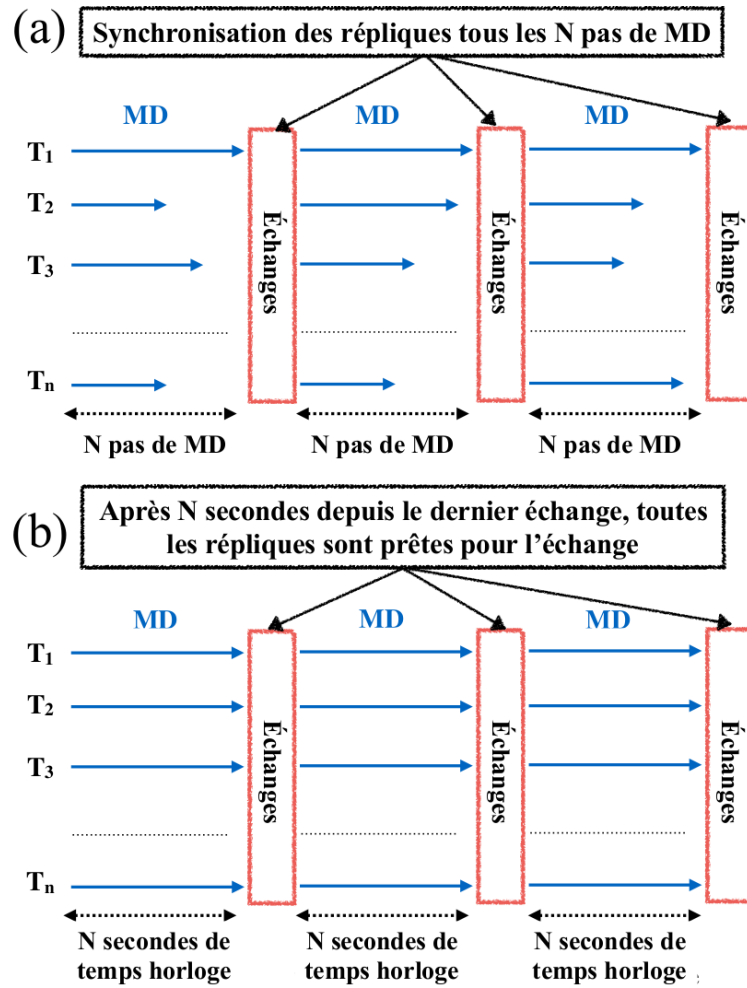


FIGURE 2.6 – (a) Schéma d’une simulation PTMD synchrone dans laquelle l’échange des répliques a lieu lorsque toutes les répliques ont réalisé un nombre prédéfini N de pas de MD depuis le dernier échange. (b) Schéma d’une simulation PTMD dans laquelle l’échange des répliques a lieu après un certain temps horloge depuis le dernier échange quel que soit le nombre de pas de MD parcourus par les répliques.

- Les variables **PT_TIME1** et **PT_TIME2** évalueront respectivement le temps CPU après chaque échange PTMD et après chaque pas de MD.
- **MTIME** correspond à la différence entre **PT_TIME2** et **PT_TIME1** et c’est cette variable qui sera comparée à **EXTIME** pour déterminer si le code doit rentrer dans la routine d’échanges PTMD. Au premier pas de MD, comme il n’y a pas encore eu d’échange, **PT_TIME1** est initialisé à **DETIME()** et **PT_TIME2** est initialisé à 0,0.

2. IMPLÉMENTATION DE LA MÉTHODE PTMD DANS DEMON2K ET APPLICATION À L'ÉTUDE D'AGRÉGATS D'EAU

Voici comment ces différentes variables sont initialisées :

```
! *** Initialization of ASYNCPTMD variables ***

ALLOCATE (REP_FINISH (MPGNCPU (MPGALL) ) )
DO IREP=1,MPGNCPU (MPGALL)
  REP_FINISH (IREP) = .FALSE.
END DO
FINISH = .FALSE.
REAL_MDTIME_INIT = DETIME ()
REAL_MDTIME = 0.0
PT_TIME1 = DETIME ()
PT_TIME2 = 0.0
MTIME = 0.0

! *** Molecular dynamics trajectory cycle ***

DO WHILE (.NOT.FINISH)
```

Après avoir propagé les positions atomiques sur un pas de temps, le temps écoulé MTIME depuis le dernier échange (ou depuis le début de la boucle MD si il n'y a pas encore eu d'échange) est évalué et les processus MPI maîtres dans MPGPT envoient leur valeur de MTIME aux autres processus MPI dans leur groupe. Cette action est nécessaire car les processus MPI au sein d'un même groupe peuvent être légèrement déphasés et avoir une valeur de MTIME légèrement différente. Dans le cas où, au sein d'un groupe de processus traitant la MD d'une même réplique, certains processus MPI ont une valeur de MTIME supérieure à EXTIME et d'autres non, tous les processus ne rentreront pas dans la routine d'échange. L'exécution du programme n'avancera plus sans pour autant s'interrompre car différents processus MPI s'attendront mutuellement dans des communications MPI bloquantes différentes. Après avoir réalisé les échanges, chaque processus MPI évalue si sa valeur de REAL_MDTIME a dépassé la valeur de MAXTIME. L'appel à mpi_allgather permet de partager la valeur de REP_FINISH de chaque processus MPI dans MPGALL à tous les autres. Si tous les processus ont dépassé la valeur de MAXTIME, alors la variable FINISH prend la valeur .TRUE. et la boucle de MD s'arrête. Les variables PT_TIME1 et PT_TIME2 sont également réinitialisées à DETIME().

```
PT_TIME2 = DETIME ()
MTIME = PT_TIME2-PT_TIME1
CALL MP_SEND_ALL (MTIME, 1, MASTER, MP_REAL)
```

2.4 Application à l'étude du changement de phase dans l'hexamère d'eau (H₂O)₆

```
! *** Perform PTMD exchange ***

IF (MTIME>EXTIME) THEN
  NBEXCH = NBEXCH + 1
  CALL MDTRJ('FLUSH')
  CALL WRTREPGEO
  CALL PTMD EXCHANGE
  REAL_MDTIME = DETIME() - REAL_MDTIME_INIT
  IF (REAL_MDTIME < MAXTIME) THEN
    REP_FINISH(MPGME(MPGALL)+1) = .FALSE.
  ELSE
    REP_FINISH(MPGME(MPGALL)+1) = .TRUE.
  END IF
  CALL MPI_ALLGATHER(MPI_IN_PLACE, 1, MPI_LOGICAL,
$                                REP_FINISH, 1, MPI_LOGICAL,
$                                MPGCOMM(MPGALL), IERR)
  FINISH = .TRUE.
  DO IREP = 1, MPGNCPU(MPGALL)
    IF (REP_FINISH(IREP) == .FALSE.) THEN
      FINISH = .FALSE.
    END IF
  END DO
  PT_TIME1 = DETIME()
  PT_TIME2 = DETIME()
END IF
```

Ces modifications dans la boucle de MD permettent de réaliser des échanges PTMD asynchrones. En effet, on peut voir que le nombre de pas de MD n'intervient plus dans cet algorithme et que la condition d'échange ne dépend que du temps horloge écoulé entre deux échanges.

2.4 Application à l'étude du changement de phase dans l'hexamère d'eau (H₂O)₆

Dans cette section, nous allons **tester et valider l'implémentation de la méthode PTMD dans le code deMon2k** en l'utilisant pour étudier la tran-

2. IMPLÉMENTATION DE LA MÉTHODE PTMD DANS DEMON2K ET APPLICATION À L'ÉTUDE D'AGRÉGATS D'EAU

sition de phase dans l'hexamère d'eau (H_2O)₆. Comme mentionné dans les chapitres précédents, (H_2O)₆ a fait l'objet de nombreuses études et la nature de son isomère le plus stable a été à l'origine d'un long débat puisque deux structures avait été proposées pour tenir ce rôle : l'isomère *cage* et l'isomère *prism*. D'une part, les études théoriques réalisées à l'aide de méthodes de différents niveaux de précision ont conduit à des conclusions opposées [279, 280, 281, 282, 283, 284]. D'autre part, des études expérimentales ont récemment conclu que l'isomère *cage* est l'isomère de plus faible énergie bien que cette structure semble coexister avec d'autres structures même à très basse température [35, 36, 33, 285, 286]. A noter également que J. Marc Pedulla et Kenneth D. Jordan ont étudié en détails l'influence de la nature du champ de force utilisé pour le calcul de la courbe de capacité calorifique de (H_2O)₆ [287]. Ils ont pu observer une variation d'environ 100 K sur le calcul de la température de transition de phase de cet agrégat soulignant ainsi la forte sensibilité de cette propriété vis à vis du potentiel utilisé. Dans cette section, nous allons également tenter de caractériser la transition de phase de (H_2O)₆ à partir de l'analyse de la **courbe de capacité calorifique** dont les variations seront corrélées à la **population d'isomères de basse énergie** en fonction de la température. À des fins de comparaison, nous avons également réalisé la même étude avec un potentiel SCC-DFTB, comme cela avait été fait auparavant par Luiz F. Oliveira *et al.* [191]. Les différences observées au niveau de la transition de phase entre les deux potentiels, DFT et SCC-DFTB, seront rationalisées par une analyse énergétique des différents isomères de basse énergie et corrélées à la densité de minima local au voisinage du minimum global. Enfin, pour les deux potentiels, des **simulations PTMD synchrones et asynchrones** ont été mises en œuvre pour discuter de l'efficacité parallèle de ces deux implémentations et pour évaluer les différences qu'elles peuvent engendrer quant aux propriétés thermodynamiques et structurales obtenues.

2.4.1 Méthodes employées et paramètres de calcul

Paramètres des calculs DFT. Tous les calculs DFT ont été réalisés avec la version développeur 6.0.8 de deMon2k [98]. L'approche de DFT employée est l'approche ADFT que nous avons décrite dans la section 1.2.3 de ce manuscrit. Nous avons utilisé la base auxiliaire automatiquement générée GEN-A2 [98, 288, 289] en combinaison avec une base double zêta à valence séparée (DZVP) [290] pour

2.4 Application à l'étude du changement de phase dans l'hexamère d'eau (H₂O)₆

tous les atomes. La fonctionnelle d'échange-corrélation GGA BLYP [139, 136] a été employée. Le potentiel d'échange-corrélation a été intégré numériquement sur une grille variable avec un seuil de précision de 10^{-5} u.a. Enfin, les intégrales de répulsion électronique (ERIs) à trois centres ont été calculées avec l'option CONVENTIONAL disponible dans deMon2k, c'est à dire qu'elle ont été calculées au début de chaque cycle SCF et gardées en mémoire vive durant l'intégralité de chaque cycle. Ceci permet une plus grande performance computationnelle au prix d'une plus grande occupation de la mémoire.

Paramètres des calculs SCC-DFTB. Les calculs SCC-DFTB ont été réalisés avec le code deMonNano [97]. La méthode SCC-DFTB a été décrite dans la section 1.2.3. Nous avons utilisé l'ensemble mio des tables d'intégrales de Slater-Koster [72]. Cet ensemble d'intégrales a été spécialement mis au point pour réaliser des simulations de molécules organiques. Afin d'améliorer la description des interactions intermoléculaires, un terme de dispersion empirique a été rajouté au potentiel SCC-DFTB. Enfin, nous avons remplacé les charges de Mulliken par des charges *Class IV-Charge Model 3* [142] pour le traitement des interactions électrostatiques. La nécessité d'inclure tous ces termes pour décrire correctement les agrégats moléculaires a été démontrée dans des travaux précédents [291, 292, 293]. La procédure auto-cohérente sur les charges a été considérée comme convergée et donc stoppée lorsque les fluctuations des charges atomiques étaient inférieures à 10^{-9} u.a.

Exploration de la surface d'énergie potentielle. Toutes les simulations ont été réalisées sur le supercalculateur Olympe du mésocentre de calcul CALMIP équipé de nœuds composés de deux processeurs *Intel®Skylake 6140* de 18 cœurs chacun, c'est à dire 36 cœurs par nœud. Les simulations DFT ont été parallélisées sur 64 cœurs de manière à ce que le calcul de la structure électronique de chaque réplique soit parallélisé sur 2 cœurs. Les simulations SCC-DFTB ont quant à elles été parallélisées sur 32 cœurs en assignant un seul cœur par réplique. Dans le cas des simulations PTMD synchrones, les paramètres de dynamique moléculaire et les échanges de répliques sont les mêmes pour les deux potentiels employés. L'exploration de la surface d'énergie potentielle s'est faite dans l'ensemble canonique. Pour cela, la température des systèmes est maintenue constante grâce à une chaîne de thermostats de Nosé-Hoover [187, 188, 190]. Nous avons fixé la longueur de cette dernière

2. IMPLÉMENTATION DE LA MÉTHODE PTMD DANS DEMON2K ET APPLICATION À L'ÉTUDE D'AGRÉGATS D'EAU

à 7 thermostats dont les masses ont été choisies de manière à ajuster la fréquence de résonance du thermostat à 800 cm^{-1} . Enfin, nous avons utilisé 32 températures distribuées linéairement entre 20 et 300 K et les équations du mouvement ont été intégrées avec un pas de temps de 1.0 fs. Dans le cas des **simulations synchrones**, les tentatives d'échange ont eu lieu tous les 100 pas de MD. Pour les **simulations asynchrones**, nous avons fixé le temps horloge de manière à ce que le nombre moyen de pas de MD parcouru par toutes les répliques entre deux échanges consécutifs soit égal à ~ 100 . Rappelons que ce temps horloge dépend de la compilation du code, de l'*hardware* de la machine de calcul ainsi que du nombre de cœurs utilisés. Dans notre cas, avec la configuration utilisée, le temps horloge est de ~ 344 secondes pour les simulations DFT et de $\sim 0,625$ secondes pour les simulations SCC-DFTB. Pour toutes les simulations, les 10 pas de MD qui suivent un échange accepté sont considérés comme une « étape d'équilibration » et n'ont donc pas été pris en compte pour les analyses statistiques. Dans les simulations DFT, pour éviter l'évaporation du système, ce dernier a été confiné dans une sphère d'Anderson de rayon 5.0 Å ayant pour centre le centre de masse de l'agrégat $(\text{H}_2\text{O})_6$. En SCC-DFTB le système a été confiné dans un potentiel sphérique de forme $V(r) = 0,008 \times (r - r_0)^4$ (en u.a.), de rayon $r_0 = 5,0\text{ Å}$ centré sur le centre de masse de l'agrégat $(\text{H}_2\text{O})_6$.

Caractérisation de la transition de phase. Les courbes de capacité calorifique ont été calculées à l'aide de la **méthode des histogrammes multiples** présentée dans la section 1.3.5. Afin de corrélérer les courbes de capacité calorifique aux changements structuraux intervenant dans $(\text{H}_2\text{O})_6$ lorsque la température augmente, nous avons étudié la **population relative d'isomères en fonction de la température**. Pour pouvoir se faire, il faut pouvoir différencier les différents isomères de $(\text{H}_2\text{O})_6$ par une approche automatisée. Ainsi, comme dans des études précédentes, les différents isomères de $(\text{H}_2\text{O})_6$ ont été discriminés en fonction du **nombre de cycles oxygène-oxygène et de la taille de ces cycles**. Par taille de cycle nous entendons ici le nombre de molécule d'eau constitutif d'un cycle. Nous avons appliqué pour cela l'algorithme des cycles primitifs développé par Xianglong Yuan et Alastair Cormack [250] en considérant que deux molécules d'eau sont liées si la distance entre leurs atomes d'oxygène est inférieure à $3,25\text{ Å}$. Ainsi, dans ce qui suit, chaque agrégat sera identifié par quatre nombres entiers entre crochets. Chaque nombre correspond au nombre de cycle d'une certaine taille présent dans l'agrégat. Ainsi, le premier chiffre correspond au nombre de cycles composés de 3 molécules,

2.4 Application à l'étude du changement de phase dans l'hexamère d'eau (H₂O)₆

les deuxième, troisième et quatrième chiffres correspondent au nombre de cycles à 4, 5 et 6 molécules d'eau, respectivement. Par exemple, la nomenclature [2100] désigne l'isomère de (H₂O)₆ composé de deux cycles contenant trois molécules d'eau et d'un cycle contenant quatre molécules d'eau. Bien sûr, toute approche de classification d'isomère comporte une part de subjectivité, dans la méthode elle-même ou dans les paramètres qui la définissent, ici le paramètre de coupure. Les courbes de distribution d'isomère sont donc dépendantes de la méthode choisie pour faire cette classification.

2.4.2 Résultats et discussion

Convergence statistique. Avant de discuter des résultats obtenus, il faut s'assurer de la convergence statistique des simulations réalisées. Au total, pour chaque potentiel et chaque type d'échange (synchrone ou asynchrone), deux simulations PTMD indépendantes de 13 ns chacune ont été réalisées. A chaque fois, nous avons considéré la première nanoseconde comme temps d'équilibration et ne l'avons donc pas pris en compte pour l'analyse. Les 12 ns suivantes ont été utilisées pour accumuler positions et énergies afin de pouvoir calculer la courbe de capacité calorifique et les distributions d'isomères en fonction de la température. Pour les quatre types de simulations réalisés (DFT synchrone, DFT asynchrone, SCC-DFTB synchrone et SCC-DFTB asynchrone), les courbes de capacité calorifique et de distribution d'isomères issues des deux simulations PTMD indépendantes sont identiques. **Ceci démontre que 13 ns sont suffisantes pour atteindre la convergence statistique de ces deux propriétés.** La figure 2.7 présente les courbes de capacité calorifique obtenues pour différents temps de simulation. On remarque sur cette figure que les courbes de capacité calorifique peuvent être considérées convergées dès $\sim 4,9$ ns après l'équilibration pour les simulations DFT et dès $\sim 2,3$ ns après l'équilibration pour les simulations SCC-DFTB. Sur cette figure, on voit également que **la convergence des courbes en fonction du temps de simulation ne semble pas dépendre de la manière dont sont réalisés les échanges**, synchrone ou asynchrone.

Changement de phase et influence du potentiel. La figure 2.8 représente les courbes de capacité calorifiques issues des simulations PTMD synchrone (lignes

2. IMPLÉMENTATION DE LA MÉTHODE PTMD DANS DEMON2K ET APPLICATION À L'ÉTUDE D'AGRÉGATS D'EAU

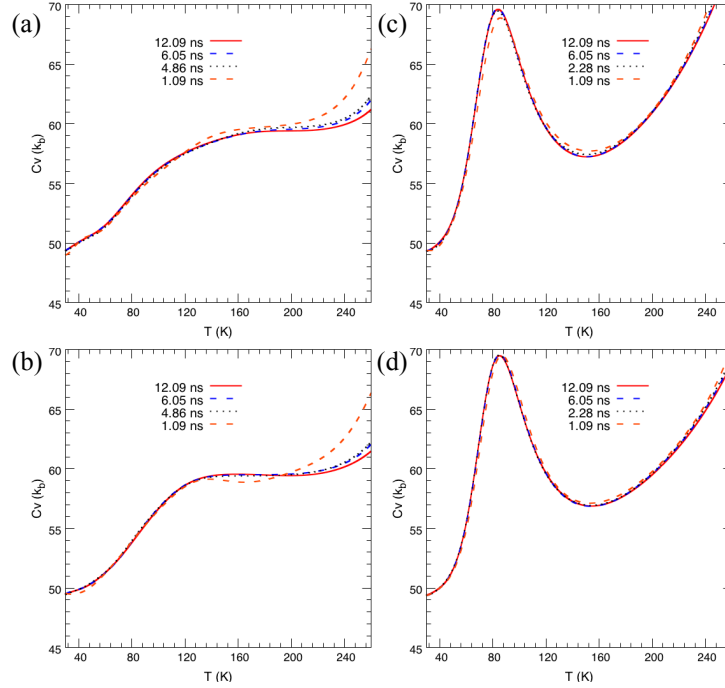


FIGURE 2.7 – Courbes de capacité calorifique obtenues pour différents temps de calcul pour des simulations (a) DFT synchrone, (b) DFT asynchrone, (c) SCC-DFTB synchrone et (d) SCC-DFTB asynchrone.

pleines) et asynchrone (lignes discontinues) en DFT (lignes rouges) et SCC-DFTB (lignes vertes). Attardons nous tout d'abord sur les différences entre les courbes issues des simulations DFT et SCC-DFTB synchrones uniquement. On remarque tout d'abord que **les courbes diffèrent de façon significative selon le potentiel employé**, DFT ou SCC-DFTB. Ce résultat est en accord avec ce que nous avons mentionné en introduction sur la forte dépendance de la courbe de capacité calorifique de $(\text{H}_2\text{O})_6$ vis à vis du potentiel employé. La courbe issue de la simulation SCC-DFTB présente un pic bien défini caractérisé par un maximum à 84,5 K. Au contraire, la courbe issue de la simulation DFT ne présente pas de pic bien marqué mais un léger changement de pente entre 60 et 80 K. Il a été montré que la présence ou l'absence de structure dans la courbe de capacité calorifique dépend de la densité de minima locaux proches du minimum global ainsi que de l'écart énergétique entre les deux structures de plus basse énergie [294, 287]. Ces propriétés sont bien entendues fortement dépendantes du potentiel employé et d'autant plus pour $(\text{H}_2\text{O})_6$ pour qui plusieurs isomères sont très proches en énergie et leur ordre relatif peut varier d'un potentiel à l'autre.

2.4 Application à l'étude du changement de phase dans l'hexamère d'eau (H_2O)₆

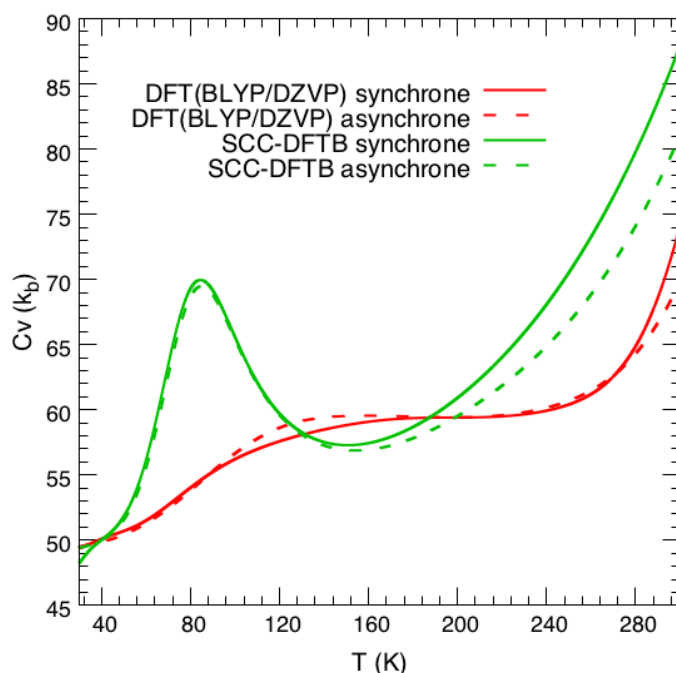


FIGURE 2.8 – Courbes de capacité calorifique obtenues pour des simulations DFT synchrone (courbe rouge pleine) et asynchrone (courbe rouge pointillée), et SCC-DFTB synchrone (courbe verte pleine) et asynchrone (courbe verte pointillée).

Ainsi, afin de mieux comprendre les différences entre les deux courbes pleines de la figure 2.8, nous avons calculé les énergies de liaison relatives des huit isomères de plus basse énergie de (H_2O)₆ aux niveaux DFT et SCC-DFTB. Les géométries de ces isomères sont présentées sur la figure 2.9 et les valeurs des énergies de liaison relatives sont reportées dans le Tableau 2.1. A titre de comparaison, nous avons aussi reporté les valeurs calculées au niveau de théorie CCSD(T) avec une extrapolation à la limite de la base complète (CBS) [284].

Le Tableau 2.1 montre que l'isomère le plus stable est l'isomère *prism* avec les trois méthodes. Les énergies de liaison des autres isomères sont donc relatives à celle de l'isomère *prism*. Au niveau DFT, les isomères *cage*, *inter*, et *book* sont presque dégénérés avec des énergies de 2,7, 2,4 et 2,7 kJ.mol⁻¹, respectivement. En revanche, en SCC-DFTB, l'isomère *book* est beaucoup plus haut en énergie avec une valeur de 5,5 kJ.mol⁻¹ ce qui est deux fois plus élevé qu'au niveau CCSD(T)/CBS. De plus, l'écart énergétique entre l'isomère le plus stable (*prism*) et le suivant (*cage* en SCC-DFTB et en CCSD(T)/CBS et *inter* en DFT) est plus élevé au niveau SCC-DFTB qu'au niveau DFT, 3,4 kJ.mol⁻¹ contre 2,4 kJ.mol⁻¹. Ces légères diffé-

2. IMPLÉMENTATION DE LA MÉTHODE PTMD DANS DEMON2K ET APPLICATION À L'ÉTUDE D'AGRÉGATS D'EAU

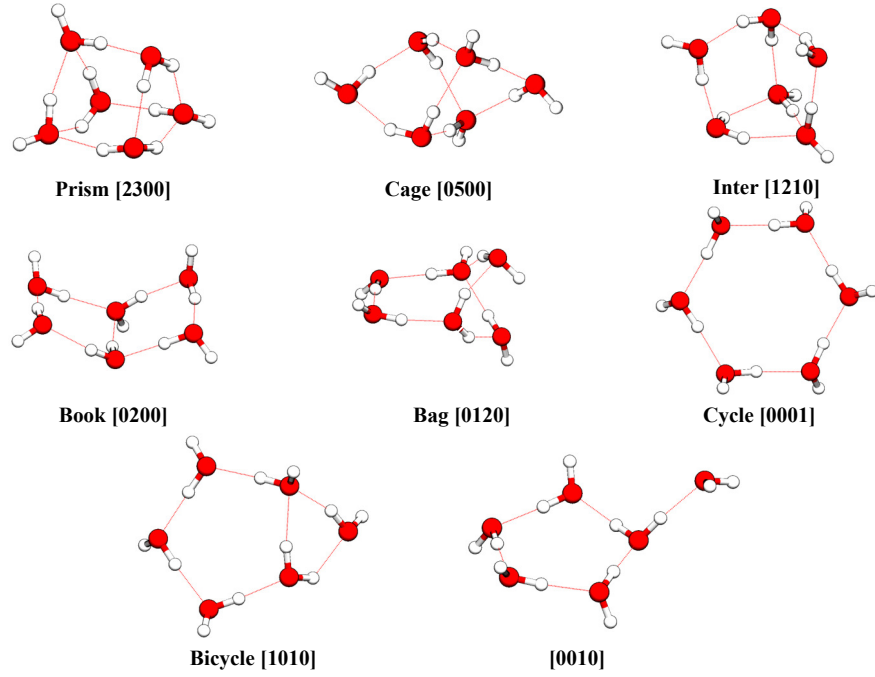


FIGURE 2.9 – Isomères de plus basse énergie de $(\text{H}_2\text{O})_6$. Toutes les structures ont été optimisées et les valeurs des énergies de liaison relatives sont reportées dans le Tableau 2.1.

rences énergétiques sont à l'origine des disparités entre les deux courbes de capacité calorifique. En particulier, l'écart énergétique entre les deux isomères de plus basse énergie joue un rôle déterminant dans la présence ou non d'un pic bien défini dans la courbe de capacité calorifique. Par exemple, cet écart est plus grand dans les agrégats $(\text{H}_2\text{O})_8$ et $(\text{H}_2\text{O})_{21}\text{H}^+$ dont les courbes de capacité calorifique font apparaître un pic bien prononcé [191, 59, 294, 287]. Dans le cas présent, ce **faible écart énergétique entre les isomères *prism* et *inter* en DFT, engendre une courbe plus monotone que celle prédite au niveau SCC-DFTB.**

Une autre propriété qu'il convient d'étudier pour caractériser la transition de phase dans $(\text{H}_2\text{O})_6$ est la population d'isomères de basse énergie en fonction de la température. Les figures 2.10 (a) et (b) comparent ces populations obtenues aux niveaux DFT(BLYP/DZVP) et SCC-DFTB, respectivement. Les différences entre les deux populations sont reportées sur la figure 2.10 (c). Dans le cas de la DFT, l'isomérisation de *prism* en *inter* débute vers ~ 40 K et celle de *book* vers *cage* commence à apparaître vers ~ 60 K. Au delà de ~ 170 K, le schéma principal d'isomérisation est *prism* \rightarrow *inter* \rightarrow *book*. Comme les isomères *inter* et *book* sont presque dégénérés

2.4 Application à l'étude du changement de phase dans l'hexamère d'eau (H₂O)₆

TABLE 2.1 – Énergies de liaison relatives (en kJ.mol⁻¹) des huit isomères de plus basse énergie de (H₂O)₆ obtenues en SCC-DFTB et DFT(BLYP/DZVP). Les valeurs obtenues au niveau CCSD(T)/CBS proviennent de la référence [284]. Les géométries des isomères sont représentées sur la figure 2.9.

Isomères	SCC-DFTB	DFT	CCSD(T)/CBS [284]
<i>prism</i> [2300]	0,0	0,0	0,0
<i>cage</i> [0500]	3,4	2,7	1,1
<i>inter</i> [1210]	3,5	2,4	-
<i>book</i> [0200]	5,5	2,7	3,0
<i>bag</i> [0120]	7,5	7,4	6,8
<i>cycle</i> [0001]	8,9	9,5	7,5
<i>bicycle</i> [1010]	9,7	11,5	-
[0010]	23,2	25,5	-

et proches en énergie de l'isomère *prism*, le schéma d'isomérisation décrit précédemment engendre une courbe de capacité calorifique relativement plate. L'impact de l'isomère *cage* est limité car il n'apparaît qu'en petites proportions et à des températures supérieures à ~ 100 K. Sur la figure 2.10 (b), on peut voir qu'au niveau SCC-DFTB, l'isomérisation de *prism* en *book* est plus abrupte et ne passe pas par l'isomère *inter* ce qui explique la transition de phase plus prononcée que l'on observe sur la courbe de capacité calorifique au niveau SCC-DFTB.

Efficacité parallèle λ . Nous allons à présent discuter de l'efficacité parallèle de nos implémentations de la méthode PTMD dans les codes deMon2k et deMonNano. Pour cela, nous allons comparer le nombre de pas de MD réalisés lors de simulations synchrone et asynchrone pour un même temps horloge (24 heures). Pour chaque simulation, le nombre total de pas de MD, le nombre de pas parcourus à la température la plus basse et à la température la plus haute ainsi que l'efficacité parallèle λ en fonction du nombre de répliques sont reportés dans le Tableau 2.2. L'efficacité parallèle λ est définie comme le ratio entre le nombre total de pas de MD réalisé par toutes les répliques (c'est à dire la somme des pas de MD réalisés à chaque température) lors de simulation PTMD avec et sans échange pour un temps horloge donné. La valeur de λ est donc une mesure directe de la perte d'efficacité due aux communications MPI bloquantes lors de l'étape d'échange de répliques. Un λ de 1,0 signifie que l'étape d'échange de répliques n'engendre aucune perte d'efficacité de l'algorithme.

2. IMPLÉMENTATION DE LA MÉTHODE PTMD DANS DEMON2K ET APPLICATION À L'ÉTUDE D'AGRÉGATS D'EAU

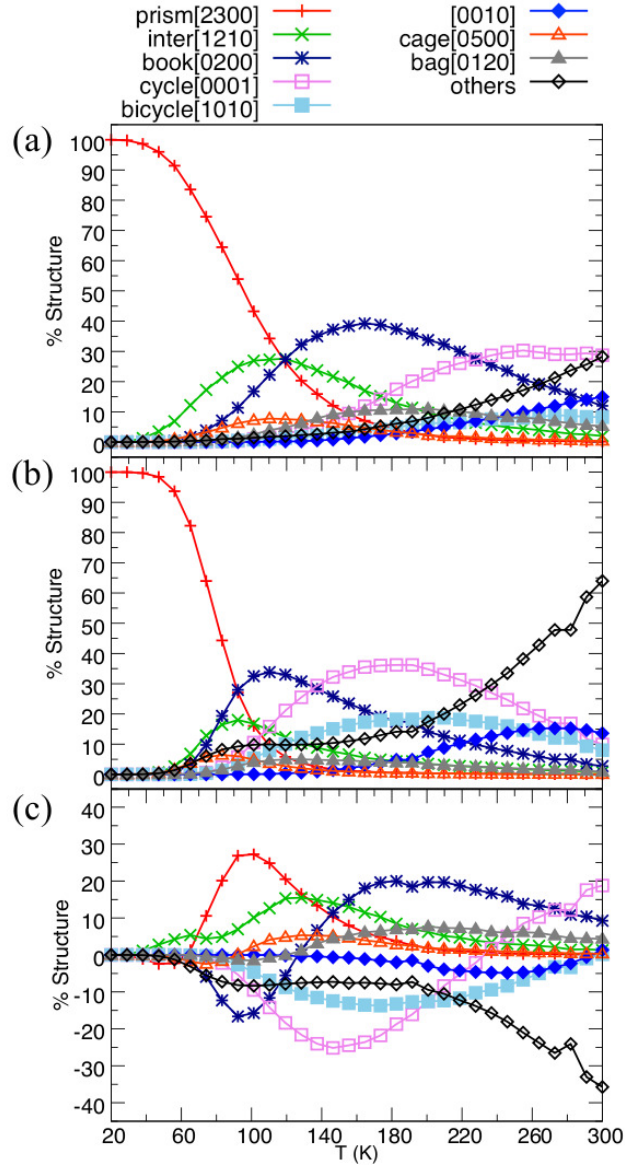


FIGURE 2.10 – Populations des isomères de plus basse énergie de $(\text{H}_2\text{O})_6$ en fonction de la température issues de simulations PTMD aux niveaux (a) DFT et (b) SCC-DFTB. (c) Différences entre les deux ensembles de populations. La partie positive correspond à une proportion d'isomères plus importante en DFT tandis que la partie négative correspond à une proportion d'isomères plus grande en SCC-DFTB.

Les données du Tableau 2.2 montrent que pour les deux codes, **les implémentations synchrones et asynchrones de l'algorithme PTMD présentent une très bonne efficacité parallèle**, proche de 90% jusqu'à 128 répliques en DFT et supérieure à 95% en SCC-DFTB. Ce résultat est en contradiction avec les résultats

2.4 Application à l'étude du changement de phase dans l'hexamère d'eau (H₂O)₆

TABLE 2.2 – Nombre total de pas de MD, nombre de pas parcourus à la plus basse et à la plus haute température et efficacité parallèle λ en fonction du nombre de répliques utilisées dans les simulations PTMD synchrones (synch.) et asynchrones (asynch.) réalisées en DFT à l'aide du code deMon2k et en SCC-DFTB avec deMonNano. La dynamique moléculaire de chaque réplique est réalisée par un seul processus MPI. La colonne no-exch correspond aux simulations sans l'étape d'échange de répliques.

		DFT			SCC-DFTB		
Réplique		Synch.	Asynch.	No-exch.	Synch.	Asynch.	No-exch.
32	Nombre de pas total	442160	430040	488710	423001600	437669460	439350840
	T° la plus basse	13820	14690	16490	13218800	13997640	13461000
	T° la plus haute	13820	13570	15420	13218800	13695660	13636520
	λ	0,90	0,88	-	0,96	1,00	-
	Nombre de pas total	847190	892420	967650	840879360	871223720	876350700
64	T° la plus basse	13240	15250	16300	13138740	13932000	13368800
	T° la plus haute	13240	14010	15880	13138740	13612440	13569800
	λ	0,88	0,92	-	0,96	0,99	-
	Nombre de pas total	1663840	1677720	1921470	1669150720	1738091540	1754380820
	T° la plus basse	13000	14310	14600	13040240	13971120	13748400
128	T° la plus haute	13000	13210	13300	13040240	13618440	13615460
	λ	0,87	0,87	-	0,95	0,99	-

de Giovanni Bussi qui a observé une baisse drastique d'efficacité parallèle au delà de 100 répliques [278]. De notre point de vue, cette différence peut provenir de deux sources : d'une **implémentation très inefficace de la version synchrone** utilisée lors de ces travaux ou d'une forte dépendance de λ au système étudié. Dans un cas comme dans l'autre, nos résultats montrent que pour l'étude d'agrégats d'eau, **nos implémentations peuvent efficacement être employées sur des super-ordinateurs** jusqu'à au moins 128 répliques dans l'intervalle de températures que nous avons considéré, c'est à dire 20-300 K. A noter que les applications auxquelles ces implémentations sont vouées ne devraient pas dépasser voire même atteindre cette quantité de répliques. Au vu des remarques ci-dessus, on peut cependant penser que λ pourrait être affecté par **la nature du système chimique étudié** ou par **l'utilisation d'un intervalle de températures beaucoup plus grand**. Ce point sera à tester lors d'applications futures. Étonnamment, le Tableau 2.2 montre aussi que si l'implémentation asynchrone est plus performante que l'implémentation synchrone en SCC-DFTB, cela quelque soit le nombre de répliques, ce n'est pas toujours le cas en DFT. En effet, pour les simulations à 32 répliques, l'efficacité parallèle λ est de 90% pour les échanges synchrones et 88% pour les échanges asynchrones. Avec 64 répliques, λ est plus grand dans le cas de l'implémentation asynchrone (92%

2. IMPLÉMENTATION DE LA MÉTHODE PTMD DANS DEMON2K ET APPLICATION À L'ÉTUDE D'AGRÉGATS D'EAU

contre 88% en synchrone) et à 128 répliques, l'efficacité parallèle est de 87% pour les deux implémentations. Ces différences sont faibles et montrent à nouveau que pour la présente étude, la mise en place d'échanges asynchrones n'est pas synonyme d'un gain en performances. Elles montrent également qu'**il existe d'autres sources de perte d'efficacité** autre que les communications bloquantes lors des échanges.

Cependant, déterminer l'origine de ces sources est un travail complexe puisqu'elles peuvent dépendre aussi bien du *hardware* que du *software*. Pour mettre cela en évidence, la figure 2.11 présente le nombre de pas de MD réalisé par chaque réplique lors des trois simulations PTMD sans échanges réalisées avec 32, 64 et 128 répliques au niveau SCC-DFTB et un temps horloge identique de 24h. Si l'on ne considère pas les points aberrants de ces courbes, **elles font clairement apparaître une diminution du nombre de pas réalisé à mesure que la température de la réplique augmente**. Ceci va dans le sens de l'hypothèse que nous avons posée en introduction sur des cycles de convergence SCF plus longs aux températures les plus hautes. L'existence claire de cette désynchronisation confirme l'intérêt d'une implémentation asynchrone des échanges de répliques. On voit également apparaître sur ces courbes **des répliques qui réalisent significativement moins de pas que les autres**. Étrangement, ces répliques présentent toujours le même rang quelque soit leur nombre total, et ce rang présente également une périodicité qui est la même que celle du nombre de processeurs par nœud du supercalculateur Olympe, c'est à dire 18. Nous n'avons à ce jour pas trouvé l'origine de ce comportement mais il nous apparaît clairement provenir du *hardware* et peut donc expliquer en partie le comportement quelque peu erratique de la valeur de λ dans le Tableau 2.2.

Pour terminer, la figure 2.8 présente les courbes de capacité calorifique obtenues pour les implémentations synchrone et asynchrone de l'algorithme PTMD dans les codes deMon2k et deMonNano. On remarque sur ces courbes une très légère différence entre les courbes issues de simulations synchrones et asynchrones, et ce pour les deux potentiels, DFT et SCC-DFTB. Au vu des courbes de convergence de la figure 2.7, ces différences ne peuvent pas être attribuées à un défaut de convergence. La nature des échanges, synchrone ou asynchrone, induit donc une légère différence dans les densités d'états calculées et donc dans les courbes de capacité calorifique ce qui se traduit par une très légère différence dans les populations d'isomères comme le montre la figure 2.12. **Ce résultat est particulièrement inattendu et sur-**

2.5 Application à l'étude du changement de phase dans l'octamère d'eau $(\text{H}_2\text{O})_8$

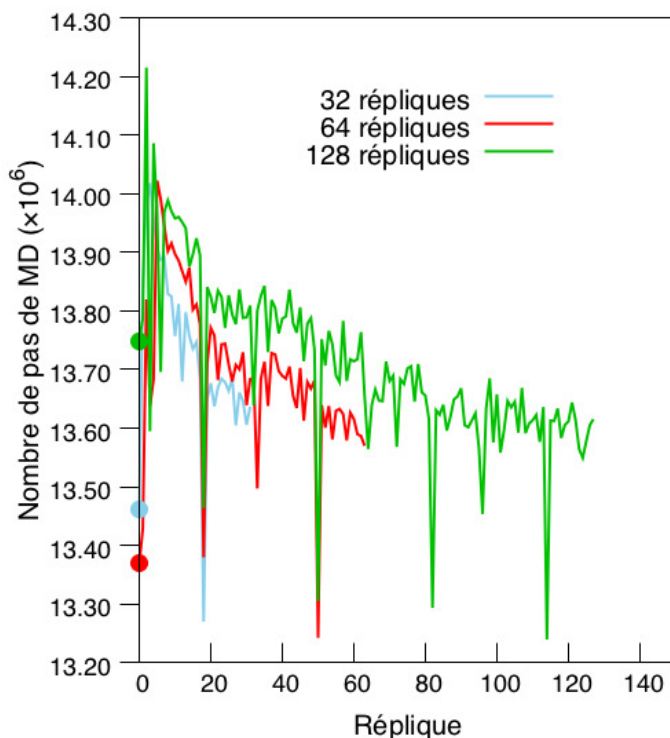


FIGURE 2.11 – Nombre de pas de MD réalisé par chaque réplique lors des trois simulations PTMD sans échanges réalisées avec 32 (bleu), 64 (rouge) et 128 (vert) répliques au niveau SCC-DFTB pour un temps horloge de 24h.

prenant. En effet, notre implémentation d'échanges asynchrones n'a pour effet que de différencier légèrement le nombre de pas de MD réalisé à chaque température. Cette différence de nombre de pas ne devrait, selon nous, n'avoir aucune influence sur les propriétés calculées. L'origine de ces différences, même très faibles, reste non comprise et reste donc un sujet d'étude

2.5 Application à l'étude du changement de phase dans l'octamère d'eau $(\text{H}_2\text{O})_8$

A la suite de notre étude sur l'agrégat $(\text{H}_2\text{O})_6$, nous avons également souhaité étudier l'agrégat d'eau neutre $(\text{H}_2\text{O})_8$. Tout comme $(\text{H}_2\text{O})_6$, ce dernier a également été **largement étudié d'un point de vue expérimental** [295] et **théorique** [287, 296, 297, 298, 299, 294, 300, 301, 302, 303, 191]. D'un point de vue

2. IMPLÉMENTATION DE LA MÉTHODE PTMD DANS DEMON2K ET APPLICATION À L'ÉTUDE D'AGRÉGATS D'EAU

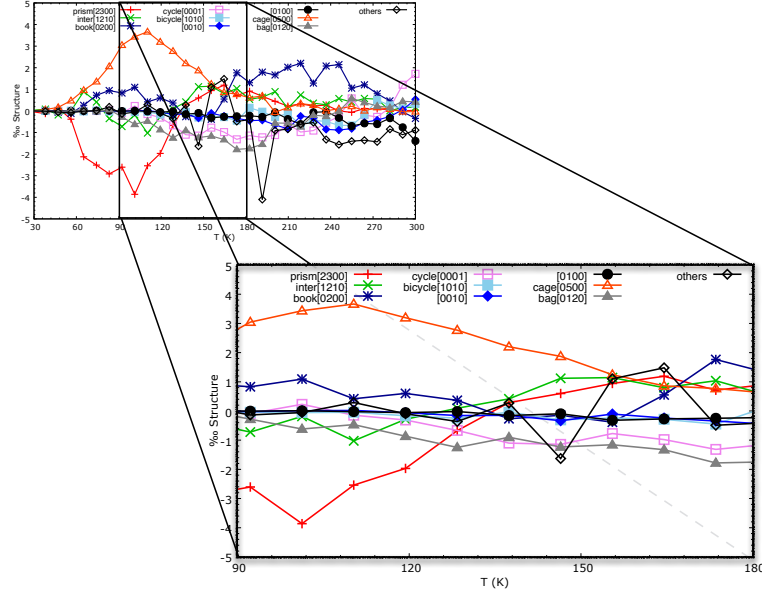


FIGURE 2.12 – Différence entre les populations d'isomères de $(\text{H}_2\text{O})_6$ issues de simulations avec échanges synchrones et asynchrones en DFT(BLYP/DZVP). Les parties positives et négatives correspondent respectivement à une majorité d'isomères dans les simulations synchrone et asynchrone.

des études théoriques, certaines se sont concentrées sur **la caractérisation du minimum global** [296, 297, 300, 301] qui dans le cas de $(\text{H}_2\text{O})_8$ a conduit à moins de débats que pour $(\text{H}_2\text{O})_6$. En effet, le minimum global de cet agrégat est sans conteste **la structure cubique** qui est représentée sur la figure 2.14. En plus de ce minimum global clairement défini, ce qui le différencie donc de $(\text{H}_2\text{O})_6$, $(\text{H}_2\text{O})_8$ présente également la particularité d'avoir un écart énergétique élevé entre cet isomère et les isomères de plus haute énergie. Là encore, ceci représente une différence notable avec $(\text{H}_2\text{O})_6$. Les autres études théoriques ont quant à elle porté sur le comportement de $(\text{H}_2\text{O})_8$ en fonction de la température au travers des mêmes propriétés que nous avons étudié dans le cas de $(\text{H}_2\text{O})_6$, c'est à dire la courbe de capacité calorifique et les distributions d'isomères en fonction de la température [287, 298, 299, 294, 302, 303, 191]. Ces études ont montré que la courbe de capacité calorifique de $(\text{H}_2\text{O})_8$ est moins sensible au potentiel utilisé car elle présente dans tous les cas un pic bien marqué et intense avec un changement abrupte de la pente de la courbe à la température de début de transition. Cette dernière est quant à elle dépendante du potentiel utilisé, de même que l'intensité du pic, et peut varier de plusieurs dizaines de Kelvin en fonction du potentiel [287]. Cependant,

2.5 Application à l'étude du changement de phase dans l'octamère d'eau $(\text{H}_2\text{O})_8$

comme pour $(\text{H}_2\text{O})_6$, l'intégralité de ces études ont été réalisées à l'aide de champs de force et **aucun potentiel de niveau théorique supérieur n'a été mis en œuvre**. Dans cette section, nous avons donc souhaité **revisiter l'étude de l'agrégat $(\text{H}_2\text{O})_8$** à l'aide des nouveaux outils que nous avons développés. L'objectif est de déterminer comment un potentiel DFT décrit les propriétés structurales et dynamiques de cette espèce comparativement à des champs de force. De même, grâce à l'**efficacité de notre implémentation et du code deMon2k**, il est possible d'aborder une telle étude à l'aide de fonctionnelles plus coûteuses, mais aussi plus précises a priori, que des fonctionnelles GGA. Nous avons donc réalisé cette étude à l'aide de deux fonctionnelles, la **fonctionnelle BLYP** par souhait de cohérence avec notre étude précédente mais également la **fonctionnelle PBE0** pour tester l'impact d'une fonctionnelle hybride sur les propriétés des agrégats de molécules d'eau. Comme précédemment, nous avons également réalisé les mêmes simulations au niveau SCC-DFTB pour comparaison. Il est important de noter que c'est la **première étude de ce type réalisée à ce niveau de théorie**.

2.5.1 Méthodes employées et paramètres de calcul

La plupart des paramètres de calcul utilisés pour cette étude sont identiques à ceux utilisés pour l'étude de $(\text{H}_2\text{O})_6$. Nous les répétons cependant ici par souci de clarté.

Paramètres des calculs DFT. Tous les calculs DFT ont été réalisés avec la version développeur 6.0.8 de deMon2k [98]. L'approche de DFT employée est l'approche ADFT que nous avons décrite dans le chapitre 1 de ce manuscrit. Nous avons utilisé la base auxiliaire automatiquement générée GEN-A2+ [98, 288, 289], qui consiste à utiliser la base GEN-A2 durant les cycles SCF et la base GEN-A2* pour le calcul de l'énergie finale une fois l'énergie SCF convergée. Tous les calculs ont été réalisés avec une base double zêta à valence séparée (DZVP) [290] pour tous les atomes. Les fonctionnelles d'échange-corrélation GGA BLYP [139, 136] et hybride PBE0 [140, 141] ont été employées. Le potentiel d'échange-corrélation a été intégré numériquement sur une grille variable avec un seuil de précision de 10^{-5} u.a. Enfin, les intégrales de répulsion électronique (ERIs) à trois centres ont été calculées avec l'option MIXED disponible dans deMon2k.

2. IMPLÉMENTATION DE LA MÉTHODE PTMD DANS DEMON2K ET APPLICATION À L'ÉTUDE D'AGRÉGATS D'EAU

Paramètres des calculs SCC-DFTB. Les calculs SCC-DFTB ont été réalisés avec le code deMonNano [97]. La méthode SCC-DFTB a été décrite dans la section 1.2.3. Nous avons utilisé l'ensemble mio des tables d'intégrales de Slater-Koster [72]. Cet ensemble d'intégrales a été spécialement mis au point pour réaliser des simulations de molécules organiques. Afin d'améliorer la description des interactions intermoléculaires, un terme de dispersion empirique a été rajouté au potentiel SCC-DFTB. Enfin, nous avons remplacé les charges de Mulliken par des charges *Class IV-Charge Model 3* [142] pour le traitement des interactions électrostatiques. La procédure auto-cohérente sur les charges a été considérée comme convergée et donc stoppée lorsque les fluctuations des charges atomiques étaient inférieures à 10^{-9} u.a.

Exploration de la surface d'énergie potentielle. Les simulations DFT ont été réalisées sur la partition AMD Irene ROME du supercalculateur Joliot-Curie du TGCC composée de noeuds contenant deux processeurs AMD EPYC 7H12 composés de 64 cœurs chacun. Chaque simulation a ainsi été parallélisée sur 128 cœurs de manière à ce que le calcul de la structure électronique de chaque réplique soit parallélisé sur 4 cœurs. Les simulations SCC-DFTB ont quant à elles été parallélisées sur 32 cœurs en assignant un seul cœur par réplique. Toutes les simulations PTMD ont été réalisées avec des échanges de répliques synchrones. L'exploration de la surface d'énergie potentielle s'est faite dans l'ensemble canonique. Pour cela, la température des systèmes est maintenue constante grâce à une chaîne de thermostats de Nosé-Hoover [187, 188, 190]. Nous avons fixé la longueur de cette dernière à 7 thermostats dont les masses ont été choisies de manière à ajuster la fréquence de résonance du thermostat à 800 cm^{-1} . Enfin, nous avons utilisé 32 températures distribuées linéairement entre 20 et 400 K et les équations du mouvement ont été intégrées avec un pas de temps de 1.0 fs. La durée des simulations est de 400 ps, dont 200 ps considérées comme de l'équilibration. Ce temps de simulation est beaucoup plus faible que celui des simulations précédentes de $(\text{H}_2\text{O})_6$. Il apparaît cependant que pour cet agrégat, la convergence des propriétés structurales et thermodynamiques est beaucoup plus rapide et donc que 200 ps de production sont suffisantes. Les tentatives d'échange ont eu lieu tous les 100 pas de MD. Pour toutes les simulations, les 10 pas de MD qui suivent un échange accepté sont considérées comme une « étape d'équilibration » et n'ont donc pas été prises en compte pour les analyses statistiques. Finalement, afin d'éviter l'évaporation du système, ce dernier a été confiné dans un potentiel sphé-

2.5 Application à l'étude du changement de phase dans l'octamère d'eau $(\text{H}_2\text{O})_8$

rique de forme $V(r) = 0,008 \times (r - r_0)^4$ (en u.a.), de rayon $r_0 = 5,5 \text{ \AA}$ centré sur le centre de masse de l'agrégat $(\text{H}_2\text{O})_8$.

2.5.2 Résultats et discussion

Les résultats obtenus sont présentés sur la figure 2.13. Comme on peut le voir, **quelque soit le potentiel utilisé, la courbe de capacité calorifique présente un pic net avec une température de début de transition et un maximum bien définis**. Ceci est en accord avec les travaux théoriques de la littérature [287, 298, 299, 294, 302, 303, 191]. Les courbes obtenues aux niveaux DFT et SCC-DFTB présentent cependant des différences significatives. En effet, dans le second cas, la courbe obtenue présente une température de début de transition à $\sim 145 \text{ K}$ et un maximum à $\sim 175 \text{ K}$. Ces valeurs sont en accord qualitatif avec les résultats obtenus avec divers champs de force par J. Marc Pedulla et Kenneth D. Jordan [287] qui observaient une température de début de transition comprise entre ~ 145 et $\sim 180 \text{ K}$, excepté pour le champ de force TIP3P qui conduit à une valeur de $\sim 80 \text{ K}$. Les deux courbes DFT présentent quant à elles une température de début de transition à $\sim 208 \text{ K}$ et un maximum à $\sim 245 \text{ K}$, la seule différence provenant de l'intensité du pic. Ainsi, comme pour $(\text{H}_2\text{O})_6$, **l'utilisation d'un potentiel DFT modifie significativement les propriétés thermodynamiques calculées pour $(\text{H}_2\text{O})_8$** . Pour comprendre ces différences, le tableau 2.3 présente les énergies de liaison relatives des isomères de basse énergie de $(\text{H}_2\text{O})_8$, les structures correspondantes ainsi que leur nomenclature sont données sur la figure 2.14. Comme attendu, **l'isomère [0604], de géométrie cubique, est l'isomère de plus basse énergie** quelque soit l'approche considérée. De plus, contrairement à $(\text{H}_2\text{O})_6$, les isomères d'énergie supérieure sont beaucoup plus hauts en énergie : 15,3, 30,3 et 31,1 kJ.mol^{-1} aux niveaux SCC-DFTB, BLYP et PBE0, respectivement. Cela induit **une densité d'isomères quasi nulle à proximité du minimum global** ce qui est caractéristique d'un pic bien défini dans la courbe de capacité calorifique [287, 294]. Cette différence d'énergie est cependant bien plus élevée au niveau DFT, approximativement le double, qu'au niveau SCC-DFTB, expliquant ainsi le décalage vers les hautes températures de la transition de phase. Pour les isomères considérés, **les fonctionnelles BLYP et PBE0 conduisent qualitativement aux mêmes énergies de liaison relatives** ce qui explique les courbes de capacité

2. IMPLÉMENTATION DE LA MÉTHODE PTMD DANS DEMON2K ET APPLICATION À L'ÉTUDE D'AGRÉGATS D'EAU

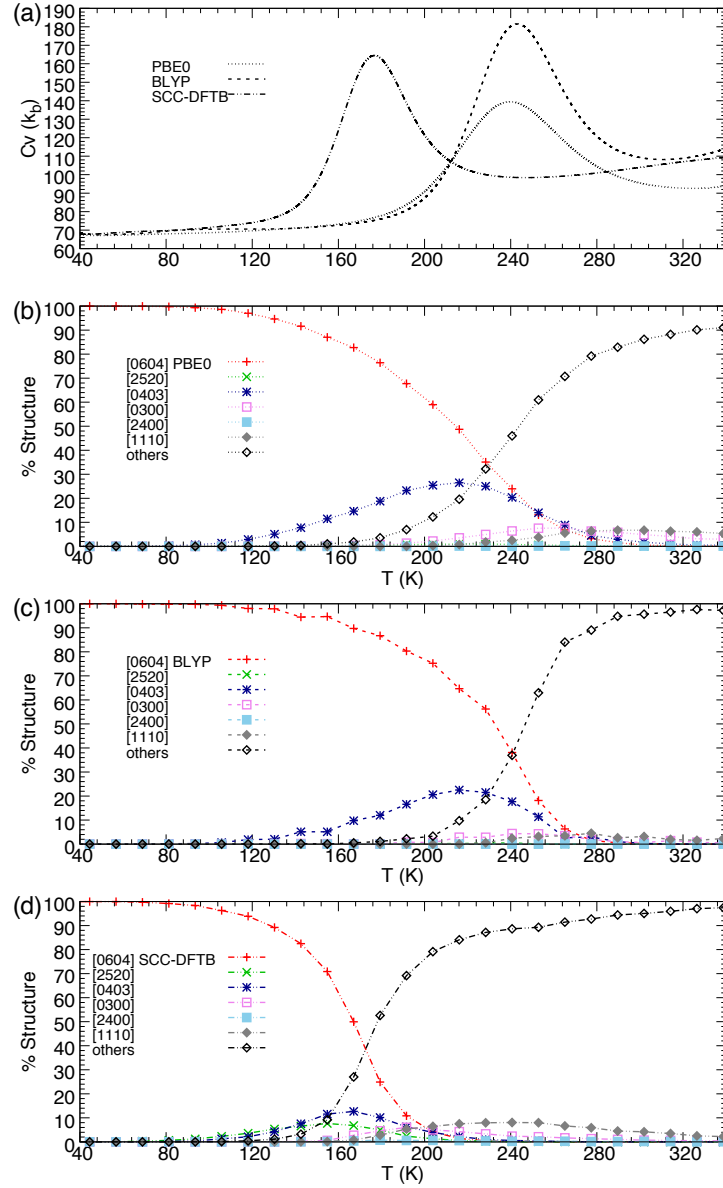


FIGURE 2.13 – (a) Courbes de capacité calorifique de $(H_2O)_8$ obtenues pour des simulations DFT avec les fonctionnelles PBE0 (courbe noire pointillée) et BLYP (courbe noire avec tirés) et avec une simulation SCC-DFTB (courbe noire avec tirés et pointillés). Populations relatives des isomères de plus basse énergie de $(H_2O)_8$ en fonction de la température issues de simulations PTMD aux niveaux DFT avec les fonctionnelles (b) PBE0 et (c) BLYP et au niveau (d) SCC-DFTB.

calorifique similaires obtenues avec ces deux potentiels.

Comme pour $(H_2O)_6$, les figures 2.13 (b), (c) et (d) décrivent les populations

2.5 Application à l'étude du changement de phase dans l'octamère d'eau $(\text{H}_2\text{O})_8$

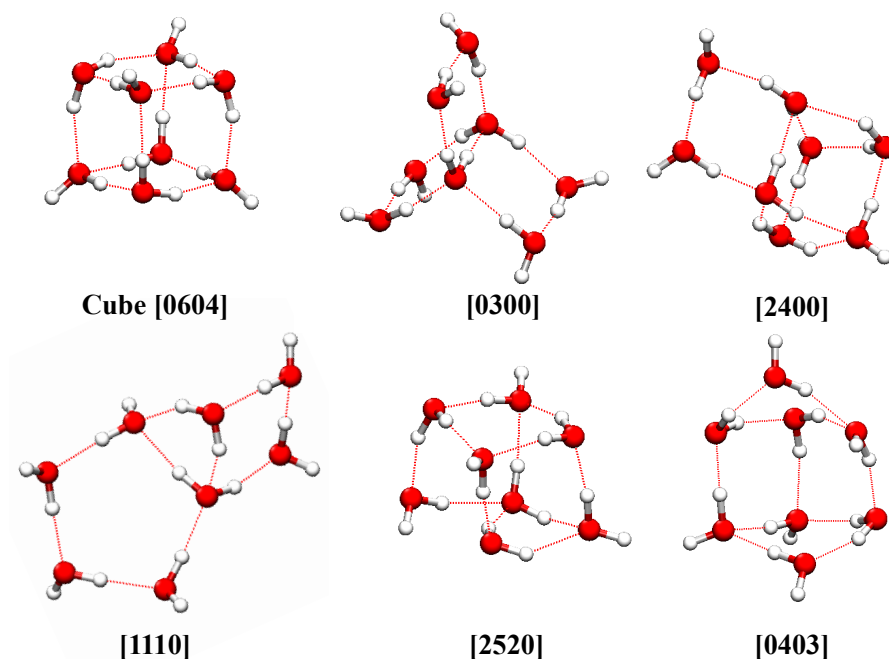


FIGURE 2.14 – Isomères de plus basse énergie de $(\text{H}_2\text{O})_8$. Toutes les structures ont été optimisées et les valeurs des énergies de liaison relatives sont reportées dans le Tableau 2.3.

TABLE 2.3 – Énergies de liaison relatives (en $\text{kJ}\cdot\text{mol}^{-1}$) des cinq isomères de plus basse énergie de $(\text{H}_2\text{O})_8$ obtenues en SCC-DFTB et DFT avec les fonctionnelles BLYP et PBE0. Les géométries des isomères sont représentées sur la figure 2.14.

Isomères	SCC-DFTB	BLYP	PBE0
<i>cube</i> [0604]	0,0	0,0	0,0
[0300]	24,7	30,3	31,1
[2400]	19,8	34,8	33,3
[1110]	33,4	42,0	40,1
[2520]	15,3	-	-

relatives des isomères de plus basse énergie de $(\text{H}_2\text{O})_8$, fournissant ainsi une analyse plus structurale de la transition de phase. Notons ici une différence importante entre $(\text{H}_2\text{O})_6$ et $(\text{H}_2\text{O})_8$. $(\text{H}_2\text{O})_6$ présente un nombre fini et bien déterminé d'isomères de basse énergie. $(\text{H}_2\text{O})_8$ présente quant à lui beaucoup plus d'isomères de basse énergie qu'il est plus difficile de différencier par notre approche de classification par nombre de cycle de molécules d'eau. De plus, hormis les isomères [0604] et [0403], aucun isomère ne présente de population supérieure à 10%. Ainsi, pour plus

2. IMPLÉMENTATION DE LA MÉTHODE PTMD DANS DEMON2K ET APPLICATION À L'ÉTUDE D'AGRÉGATS D'EAU

de clarté, la majorité des isomères formés sont regroupés dans la courbe *others* sur les figures 2.13 (b), (c) et (d). Comme on peut le voir, au niveau SCC-DFTB, la proportion de l'isomère [0604] commence à décroître à partir de ~ 80 K. La proportion de cet isomère croise ensuite celle des *others* à ~ 172 K, soit pour une température plus élevée de ~ 92 K que la température de début de transition. Au niveau DFT, quelque soit la fonctionnelle, la proportion de l'isomère [0604] commence à décroître à partir de ~ 96 et ~ 104 K aux niveaux PBE0 et BLYP, respectivement. La proportion de cet isomère croise ensuite celle des *others* à ~ 228 et ~ 240 K, aux niveaux PBE0 et BLYP, respectivement, soit pour des températures plus élevées de ~ 132 et ~ 136 K que les températures de début de transition. **Ainsi l'isomérisation de [0604] démarre à plus haute température au niveau DFT et se prolonge sur une gamme de températures plus élevée que ce qui est obtenu au niveau SCC-DFTB.**

2.6 Conclusions et perspectives

En conclusions, ce chapitre a présenté en détails **une nouvelle implémentation de l'algorithme PTMD** dans le code deMon2k. A notre connaissance, il s'agit de la **première implémentation native de cet algorithme** dans un code DFT à fonctions de base localisées. Cet algorithme a été efficacement parallélisé par l'utilisation de bibliothèques MPI. Nous avons introduit pour cela **une double couche de parallélisation MPI** qui nous a permis d'implémenter l'algorithme PTMD sans modifier l'architecture MPI existante pour le calcul de la structure électronique. L'absence de code DFT intégrant l'algorithme PTMD, la puissance des supercalculateurs actuels, et la large gamme d'applications potentielles de cet algorithme nous laissent penser qu'il sera largement utilisé dans les années à venir par la communauté des chimistes théoriciens.

Il est important de noter que l'implémentation de la structure MPI en deux couches que nous avons réalisée pourra à l'avenir **servir à paralléliser de nombreuses autres fonctionnalités du code deMon2k**. En effet, toute approche de calcul qui consiste à répéter une opération identique en changeant un nombre fini de paramètres peut tirer parti de cette double couche MPI. C'est le cas par exemple d'une propriété qui nécessiterait le calcul d'une dérivée numérique par rapport aux

positions atomiques. Dans ce cas, chaque déplacement atomique pourrait être pris en charge par un sous-groupe de processus et non pas par l'ensemble des processus disponibles pour le calcul. Cela permet **une meilleure gestion des ressources de calcul** en maximisant l'efficacité parallèle d'une tâche en lui allouant une quantité optimale de processus. On peut également imaginer l'utilisation de la structure MPI en deux couches pour implémenter nativement l'algorithme PIMD dans deMon2k (voir chapitre 3 pour plus de discussions sur ce point) ou tout autre algorithme multi-réplique. Pour permettre cela, nous avons obtenu en 2019 **une bourse de mobilité de 3000 € du Labex NEXT/NanoX** pour réaliser une mobilité de trois mois dans le groupe du Pr. Andreas M. Köster à Mexico. L'objectif de cette mobilité était d'utiliser la structure MPI en deux couches nouvellement développée pour paralléliser le calcul des couplage spin-spin dans deMon2k. Malheureusement, les conditions sanitaires n'ont pas permis de réaliser cette mobilité et le projet envisagé n'a donc pas pu être réalisé pour le moment.

A l'avenir, l'étude de deux types de systèmes chimiques est envisagée avec cette approche : des **agrégats d'aluminium** et des **agrégats d'eau protonés et déprotonés**. Dans le premier cas, il s'agit d'une étude déjà en cours en collaboration avec Andreas M. Köster sur la transition de phase dans les agrégats Al_{27}^+ et Al_{28}^+ . Ces derniers ont été étudiés expérimentalement par Colleen M. Neal *et al.* en 2007 [304] et une première étude théorique a été proposée en 2013 par José-Manuel Vásquez-Pérez *et al.* [305]. Dans cette étude, les auteurs ont pu obtenir des courbes de capacité calorifique théoriques pour Al_{27}^+ et Al_{28}^+ . Dans ce second cas, la température de transition obtenue par le calcul est beaucoup plus haute que celle déterminée expérimentalement. Cependant, ces résultats ont été obtenus par des simulations relativement courtes, sans outils d'échantillonnage accéléré. Notre objectif est de revisiter ce travail pour aboutir à des résultats proprement convergés et ainsi décrire le plus précisément possible la transition de phase dans ces deux espèces. Concernant les agrégats d'eau protonés et déprotonés, ces derniers sont très mal décrits au niveau SCC-DFTB, en tout cas pour les petites tailles [306, 307, 308, 309]. Les transferts de protons intervenant dans ces espèces ainsi que les phénomènes de polarisation au niveau du proton ou de l'hydroxyde, limitent également l'utilisation de champ de force pour leur description. A l'inverse, un potentiel DFT avec fonctionnelle hybride est bien adapté à la description de ces systèmes [310, 311]. Il s'agira donc de revisiter les propriétés des agrégats $(\text{H}_2\text{O})_n\text{H}^+$ et $(\text{H}_2\text{O})_{n-1}\text{OH}^-$ à ce niveau de théorie. Notons

2. IMPLÉMENTATION DE LA MÉTHODE PTMD DANS DEMON2K ET APPLICATION À L'ÉTUDE D'AGRÉGATS D'EAU

pour finir que ces espèces présentent des configurations beaucoup plus ouvertes que les agrégats d'eau neutres [312, 313, 314]. Par conséquent, une analyse des cycles de molécules d'eau n'est pas pertinente dans ce cas. Il s'agira donc de définir des indicateurs structuraux mieux adaptés à ces espèces, quitte à utiliser des indicateurs plus abstraits tels que SOAP (pour *Smooth Overlap of Atomic Positions*) [255].

Impact des effets quantiques
nucléaires sur l'écrantage isotrope
 ^{13}C de carboxylates linéaires

Chapitre 3

Impact des effets quantiques nucléaires sur l'écrantage isotrope ^{13}C de carboxylates linéaires

3.1 Introduction

Plusieurs études ont montré que la prise en compte des **fluctuations thermiques et quantiques des noyaux légers** améliorerait considérablement l'évaluation théorique de paramètres RMN tels que les déplacements chimiques isotropes (δ_{iso}) et permettrait une comparaison quantitative de ces derniers aux valeurs expérimentales [315, 316, 317, 318, 319, 320, 321, 322, 323, 324, 325, 326, 327, 328, 329, 89, 330, 323, 324, 325, 326, 327]. Les **fluctuations thermiques** peuvent être prises en compte en réalisant des moyennes d'ensemble du déplacement chimique sur une trajectoire réalisée en dynamique moléculaire (MD) ou en Monte-Carlo (MC). Pour prendre en compte **les effets quantiques des noyaux (NQEs)**, il faut recourir à des approches plus élaborées, en particulier la méthode basée sur les intégrales de chemin de Feynman, l'approche *Path-Integral Molecular Dynamics* (PIMD) décrite en section 1.3.3 [219]. Dans les deux cas, MD classique et PIMD, la surface d'énergie potentielle peut être calculée par une approche de type champ de force ou une approche *ab-initio*. Les paramètres RMN en revanche ne peuvent être calculés que par une méthode traitant explicitement la structure électronique comme la DFT.

3. IMPACT DES EFFETS QUANTIQUES NUCLÉAIRES SUR L'ÉCRANTAGE ISOTROPE ^{13}C DE CARBOXYLATES LINÉAIRES

Si l'on trouve dans la littérature de nombreux travaux traitant de l'importance de la prise en compte des fluctuations thermiques pour une description théorique précise de paramètres RMN [315, 316, 317, 318, 319, 320, 321, 322, 331, 332, 333, 334, 335, 336, 337, 338], on ne retrouve que **peu d'études dédiées à l'impact des NQEs** [328, 329, 89, 330, 323, 324, 325, 326, 327]. Parmi ces travaux, Michael C. Böhm et Joachim Schulte ont étudié l'impact des NQEs sur les déplacements chimiques ^1H et ^{13}C de divers hydrocarbures [328, 325, 326]. Concernant les systèmes aqueux, on retrouve l'article de John Lobaugh et Gregory A. Voth [324] qui a porté sur le temps de corrélation rotationnel d'une molécule d'eau dans de l'eau liquide et les travaux de Ali A. Hassanali *et al.* [329] et de Michele Ceriotti *et al.* [89] qui ont respectivement déterminé **les effets de la délocalisation quantique des protons** sur leur déplacement chimique dans des hydrates de chlorure d'hydrogène et dans de l'eau liquide. Enfin, pour ce qui est des systèmes biologiques, l'impact des NQEs a été étudié par Sishi Tang et David A. Case sur le déplacement chimique ^{15}N de peptides [327] et par l'équipe de Martin Dračinský sur les paramètres RMN ^1H , ^{13}C et ^{15}N de différentes molécules d'intérêt biologique [330, 339, 323, 340]. Ces derniers comportent systématiquement une part expérimentale qui a permis une comparaison des résultats théoriques et donc leur validation. Ces comparaisons mettent en évidence **l'amélioration systématique des paramètres RMN théoriques grâce à la prise en comptes des NQEs** par l'approche PIMD. Concernant les études dédiées aux déplacements chimiques ^{13}C , il a été montré que **le traitement quantique des noyaux induit un déblindage systématique des carbones**. Ce déblindage résulte de la délocalisation plus marquée des hydrogènes liés aux carbones en PIMD qu'en MD classique. Cette délocalisation a pour effet d'augmenter les distances moyennes C-H et donc de diminuer l'écrantage isotrope ^{13}C .

Le faible nombre de travaux concernant l'influence des NQEs sur le calcul de paramètres RMN est dû à deux raisons principales : (i) le **coût calculatoire élevé des simulations PIMD** et (ii) la **complexité de l'implémentation de cette approche**. Concernant le premier point, l'intérêt croissant porté aux méthodes de type PIMD a donné lieu à **de nombreux développements méthodologiques** [341, 208, 342, 343, 344, 345] qui ont permis de **réduire considérablement le coût de calcul de telles simulations** allant parfois jusqu'à en faire des simulations de routine. Parmi ces améliorations, on compte par exemple les **thermostats de type GLE** que nous avons présenté dans la section 1.3.2 et

que nous avons utilisé dans les travaux de ce chapitre. Concernant le second point, **Michele Ceriotti *et al.* ont développé récemment le programme i-PI [99]** dédié aux simulations de type PIMD. L'intérêt de i-PI réside dans le fait qu'il **peut être interfacé avec n'importe quel programme dédié au calcul d'énergie et de forces**, que ce soit *ab-initio* ou champ de force. Une fois l'interface mise au point, i-PI prendra en charge la propagation des équations du mouvement, le second programme sera lui en charge du calcul de l'énergie et des forces. Un tel programme permet donc d'éviter de reprogrammer l'ensemble des algorithmes nécessaires aux simulations PIMD comme nous l'avons déjà mentionné en introduction du chapitre 2. C'est **cette approche que nous avons privilégié au cours de cette thèse** plutôt que coder nativement l'approche PIMD dans deMon2k, ce qui aurait représenté un effort considérable pour un gain limité. i-PI a d'ores et déjà été interfacé avec de nombreux programmes de DFT en conditions aux limites périodiques tels que Siesta [346], CP2k [273], ou Quantum-Espresso [347] mais pas seulement [348]. En revanche, à notre connaissance, **aucun code de DFT dédié à l'étude des systèmes moléculaires n'a été interfacé avec i-PI** jusqu'à présent. C'est dans ce contexte que nous avons interfacé i-PI avec le code deMon2k [98], que nous avons présenté dans le chapitre 2. **L'implémentation de l'interface avec deMon2k** ainsi que le **fonctionnement du système de communication** entre ces deux codes seront présentés dans ce chapitre en section 3.2.

En plus de présenter notre interface entre les codes i-PI et deMon2k, ce chapitre est également dédié à l'application de cette interface à l'étude de l'influence des fluctuations thermiques et quantiques sur le déplacement chimique ^{13}C de carboxylates linéaires de formule chimique : HCOO^- et $\text{CH}_3(\text{CH}_2)_n\text{COO}^-$ ($n = 0-2, 6$), explicitement solvatés par un agrégat d'eau. À notre connaissance, il s'agit de la première étude à **prendre en compte à la fois les effets de la solvation explicites par de l'eau et les NQEs**. Les ions carboxylates sont les bases conjugués d'acides carboxyliques. Ces molécules sont particulièrement intéressantes car elles possèdent une grande solubilité dans des solvants polaires (eau, alcools, ...). Leur extrémité polaire, composée d'un groupe $-\text{COOH}$, est hautement réactive avec des composés polaires. En raison de ces caractéristiques, les acides carboxyliques trouvent de nombreuses applications en synthèse organique, en chimie des matériaux nanocomposites hybrides organique-inorganique, en chimie des polymères ou comme modificateurs de surface [349].

3. IMPACT DES EFFETS QUANTIQUES NUCLÉAIRES SUR L'ÉCRANTAGE ISOTROPE ^{13}C DE CARBOXYLATES LINÉAIRES

L'influence de l'interaction entre un composé solvaté et l'eau sur les paramètres RMN n'est pas évidente à rationaliser en raison de **l'influence mutuelle du composé solvaté et de l'eau**. En effet, l'évolution dynamique des conformations du soluté modifie l'organisation des molécules d'eau dans sa première sphère de coordination qui en retour vont avoir un impact différent sur les paramètres RMN du soluté. Pour décrire cela, la plupart des études théoriques utilisent un modèle de solvant implicite combiné à une description explicite du solvant pour la première sphère de coordination du soluté uniquement. Cette description inclue donc quelques molécules d'eau pouvant évoluer en MD ce qui permet d'inclure les effets de fluctuations thermiques dans une certaine mesure [350, 334, 335, 336, 337]. Dans la présente étude, nous avons privilégié **une description purement explicite de l'interaction entre les ions carboxylates et le solvant** pour permettre une meilleure description des effets de solvatation. La prise en compte des fluctuations thermiques en traitant explicitement les molécules d'eau en dynamique moléculaire classique et le calcul des paramètres RMN au niveau DFT s'est révélé être une approche de choix pour le calcul du tenseur d'écrantage ^{13}C de molécules amphiphiles [336]. Cependant **de telles simulations peuvent devenir rapidement coûteuses**, en particulier si l'on souhaite y inclure également les NQEs. Pour réduire ce coût, **il est possible d'utiliser un potentiel QM/MM**, voir section 1.2.5) pour la partie dynamique moléculaire. C'est le choix que nous avons fait pour la présente étude que ce soit pour les simulations de MD classique ou les simulations PIMD. Les molécules d'eau du solvant ont été traitées par un potentiel de type champ de force alors que les ions carboxylates ont été décrits au niveau DFT. Les interactions entre les parties QM et MM ont été décrites par un couplage électrostatique [276, 351].

Le chapitre est organisé comme suit : **dans une première partie**, nous présenterons l'implémentation de l'interface entre les codes deMon2k et i-PI et nous expliquerons le fonctionnement des communications entre ces deux codes. **Dans une deuxième partie**, nous présenterons les méthodes ainsi que les paramètres de calcul utilisés pour réaliser cette étude. **La troisième partie** sera dédiée à la validation de ces méthodes. Nous présenterons les résultats de cette étude dans **la quatrième partie**. Les conclusions et perspectives de ce travail seront finalement présentées dans une **cinquième partie**.

3.2 Implémentation de l'interface entre deMon2k et i-PI

3.2.1 Fonctionnement du programme i-PI

i-PI [99] est un programme codé en langage python dédié initialement aux simulations PIMD mais qui peut-être également utilisé pour des simulations de MD classique. La principale caractéristique de ce programme est de pouvoir être facilement interfacé à n'importe quel autre programme dédié au calcul d'énergie et de forces que ce soit aux niveaux *ab-initio*, semi-empirique ou empirique. Ceci permet de réaliser des simulations PIMD sans avoir à implémenter tout le formalisme associé. **i-PI est un logiciel libre** qui peut-être téléchargé depuis un répertoire git : <https://github.com/i-pi/i-pi>. Une communauté internationale de développeurs participe activement à son développement, à l'implémentation de nouvelles approches ainsi qu'à la résolution de bogues. Le fonctionnement schématique d'une simulation de dynamique avec i-PI est représenté en figure 3.1. Sur ce schéma, le programme i-PI (« *i-PI server* » sur la figure) est connecté à n instances d'un code de calcul de structure électronique quelconque (« *client 1,n* » sur la figure). Comme nous l'avons dit précédemment, i-PI est en charge de la propagation des positions atomiques, que ce soit dans un formalisme MD, CMD, RPMD ou autre, tandis que le calcul de l'énergie et des forces est totalement pris en charge par le code client. A présent voyons en détails les différentes étapes d'une simulation avec cette installation.

1. Le code client ouvre une socket. Une socket est une interface de connexion au travers de laquelle peuvent communiquer plusieurs applications. Cette socket sera à la base de la communication entre i-PI et les n instances du code client. Une fois la socket ouverte, le code client y écrit un message spécifiant qu'il est prêt à recevoir les coordonnées atomiques \mathbf{x} . Il se met à lire régulièrement la socket pour déterminer si elles y sont écrites. Pendant ce temps, i-PI génère les structures initiales (une seule pour une simulation MD classique et n pour une simulation PIMD, n étant le nombre de répliques dans le polymère cyclique). Ces structures initiales sont composées des coordonnées \mathbf{x} de chaque réplique ainsi que des dimensions de la maille \mathbf{h} . i-PI lit ensuite le message écrit par

3. IMPACT DES EFFETS QUANTIQUES NUCLÉAIRES SUR L'ÉCRANTAGE ISOTROPE ^{13}C DE CARBOXYLATES LINÉAIRES

- le code client sur la socket. Si ce message précise que le code client est prêt à recevoir les structures, i-PI les écrit. Ensuite, i-PI écrit un message précisant qu'il est prêt à recevoir les forces et autres informations et lit périodiquement le message écrit sur la socket.
2. Chaque instance du code client lit les coordonnées d'une structure sur la socket et calcule l'énergie et les forces correspondantes. Il lit ensuite le message écrit sur la socket. Si ce message spécifie que i-PI est prêt à recevoir les forces, alors le code client écrit les forces sur la socket puis écrit un message disant qu'il est prêt à recevoir une nouvelle structure.
 3. Quand i-PI a reçu les forces de toutes les répliques du polymère cyclique (une seule dans le cas d'une dynamique moléculaire classique), il intègre les équations du mouvement, propage les positions atomiques, calcule les propriétés nécessaires puis écrit à nouveau les coordonnées atomiques sur la socket.
 4. Le code client lit les nouvelles coordonnées d'une structure sur la socket, calcule l'énergie et les forces et le processus se répète.

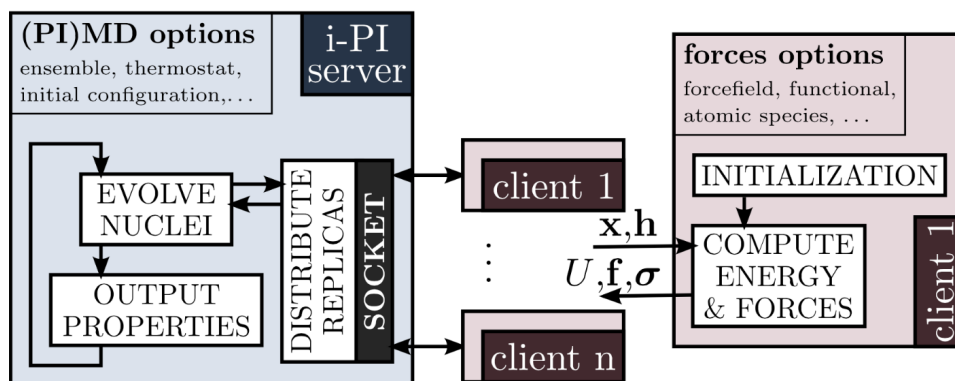


FIGURE 3.1 – Tiré de la référence [99]. Représentation schématique du fonctionnement d'un calcul de dynamique moléculaire réalisé par le programme i-PI (« *i-PI server* ») connecté à n instances d'un code de calcul de structure électronique quelconque (« *client 1, n* »). i-PI génère les structures initiales puis envoie les coordonnées atomiques \mathbf{x} ainsi que les dimensions de la maille \mathbf{h} à chaque instance du code client connecté. Chacune de ces instances calcule l'énergie potentielle U , les forces \mathbf{f} , le tenseur de contraintes $\boldsymbol{\sigma}$ et éventuellement une autre propriété (moment dipolaire, charges atomiques partielles, ...) puis les renvoie à i-PI. A partir de ces données, i-PI intègre les équations du mouvement, fait évoluer les positions atomiques et calcule les propriétés désirées. Il renvoie les nouvelles positions au code client et ainsi de suite.

3.2 Implémentation de l'interface entre deMon2k et i-PI

Pour qu'une simulation i-PI soit la plus efficace en terme de temps de calcul, **il est préférable d'exécuter autant d'instances du code client qu'il y a de répliques n** dans le polymère cyclique. Le nombre optimal de cœurs disponible pour le calcul est donc égal à $N = 1 + n \times p$ où 1 vaut pour i-PI, n pour le nombre d'instances du code client à exécuter et p pour le nombre de cœurs requis par code client lors d'une exécution parallèle de ce dernier. En effet, selon la taille du système étudié et la méthode employée pour le calcul de l'énergie et des forces, il peut être préférable d'exécuter chaque code client en parallèle sur plusieurs cœurs de calcul. C'est notamment le cas lorsque le calcul de l'énergie et des forces est réalisé en DFT. Le nombre total de cœurs N ainsi requis pour une simulation décidera du type de communication à employer entre i-PI et le programme client.

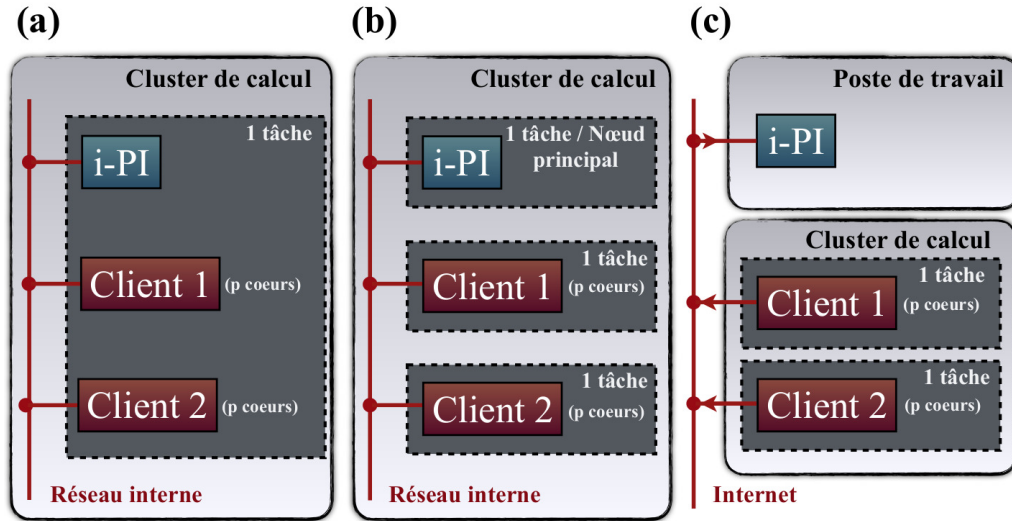


FIGURE 3.2 – Représentation schématique des différentes configurations possibles pour soumettre un calcul i-PI + code client sur un supercalculateur. (a) i-PI et les n instances du code client sont exécutées en une seule tâche sur un même nœud de calcul. La communication se fait au travers du réseau local avec une socket UNIX. (b) i-PI est exécuté sur le nœud principal ou bien sur un nœud de calcul du supercalculateur en une tâche. Les n instances du code client sont exécutées en n tâches sur des nœuds de calcul. La communication se fait au travers du réseau local avec une socket INET. (c) i-PI est exécuté sur un poste de travail local. Les n instances du code client sont exécutées en n tâches sur des nœuds de calcul. La communication se fait au travers du réseau internet avec une socket INET. Dans les trois cas, l'exécution de chaque instance du code client peut être parallélisée sur p cœurs.

3. IMPACT DES EFFETS QUANTIQUES NUCLÉAIRES SUR L'ÉCRANTAGE ISOTROPE ^{13}C DE CARBOXYLATES LINÉAIRES

Comme nous l'avons dit précédemment, les communications se font au travers d'une socket. **Il existe deux types de socket, UNIX et INET.** Les communications avec une socket UNIX sont plus rapides. Toutefois elles ne permettent pas de communications inter-nœuds. i-PI et toutes les instances du code client doivent donc être exécutées sur le même nœud de calcul. L'utilisation d'une socket INET peut donc s'avérer nécessaire lorsque le calcul nécessite plusieurs nœuds de calcul ce qui est le cas lorsque N est supérieur au nombre de cœurs physiques sur un nœud de calcul. La figure 3.2 présente de façon schématique les différentes approches possibles pour exécuter un calcul i-PI + n codes clients. Voici **les trois stratégies possibles pour réaliser un tel calcul** :

- **N est inférieur au nombre de cœurs disponibles sur un nœud de calcul.** On exécute alors i-PI et les n instances du code client sur le même nœud et on utilise une socket UNIX pour la communication, voir figure 3.2 (a). Un script de soumission d'un calcul selon cette stratégie est donné en annexes.
- **N est supérieur au nombre de cœurs disponibles sur un nœud.** On peut alors exécuter i-PI sur un nœud de calcul ou bien sur le nœud principal, c'est à dire la frontale de connexion, et les codes clients sur un ou différents nœuds de calcul et utiliser une socket INET pour la communication, voir figure 3.2 (b). Cette approche est compliquée à mettre en place car elle nécessite d'ouvrir un tunnel *ssh* persistant entre le nœud de calcul sur lequel est exécuté i-PI et les nœuds de calcul sur lesquelles sont exécutés les codes clients. Un effort supplémentaire d'implémentation en *bash* est nécessaire pour automatiser la soumission d'un tel calcul sur un supercalculateur. De plus, les scripts et commandes peuvent varier d'un système d'exploitation à un autre. Les scripts pour la soumission d'un calcul selon cette stratégie sont donnés en annexes.
- **N est supérieur au nombre de cœurs disponibles sur un nœud.** On peut exécuter i-PI sur un poste de travail et les n instances du code client sur les nœuds de calcul d'un cluster et utiliser une socket INET, voir figure 3.2 (c). Cette configuration offre une grande flexibilité pour l'exécution parallèle du code client. En pratique, il convient d'exécuter i-PI sur le poste de travail avec un *wall-clock time* très grand. On peut ensuite exécuter n instances du code client avec le schéma de parallélisation souhaité. Dans le cas où un code client viendrait à s'arrêter pour une raison quelconque, il est possible de simplement le relancer et il se reconnectera à i-PI automatiquement. Comme

3.2 Implémentation de l'interface entre deMon2k et i-PI

pour l'approche précédente, l'inconvénient de cette approche réside dans l'effort d'implémentation *bash* nécessaire pour créer des tunnels *ssh* permanents. De plus, dans ce cas, il faut parfois avoir à ouvrir des ports TCP selon les paramètres de coupe-feu de la station de travail. Les scripts pour la soumission d'un calcul selon cette stratégie sont les même que pour la méthode précédente et sont donnés en annexes.

3.2.2 Implémentation de l'interface entre i-PI et deMon2k

L'implémentation de l'interface entre i-PI et deMon2k nécessite trois routines. Deux de ces routines sont disponibles dans le package i-PI. Une routine écrite en C contient toutes les opérations d'I/O (ouverture/écriture/lecture/fermeture) de la socket et une routine en FORTRAN 90 sert d'interface à l'appel des routines en C depuis le code client. La dernière routine est à implémenter dans le code client pour que celui-ci puisse communiquer avec la socket. Une routine modèle est fournie dans le package i-PI. Nous sommes donc partis de cette routine modèle et l'avons adaptée à la structure de deMon2k. Voici un pseudocode qui présente le fonctionnement de cette routine :

Algorithm 1 Algorithme de la routine d'interface entre deMon2k et i-PI.

```
1: if MASTER then                                ▷ MPI master opens the socket
2:   open socket
3: end if
4: while true do                                    ▷ Infinite loop. Wall-clock time set by i-PI
5:   if master then
6:     READ message FROM SOCKET
7:     if message="STATUS" then
8:       stat=1                                       ▷ Write deMon2k status on socket
9:     else if message="POSDATA" then
10:      stat=2                                       ▷ Read structure on socket. Compute energy and forces
11:    else if message="GETFORCE" then
12:      stat=3                                       ▷ Write energy and forces on socket
13:    end if
14:    SEND stat TO MPI SLAVES PROCESS
```

3. IMPACT DES EFFETS QUANTIQUES NUCLÉAIRES SUR L'ÉCRANTAGE ISOTROPE ^{13}C DE CARBOXYLATES LINÉAIRES

```
15:  end if
16:  if stat=1 then
17:    if master then
18:      if hasdata then
19:        WRITE "HAVEDATA" TO SOCKET  ▷ deMon2k is processing
20:      else
21:        WRITE "READY" TO SOCKET      ▷ deMon2k is free
22:      end if
23:    end if
24:  else if stat=2 then
25:    if master then
26:      READ box FROM SOCKET
27:      READ geometry FROM SOCKET
28:    end if
29:    COMPUTE potential energy
30:    COMPUTE forces
31:    APPLY spherical potential for confinement
32:    hasdata = true
33:  end if
34:  if stat=3 then
35:    if master then
36:      WRITE potential energy TO SOCKET
37:      WRITE forces TO SOCKET
38:      hasdata = false
39:    end if
40:  end if
41: end while
```

Dans cette routine, seul le processus MPI maître est responsable des I/O sur la socket. deMon2k ouvre tout d'abord la socket. Puis, dans une boucle infinie, deMon2k lit le message laissé par i-PI sur la socket. Si le message de i-PI sur la socket est : (i) "STATUS", i-PI demande à deMon2k s'il traite une structure. Si c'est le cas, deMon2 écrit "HAVEDATA" sur la socket. Sinon, il écrit "READY" pour spécifier qu'il est disponible pour recevoir des coordonnées atomiques. (ii) "POSDATA", alors deMon2k va lire un jeu de coordonnées atomiques et des paramètres de maille sur

la socket et calculer l'énergie potentielle et les forces correspondantes. (iii) "GET-FORCE" deMon2k écrit l'énergie potentielle ainsi que les forces provenant des coordonnées qui ont été traitées. Le type de socket (UNIX ou INET) ainsi que l'adresse IP de la machine sur laquelle i-PI est exécuté et le port sont spécifiés dans le fichier input de deMon2k.

3.3 Méthodes employées et paramètres de calcul

Tous les calculs présentés dans ce chapitre ont été réalisés avec le programme deMon2k dans sa version développeur 6.0.2 [98], seul ou combiné au code i-PI version 2.0 [99]. Tous les systèmes étudiés sont composés d'un ion carboxylate explicitement solvato par 34 molécules d'eau en présence ou non d'un contre-cation Na^+ . Lorsque le contre-cation est présent, le système est neutre tandis que lorsqu'il est absent le système est chargé -1. Le choix qui a été fait de ne pas totalement solvater les ions carboxylates par de l'eau liquide dans un système périodique mais uniquement par un agrégat de 34 molécules d'eau s'appuie sur deux arguments. Premièrement, une solvation totale aurait requis de trop grande ressources computationnelles. Deuxièmement, l'influence de la solvation par les molécules d'eau provient principalement de leur interaction avec la partie hydrophile du carboxylate (le groupement $-\text{COO}^-$) et une solvation totale de la chaîne alkyle n'est donc pas nécessaire.

Calculs DFT. Comme nous l'avons présenté dans la section 1.2.3, deMon2k utilise un ensemble de fonctions de bases auxiliaires pour le calcul de l'ajustement variationnel du potentiel de Coulomb [352]. Dans les travaux présentés dans ce chapitre, nous avons utilisé la base auxiliaire automatiquement générée GEN-A2 [98, 288, 289] en combinaison avec une base double zêta à valence séparée (DZVP) pour tous les atomes traités en DFT [290]. La fonctionnelle d'échange-corrélation GGA BLYP [139, 136] a été employée. Le potentiel d'échange-corrélation a été intégré numériquement sur une grille adaptative de précision « MEDIUM » (10^{-4} u.a.) [353]. La tolérance pour la convergence de la densité auxiliaire et pour la convergence des cycles SCF ont été respectivement fixées à $5 \cdot 10^{-4}$ et 10^{-5} u.a.

Calculs QM/MM. Pour tous les calculs QM/MM présentés dans ce chapitre, la partie QM composée de l'ion carboxylate et du sodium a été traitée au niveau

3. IMPACT DES EFFETS QUANTIQUES NUCLÉAIRES SUR L'ÉCRANTAGE ISOTROPE ^{13}C DE CARBOXYLATES LINÉAIRES

DFT avec les mêmes paramètres que ceux décrits plus haut. Les molécules d'eau ont été traitées en mécanique moléculaire. **deMon2k possède une implémentation native de la méthode QM/MM** qui permet son utilisation en combinaison avec de nombreux algorithmes de modélisation : l'optimisation de géométrie, la dynamique moléculaire classique, la dynamique de Ehrenfest, la dynamique électronique en temps réel à l'échelle de l'attoseconde [276]. Dans cette étude, nous avons utilisé le champ de force TIP3P pour décrire les molécules d'eau [145]. Pour décrire l'interaction des atomes des ions carboxylates dans la partie MM, nous avons choisi les paramètres du champ de force OPLS-AA [143, 144]. Les termes de couplage entre les parties QM et MM ont été traités par une approche d'incorporation électrostatique (*electrostatic embedding* en anglais).

Dynamique moléculaire classique. Toutes les simulations ont été réalisées dans l'ensemble canonique à une température de 300 K. Le contrôle de la température a été réalisé à l'aide d'un thermostat Nosé-Hoover composé d'une chaîne de 3 thermostats avec une fréquence de 100 cm^{-1} [187, 188, 190]. Les équations du mouvement ont été intégrées avec un pas de temps de 1,0 fs. Le temps total de chaque simulation est de 120 ps. Les 20 premières picosecondes ont été considérées comme du temps d'équilibration et n'ont pas été prises en compte dans les analyses. Pour éviter que des molécules d'eau ne s'évaporent lors des simulations, nous avons confiné le système dans un potentiel sphérique de forme $V(r) = 0,008 \times (r - r_0)^4$ (en u.a.), où r_0 est le rayon de la sphère centrée sur le centre de masse du système. Nous avons utilisé un rayon r_0 de 16 Å pour tous les systèmes étudiés dans ce chapitre.

Dynamique moléculaire PIMD. Les simulations PIMD ont été réalisées à l'aide du programme i-PI v. 2.0 interfacé avec deMon2k. Les simulations PIMD ont également été réalisées dans l'ensemble canonique à une température de 300 K. Pour les simulations QM, 6 répliques ont été utilisées dans le polymère cyclique (RP). Pour les simulations QM/MM, plusieurs combinaisons entre type de thermostat (PILE-L, PILE-G, PIGLET) et nombre de répliques dans le RP ont été testées. La convergence des NQEs pour ces différentes combinaisons est présentée dans la section 3.4. Finalement, nous avons choisi d'utiliser 12 répliques dans le RP et le thermostat PIGLET basé sur l'équation de Langevin généralisée [204, 208]. Ce thermostat permet de converger les NQEs pour un nombre de répliques réduit dans le RP. Pour toutes les simulations PIMD, les équations du mouvement ont été intégrées avec

3.3 Méthodes employées et paramètres de calcul

un pas de temps de 0,5 fs. Pour les simulations utilisant les thermostats PILE-L et PILE-G, un coefficient de friction de 100 fs a été utilisé. Pour les simulations utilisant un thermostat PIGLET, les matrices de friction et de diffusion ont été générées automatiquement sur le site web GLE4MD [207] de sorte à ce que les fréquences inférieures à 4170 cm^{-1} soient efficacement thermostatées. Le temps total de chaque simulation PIMD est de 120 ps et les 20 premières picosecondes ont été considérées comme temps d'équilibration et n'ont pas été prises en compte pour les analyses.

Configurations initiales. Pour générer les configurations initiales, chaque ion carboxylate a été préalablement placé au centre d'une maille cubique de 10 Å de côté et solvatés par de l'eau. Les structures ont été optimisées en MM avec un potentiel OPLS-AA pour les atomes des carboxylates et TIP3P pour les molécules d'eau. L'ion carboxylate ainsi que les 34 molécules d'eau les plus proches de la tête polaire ont ensuite été extraites et utilisées comme configuration initiale des simulations de MD et PIMD. La figure 3.3 représente une configuration du système octanoate- Na^+ solvaté par 34 molécules d'eau.

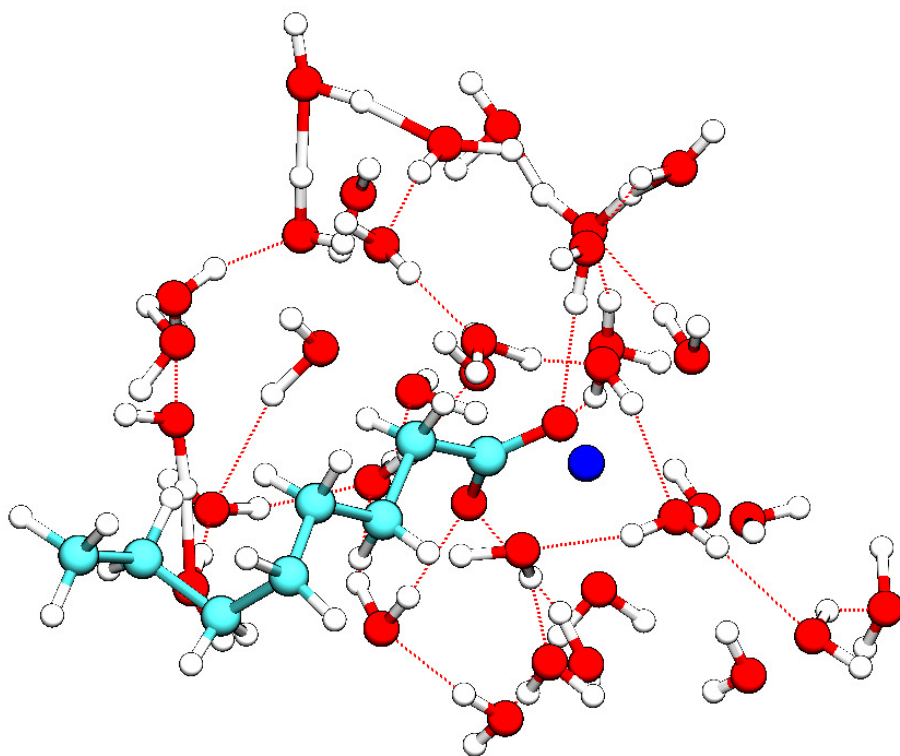


FIGURE 3.3 – Configuration du système octanoate- Na^+ solvaté par 34 molécules d'eau issue d'une simulation de MD.

3. IMPACT DES EFFETS QUANTIQUES NUCLÉAIRES SUR L'ÉCRANTAGE ISOTROPE ^{13}C DE CARBOXYLATES LINÉAIRES

Calcul des paramètres RMN. Les écrantages isotropes (σ_{iso}) ^{13}C ont été calculés dans le formalisme ADFT-GIAO implémenté dans deMon2k [354]. Pour ces calculs, nous avons utilisé la base auxiliaire automatiquement générée GEN-A2* [98, 288, 289] en combinaison avec une base double zêta à valence séparée (DZVP) pour tous les atomes [290]. La tolérance pour l'intégration numérique du potentiel d'échange-corrélation sur la grille, ainsi que celle de la convergence de la densité auxiliaire et de la convergence SCF ont respectivement été fixés à 10^{-5} , 10^{-4} et 10^{-7} u.a. L'énergie et le potentiel d'échange-corrélation ont été décrits à l'aide de la fonctionnelle BLYP [139, 136] .

Les déplacements chimiques isotropes ^{13}C δ_{iso} ont été calculés relativement au ^{13}C σ_{iso} du tétraméthylsilane (TMS) selon l'expression $\delta_{iso,i} = \sigma_{iso,TMS} - \sigma_{iso,i}$. Les valeurs de $\sigma_{iso,TMS}$ ont été obtenues à partir de simulations de MD et PIMD d'une molécule de TMS seule dans le vide utilisant les mêmes paramètres que ceux décrits plus haut. Ces valeurs ont été moyennées sur les quatre carbones du TMS et sont égales à $176,4 \pm 0,3$ en MD et $172,8 \pm 0,2$ en PIMD. Chaque valeur de ^{13}C σ_{iso} a été moyennée sur 1000 structures linéairement réparties le long d'une trajectoire de 100 ps. Dans le cas des simulations PIMD, pour chaque structure, la valeur de ^{13}C σ_{iso} a également été moyennée sur toutes les répliques du RP. Pour chaque valeur moyenne de ^{13}C σ_{iso} , l'erreur statistique sur la moyenne est évaluée comme étant l'écart-type de la moyenne glissante du ^{13}C σ_{iso} . Pour les valeurs issues de différences de ^{13}C σ_{iso} , comme par exemple la différence entre le ^{13}C σ_{iso} d'un carbone d'un carboxylate avec le ^{13}C σ_{iso} du TMS, l'erreur statistique a été évaluée comme étant l'écart-type de la moyenne glissante de la différence entre les moyennes glissantes des deux valeurs de ^{13}C σ_{iso} .

3.4 Validation des méthodes

Avant de rentrer dans les détails de la discussion des résultats, nous allons d'abord discuter de la validité du potentiel QM/MM utilisé en comparant les résultats obtenus par des simulations QM/MM à ceux obtenus par des simulations QM. Pour ce faire, nous allons comparer les résultats issus de simulations MD et PIMD chacune réalisée en QM et en QM/MM pour le système méthanoate (HCOO^-) sans contre-ion Na^+ . En nous basant sur le même système, nous discuterons ensuite de la

convergence des NQEs en fonction du thermostat employé et du nombre de répliques utilisées pour décrire le polymère cyclique. Enfin, nous testerons la convergence statistique des résultats vis à vis de la durée des simulations et du nombre de structures utilisées pour évaluer les valeurs moyennes de $^{13}\text{C } \sigma_{iso}$.

3.4.1 Validation de la méthode QM/MM

Le grand nombre d'atomes présent dans les systèmes étudiés ainsi que le temps de simulation nécessaire pour assurer une convergence statistique des résultats rend difficile en terme de coût de calcul le traitement explicite de la structure électronique en DFT du carboxylate et des 34 molécules d'eau. Si c'est le cas en MD, ça l'est d'autant plus en PIMD où le calcul de l'énergie et des forces est multiplié par le nombre de répliques dans le polymère cyclique. **L'utilisation d'un potentiel QM/MM apparaît alors comme un bon compromis entre précision et coût de calcul.** Cependant, il faut s'assurer que l'utilisation d'un potentiel QM/MM n'introduit pas d'erreur sur l'évaluation des écrantages isotropes ^{13}C . Pour valider l'utilisation du potentiel QM/MM décrit plus haut, nous avons comparé le décalage induit par le traitement quantique des noyaux en PIMD sur les $^{13}\text{C } \sigma_{iso}$ en QM et en QM/MM. Si cet écart est le même pour les deux potentiels, alors le potentiel QM/MM sera validé. Pour décrire correctement les NQEs avec un nombre de répliques suffisamment faible pour traiter l'ensemble du système au niveau DFT, nous avons utilisé le thermostat PIGLET avec 6 répliques. Les valeurs moyennes $^{13}\text{C } \sigma_{iso}$ issues des simulations MD et PIMD en QM et QM/MM sont présentées dans le tableau 3.4.1. On remarque que quelque soit le potentiel employé, l'inclusion des NQEs par l'approche PIMD induit un décalage vers le bas des valeurs de $^{13}\text{C } \sigma_{iso}$ ($-1,6 \pm 0,6$ ppm en QM et $-1,4 \pm 0,8$ ppm en QM/MM). Bien que ce décalage ne soit pas exactement le même avec les deux potentiels, la différence est inférieure à l'incertitude statistique. Par conséquent, comme **l'utilisation du potentiel QM/MM ne modifie pas le décalage des valeurs de $^{13}\text{C } \sigma_{iso}$ induit par les NQEs** et qu'il permet de réduire considérablement les temps de calcul, **nous avons employé ce potentiel pour le reste des simulations réalisées dans ce chapitre.** Pour la description MM du cation Na^+ , nous avons testé les paramètres des champ de force GBSA [355, 356] et OPLS-AA en MD. Les valeurs correspondantes des $^{13}\text{C } \sigma_{iso}$ sont reportées dans le tableau 3.3. On peut voir que la différence de la valeur de $^{13}\text{C } \sigma_{iso}$ calculée avec les

3. IMPACT DES EFFETS QUANTIQUES NUCLÉAIRES SUR L'ÉCRANTAGE ISOTROPE ^{13}C DE CARBOXYLATES LINÉAIRES

TABLE 3.1 – Comparaison des moyennes thermiques du ^{13}C σ_{iso} (en ppm) du méthanoate issue des simulations MD et PIMD en QM et QM/MM. Les valeurs de $\Delta\sigma_{iso}$ sont les différences entre ces deux valeurs. L'incertitude statistique est calculée selon la procédure décrite dans la section 3.3.

	QM	QM/MM
$\sigma_{iso,MD}$	$2,8 \pm 0,3$	$-0,6 \pm 0,5$
$\sigma_{iso,PIMD}$	$1,2 \pm 0,8$	$-2,0 \pm 0,8$
$\Delta\sigma_{iso}$	$-1,6 \pm 0,6$	$-1,4 \pm 0,8$

deux champ de force est faible, $-2,3 \pm 0,5$ et $-2,9 \pm 0,8$ avec les paramètres GBSA et OPLS-AA, respectivement. Une étude comparative détaillée de ces deux champ de force n'étant pas l'objet de ce travail, nous avons arbitrairement décidé d'utiliser les paramètres OPLS-AA pour la description du Na^+ dans le reste de cette étude.

3.4.2 Convergence des NQEs

Maintenant que l'utilisation du potentiel QM/MM a été validée, il faut s'assurer que les NQEs dans les simulations PIMD puissent être convergés. **La convergence des NQEs dépend principalement de deux paramètres : le type de thermostat utilisé et le nombre de répliques** dans le polymère cyclique. Nous avons étudié ces deux paramètres en détails pour la valeur de ^{13}C σ_{iso} du méthanoate en utilisant plusieurs combinaisons de thermostats et de nombre de répliques. Les résultats sont présentés sur la figure 3.4 et dans le tableau 3.2. Le tableau 3.2 présente le décalage $\Delta\sigma_{iso}$ des valeurs de ^{13}C σ_{iso} résultant de la prise en compte des NQEs par l'approche PIMD pour différentes combinaisons de thermostat et de nombre de répliques. On observe qu'en PIMD, les valeurs de ^{13}C σ_{iso} sont systématiquement inférieures à la valeur issue d'une simulation de MD classique (BOMD sur la figure 3.4). On peut regrouper les valeurs obtenues par des simulations PIMD en deux groupes : (i) PILE-L avec 32 répliques, PILE-L avec 16 répliques et PIGLET avec 12 répliques qui donnent des valeurs de ^{13}C σ_{iso} proches de $-2,0$ ppm ; (ii) PILE-G avec 16 répliques et PIGLET avec 6 répliques qui donnent des valeurs de ^{13}C σ_{iso} égales à $-1,4$ ppm. D'après la figure 3.4, on voit que ces valeurs sont à chaque fois convergée vis à vis du nombre de structures prise en compte pour calculer la moyenne. Les simulations PIGLET avec 12 répliques, PILE-G avec 16 répliques et

PILE-L avec 32 répliques sont celles qui convergent le plus vite avec une valeur moyenne relativement stable à partir de 600 structures. La valeur de PILE-G est même relativement stable à partir de 200 structures. Après cette analyse, nous pouvons justifier l'**utilisation du thermostat PIGLET avec 12 répliques pour le reste des simulations**. En effet cette combinaison converge vers la même valeur de $^{13}\text{C } \sigma_{iso}$ qu'avec le thermostat PILE-L et 32 répliques mais elle permet de réduire drastiquement le coût de calcul.

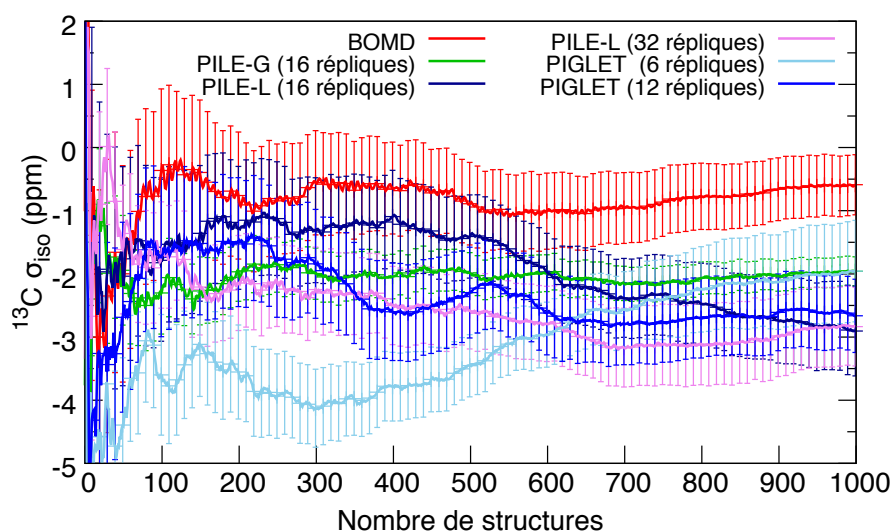


FIGURE 3.4 – Moyennes glissantes des valeurs de $^{13}\text{C } \sigma_{iso}$ (en ppm) issues d'une simulation de MD classique et de cinq simulations PIMD avec différentes combinaisons de thermostats et de nombre de répliques. Les barres verticales correspondent à l'incertitude statistique calculée selon la procédure décrite dans la section 3.3.

3.4.3 Convergence statistique

Les structures des cinq carboxylates étudiés et la numérotation des carbones est présentée sur la figure 3.5. En allant du méthanoate vers l'octanoate, l'allongement de la chaîne carbonée peut introduire des **problèmes de convergence en raison des changements conformationnels qui peuvent intervenir autour de chaque atome de carbone**. Si le méthanoate est exempt de tels changements, il peut y avoir plusieurs rotations le long des liaisons C–C dans l'octanoate. Cependant, comme on peut le voir sur la figure 3.6, les valeurs de $^{13}\text{C } \sigma_{iso}$ des atomes de carbones allant de C1 à C8 convergent rapidement avec le nombre de structures

3. IMPACT DES EFFETS QUANTIQUES NUCLÉAIRES SUR L'ÉCRANTAGE ISOTROPE ^{13}C DE CARBOXYLATES LINÉAIRES

TABLE 3.2 – Valeurs ^{13}C σ_{iso} moyennes (en ppm) du méthanoate issues de cinq simulations PIMD réalisées avec différentes combinaisons de thermostat et de nombre de répliques et exprimées par rapport à la valeur obtenue par une simulation de MD classique. La différence est notée $\Delta\sigma_{iso}$. Toutes les simulations ont été réalisées en QM/MM. L'incertitude statistique est évaluée selon la procédure décrite dans la section 3.3.

	PILE-L (32 répliques)	PILE-L (16 répliques)	PILE-G (16 répliques)
$\Delta\sigma_{iso}$	$-2,2 \pm 0,8$	$-2,3 \pm 0,9$	$-1,4 \pm 0,6$
	PIGLET (6 répliques)	PIGLET (12 répliques)	
$\Delta\sigma_{iso}$	$-1,4 \pm 0,8$	$-2,1 \pm 0,6$	

prises en compte pour l'analyse. L'incertitude statistique est très faible par rapport aux valeurs moyennes de ^{13}C σ_{iso} que ce soit MD ou en PIMD. Quelque soit le carbone pris en compte, l'erreur est inférieure à 1,0 ppm. Sur la figure 3.6, on peut aussi voir qu'en dehors des carbones C4 et C5, la différence de valeur de ^{13}C σ_{iso} entre chaque carbone est systématiquement supérieure à l'incertitude statistique évaluée pour chaque carbone. L'ordre des carbones selon leur valeur de ^{13}C σ_{iso} est donc convergé excepté pour les carbones C4 et C5. Cette ordre est le même en MD qu'en PIMD. Ainsi, les NQEs n'ont pas d'impact sur l'ordre des valeurs de ^{13}C σ_{iso} . En ce qui concerne les valeurs de ^{13}C σ_{iso} des carbones C4 et C5, elles sont très proches l'une de l'autre en MD comme en PIMD. Ce résultat est aussi observé expérimentalement. Dans ce cas, l'attribution des deux signaux correspondants n'est pas faite [357].

D'autres observations peuvent être faites d'après la figure 3.6. Tout d'abord, on peut remarquer une grande différence entre la valeur de ^{13}C σ_{iso} (> 140 ppm) du le carbone C1 et celles des autres carbones. Ce résultat n'est pas surprenant étant donné que le carbone C1 est lié à 2 atomes d'oxygène qui vont fortement le déblinder. Les valeurs de ^{13}C σ_{iso} des carbones C2 à C7 sont relativement similaires alors que le C8 est plus blindé. Ce résultat est cohérent avec la nature des groupements, $-\text{CH}_2-$ pour les carbones C2 à C7 et $-\text{CH}_3$ pour le carbone C8. Si l'on compare les valeurs de la figure 3.6 (a) et (b), on peut voir que la prise en compte des NQEs décale systématiquement les valeurs de ^{13}C σ_{iso} vers le bas d'environ 5-7 ppm pour tous

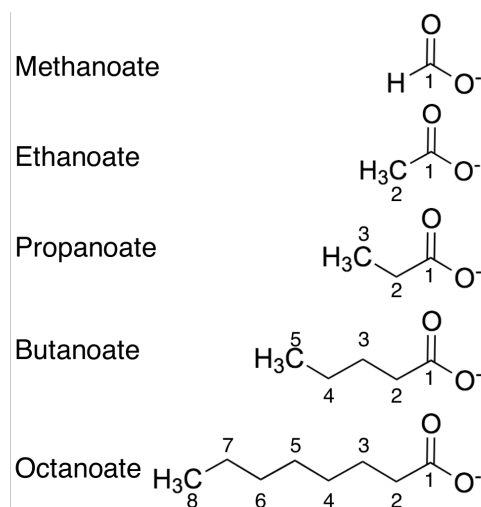


FIGURE 3.5 – Représentations schématique des ions carboxylates étudiés dans ces travaux et numérotation des carbones.

les carbones hormis le C1, ce qui a pour effet de diminuer tous les écarts de ^{13}C σ_{iso} entre le C1 et les autres carbones. Le décalage des ^{13}C σ_{iso} induit par les NQEs est largement supérieur à l'incertitude statistique sur les valeurs moyennes. **On peut donc affirmer que les NQEs ont un impact important sur les valeurs de ^{13}C σ_{iso} .**

3.5 Résultats et discussion

3.5.1 Discussion sur les valeurs absolues de ^{13}C σ_{iso}

Les valeurs moyennes des ^{13}C σ_{iso} de tous les carbones des cinq carboxylates étudiés en présence ou non du contre-ion Na^+ sont reportées dans le tableau 3.3. Comme cela a été dit précédemment, **la valeur de ^{13}C σ_{iso} du carbone C1 est beaucoup plus basse que celle des autres carbones**, et cela quelque soit le composé. De plus, cette valeur décroît lorsque la taille de la chaîne carbonée augmente. **La prise en compte des NQEs décale systématiquement la valeur de ^{13}C σ_{iso} de tous les carbones vers le bas** jusqu'à 6,8 ppm pour le C5 de l'octanoate en présence de Na^+ . Bien que le lien entre la prise en compte des NQEs et leur impact sur une propriété donnée ne soit pas toujours évident à déterminer, deux

3. IMPACT DES EFFETS QUANTIQUES NUCLÉAIRES SUR L'ÉCRANTAGE ISOTROPE ^{13}C DE CARBOXYLATES LINÉAIRES

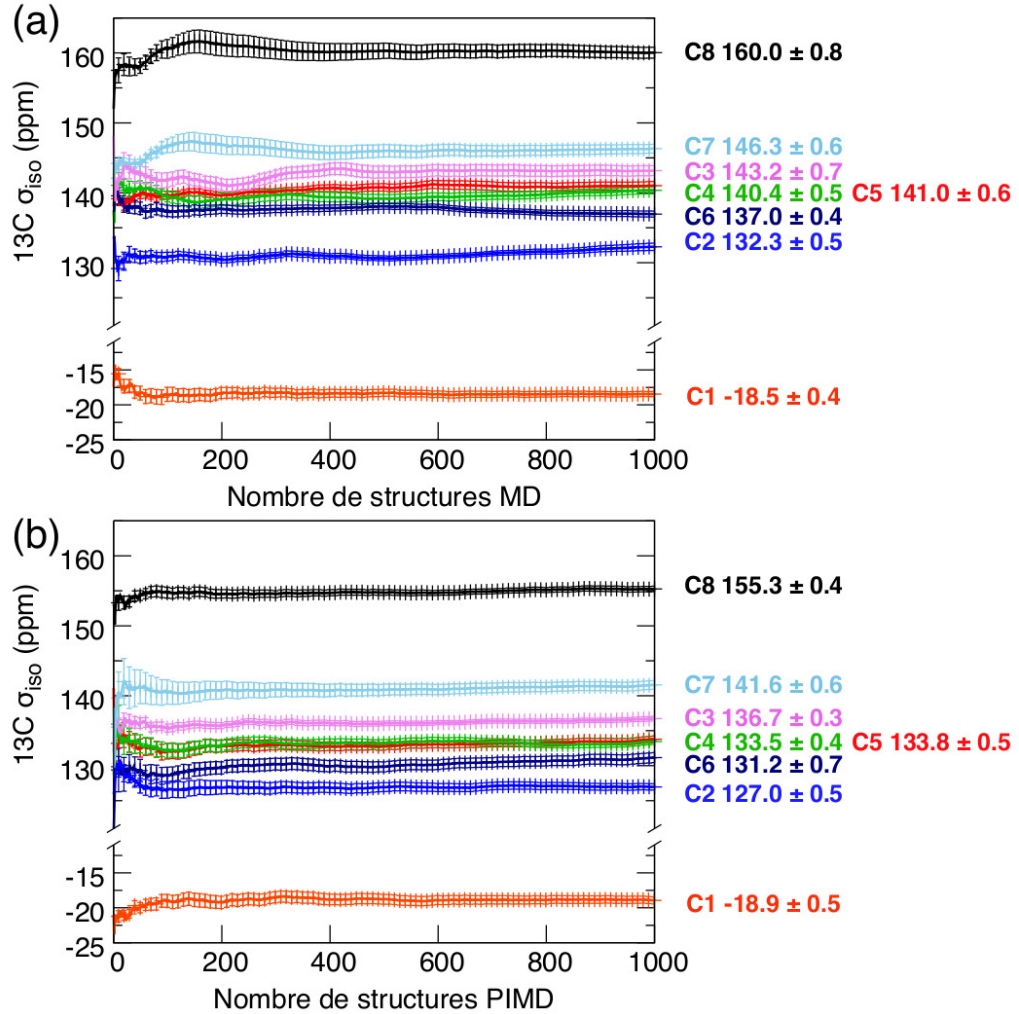


FIGURE 3.6 – Moyennes glissantes des valeurs de $^{13}\text{C } \sigma_{iso}$ des carbones C1 à C8 (voir figure 3.5 pour la numérotation) du système octanoate- Na^+ obtenus à partir de simulations de (a) MD classique et (b) PIMD. L'incertitude statistique est calculée selon la procédure décrite dans la section 3.3.

raisons distinctes peuvent être proposées au décalage vers le bas des valeurs de $^{13}\text{C } \sigma_{iso}$ en PIMD. **La première raison** est que les fluctuations quantiques des liaisons C–H modifient les distances C–H moyennes qui diffèrent donc entre les simulations MD et PIMD. Ces fluctuations quantiques conduisent alors à des distances C–H moyennes plus courtes ou plus longues que celles obtenues par une simulation de MD classique. Dans le cas d'un système purement harmonique, la valeur moyenne serait identique quelque soit les fluctuations autour de la position d'équilibre mais ce n'est pas le cas ici. La figure 3.7 montre les différences de fluctuations des dis-

tances C–H pour les carbones C2 et C3 du propanoate entre une simulation de MD classique et une simulation PIMD. Cette figure montre qu’en PIMD, les distances extrêmes sont plus petites et plus grandes qu’en MD, conduisant respectivement à des valeurs plus grandes et plus petites de $^{13}\text{C } \sigma_{iso}$. Moyennées, les distances C–H sont légèrement plus grandes en PIMD. Par conséquent, les valeurs de $^{13}\text{C } \sigma_{iso}$ sont légèrement plus faibles en PIMD, comme cela a été dit précédemment. Pour le carbone C1, cet effet n’est présent que dans le méthanoate. En effet, pour les autres carboxylates, le carbone C1 n’est pas lié à un atome d’hydrogène. **La seconde raison** est que les fluctuations quantiques des liaisons O–H des molécules d’eau liées aux atomes d’oxygène du groupement CHOO^- perturbent la structure électronique du carboxylate et peuvent donc modifier les valeurs de $^{13}\text{C } \sigma_{iso}$. On s’attend à ce que le carbone C1 soit le plus impacté par cet effet. **Les variations des valeurs de $^{13}\text{C } \sigma_{iso}$ présentées dans le tableau 3.3 peuvent être expliquées par la combinaison de ces deux effets.** À part pour le méthanoate pour lequel les NQEs décalent le $^{13}\text{C } \sigma_{iso}$ du C1 de 2,1 et 1,5 ppm selon que le Na^+ soit présent ou pas, les carbones C1 sont les moins impactés par les NQEs avec un maximum de 1,4 ppm dans le cas de l’éthanoate sans le Na^+ . Les valeurs moyennes des $^{13}\text{C } \sigma_{iso}$ des autres carbones sont diminuées de plus de 3,0 ppm par les NQEs avec un maximum de 6,8 ppm pour le C5 de l’octanoate en présence du Na^+ . On peut en conclure que **la modification des valeurs de $^{13}\text{C } \sigma_{iso}$ due aux NQEs est principalement liée aux fluctuations quantiques des liaisons C–H** tandis que les fluctuations quantiques des molécules d’eau n’ont qu’un impact limité sur les valeurs de $^{13}\text{C } \sigma_{iso}$.

Le tableau 3.3 met aussi en évidence l’impact de la présence du contre-ion Na^+ sur les valeurs de $^{13}\text{C } \sigma_{iso}$. Dans le cas du méthanoate, sa présence induit un décalage vers le bas de la valeur de $^{13}\text{C } \sigma_{iso}$ du C1 de 2,3 et 1,7 ppm en MD et PIMD, respectivement. En revanche, pour l’éthanoate, le propanoate et le butanoate, la présence du Na^+ n’a pratiquement aucune influence sur les valeurs de $^{13}\text{C } \sigma_{iso}$ à part pour le carbone C1. Même dans ce cas, son influence est relativement faible avec un maximum de -2,0 ppm pour l’éthanoate en MD. De façon surprenante, l’influence du Na^+ est plus marquée dans le cas de l’octanoate. Sa présence induit un décalage vers le bas de 1,0 - 2,0 ppm des carbones C1 à C5. Les carbones C6, C7 et C8 quant à eux, ne sont pas affectés par la présence ou l’absence du Na^+ . De façon générale, ces résultats montrent que le contre-ion n’a qu’un impact limité sur les valeurs de $^{13}\text{C } \sigma_{iso}$. Ces valeurs de $^{13}\text{C } \sigma_{iso}$ diffèrent considérablement des valeurs

3. IMPACT DES EFFETS QUANTIQUES NUCLÉAIRES SUR L'ÉCRANTAGE ISOTROPE ^{13}C DE CARBOXYLATES LINÉAIRES

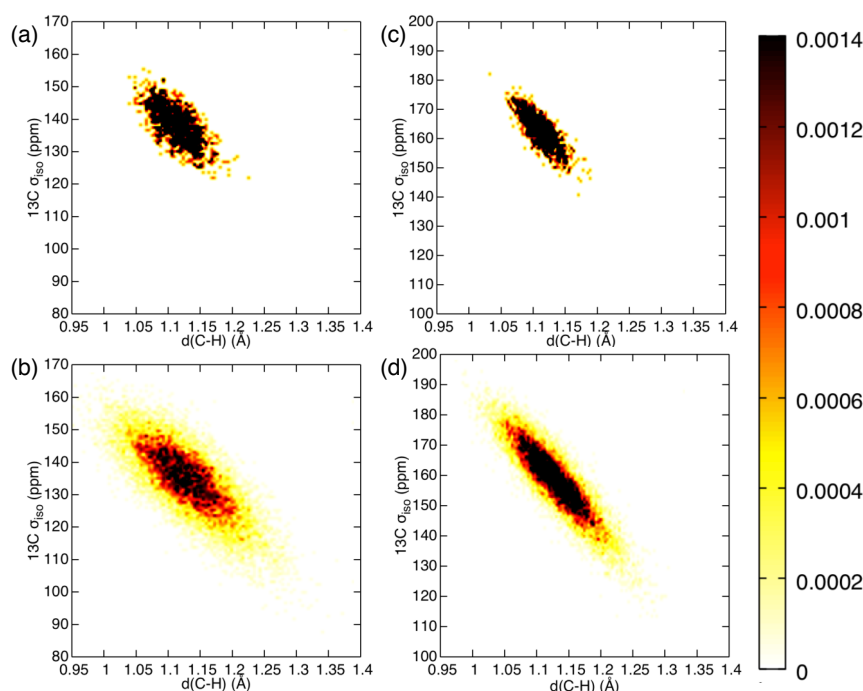


FIGURE 3.7 – Densité de probabilité des valeurs de ^{13}C σ_{iso} instantanées du carbone C2 du groupement CH_2 du propanoate issues d’une simulation de (a) MD classique (b) et d’une simulation PIMD en fonction de la distance moyenne des liaisons C–H de ce carbone. Même données pour le carbone C3 du groupement CH_3 pour une simulation de (c) MD classique (d) et d’une simulation PIMD.

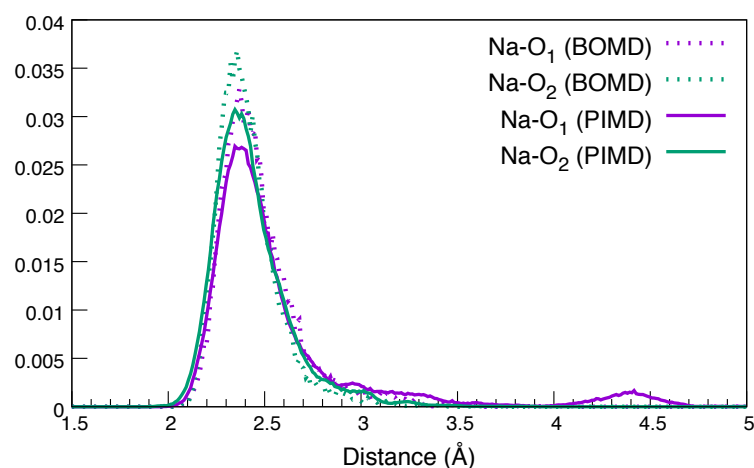


FIGURE 3.8 – Densité de probabilité des distances Na–O₁ (violet) et Na–O₂ (vert) dans le système octanoate- Na^+ en MD (pointillés) et PIMD (lignes pleines). O₁ et O₂ désignent les deux oxygènes du groupement $-\text{COO}^-$.

3.5 Résultats et discussion

TABLE 3.3 – Valeurs moyennes de ^{13}C σ_{iso} (en ppm) de tous les carbones des cinq carboxylates issues des simulation de MD classique et PIMD. Pour chaque carboxylate, les valeurs ont été calculées pour le système chargé négativement en interaction avec l’agrégat de 34 molécules d’eau et pour le système en présence du contre-ion Na^+ et de l’agrégat de 34 molécules d’eau. Dans le deuxième cas, nous avons systématiquement utilisé les paramètres OPLS-AA pour la description du Na^+ excepté pour le méthanoate où les paramètres GBSA ont aussi été testés. Pour le propanoate, nous avons aussi réalisé une simulation en phase gaz sans le Na^+ . L’incertitude statistique est calculée selon la procédure décrite dans la section 3.3.

	C1	C2	C3	MD C4	C5	C6	C7	C8
Methanoate	$-0,6 \pm 0,5$	–	–	–	–	–	–	–
Methanoate+ Na^+ (GBSA)	$-2,3 \pm 0,5$	–	–	–	–	–	–	–
Methanoate+ Na^+	$-2,9 \pm 0,8$	–	–	–	–	–	–	–
Ethanoate	$-11,9 \pm 0,6$	$149,7 \pm 0,3$	–	–	–	–	–	–
Ethanoate+ Na^+	$-13,9 \pm 0,8$	$149,4 \pm 0,5$	–	–	–	–	–	–
Propanoate	$-17,0 \pm 0,6$	$137,8 \pm 0,3$	$162,7 \pm 0,5$	–	–	–	–	–
Propanoate (gaz)	$1,6 \pm 0,4$	$140,5 \pm 0,4$	$162,9 \pm 0,2$	–	–	–	–	–
Propanoate+ Na^+	$-18,2 \pm 0,8$	$138,5 \pm 0,3$	$162,6 \pm 0,5$	–	–	–	–	–
Butanoate	$-16,5 \pm 0,7$	$128,1 \pm 0,4$	$149,3 \pm 0,5$	$159,9 \pm 0,3$	–	–	–	–
Butanoate+ Na^+	$-17,5 \pm 0,5$	$129,8 \pm 0,5$	$149,6 \pm 0,4$	$160,2 \pm 0,5$	–	–	–	–
Octanoate	$-16,7 \pm 0,6$	$130,0 \pm 0,4$	$141,8 \pm 0,7$	$138,9 \pm 0,7$	$138,5 \pm 0,4$	$136,7 \pm 0,4$	$147,4 \pm 0,9$	$159,9 \pm 0,6$
Octanoate- Na^+	$-18,5 \pm 0,4$	$132,3 \pm 0,5$	$143,2 \pm 0,7$	$140,4 \pm 0,5$	$141,0 \pm 0,6$	$137,0 \pm 0,4$	$146,3 \pm 0,6$	$160,0 \pm 0,8$
	C1	C2	C3	PIMD C4	C5	C6	C7	C8
Methanoate	$-2,7 \pm 0,6$	–	–	–	–	–	–	–
Methanoate+ Na^+ (GBSA)	–	–	–	–	–	–	–	–
Methanoate+ Na^+	$-4,4 \pm 0,6$	–	–	–	–	–	–	–
Ethanoate	$-13,3 \pm 0,7$	$145,9 \pm 0,6$	–	–	–	–	–	–
Ethanoate+ Na^+	$-14,9 \pm 0,7$	$146,2 \pm 0,5$	–	–	–	–	–	–
Propanoate	$-17,4 \pm 0,6$	$134,3 \pm 0,5$	$158,8 \pm 0,4$	–	–	–	–	–
Propanoate (gaz)	$0,1 \pm 0,6$	$136,5 \pm 0,6$	$157,6 \pm 1,3$	–	–	–	–	–
Propanoate+ Na^+	$-19,5 \pm 0,4$	$134,8 \pm 0,6$	$158,8 \pm 0,6$	–	–	–	–	–
Butanoate	$-17,6 \pm 0,2$	$123,8 \pm 0,4$	$144,0 \pm 0,7$	$155,0 \pm 0,5$	–	–	–	–
Butanoate+ Na^+	$-18,7 \pm 0,6$	$124,8 \pm 0,6$	$144,3 \pm 0,7$	$155,4 \pm 0,3$	–	–	–	–
Octanoate	$-17,8 \pm 0,7$	$125,1 \pm 0,7$	$135,4 \pm 0,3$	$132,1 \pm 0,5$	$131,9 \pm 1,1$	$130,7 \pm 1,1$	$141,1 \pm 0,6$	$155,2 \pm 0,4$
Octanoate- Na^+	$-18,9 \pm 0,5$	$127,0 \pm 0,5$	$136,7 \pm 0,3$	$133,5 \pm 0,4$	$133,8 \pm 0,5$	$131,2 \pm 0,7$	$141,6 \pm 0,6$	$155,3 \pm 0,4$

obtenues par Tzonka Mineva *et al.* [336] sur une structure unique d’un octanoate en conformation *trans* optimisé dans un minimum local. En effet, dans cet article, les auteurs observent un décalage vers le bas de 5,8 ppm entre une structure où le Na^+ est éloigné du carbone C1 ($d(\text{Na}^+-\text{C1}) = 13,54 \text{ \AA}$) et où le groupement $-\text{COO}^-$ est entièrement solvaté par de l’eau et une structure où le Na^+ est proche du carbone C1 ($d(\text{Na}^+-\text{C1}) = 2,77 \text{ \AA}$) et où le Na^+ est lié aux oxygènes du groupement $-\text{COO}^-$. Dans ce même article, une étude sur l’effet de la base montre également que l’impact de la présence du Na^+ sur les valeurs des ^{13}C σ_{iso} est diminué lorsque la taille de la base augmente. Cependant, les auteurs de cet article mettent en avant la nécessité de considérer plus de conformations que seulement une structure *trans* optimisée. La figure 3.8 montre que dans une simulation MD comme dans une simulation PIMD, le Na^+ est fortement localisé entre les deux oxygènes du groupement $-\text{COO}^-$ ce qui

3. IMPACT DES EFFETS QUANTIQUES NUCLÉAIRES SUR L'ÉCRANTAGE ISOTROPE ^{13}C DE CARBOXYLATES LINÉAIRES

engendre des distributions de distance Na–O1 et Na–O2 très similaires. La prise en compte des NQEs ne modifie pratiquement pas cette distribution bien qu'elle engendre une légère rotation du Na^+ autour du O2. Soulignons enfin que l'effet de la solvation explicite par l'agrégat d'eau est nettement plus important que la présence du Na^+ . Si l'on regarde dans le tableau 3.3 les valeurs de ^{13}C σ_{iso} du propanoate en phase gaz, on s'aperçoit que pour les carbones C2 et C3, elles ne sont pratiquement pas modifiées par la présence de l'agrégat ($\leq 2,2$ ppm) que ce soit en MD ou en PIMD. En revanche, dans le cas du carbone C1, des différences de 18,6 et 17,5 ppm sont observées en MD et PIMD, respectivement. Cette observation montre qu'une **description correcte des paramètres RMN du carbone C1 nécessitent une solvation explicite du groupement $-\text{COO}^-$** du carboxylate.

3.5.2 Discussion sur les valeurs relatives de ^{13}C σ_{iso}

En plus de la discussion sur les valeurs absolues de ^{13}C σ_{iso} , il est important de regarder si la prise en compte des NQEs améliore la comparaison des valeurs calculées de ^{13}C σ_{iso} avec les valeurs expérimentales. Pour cela, **deux approches peuvent être considérées** : (i) comparer les valeurs des déplacements chimiques ^{13}C δ_{iso} aux valeurs expérimentales ou (ii) comparer les écarts relatifs des ^{13}C σ_{iso} entre différents atomes d'une même molécule. Les valeurs correspondant à la première approche sont reportées dans le tableau 3.4. Comme on peut le voir, les valeurs issues des simulations PIMD sont parfois plus proches des valeurs expérimentales que celles issues d'une simulation de MD classique, mais ce n'est pas toujours le cas. Par exemple, pour le système octanoate- Na^+ , les valeurs de ^{13}C δ_{iso} issues de la simulation MD sont plus proches des valeurs expérimentales que celles issues de la simulation PIMD pour tous les carbones à l'exception du carbone C1. Dans leur article, Tzonka Mineva *et al.* montrent que pour le système octanoate- Na^+ dans une boîte composée de 226 molécules d'eau, les valeurs moyennes de ^{13}C δ_{iso} sont systématiquement surestimées par rapport aux valeurs expérimentales, excepté pour les carbones C1 et C8 [336]. En réalité, cette approche pour comparer les valeurs expérimentales et théoriques dépend fortement de la précision du ^{13}C σ_{iso} calculé pour le TMS. En effet, la moindre erreur sur cette valeur induit un décalage artificiel de toutes les valeurs de δ_{iso} et biaise la comparaison avec l'expérience. C'est pourquoi il convient d'utiliser une valeur de référence interne au système étudié et de comparer

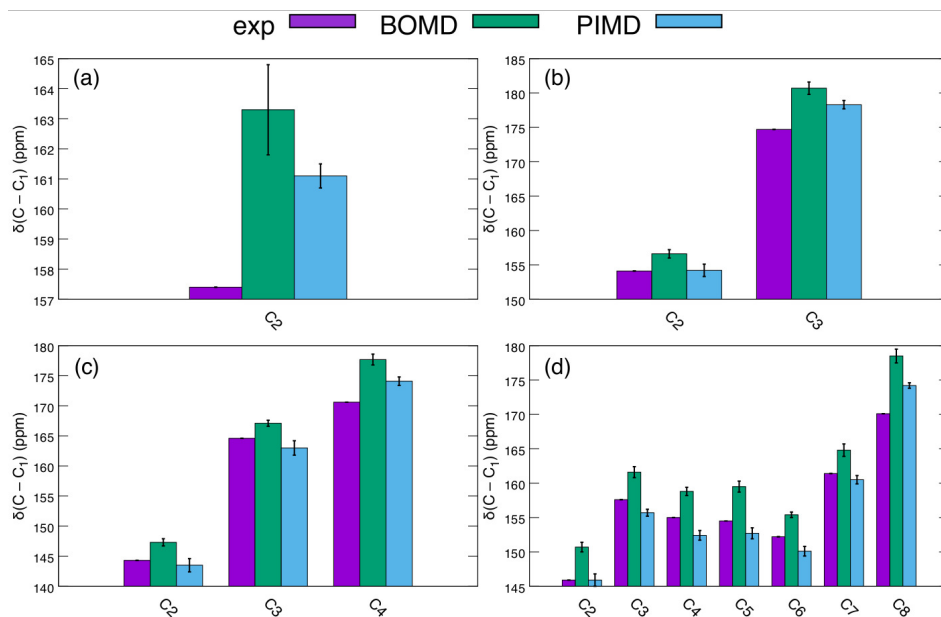


FIGURE 3.9 – Valeurs moyennes de ^{13}C $\delta_{iso,i}$ (en ppm) relatives au ^{13}C σ_{iso} du C1 pour les composés éthanoate (a), propanoate (b), butanoate (c) et octanoate (d). Les valeurs expérimentales sont en violet, les valeurs issues des simulations MD en vert et celles issues des simulations PIMD en bleu.

non pas les ^{13}C δ_{iso} mais les écarts de ^{13}C δ_{iso} (ou de façon équivalente les écarts de ^{13}C σ_{iso}) au sein d'une même molécule. Pour tous les composés étudiés, nous avons donc choisi d'utiliser la valeur de ^{13}C σ_{iso} du C1 comme valeur de référence tel que cela a été fait par Tzonka Mineva *et al.* [336]. Les déplacements chimiques relatifs $\delta_{iso,i}$ sont donc calculés selon l'expression $\delta_{iso,i} = \sigma_{iso,i} - \sigma_{iso,C1}$. Les valeurs ainsi calculées sont reportées dans le tableau 3.5. On notera tout d'abord que les valeurs issues de simulations de MD classique sont légèrement plus proches de quelques ppm des valeurs expérimentales que celles de la référence [336]. De plus, concernant l'inclusion des NQEs, on voit sur la figure 3.9 qu'elle améliore systématiquement la comparaison avec les résultats expérimentaux. Ainsi, pour les carboxylates étudiés ici, **la prise en compte des NQEs permet d'avoir des résultats en très bon accord avec les résultats expérimentaux** si l'on considère les écarts relatifs $\delta_{iso,i}$ au sein d'un même composé. De plus, l'inclusion des NQEs provoque un décalage vers le bas (déblindage) des $\delta_{iso,i}$. Cette tendance a déjà été observée dans des cristaux d'acides aminés [330].

3. IMPACT DES EFFETS QUANTIQUES NUCLÉAIRES SUR L'ÉCRANTAGE ISOTROPE ^{13}C DE CARBOXYLATES LINÉAIRES

TABLE 3.4 – Valeurs moyennes de ^{13}C δ_{iso} (en ppm) relatives au TMS de tous les carbones des cinq carboxylates issues de simulations de MD classique et PIMD. Pour chaque carboxylate, les valeurs ont été calculées pour le système chargé négativement en interaction avec l'agrégat de 34 molécules d'eau et pour le système en présence du contre-ion Na^+ et de l'agrégat de 34 molécules d'eau. Dans le deuxième cas nous avons systématiquement utilisé les paramètres OPLS-AA pour la description du Na^+ excepté pour le méthanoate où les paramètres GBSA ont aussi été testés. Pour le propanoate nous avons aussi réalisé une simulation en phase gaz sans le Na^+ . Les valeurs expérimentales ont été prises sur le site web [357]. L'incertitude statistique est calculée selon la procédure décrite dans la section 3.3.

	C1	C2	C3	MD C4	C5	C6	C7	C8
Methanoate	177,0 \pm 0,5	–	–	–	–	–	–	–
Methanoate+ Na^+ (GBSA)	178,8 \pm 0,6	–	–	–	–	–	–	–
Methanoate+ Na^+	179,4 \pm 0,9	–	–	–	–	–	–	–
Ethanoate	188,3 \pm 0,4	26,8 \pm 0,1	–	–	–	–	–	–
Ethanoate+ Na^+	190,4 \pm 1,2	27,1 \pm 0,1	–	–	–	–	–	–
Propanoate	193,5 \pm 0,4	38,7 \pm 0,4	13,8 \pm 0,5	–	–	–	–	–
Propanoate(gaz)	174,8 \pm 0,7	35,9 \pm 0,1	13,6 \pm 0,4	–	–	–	–	–
Propanoate+ Na^+	194,7 \pm 1,0	38,0 \pm 0,7	13,9 \pm 0,9	–	–	–	–	–
Butanoate	192,9 \pm 1,1	48,4 \pm 0,4	27,2 \pm 0,4	16,5 \pm 0,4	–	–	–	–
Butanoate+ Na^+	194,0 \pm 0,6	46,7 \pm 0,4	26,9 \pm 0,3	16,3 \pm 0,5	–	–	–	–
Octanoate	193,1 \pm 0,6	46,5 \pm 0,4	34,7 \pm 1,1	37,6 \pm 0,8	38,0 \pm 0,2	39,8 \pm 0,6	30,1 \pm 0,9	16,6 \pm 0,6
Octanoate- Na^+	195,0 \pm 0,5	44,1 \pm 0,2	33,3 \pm 0,3	36,2 \pm 0,6	35,5 \pm 0,2	39,5 \pm 0,4	30,2 \pm 0,8	16,5 \pm 1,1
Octanoate- Na^+ (ref [336])	178,1	41,2	30,2	33,6	34,6	35,8	26,5	14,6
	C1	C2	C3	PIMD C4	C5	C6	C7	C8
Methanoate	175,4 \pm 0,3	–	–	–	–	–	–	–
Methanoate+ Na^+ (GBSA)	–	–	–	–	–	–	–	–
Methanoate+ Na^+	177,2 \pm 0,8	–	–	–	–	–	–	–
Ethanoate	186,1 \pm 0,2	26,8 \pm 0,3	–	–	–	–	–	–
Ethanoate+ Na^+	187,7 \pm 0,8	26,6 \pm 0,5	–	–	–	–	–	–
Propanoate	190,2 \pm 0,9	38,5 \pm 0,6	14,9 \pm 0,2	–	–	–	–	–
Propanoate(gaz)	172,7 \pm 0,5	36,3 \pm 0,5	15,1 \pm 1,5	–	–	–	–	–
Propanoate+ Na^+	192,3 \pm 0,2	38,0 \pm 0,9	14,0 \pm 0,5	–	–	–	–	–
Butanoate	190,4 \pm 0,4	49,0 \pm 0,3	28,7 \pm 0,5	17,8 \pm 0,4	–	–	–	–
Butanoate+ Na^+	191,5 \pm 0,6	48,0 \pm 0,4	28,5 \pm 0,9	17,4 \pm 0,3	–	–	–	–
Octanoate	190,6 \pm 0,8	47,7 \pm 0,8	37,4 \pm 0,4	40,7 \pm 0,7	40,9 \pm 1,3	42,1 \pm 1,2	31,7 \pm 0,5	17,6 \pm 0,4
Octanoate- Na^+	191,7 \pm 0,4	45,8 \pm 0,7	36,1 \pm 0,1	39,3 \pm 0,3	39,0 \pm 1,3	41,6 \pm 0,4	31,2 \pm 0,5	17,5 \pm 0,2
	C1	C2	C3	Exp. (ref [357]) C4	C5	C6	C7	C8
Methanoate+ Na^+	173,8	–	–	–	–	–	–	–
Ethanoate+ Na^+	181,9	24,5	–	–	–	–	–	–
Propanoate+ Na^+	185,7	31,6	11,0	–	–	–	–	–
Butanoate+ Na^+	184,8	40,5	20,2	14,2	–	–	–	–
Octanoate- Na^+	184,5	38,6	26,9	29,5	30,0	32,3	23,1	14,4

3.6 Conclusions et perspectives

Nous avons présenté dans ce chapitre la récente implémentation de l'interface entre les codes i-PI et deMon2k. Pour une première application, nous avons calculé les ^{13}C σ_{iso} et étudié les déplacements chimiques ^{13}C δ_{iso} des atomes de carbone d'une série de carboxylates linéaires. Le **caractère inédit de cette étude** (molé-

3.6 Conclusions et perspectives

TABLE 3.5 – Valeurs moyennes de ^{13}C $\delta_{iso,i}$ (en ppm) relatives au ^{13}C σ_{iso} du C1 de tous les carbones des carboxylates (hormis le méthanoate) issues de simulation de MD classique et PIMD. Pour chaque carboxylate, les valeurs ont été calculées pour le système chargé négativement en interaction avec l’agrégat de 34 molécules d’eau et pour le système en présence du contre-ion Na^+ et de l’agrégat de 34 molécules d’eau. Dans le deuxième cas, nous avons systématiquement utilisé les paramètres OPLS-AA pour la description du Na^+ . Pour le propanoate, nous avons aussi réalisé une simulation en phase gaz sans le Na^+ . Les valeurs expérimentales ont été prises sur le site web [357]. L’incertitude statistique est calculée selon la procédure décrite dans la section 3.3.

	C2	C3	C4	MD C5	C6	C7	C8
Ethanoate	161,6 ± 0,3	–	–	–	–	–	–
Ethanoate+ Na^+	163,3 ± 1,5	–	–	–	–	–	–
Propanoate	154,8 ± 0,4	179,7 ± 0,6	–	–	–	–	–
Propanoate(gaz)	138,9 ± 0,7	161,3 ± 0,4	–	–	–	–	–
Propanoate+ Na^+	156,7 ± 0,5	180,8 ± 0,8	–	–	–	–	–
Butanoate	144,6 ± 1,4	165,8 ± 0,9	176,4 ± 0,8	–	–	–	–
Butanoate+ Na^+	147,3 ± 0,3	167,1 ± 0,3	177,7 ± 1,0	–	–	–	–
Octanoate	146,7 ± 0,4	158,5 ± 1,5	155,6 ± 0,6	155,2 ± 0,4	153,4 ± 0,9	163,1 ± 0,4	176,6 ± 0,3
Octanoate+ Na^+	150,8 ± 0,5	161,7 ± 0,5	158,8 ± 0,4	159,5 ± 0,7	155,5 ± 0,2	164,8 ± 1,0	178,5 ± 1,3
Octanoate- Na^+ (ref [336])	136,9	147,9	144,5	143,5	142,3	151,6	163,5
	C2	C3	C4	PIMD C5	C6	C7	C8
Ethanoate	159,1 ± 0,3(–2,5)	–	–	–	–	–	–
Ethanoate+ Na^+	161,1 ± 0,4(–2,2)	–	–	–	–	–	–
Propanoate	151,7 ± 0,4(–3,1)	176,2 ± 0,8(–3,5)	–	–	–	–	–
Propanoate(gaz)	136,4 ± 0,8(–2,5)	157,6 ± 1,8(–3,7)	–	–	–	–	–
Propanoate+ Na^+	154,3 ± 1,1(–2,4)	178,3 ± 0,4(–2,5)	–	–	–	–	–
Butanoate	141,4 ± 0,6(–3,2)	161,6 ± 0,8(–4,2)	172,6 ± 0,4(–3,8)	–	–	–	–
Octanoate	142,9 ± 0,3(–3,8)	153,2 ± 0,8(–5,3)	149,9 ± 0,9(–5,7)	149,7 ± 0,9(–5,5)	148,5 ± 0,6(–4,9)	158,9 ± 0,7(–4,2)	173,0 ± 0,6(–3,6)
Octanoate+ Na^+	145,9 ± 1,0(–4,9)	155,6 ± 0,5(–6,1)	152,4 ± 0,6(–6,4)	152,7 ± 0,9(–6,8)	150,1 ± 0,4(–5,4)	160,5 ± 0,5(–4,3)	174,2 ± 0,3(–4,3)
	C2	C3	C4	Exp. (ref [357]) C5	C6	C7	C8
Ethanoate+ Na^+ (exp)	157,4	–	–	–	–	–	–
Propanoate+ Na^+ (exp)	154,1	174,7	–	–	–	–	–
Butanoate+ Na^+ (exp)	144,3	164,6	170,6	–	–	–	–
Octanoate+ Na^+ (exp)	145,9	157,6	155,0	154,5	152,2	161,4	170,1

cule organique explicitement solvaté) a nécessité de nombreuses étapes de validation de la méthodologie. Nous avons validé dans un premier temps l’**utilisation d’un potentiel QM/MM** pour pouvoir réduire drastiquement le coût de calcul des simulations sans pour autant impacter la précision de ces calculs. Nous avons ainsi pu traiter les solutés (carboxylates) en DFT et le solvant (molécules d’eau) en MM. Dans un deuxième temps, nous avons testé **plusieurs combinaisons de thermostat et de nombre de répliques** dans le polymère cyclique afin de déterminer quelle stratégie nous permettait de converger les effets quantiques nucléaires avec un coût calculatoire le plus faible possible. Finalement, nous avons choisi d’utiliser le thermostat PIGLET basé sur l’équation de Langevin généralisée et un polymère cyclique composé de 12 répliques. Une fois la méthodologie validée, nous avons étudié l’impact de la prise en compte des NQEs sur les valeurs de ^{13}C σ_{iso} en comparant les résultats issus de simulations de MD classique et PIMD aux valeurs expérimentales. Nous avons montré que l’impact des NQEs sur les valeurs de ^{13}C σ_{iso} résulte princi-

3. IMPACT DES EFFETS QUANTIQUES NUCLÉAIRES SUR L'ÉCRANTAGE ISOTROPE ^{13}C DE CARBOXYLATES LINÉAIRES

pablement des fluctuations quantiques des liaisons C-H qui ont pour effet d'augmenter la distance moyenne de ces liaisons. Ainsi, **les NQEs déblindent systématiquement les valeurs théoriques**. La comparaison des valeurs de ^{13}C δ_{iso} théoriques aux valeurs expérimentales est délicate si l'on calcule les valeurs de ^{13}C δ_{iso} relativement à la valeur de ^{13}C σ_{iso} du TMS car la qualité de cette dernière introduit un décalage artificiel de tous les signaux dans les composés étudiés. En revanche, si l'on s'intéresse aux écart relatifs au sein d'un même composé on évite d'introduire un quelconque biais et l'on obtient un accord qualitatif avec les valeurs expérimentales ainsi qu'**une amélioration systématique des valeurs théoriques grâce à la prise en compte des NQEs**. Ce travail ouvre la voie à de futures études de composés organiques et inorganiques pour lesquelles les NQEs devraient jouer un rôle important dans la description de certaines propriétés. Par ailleurs, **cette étude valide l'implémentation de l'interface entre les codes i-PI et deMon2k**. Il est donc désormais possible de réaliser des simulations PIMD avec un potentiel ADFT et nous sommes convaincus que **cette implémentation sera utile pour de nombreux travaux dans le futur**.

Impact des effets quantiques
nucléaires sur la distribution
d'isomères dans les agrégats d'eau
 $(\text{H}_2\text{O})_6$ et $(\text{H}_2\text{O})_8$

Chapitre 4

Impact des effets quantiques nucléaires sur la distribution d'isomères dans les agrégats d'eau $(\text{H}_2\text{O})_6$ et $(\text{H}_2\text{O})_8$

4.1 Introduction

Comme nous l'avons mentionné dans la section 2.1, l'hexamère d'eau $(\text{H}_2\text{O})_6$ **occupe une place particulière dans les études sur les agrégats d'eau** tant au niveau expérimentale que théorique. $(\text{H}_2\text{O})_6$ présente la particularité d'être l'agrégat d'eau le plus petit qui présente différents isomères de basse énergie à 3 dimensions. Expérimentalement, de nombreuses techniques spectroscopiques ont été utilisées afin de caractériser sous quelle forme se présente cet agrégat [22, 33, 36, 37, 38, 39, 286, 358, 359]. Ces travaux ont montré que l'hexamère d'eau pouvait se présenter sous la forme d'un mélange de différents isomères et cela même à très basse température. Certaines de ces études ont en effet montré que les isomères *cage*, *prism* et *book* coexistaient même à très basse température, 6 K. De nombreuses études théoriques basées sur le calcul de la structure électronique ont également été réalisées [32, 279, 280, 281, 282, 283, 284, 358] parfois en collaboration directe avec les expérimentateurs. Ces études utilisant des méthodes et des potentiels très précis ont permis d'affiner la description que l'on avait de cet agrégat. D'après ces résultats expérimentaux et théoriques, on sait maintenant

4. IMPACT DES EFFETS QUANTIQUES NUCLÉAIRES SUR LA DISTRIBUTION D'ISOMÈRES DANS $(\text{H}_2\text{O})_6$ et $(\text{H}_2\text{O})_8$

que les isomères de plus basse énergie de $(\text{H}_2\text{O})_6$, notamment les isomères *cage*, *prism* et *book* sont très proches en énergie. Les calculs les plus précis donnent l'isomère *prism* comme étant le plus stable. Cependant, ces études réalisées sur des structures fixes ne tiennent généralement pas compte des effets thermiques. Or la différence d'énergie entre ces isomères est plus faible que l'énergie résultant de l'agitation thermique même à basse température. Il est donc important de prendre en compte ces effets thermiques au travers de simulations de dynamique moléculaire (MD) ou Monte-Carlo (MC) comme cela a été fait dans un certain nombre d'études [287, 191, 359, 360, 361, 362, 363, 364, 365]. Bien entendu, comme nous l'avons déjà présenté dans le chapitre 2, de telles simulations doivent être ergodiques et donc pour cela être couplées à des outils d'échantillonnage accéléré tel que la méthode *Parallel-Tempering Molecular Dynamics* (PTMD).

Toutefois, **si les effets thermiques sont pris en compte dans certaines de ces études, les noyaux restent décrits comme des particules classiques.** Comme nous l'avons discuté au chapitre 3, **les effets quantiques nucléaires à basse température dans les agrégats d'eau ne sont pas négligeables** et on peut s'attendre à ce qu'ils jouent un rôle clef dans la description des distributions d'isomères de $(\text{H}_2\text{O})_6$ à basse température. Ceci est d'autant plus attendu qu'il a été montré que les NQEs jouent un rôle important dans la structuration de l'eau liquide même à température ambiante [85, 86, 87, 88, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217, 218]. Ainsi, pour avoir une description plus complète des distributions d'isomères de $(\text{H}_2\text{O})_6$ à basse température, il faut utiliser une méthode qui permet de prendre en compte les NQEs. A notre connaissance, **la seule simulation de ce type a été réalisée par Yimin Wang *et al.* [361].** Ces derniers ont réalisé une simulation PTPIMD (pour *Parallel-Tempering Path-Integral Molecular Dynamics*) avec le champ de force WHBB [366], qui leur a permis de retrouver le résultat expérimental de la coexistence de différents isomères de $(\text{H}_2\text{O})_6$ à basse température. Dans cet article, les auteurs montrent que les NQEs induisent une modification remarquable des populations relatives d'isomères en dessous de 100 K comparative-ment à une simulation classique. A la température la plus basse de la simulation, c'est à dire 30 K, la simulation PTMD conduit à une population relative de 100% d'isomère *prism* tandis qu'une simulation PTPIMD conduit à la coexistence des isomères *prism* et *cage* dans un ratio $\sim 60\%-40\%$. Dans un autre article [362], Volodymyr Babin et Francesco Paesani montrent que ce résultat dépend grandement

du potentiel employé en comparant quatre champs de force : q-TIP4P/F [213], TTM3-F [367], WHBB [366] et HBB2-pol [368, 369]. Le premier est un potentiel empirique développé pour réaliser des simulations PIMD. Le deuxième est un potentiel polarisable paramétré pour reproduire à la fois le spectre vibrationnel de l'eau liquide et d'agrégats d'eau. Enfin, les champs de force WHBB et HBB2-pol sont des potentiels paramétrés sur des données *ab-initio*. Le potentiel HBB2-pol a été développé dans le but d'être combiné à des méthodes incluant les NQEs telles que les simulations PIMD. Dans cet article, Volodymyr Babin et Francesco Paesani comparent les distributions relatives d'isomères en fonction de la température obtenues à partir de simulations PTMD et PTPIMD avec ces quatre potentiels. Avec le potentiel q-TIP4P/F, l'isomère *cage* est le seul présent à 30 K, que les NQEs soient incluent ou non. Toutefois, dans la simulation PTPIMD, on voit que l'isomérisation de l'isomère *cage* en *prism*, *book* et *prism-book* (*inter* dans notre nomenclature) a lieu à plus basse température que dans le cas classique. L'isomérisation commence quelques kelvin au dessus de 30 K tandis qu'elle débute à ~ 60 K en PTMD. La part de l'isomère *prism* ne dépasse en revanche jamais plus que quelques %, que les NQEs soient pris en compte ou pas. Avec le potentiel TTM3-F, le comportement est totalement différent. En PTMD, à 30 K les isomères *prism* et *cage* coexistent dans un ratio 90%-10% tandis qu'en PTPIMD ce sont les isomères *cage* et *book* qui sont présents dans un ratio de 70%-20% avec une faible population des isomères *prism* et *prism-book* (*i.e.* *inter*). Enfin, avec les champs de force WHBB et HBB2-pol, les simulations PTMD prédisent 100% d'isomère *prism* à 30 K tandis que les simulations PTPIMD prédisent la coexistence des isomères *prism* et *cage* dans des proportions de 60%-40% en WHBB et 30%-65% et quelques % de *book* et *prism-book* (*i.e.* *inter*) en HBB2-pol. Ainsi, dans ces travaux [361, 362], **les auteurs ont reproduit l'observation expérimentale de la coexistence des isomères *prism* et *cage* à basse température.** Ce résultat a pu être obtenu grâce à la combinaison de champs de force adroitement paramétrés pour décrire les interactions moléculaires dans des agrégats d'eau et de la prise en compte des NQEs par l'approche PTPIMD.

Cependant, comme nous l'avons mentionné plus haut, ce type d'étude n'a porté que sur l'agrégat $(\text{H}_2\text{O})_6$ et rien n'a été fait sur des systèmes aqueux contenant des impuretés tels que des agrégats protonés ou des agrégats contenant d'autres molécules (NH_3 , NH_4^+ , CH_3OH , SO_4^{2-} , ...). Les agrégats contenant ces molécules sont cependant des espèces très étudiées en raison de leur rôle en chimie atmosphé-

4. IMPACT DES EFFETS QUANTIQUES NUCLÉAIRES SUR LA DISTRIBUTION D'ISOMÈRES DANS (H₂O)₆ et (H₂O)₈

rique [41, 44, 370, 371, 372, 373]. En effet, ces impuretés présentes dans l'atmosphère jouent un rôle important dans de nombreux phénomènes météorologiques comme les pluies acides ou la nucléation de micro-gouttelettes. Des études théoriques concernant les propriétés structurales et thermodynamiques d'agrégats d'eau contenant de telles molécules ont déjà fait l'objet de plusieurs études en MD [374, 375, 376], en PTMD [59, 377, 378], en PTMC [379], et en PIMD [380, 381]. En revanche, à notre connaissance, **aucune étude sur ces systèmes n'a été réalisée en utilisant une combinaison des méthodes PTMD et PIMD** comme cela a été fait pour l'hexamère d'eau. Cette absence d'étude PTPIMD sur de tels systèmes peut provenir de plusieurs raisons. La première et la plus évidente est le coût de calcul très élevé d'une simulation PTPIMD. La seconde est l'absence de programmes de dynamique moléculaire permettant de combiner des simulations PTMD et PIMD. La troisième raison expliquant le peu d'études PTPIMD sur les agrégats d'eau chargés ou contenant des impuretés est le manque de potentiel précis permettant leur étude. L'utilisation d'un potentiel *ab-initio* pourrait permettre de résoudre ce problème mais le coût de calcul associé rend impossible des simulations PTPIMD. Rappelons ici que le nombre de calculs d'énergie et de forces à réaliser en PTPIMD pour la réalisation d'un pas de temps est égal au produit du nombre de répliques dans le polymère cyclique et du nombre de répliques dans l'algorithme PTMD. Dans les travaux de Volodymyr Babin et Yimin Wang, le nombre de répliques dans le polymère cyclique était de 32 ou 64 et le nombre de répliques dans l'approche PTMD de 32 ce qui implique **1024 ou 2048 calculs d'énergie et de forces pour réaliser un pas de temps** ce qui rend une approche de calcul telle que la DFT inutilisable. L'utilisation d'un potentiel semi-empirique comme la **SCC-DFTB apparaît alors comme une solution de choix pour réaliser ce type de simulations**. En raison des approximations sous-jacentes à cette méthode, qui ont été présentées dans la section 1.2.3, le coût calculatoire de cette méthode est beaucoup plus faible que celui de la DFT. De plus, les jeux de paramètres déjà existants pour un grand nombre d'atomes, ajouté au traitement explicite de la structure électronique offre une meilleure transférabilité que les potentiels de type champ de force.

Notre **objectif dans ce chapitre a donc été d'utiliser l'approche PTPIMD combinée à un potentiel SCC-DFTB** pour étudier des agrégats d'eau protonés ou contenant des impuretés chimiques. Pour cela, un des objectifs était l'implémentation de la méthode PIMD dans le code deMonNano afin de pou-

4.2 Méthodes employées et paramètres de calcul

voir la combiner à l’approche PTMD déjà implémentée dans ce code et ainsi pouvoir réaliser de telles simulations. Cependant, au moment où a commencé cette thèse, la parution de la version 2.0 du code i-PI proposait cette nouvelle fonctionnalité. Comme nous l’avons vu dans le chapitre 3, le programme i-PI peut être facilement interfacé avec n’importe quel programme dédié au calcul de l’énergie et des forces. Il nous est donc apparu plus judicieux d’**implémenter l’interface avec i-PI dans deMonNano** plutôt que de re-implémenter toute l’approche PIMD à partir de zéro. **Pour valider cette nouvelle méthode, nous avons dans un premier temps choisi de l’appliquer à l’étude des agrégats $(\text{H}_2\text{O})_6$ et $(\text{H}_2\text{O})_8$.** $(\text{H}_2\text{O})_6$ a été choisi dans le but de comparer le potentiel SCC-DFTB aux potentiels WHBB et HBB2-pol des articles de Volodymyr Babin et Yimin Wang $(\text{H}_2\text{O})_8$ a quant à lui été choisi car, contrairement à $(\text{H}_2\text{O})_6$, cet agrégat ne présente pratiquement qu’un seul isomère à basse température et jusqu’à la transition de phase, voir section 2.5. Il apparaissait donc intéressant de voir comment la prise en compte des NQEs modifierait la distribution d’isomères et la transition de phase pour cet agrégat. Finalement nous n’avons étudié que ces deux agrégats et nous n’avons pas pu étendre notre étude aux systèmes protonés ou contenant des impuretés. Ce travail pourra être mené dans les mois prochains.

Dans une première partie, nous présenterons les outils méthodologiques employés dans cette étude. Dans une deuxième partie, nous présenterons les résultats obtenus et discuterons de l’impact des NQEs sur la distribution d’isomères en fonction de la température et sur la transition de phase dans les agrégats $(\text{H}_2\text{O})_6$ et $(\text{H}_2\text{O})_8$. Enfin dans une dernière partie, nous présenterons les conclusions de ce travail et les perspectives concernant ce type d’étude pour le futur.

4.2 Méthodes employées et paramètres de calcul

Toutes les simulations ont été réalisées à l’aide de l’interface entre deMonNano et i-PI v 2.0. Le calcul des énergies et des forces est délégué à deMonNano tandis que i-PI est en charge de la propagation des équations du mouvement et de l’échange des répliques dans l’approche PTMD.

Calcul de la structure électronique en SCC-DFTB. Le calcul de l’éner-

4. IMPACT DES EFFETS QUANTIQUES NUCLÉAIRES SUR LA DISTRIBUTION D'ISOMÈRES DANS $(\text{H}_2\text{O})_6$ et $(\text{H}_2\text{O})_8$

gie et des forces a été réalisé avec un potentiel SCC-DFTB tel que décrit dans la section 1.2.3. Nous avons utilisé l'ensemble *mio* des tables d'intégrales de Slater-Koster [72]. Nous avons utilisé un terme de dispersion empirique afin d'améliorer la description des interactions intermoléculaires et nous avons utilisé des charges *Class IV-Charge Model 3* [142] pour le traitement des interactions de Coulomb. La procédure auto-cohérente est considérée comme convergée et stoppée quand les fluctuations de charges atomiques sont inférieures à 10^{-9} u.a. Pour éviter que les agrégats ne s'évaporent, nous avons confiné le système dans un potentiel sphérique de forme $V(r) = 0,008 \times (r - r_0)^4$ (en u.a.), où r_0 est le rayon de la sphère. Nous avons utilisé un rayon r_0 de 5,0 Å pour $(\text{H}_2\text{O})_6$ et 5,5 Å pour $(\text{H}_2\text{O})_8$ comme dans la référence [191].

Parallel-Tempering Path-Integral Molecular Dynamics. Pour $(\text{H}_2\text{O})_6$, nous avons réalisé une simulation PTMD et deux simulations PTPIMD avec 16 et 32 répliques dans le polymère cyclique. Pour les trois simulations, nous avons utilisé 32 températures dans l'algorithme PTMD réparties entre 30 et 150 K selon une distribution exponentielle comme dans les références [361] et [362]. Il faut savoir que l'augmentation du nombre de répliques dans le polymère cyclique rend ce dernier de plus en plus rigide et donc l'exploration de l'espace des phases beaucoup moins efficace. Cela a pour conséquence de réduire fortement la largeur des densités d'état et donc le taux d'acceptation des échanges dans l'algorithme PTMD. Dans la référence [362], les auteurs mentionnent avoir utilisé un polymère cyclique composé de 64 répliques pour tester la convergence des NQEs mais ils ont alors dû utiliser une distribution des températures comprise entre 50 et 150 K à cause d'un mauvais taux d'acceptation des échanges à basse température. Pour $(\text{H}_2\text{O})_8$, nous avons réalisé une simulation PTMD et une seule simulation PTPIMD avec 32 répliques dans le polymère cyclique. Dans les deux cas, 32 températures ont été distribuées exponentiellement entre 100 et 300 K. Les simulations PTMD et PTPIMD ont respectivement été thermostatées avec les thermostats *stochastic velocity rescale* et PILE-L avec une constante de temps de 0,1 ps dans les deux cas. Ces deux thermostats ont été décrits dans les sections 1.3.2 et 1.3.3. Dans toutes les simulations, le centre de masse de l'agrégat a été fixé et les équations du mouvement ont été intégrées avec un pas de temps de 0,2 fs. La procédure pour l'échange des répliques utilise toujours le critère de Metropolis-Hasting mais l'algorithme est légèrement différent de celui qui est présenté dans les sections 1.3.4 et 2.3.2. En ef-

fet, dans les simulations réalisées dans ce chapitre, nous avons choisi d'échanger les répliques tous les 500 pas de temps. L'échange n'aura pas lieu entre répliques adjacentes tous les 500 pas de temps, comme cela serait le cas avec l'algorithme présenté plus tôt, mais à chaque pas de temps, tous les échanges possibles sont tentés avec une probabilité de $1/500$. En moyenne, chaque réplique sera donc échangée tous les 500 pas de temps. Pour les deux agrégats, $(\text{H}_2\text{O})_6$ et $(\text{H}_2\text{O})_8$, nous avons réalisé une simulation de 1,6 ns. Dans les deux cas, la première nanoseconde a été considérée comme temps d'équilibration et n'a pas été prise en compte dans les statistiques. Cette nanoseconde correspond au temps nécessaire pour que chaque réplique ait pu réaliser plusieurs aller-retour entre la température la plus basse et la température la plus haute. Pour chaque simulation, nous avons utilisé l'isomère de plus basse énergie optimisé au niveau SCC-DFTB comme structure de départ pour éviter du temps d'équilibration supplémentaire.

Distribution d'isomères. Nous avons étudié la distribution relative d'isomères en fonction de la température. Pour pouvoir se faire, il faut pouvoir différencier les différents isomères de $(\text{H}_2\text{O})_6$ et $(\text{H}_2\text{O})_8$ par une approche automatisée. Ainsi, comme dans le chapitre 2, les différents isomères de $(\text{H}_2\text{O})_6$ et $(\text{H}_2\text{O})_8$ ont été discriminés en fonction du nombre de cycles oxygène-oxygène et de la taille de ces cycles. Par taille de cycle nous entendons ici le nombre de molécule d'eau constitutif d'un cycle. Nous avons appliqué pour cela l'algorithme des cycles primitifs développé par Xianglong Yuan et Alastair Cormack [250] en considérant que deux molécules d'eau sont liées si la distance entre leurs atomes d'oxygène est inférieure à 3,25 Å. La nomenclature des isomères est également identique à celle de décrite dans le chapitre 2.

4.3 Résultats et discussion

La figure 4.1 présente la distribution relative des différents isomères de $(\text{H}_2\text{O})_6$ en fonction de la température. Ces distributions ont été calculées à partir des simulations (a) PTMD, (b) PTPIMD avec 16 répliques dans le polymère cyclique et (c) PTPIMD avec 32 répliques dans le polymère cyclique. Plusieurs points peuvent être discutés en regardant des courbes. Tout d'abord, on voit qu'à 30 K, l'**isomère *prism*** est le seul isomère présent que les NQEs soient pris en compte

4. IMPACT DES EFFETS QUANTIQUES NUCLÉAIRES SUR LA DISTRIBUTION D'ISOMÈRES DANS $(\text{H}_2\text{O})_6$ et $(\text{H}_2\text{O})_8$

ou pas. Comme nous l'avons précisé en introduction, dans la référence [362], les auteurs mettent en évidence que la coexistence ou non de plusieurs isomères à 30 K dépend fortement du potentiel employé. Pour rappel, le tableau 4.2 regroupe les proportions d'isomères présents à 30 K obtenues à partir de nos simulations au niveau SCC-DFTB ainsi que celles obtenues au niveau champ de force dans la référence [362]. L'absence dans nos simulations de l'isomère *prism* à 30 K peut s'expliquer à partir des énergies relatives des isomères de basse énergie de $(\text{H}_2\text{O})_6$ qui sont rappelées dans le tableau 4.1. Dans ce tableau, l'isomère *cage* est $3,4 \text{ kJ.mol}^{-1}$ plus haut en énergie que l'isomère *prism* alors que les calculs de haute précision en CCSD(T) à la limite de la base complète prévoient une différence d'énergie de $1,1 \text{ kJ.mol}^{-1}$ seulement [284]. **Cette différence d'énergie au niveau SCC-DFTB est suffisamment grande pour que les effets quantiques des noyaux ne permettent pas de rendre compte de l'équilibre entre les isomères *prism* et *cage* à 30 K.** Il serait intéressant de connaître les énergies de liaison de ces différents isomères calculées à l'aide des potentiels WHBB et HBB2-pol utilisés par Volodymyr Babin *et al.* [362] pour confirmer qu'ils conduisent bien à une différence d'énergie entre les isomères *prism* et *cage* de l'ordre de $1,0 \text{ kJ.mol}^{-1}$. Bien que les simulations SCC-DFTB ne permettent pas de reproduire la coexistences des isomères *prism* et *cage* à basse température, l'impact des NQEs n'est pas négligeable pour autant. En effet, on peut voir sur la figure 4.1 que l'isomérisation de l'isomère *prism* en *book* se produit à une température bien plus basse dans les simulations PTPIMD que dans la simulation PTMD. En effet, elle commence dès 30 K que l'on utilise 16 ou 32 répliques dans le polymère cyclique alors qu'elle ne débute qu'à $\sim 50 \text{ K}$ en PTMD. Le point de croisement entre les distributions de ces isomères *prism* et *book* est situé à ~ 67 et $\sim 74 \text{ K}$ dans les simulations PTPIMD avec 16 et 32 répliques, respectivement alors qu'il est situé à 110 K en PTMD. Ainsi, tout le processus d'isomérisation intervient à des températures plus basses quand les NQEs sont pris en compte. De même, l'apparition des isomères *cycle* et *bicycle* intervient à beaucoup plus basse température quand les NQEs sont pris en compte. Ce comportement a été observé dans les études précédentes [361, 362] et démontre donc que **le potentiel SCC-DFTB utilisé permet de rendre compte d'une part importe des NQEs sur les distributions relatives d'isomères de $(\text{H}_2\text{O})_6$.** Il s'agit d'un résultat très positif pour de futures études. Notons enfin que les distributions obtenues en PTPIMD avec 16 et 32 répliques sont très proches mais pas strictement identiques. Bien qu'un léger défaut de convergence puisse être observé sur la simulations à 32 répliques, cette dif-

4.3 Résultats et discussion

férence démontre qu'un résultat quantitatif nécessite une simulation à 32 répliques. Ceci confirme la nécessité de disposer d'un potentiel précis mais présentant un coût de calcul réduit tel que la SCC-DFTB pour pouvoir mener à bien de telles études.

TABLE 4.1 – Énergies de liaison relative (en kJ.mol^{-1}) des huit isomères de plus basse énergie de $(\text{H}_2\text{O})_6$ obtenues en SCC-DFTB. Les valeurs obtenues au niveau CCSD(T)/CBS proviennent de la référence [284]. Les géométries des isomères sont représentées dans le chapitre 2.

Isomères	SCC-DFTB	CCSD(T)/CBS [284]
<i>prism</i> [2300]	0,0	0,0
<i>cage</i> [0500]	3,4	1,1
<i>inter</i> [1210]	3,5	-
<i>book</i> [0200]	5,5	3,0
<i>bag</i> [0120]	7,5	6,8
<i>cycle</i> [0001]	8,9	7,5
<i>bicycle</i> [1010]	9,7	-
[0010]	23,2	-

TABLE 4.2 – Proportion des isomères de $(\text{H}_2\text{O})_6$ présent à 30 K en fonction du potentiel employé. Les données obtenues à partir des potentiels q-TIP4P/F, TTM3-F, WHBB et HBB2-pol sont issus de la référence [362].

Potentiel	PTMD	PTPIMD (32 répliques)
SCC-DFTB	<i>prism</i> (100%)	<i>prism</i> (100%)
q-TIP4P/F	<i>cage</i> (100%)	<i>cage</i> (100%)
TTM3-F	<i>prism-cage</i> (92%-8%)	<i>cage-book-prism-inter</i> (78%-16%-4%-2%)
WHBB	<i>prism</i> (100%)	<i>prism-cage</i> (57%-43%)
HBB2-pol	<i>prism-cage</i> (67%-33%)	<i>cage-prism-book-inter</i> (67%-27%-3%-3%)

La distribution des isomères de $(\text{H}_2\text{O})_8$ en fonction de la température est représentée sur la figure 4.2. Dans le cas de $(\text{H}_2\text{O})_8$, contrairement à $(\text{H}_2\text{O})_6$, on peut voir que **les NQEs n'ont que très peu d'impact sur les distributions des isomères considérés et sur la température de changement de phase**. En effet, celle-ci est de ~ 155 - 165 K dans les deux cas et la température de croisement entre les isomères [0604] et *others* ne diffère que de 2 K. Comme pour d'autres propriétés, cela s'explique par la grande stabilité du minimum global de $(\text{H}_2\text{O})_8$, l'isomère [0604], par rapport aux autres isomères de basse énergie. L'inclusion des NQEs ne modifiant pas ou très peu cet écart énergétique, ils n'ont que peu d'influence sur la transition de phase. Si l'on regarde les distributions des isomères [1110] et [0403],

4. IMPACT DES EFFETS QUANTIQUES NUCLÉAIRES SUR LA DISTRIBUTION D'ISOMÈRES DANS $(\text{H}_2\text{O})_6$ et $(\text{H}_2\text{O})_8$

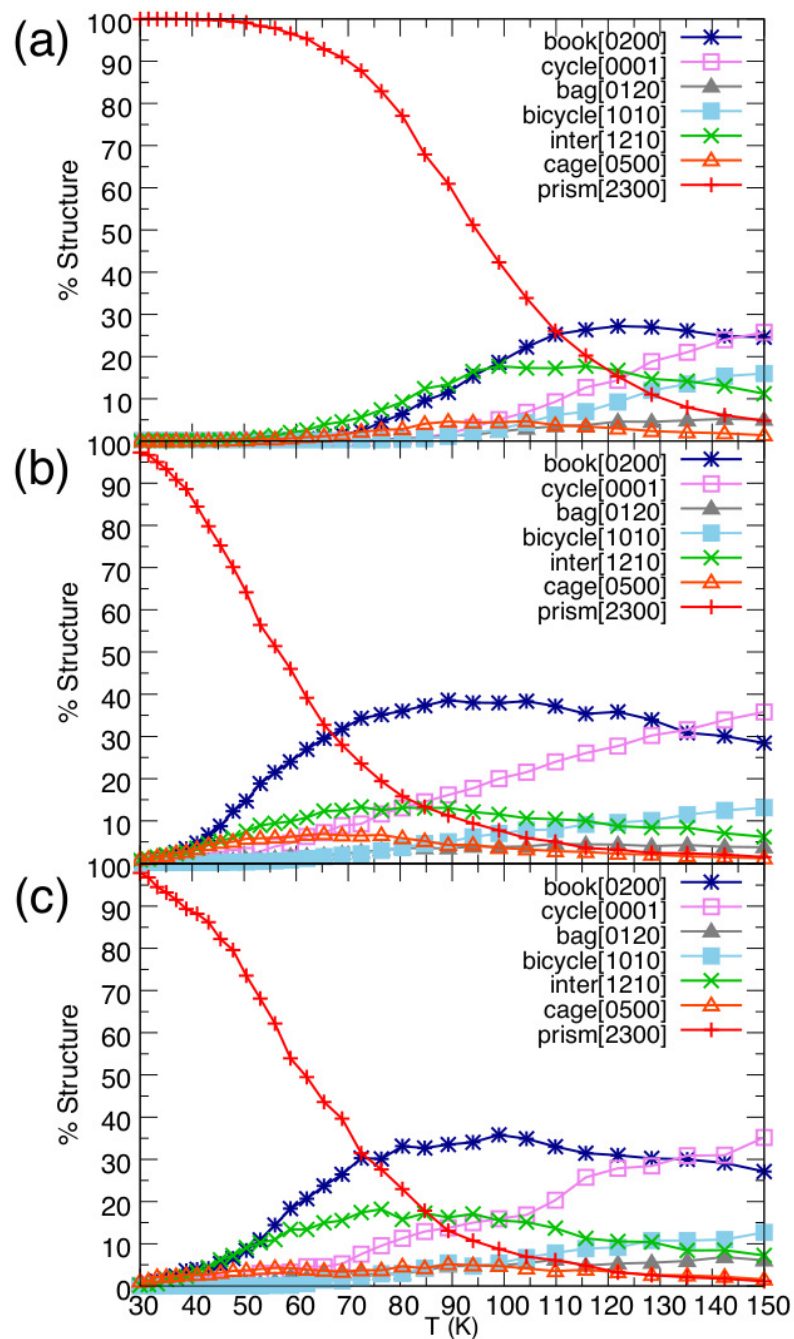


FIGURE 4.1 – Distribution des isomères de $(\text{H}_2\text{O})_6$ en fonction de la température issue de simulations (a) PTMD, (b) PTPIMD (16 répliques) et (c) PTPIMD (32 répliques).

4.4 Conclusions et perspectives

TABLE 4.3 – Énergies de liaison relatives (en kJ.mol^{-1}) des cinq isomères de plus basse énergie de $(\text{H}_2\text{O})_8$ obtenues au niveau SCC-DFTB et DFT avec les fonctionnelles BLYP et PBE0. Les géométries des isomères sont représentées dans le chapitre 2.

Isomères	SCC-DFTB	BLYP	PBE0
<i>cube</i> [0604]	0,0	0,0	0,0
[0300]	24,7	30,3	31,1
[2400]	19,8	34,8	33,3
[1110]	33,4	42,0	40,1
[2520]	15,3	-	-

ils diffèrent sensiblement entre les simulations PTMD et PTPIMD. Par conséquent, bien que les NQEs n'ont que peu d'effet sur la distribution de l'isomère [0604], on peut s'attendre à ce qu'ils perturbent sensiblement les distributions des isomères regroupés dans la courbe *others* sans que cela ait un impact sur la transition de phase. La dernière observation importante que l'on peut faire est **l'apparition à 100 K de 2% d'isomère [2520] dans la simulations PTPIMD**. Dans la simulation classique, à 100 K, seul l'isomère [0604] est présent. Cette coexistence de [0604] et [2520] est rendue possible par le potentiel SCC-DFTB qui prédit une différence d'énergie de seulement $15,3 \text{ kJ.mol}^{-1}$ entre ces deux isomères comme le montre le tableau 4.3. On peut penser que cette coexistence sera dépendante du potentiel utilisé pour la description de l'agrégat, en particulier de l'énergie relative entre [0604] et [2520]. Au niveau DFT par exemple, les données du tableau 4.3 laissent penser que cette coexistence ne sera pas observée puisque [2520] n'est pas un minium de la surface d'énergie potentielle.

4.4 Conclusions et perspectives

Dans ce chapitre, nous avons étudié l'impact des NQEs sur les distributions d'isomères en fonction de la température pour deux agrégats d'eau, $(\text{H}_2\text{O})_6$ et $(\text{H}_2\text{O})_8$. Pour cela, nous avons réalisé des simulations PTMD et PTPIMD. Dans le cas de $(\text{H}_2\text{O})_6$, nous avons comparé nos résultats à ceux obtenus dans l'article [362]. Dans cet article, les auteurs avaient utilisé quatre potentiels de type champ de force différents. Ils ont montré que deux d'entre eux, les potentiels WHBB et HBB2-pol

4. IMPACT DES EFFETS QUANTIQUES NUCLÉAIRES SUR LA DISTRIBUTION D'ISOMÈRES DANS $(\text{H}_2\text{O})_6$ et $(\text{H}_2\text{O})_8$

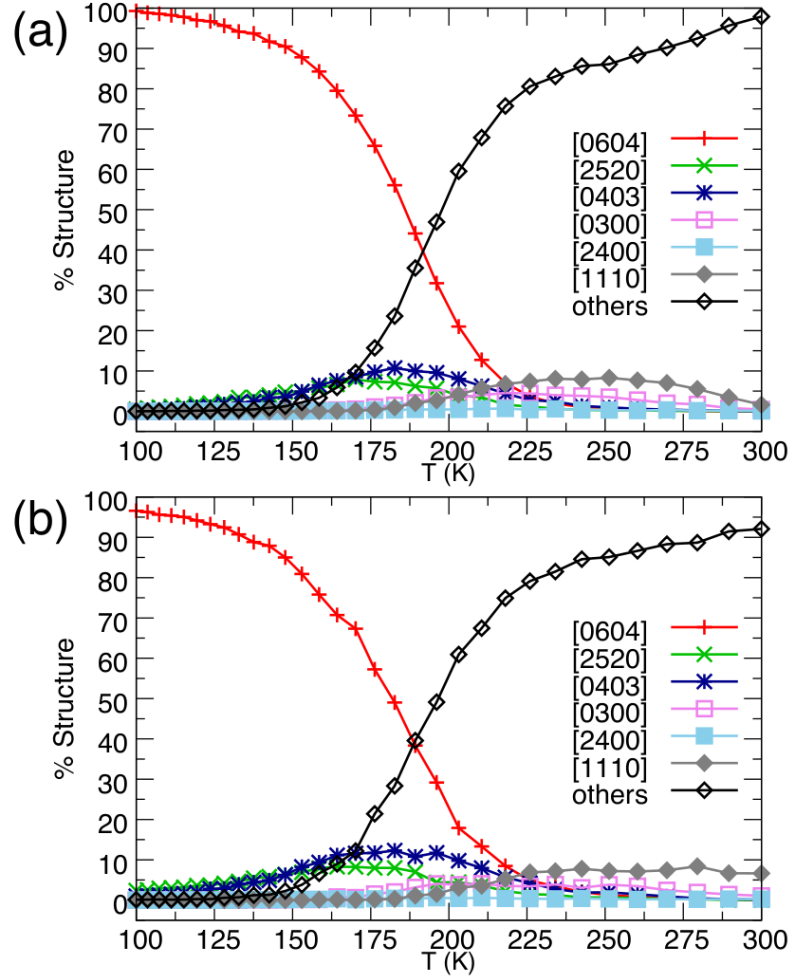


FIGURE 4.2 – Distribution des isomères de $(\text{H}_2\text{O})_8$ en fonction de la température issue de simulations (a) PTMD et (b) PTPIMD (32 répliques).

permettaient de reproduire la coexistence des deux isomères *prism* et *cage* qui avait été observée expérimentalement à basse température. Dans nos simulations, **nous n'avons pas pu retrouver cette coexistence** du fait que le potentiel SCC-DFTB utilisé prédit une énergie de liaison de l'isomère *cage* trop grande par rapport à celle de l'isomère *prism*. Cependant, l'isomérisation de l'isomère *prism* en isomère *book* que l'on observe dans la simulation PTMD, a lieu 20 K plus bas dans les simulations PTPIMD. **Les NQEs ont donc bel et bien un impact sur les distributions d'isomères en SCC-DFTB.** Pour ce qui est de $(\text{H}_2\text{O})_8$, les NQEs n'ont qu'un impact limité sur la distribution des isomères que nous avons considérés. Cela est attribuable à la grande stabilité du minimum global de cet agrégat

4.4 Conclusions et perspectives

comparativement aux autres isomères de basse énergie. Dans le futur, nous souhaiterions étudier l’impact des NQEs sur les courbes de capacité calorifique de ces deux agrégats. Malheureusement, la méthode des histogrammes multiples telle qu’elle est implémentée dans le programme que nous utilisons ne peut pas être appliquée à des simulations PIMD. Pour cela, la capacité calorifique doit être évaluée d’une autre façon comme par exemple celle proposée par Stephen F. Langley *et al.* [382]. **Nous souhaiterions également étendre cette étude à d’autres agrégats aqueux** tels que $(\text{H}_2\text{O})_7$, $(\text{H}_2\text{O})_9$ et $(\text{H}_2\text{O})_{10}$ et à l’agrégat $(\text{H}_2\text{O})_8\text{H}^+$. En effet, dans ce dernier cas le proton déstabilise la solidité de la structure cubique de l’agrégat et nous pouvons nous attendre à ce que les NQEs jouent alors un rôle plus important. Cependant, la SCC-DFTB ne permet pas de décrire correctement les agrégats d’eau protonés dans sa formulation actuelle [307, 306, 308, 309] et une étape de reparamétrisation du potentiel est à envisager. De plus, comme nous l’avons souligné dans le chapitre 2, une analyse des cycles de molécules d’eau n’est pas pertinente n’est pas adaptée aux agrégats protonés. Dans ce cas, il pourrait être intéressant d’utiliser des arbres binaires enracinés pour décrire les réseaux de liaisons hydrogènes [251, 252, 253, 254] ou alors des indicateurs plus abstraits tels que SOAP (pour *Smooth Overlap of Atomic Positions*) [255].

Impact des effets quantiques
nucléaires sur les spectres
d'absorption infrarouge de systèmes
 $(\text{H}_2\text{O})_n/\text{Ar}$ et $(\text{PAH})(\text{H}_2\text{O})/\text{Ar}$

Chapitre 5

Impact des effets quantiques nucléaires sur les spectres d'absorption infrarouge de systèmes $(\text{H}_2\text{O})_n/\text{Ar}$ et $(\text{PAH})(\text{H}_2\text{O})/\text{Ar}$

5.1 Introduction

Les hydrocarbures polycycliques aromatiques (PAHs) sont importants en astrochimie [383, 384, 385, 386, 387, 388, 389, 390, 391, 392] et en chimie atmosphérique [393, 394, 395, 396, 397, 398, 399, 400, 401, 402, 403]. Dans le contexte **astrophysique**, les PAHs ont suscité un grand intérêt depuis le milieu des années 80, lorsque les bandes infrarouges aromatiques (AIBs pour *Aromatic Infrared Bands* en anglais), un ensemble de bandes d'émission dans l'IR moyen, détectées dans de nombreuses régions du milieu interstellaire (MIS), ont été attribuées à une population de PAHs dite « AstroPAH » [404, 405]. Cette population contiendrait 10–20% du carbone interstellaire [406]. Néanmoins, jusqu'à très récemment [407], **aucune molécule de PAH spécifique n'avait été détectée**. Dans le MIS, les chimies du gaz et des grains sont étroitement liées. Dans les régions les plus froides et denses du MIS, il est envisagé que les PAHs se condensent sur des grains recouverts par une couche de glace et que la réactivité des PAHs, induite par des processus photophysiques ou d'autres processus énergétiques, avec les molécules à la surface de cette glace contribue à la complexité moléculaire du MIS. **Ce contexte a motivé**

5. IMPACT DES EFFETS QUANTIQUES NUCLÉAIRES SUR LES SPECTRES IR DE SYSTÈMES $(\text{H}_2\text{O})_n/\text{Ar}$ ET $(\text{PAH})(\text{H}_2\text{O})/\text{Ar}$

des études expérimentales en laboratoire de photoréactivité des PAHs avec des glaces d'eau en environnement cryogénique [408, 409, 410, 411]. Il a été montré que les PAHs plans réagissaient efficacement avec des molécules d'eau et s'oxydaient en alcools et quinones après irradiation par photons de basse énergie ($\lambda > 160$ nm). Une réactivité similaire a été montrée lors d'expériences en matrice de gaz rare où le PAH n'est pas en présence de glace mais en interaction avec des molécules et agrégats d'eau [412, 413, 414]. **Ces résultats ont motivé des études théoriques pour comprendre cette photoréactivité.** La première étape est de caractériser les structures des agrégats $(\text{PAH})(\text{H}_2\text{O})_n$ en matrice de gaz rare, notamment par la détermination de leurs spectres d'absorption IR qui pourront être comparés aux résultats expérimentaux. Cette étude a été effectuée dans le cas de complexes $(\text{C}_{24}\text{H}_{12})(\text{H}_2\text{O})/\text{Ar}$ où $\text{C}_{24}\text{H}_{12}$ est le coronène [415], en lien avec des résultats expérimentaux obtenus par l'équipe de Joëlle Mascetti (ISM, Bordeaux). Une étude similaire est effectuée dans le cas de $(\text{C}_{20}\text{H}_{10})(\text{H}_2\text{O})/\text{Ar}_n$ ($n=1-3$) où $\text{C}_{20}\text{H}_{10}$ est le corannulène, un PAH non plan qui lui ne photoréagit pas avec l'eau dans des conditions similaires à celles des expériences réalisées avec le coronène. Une étude est en cours pour comprendre cette différence.

Ces études théoriques sont difficiles car les systèmes $(\text{PAHs})(\text{H}_2\text{O})_n$ environnés par une matrice sont des systèmes de grande taille et il est nécessaire de décrire explicitement leur structure électronique pour décrire les spectres d'absorption IR et la réactivité des agrégats. Cependant, la description de ces systèmes dans leur ensemble (complexe $(\text{PAH})(\text{H}_2\text{O})_n$ + matrice d'argon) en traitant explicitement leur structure électronique conduit à des calculs beaucoup trop coûteux pour pouvoir mettre en œuvre des simulations de dynamique moléculaire. Pour palier à cette limitation, **un potentiel QM/MM utilisant une partie QM en SCC-DFTB pour la description du complexe $(\text{PAH})(\text{H}_2\text{O})_n$ et un champ de force [416] pour la description des atomes d'argon a donc été développé** dans l'équipe MAD du LCPQ [417]. Ce potentiel QM/MM a ensuite été utilisé pour réaliser des spectres d'absorption IR d'agrégats d'eau $(\text{H}_2\text{O})_{1-6}$ [418], de coronène $\text{C}_{24}\text{H}_{12}$ et de coronène en interaction avec une ou deux molécules d'eau $(\text{C}_{24}\text{H}_{12})(\text{H}_2\text{O})_{1,2}$ [415] en matrice d'argon à partir de simulations de dynamique moléculaire. **Les spectres d'absorption IR ainsi obtenus ont permis de corroborer les observations expérimentales** : l'environnement matriciel stabilise les structures où les molécules d'eau interagissent via leurs atomes d'oxygène avec les atomes d'hydrogène du co-

ronène, ce qui modifie l'énergie de la vibration γ_{CH} (vibration hors plan), en accord avec les résultats expérimentaux [415].

Ces expériences et ces simulations sont réalisées à une température proche de 10 K. A cette température, on peut s'attendre à ce que les effets quantiques des noyaux jouent un rôle significatif sur les propriétés de l'agrégat. **Nous avons donc décidé de reproduire ces simulations en prenant en compte les effets quantiques des noyaux.** Pour les **agrégats d'eau**, nous avons étudié les systèmes $(\text{H}_2\text{O})_{1-4}$ et l'isomère *bag* de $(\text{H}_2\text{O})_6$. Concernant les **systèmes contenant du coronène**, nous avons étudié $\text{C}_{24}\text{H}_{12}$ et $(\text{C}_{24}\text{H}_{12})(\text{H}_2\text{O})$ avec H_2O dans le plan du coronène qui est la configuration la plus stable de ce complexe [415]. Dans la suite de ce manuscrit, elle sera notée $(\text{C}_{24}\text{H}_{12})(\text{H}_2\text{O})_\sigma$. En plus de ces systèmes, nous avons également commencé l'étude de **systèmes de corannulène** seul, $\text{C}_{20}\text{H}_{10}$, et de corannulène en interaction avec une molécule d'eau, $(\text{C}_{20}\text{H}_{10})(\text{H}_2\text{O})$. Pour ces systèmes, les géométries étudiées ont été déterminées récemment.

Cependant, si les simulations de dynamique moléculaire classique pour obtenir des spectres d'absorption IR sont compliquées, la modélisation de spectres d'absorption IR dans le formalisme PIMD l'est encore plus. En effet, comme nous l'avons vu dans la section 1.3.5, les spectres d'absorption IR sont obtenus à partir de la transformée de Fourier de la fonction d'autocorrélation du moment dipolaire instantané du système. Une telle approche permet de prendre en compte les effets anharmoniques contrairement à une approche harmonique basée sur la calcul de la matrice hessienne du système. Une telle approche rend incompatible l'utilisation d'un thermostat pour ce type de simulations, car ce dernier perturbe nécessairement le comportement dynamique du système. **Ces simulations doivent donc être réalisées dans l'ensemble NVE.** L'équivalent d'une simulation NVE dans le formalisme PIMD est une **simulation RPMD**. Le problème de cette méthode pour l'étude des spectres d'absorption IR est que les premiers modes propres du polymère cyclique (RP) se trouvent dans la même gamme de fréquences que les modes de vibration du système physique. A titre d'exemple, à 10 K avec 32 répliques dans le RP, le mode $k = 0$ présente une fréquence de vibration de 87 cm^{-1} . On peut donc s'attendre à ce que les modes du système physique présentant des fréquences de vibration supérieures soient contaminés par les modes du polymère cyclique. Une solution serait d'utiliser la **méthode CMD**. Cela a déjà été fait par Florent Calvo

5. IMPACT DES EFFETS QUANTIQUES NUCLÉAIRES SUR LES SPECTRES IR DE SYSTÈMES $(\text{H}_2\text{O})_n/\text{Ar}$ ET $(\text{PAH})(\text{H}_2\text{O})/\text{Ar}$

et al. [419] qui ont étudié les spectres d'absorption IR du coronène, de l'anthracène et du pyrène à l'aide des méthodes RPMD et CMD. **La méthode CMD souffre cependant d'autres problèmes.** En effet, l'utilisation de masses fictives plus faibles que les masses physiques pour les modes normaux autres que le mode du centroïde nécessite l'emploi d'un pas de temps beaucoup plus faible pour l'intégration des équations du mouvement et la méthode CMD induit des déplacements artificiels vers le rouge résultant de ce que l'on appelle "le problème de courbure" (*curvature problem* en anglais). Finalement, une autre solution consiste à thermostat les modes propres du RP afin d'atténuer leur vibration et donc le couplage fictif avec les modes de vibration du système physique. **C'est la méthode TRPMD** proposée par Rossi *et al.* [223, 224]. Initialement, cette méthode a été proposée en utilisant un thermostat de Langevin avec un bruit blanc de type PILE [223]. Cette méthode s'est avérée efficace pour éviter les couplages fictifs entre modes normaux du RP et modes de vibration du système physique mais elle induit un élargissement des pics dans le spectre de vibration et ce d'autant plus que la température est faible. De plus, les auteurs de cet article recommandent de ne pas utiliser cette méthode pour étudier les spectres de vibration de molécules en phase gazeuse. Finalement, il a été proposé de thermostat les modes du RP avec un thermostat basé sur l'équation de Langevin généralisée présentée dans la section 1.3.3 [224]. Cette dernière approche résout le problème d'élargissement des pics en TRPMD avec un bruit blanc mais induit un déplacement vers le bleu des modes de libration. De façon générale, **les méthodes CMD ou TRPMD avec un bruit blanc devraient être utilisées pour étudier des modes à basses fréquences et les modes à hautes fréquences devraient être étudiés en TRPMD avec un thermostat GLE.** C'est cette dernière méthode que nous avons utilisée dans ce chapitre pour simuler les spectres d'absorption IR des agrégats d'eau $(\text{H}_2\text{O})_{1,4}$, $(\text{H}_2\text{O})_{6(\text{bag})}$, $\text{C}_{24}\text{H}_{12}$, $(\text{C}_{24}\text{H}_{12})(\text{H}_2\text{O})_\sigma$, $\text{C}_{20}\text{H}_{10}$ et $(\text{C}_{20}\text{H}_{10})(\text{H}_2\text{O})$ en matrice d'argon.

Le chapitre est organisé comme suit : dans **une première partie**, nous présenterons les méthodes ainsi que les paramètres de calcul utilisés pour réaliser cette étude. Dans **la seconde partie**, nous présenterons et discuterons les spectres d'absorption IR obtenus par les méthodes MD et TRPMD-GLE pour des agrégats d'eau, des espèces contenant du coronène seul ou en interaction avec de l'eau et des espèces contenant du corannulène seul ou en interaction avec de l'eau. Les conclusions et perspectives de ce travail seront finalement présentées dans une **troisième partie**.

5.2 Méthodes employées et paramètres de calcul

5.2.1 Configurations initiales

Pour simuler les matrices d'argon solides utilisées dans les expériences, nous avons utilisé des agrégats d'argon composés de 1055 atomes organisés en maille cubique à face centrée. Ce nombre d'atomes correspond à la taille minimale de l'agrégat pour que celui-ci ne soit pas perturbé par l'introduction d'un composé en son sein. Les composés étudiés ont été insérés dans les matrices d'argon en ôtant le nombre minimum d'argon nécessaire pour que l'énergie de substitution soit négative après optimisation. Pour les systèmes $(\text{H}_2\text{O})_n$, les molécules d'eau ont systématiquement été insérées à la place d'un atome d'argon. Pour les systèmes $\text{C}_{24}\text{H}_{12}/\text{Ar}$ et $(\text{C}_{24}\text{H}_{12})(\text{H}_2\text{O})_\sigma/\text{Ar}$, le coronène a systématiquement été introduit dans le plan cristallographique [111] en retirant 7 atomes d'argon. Il a été montré que cette insertion est celle qui donne une énergie de substitution la plus faible (-4814 cm^{-1}) [415]. Comme souligné en introduction, pour le complexe $(\text{C}_{24}\text{H}_{12})(\text{H}_2\text{O})/\text{Ar}$, la structure la plus stable est celle dans laquelle la molécule d'eau est dans le plan du coronène avec l'oxygène pointant entre deux liaisons C–H. La structure du système $(\text{C}_{24}\text{H}_{12})(\text{H}_2\text{O})_\sigma/\text{Ar}$ est représentée sur la figure 5.1.

Pour les systèmes $\text{C}_{20}\text{H}_{10}/\text{Ar}$ et $(\text{C}_{20}\text{H}_{10})(\text{H}_2\text{O})/\text{Ar}$, les structures les plus stables sont également celles qui présentent le $\text{C}_{20}\text{H}_{10}$ au centre de la matrice d'argon dans le plan cristallographique [111] et 7 atomes d'argon ont été retirés de la matrice. Pour les complexes $(\text{C}_{20}\text{H}_{10})(\text{H}_2\text{O})/\text{Ar}$, trois insertions possibles de la molécule d'eau ont été considérées dans cette étude. Les structures des trois complexes $(\text{C}_{20}\text{H}_{10})(\text{H}_2\text{O})/\text{Ar}$ sont représentées sur la figure 5.2. Les noms *ext*, *int* et *side* désignent respectivement les complexes de $(\text{C}_{20}\text{H}_{10})(\text{H}_2\text{O})$ avec H_2O situé sur la face convexe du corannulène (figure 5.2 (a)), sur la face concave (figure 5.2 (b)) et dans le plan des liaisons C–H (figure 5.2 (c)).

5. IMPACT DES EFFETS QUANTIQUES NUCLÉAIRES SUR LES SPECTRES IR DE SYSTÈMES $(\text{H}_2\text{O})_n/\text{Ar}$ ET $(\text{PAH})(\text{H}_2\text{O})/\text{Ar}$

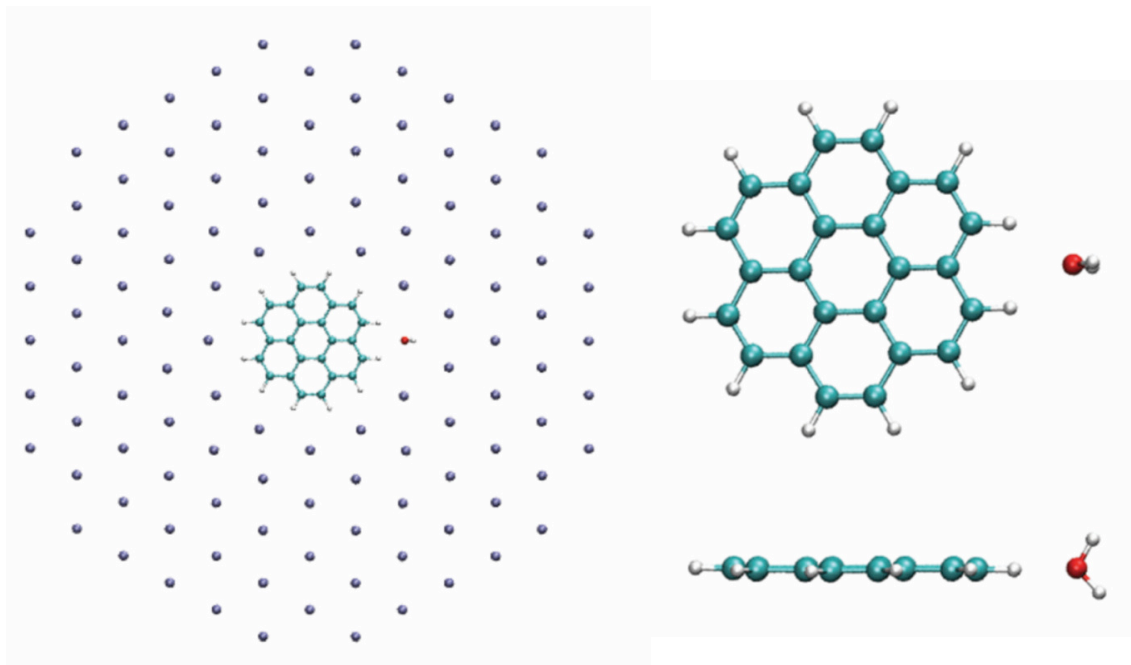


FIGURE 5.1 – Tirée de la référence [415]. Géométrie de $(\text{C}_{24}\text{H}_{12})(\text{H}_2\text{O})_\sigma/\text{Ar}$ optimisées en SCC-DFTB/FF. À gauche, le complexe $(\text{C}_{24}\text{H}_{12})(\text{H}_2\text{O})_\sigma$ est représenté vu du dessus avec un plan d’atomes d’argon. À droite, les représentations correspondent aux géométries du complexe tel qu’il est dans la matrice mais les atomes d’argon ont été retirés pour faciliter la visualisation.

5.2.2 Calcul de la structure électronique

Comme mentionné en introduction, le calcul des énergies et des forces a été réalisé à l’aide d’un potentiel QM/MM [417]. Ce potentiel consiste en une modification de la matrice hamiltonienne SCC-DFTB basée sur la théorie des perturbations au premier ordre combinée à une contribution de polarisation. En pratique, deux termes scalaires décrivent l’interaction de l’argon avec les électrons de cœur de la partie QM, PAH et molécules d’eau, et l’interaction de dispersion entre ces deux sous-systèmes. Un troisième terme non scalaire décrit l’interaction entre l’argon et les électrons de valence de la partie QM. Enfin, un terme de polarisation est ajouté pour prendre en compte la polarisation des atomes d’argon par la partie QM dans un schéma

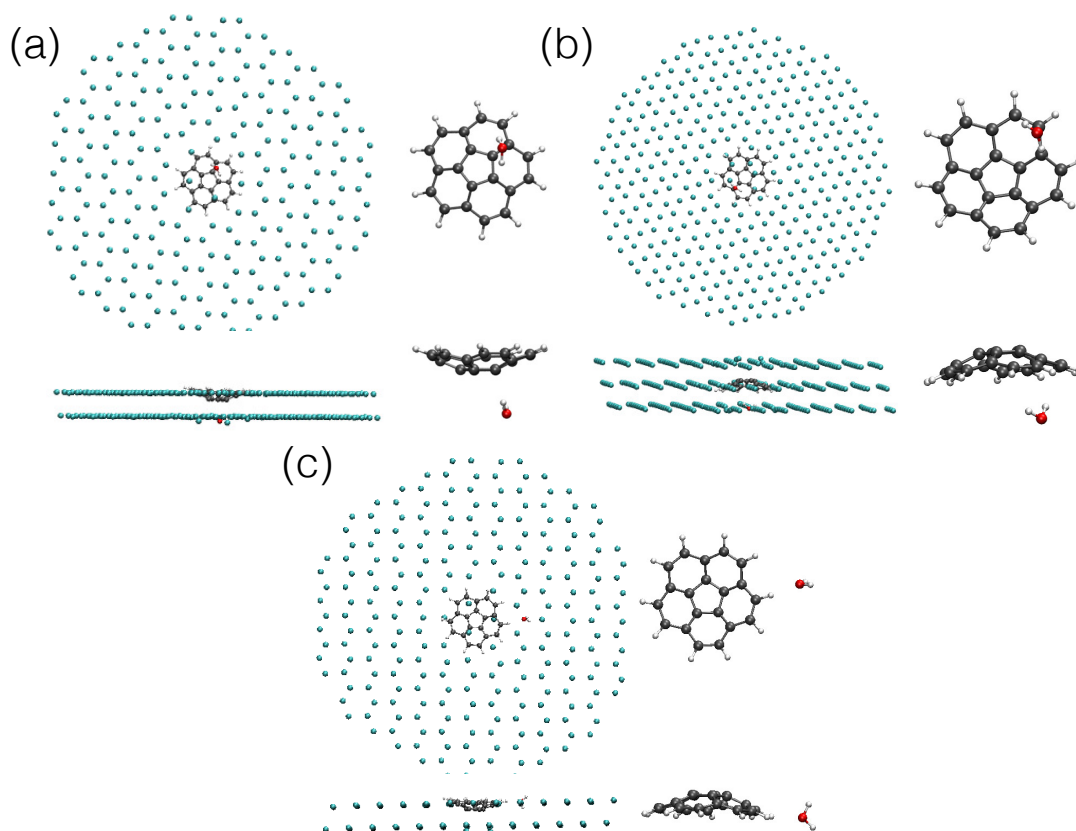


FIGURE 5.2 – Géométries de $(C_{20}H_{10})(H_2O)/Ar$ (a) *ext*, (b) *int* et (c) *side* optimisées en SCC-DFTB/FF. À gauche, les géométries sont représentées avec la matrice d’argon vues du dessus et sur la tranche. Deux plans d’argon sont représentés pour les systèmes *ext* et *side* et trois plans d’argon sont représentés pour le système *int*. À droite, les géométries sont représentées telles qu’elles sont en matrice mais les atomes d’argon ont été retirés pour faciliter la visualisation.

auto-cohérent. L’interaction entre atomes d’argon est décrite par un potentiel de Aziz [416]. Pour la partie QM, c’est à dire SCC-DFTB, nous avons utilisé l’ensemble *mio* des tables d’intégrales de Slater-Koster [72]. Nous avons utilisé un terme de dispersion empirique afin d’améliorer la description des interactions intermoléculaires et nous avons utilisé des charges *Class IV-Charge Model 3* [142] pour le traitement des interactions de Coulomb comme décrites dans la section 1.2.3. La procédure auto-cohérente est considérée comme convergée et stoppée quand les fluctuations de charges atomiques sont inférieures à 10^{-9} u.a. Les calculs SCC-DFTB/FF ont été effectués avec le code deMonNano [97]. Ce modèle SCC-DFTB/FF a été développé puis validé pour l’étude d’agrégats d’eau piégés en matrice d’argon [418].

5. IMPACT DES EFFETS QUANTIQUES NUCLÉAIRES SUR LES SPECTRES IR DE SYSTÈMES $(\text{H}_2\text{O})_n/\text{Ar}$ ET $(\text{PAH})(\text{H}_2\text{O})/\text{Ar}$

5.2.3 Exploration de la surface d'énergie potentielle

Les simulations de MD classique et de PIMD ont été réalisées à l'aide de l'interface que nous avons mis en place entre les programmes i-PI et deMonNano. Comme expliqué dans le chapitre 4, le calcul de l'énergie et des forces est délégué à deMonNano tandis que l'intégration des équations du mouvement est prise en charge par i-PI. Le moment dipolaire μ est calculé par deMonNano et envoyé à i-PI de façon identique à l'énergie et aux forces. i-PI ne procède à aucune opération sur μ et se contente de l'écrire dans un fichier d'output supplémentaire. Comme **les spectres d'absorption IR sont calculés à partir de la fonction d'autocorrélation du moment dipolaire**, ce dernier **doit être calculé à partir de configuration extraites de dynamiques NVE** pour éviter les perturbations engendrées par l'application d'un thermostat. Comme mentionné dans la section 1.3.2, les simulations de MD classiques, mais également PIMD, dans l'ensemble NVE posent des problèmes d'ergodicité. Pour remédier à ces problèmes, une stratégie consiste à réaliser plusieurs simulations dans l'ensemble NVE utilisant comme conditions initiales des configurations générées dans l'ensemble NVT. Les spectres d'absorption IR obtenus à partir de ces différentes simulations NVE sont ensuite moyennés. Cette stratégie est celle que nous avons employée pour les simulations de MD classique comme celles en PIMD. Notre méthodologie a été la suivante. Dans les deux cas, **nous avons tout d'abord réalisé 1 ns d'équilibration dans l'ensemble NVT**. Un pas de temps de 0,2 fs a été employé en MD comme en PIMD. L'énergie cinétique des systèmes a été contrôlée pour que leur température moyenne corresponde à la température cible de 10 K. Pour les simulations de MD classique, un thermostat SVR avec une constante de temps de 0,1 ps a été utilisé. Pour les simulations PIMD, un polymère cyclique composé de 32 répliques a été employé et la température a été contrôlée à l'aide d'un thermostat PI+GLE dont les paramètres ont été automatiquement générés à l'aide du site web GLE4MD [207]. Ces derniers ont été choisis de manière à couvrir les fluctuations quantiques dans une gamme de fréquence allant jusqu'à 6936 cm^{-1} , c'est à dire au delà de la plus grande fréquence de vibration physique du système ($\sim 4000\text{ cm}^{-1}$). Nous avons ensuite **extrait dix configurations des dernières 500 ps de ces simulations**. À partir de ces dix configurations initiales, **des dynamiques de 300 ps chacune ont été réalisées dans l'ensemble NVE**. Les premières 50 ps de ces simulations NVE ont été considérées comme phase

d'équilibration et le moment dipolaire μ a donc été calculé à chaque pas de temps au cours des dernières 250 ps. Nous avons utilisé pour ces simulations un pas de temps de 0,1 fs. Pour pallier aux problèmes de résonance entre les modes de vibration du polymère cyclique et les modes physiques qui nous intéressent, l'**approche TRPMD-GLE a été employée** pour les simulations PIMD. Nous avons utilisé le thermostat dans sa version exponentiellement amorti GLE(C) dont les paramètres ont été automatiquement générés à l'aide du site web GLE4MD [207].

5.2.4 Calcul des spectres d'absorption IR

Les spectres d'absorption IR ont été calculés à partir de la transformée de Fourier de la fonction d'autocorrélation du moment dipolaire μ comme décrit dans la section 1.3.5. Pour chaque simulation, la fonction d'autocorrélation est moyennée sur 100×1600 segments de chevauchement de 26 ps chacun. Les spectres d'absorption IR ainsi obtenus sont ensuite moyennés sur les 10 simulations NVE décrites ci-dessus. Cette procédure permet de minimiser l'impact des conditions initiales sur le spectre final. Pour les simulations TRPMD-GLE, la fonction d'autocorrélation a été calculée sur le moment dipolaire μ_{moy} moyenné sur les 32 répliques du polymère cyclique :

$$\mu_{moy} = \frac{1}{32} \sum_{j=1}^{32} \mu_j \quad (5.1)$$

5.3 Résultats et discussion

5.3.1 Comparaison des spectres IR de $(\text{H}_2\text{O})_n/\text{Ar}$ en MD et TRPMD-GLE

Dans cette partie, nous discuterons des spectres d'absorption IR issus de simulations MD classiques et TRPMD-GLE pour les systèmes $(\text{H}_2\text{O})_{1-4,6(bag)}/\text{Ar}$. Dans

5. IMPACT DES EFFETS QUANTIQUES NUCLÉAIRES SUR LES SPECTRES IR DE SYSTÈMES $(\text{H}_2\text{O})_n/\text{Ar}$ ET $(\text{PAH})(\text{H}_2\text{O})/\text{Ar}$

un premier temps, les spectres issus des simulations de MD classique obtenus au cours de ce travail seront comparés à ceux obtenus dans l'article [418]. L'origine des différences entre les deux séries de spectres sera discutée. Dans un second temps, les spectres MD et TRPMD-GLE seront comparés afin d'évaluer l'impact des NQEs sur les modes de vibration dans ces différents agrégats d'eau.

Analyse des spectres MD. Les spectres d'absorption IR des systèmes $(\text{H}_2\text{O})_{1-4,6(\text{bag})}/\text{Ar}$ obtenus en MD classique et en TRPMD-GLE sont présentés sur la figure 5.3. Les positions des bandes d'absorptions des modes d'élongation symétriques et asymétriques ν_1 et ν_3 et du mode de cisaillement ν_2 sont reportées dans le tableau 5.1. Nous avons également reporté les valeurs obtenues par Aude Simon *et al.* [293] ainsi que les valeurs expérimentales pour $\text{H}_2\text{O}/\text{Ar}$ [420], $(\text{H}_2\text{O})_2/\text{Ar}$ [421], $(\text{H}_2\text{O})_3/\text{Ar}$ [422] et $(\text{H}_2\text{O})_4/\text{Ar}$ [424]. En premier lieu, il faut remarquer que les fréquences obtenues en MD classique dans nos travaux ne correspondent pas exactement aux fréquences obtenues par Aude Simon *et al.* dans la référence [418]. L'origine de ces différences est multiple. Dans les simulations de la référence [418], les matrices d'argon ne possèdent que ~ 340 atomes d'argon et les positions des atomes d'argon en surface ne sont pas fixes. Dans cet article, les auteurs mentionnent qu'en utilisant une matrice de 1330 atomes d'argon la position de la bande ν_2 du monomère d'eau $\text{H}_2\text{O}/\text{Ar}$ n'est pas modifiée et que les positions des bandes ν_1 et ν_3 sont respectivement décalées de $+5$ et $+8 \text{ cm}^{-1}$. Une autre raison pour laquelle les positions des bandes sont légèrement différentes entre nos travaux et ceux de Aude Simon *et al.* est la méthodologie d'équilibration des systèmes pour échantillonner les conditions initiales. En effet, dans la référence [418], les auteurs disent avoir équilibré le système dans l'ensemble NVE tandis que dans nos travaux l'équilibration a été réalisée dans l'ensemble NVT. Comme mentionné dans la section 1.3.2, les simulations dans l'ensemble NVT sont plus ergodiques que dans l'ensemble NVE. Ainsi, les conditions initiales pour réaliser les simulations NVE à partir desquelles a été calculé le moment dipolaire devraient être mieux échantillonnées dans les simulation NVT que nous avons réalisés. Enfin, dans nos travaux, nous avons utilisé les positions issues des équilibrations NVT comme conditions initiales des simulations NVE mais les vitesses ont été générées aléatoirement à partir d'une distribution normale centrée en 10 K. Il en résulte que la température moyenne des simulations NVE à partir desquelles ont été calculé nos spectres IR est de 10 K tandis qu'elle est de 13 K dans [418]. Malgré ces différences, les positions des bandes d'absorptions dans

5.3 Résultats et discussion

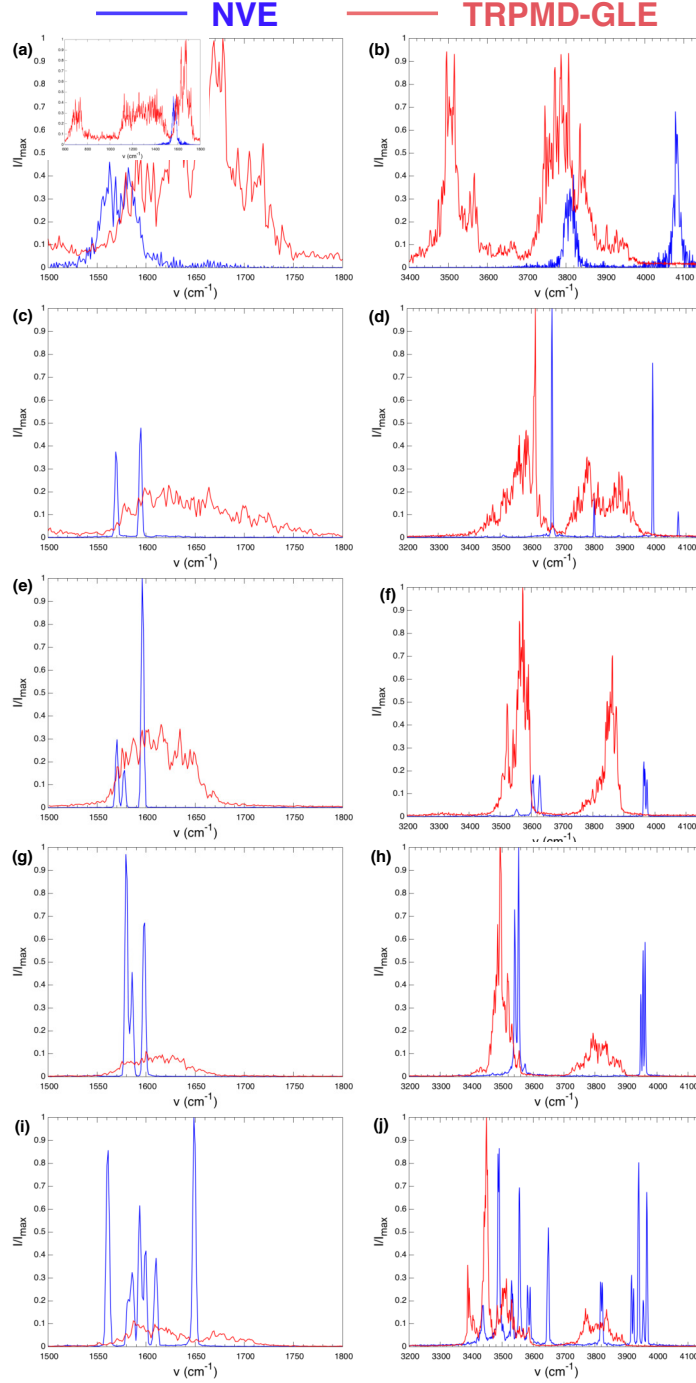


FIGURE 5.3 – Spectres IR des agrégats d’eau (a) et (b) H_2O , (c) et (d) $(\text{H}_2\text{O})_2$, (e) et (f) $(\text{H}_2\text{O})_3$, (g) et (h) $(\text{H}_2\text{O})_4$, (i) et (j) $(\text{H}_2\text{O})_{6(\text{bag})}$, en matrice d’argon. Les courbes bleus et rouges correspondent respectivement aux spectres MD et TRPMD-GLE. La partie gauche correspond aux mode de cisaillement ν_2 et la partie de droite aux elongations symétriques et asymétriques ν_1 et ν_3 .

5. IMPACT DES EFFETS QUANTIQUES NUCLÉAIRES SUR LES SPECTRES IR DE SYSTÈMES $(\text{H}_2\text{O})_n/\text{Ar}$ ET $(\text{PAH})(\text{H}_2\text{O})/\text{Ar}$

TABLE 5.1 – Positions des bandes d’absorption (en cm^{-1}) des modes de vibration de $(\text{H}_2\text{O})_{1-4,6(bag)}/\text{Ar}$ calculées en MD et TRPMD-GLE. Les valeurs issues de simulation MD de la référence [418] sont reportées ainsi que les valeurs expérimentales issues de ^a [420], ^b [421], ^c [422], ^d [423], ^e [424].

		MD	TRPMD-GLE	MD [418]	Exp
$\text{H}_2\text{O}/\text{Ar}$	ν_1	3813	3505	3783	3638,5 ^a , 3639,2 ^b
	ν_2	1569	1669	1568	1589 ^a , 1589,1 ^b
	ν_3	4080	3788	4050	3733,5 ^a , 3732,2 ^b
$(\text{H}_2\text{O})_2/\text{Ar}$	ν_1	3804	3614	3783	3633 ^c
		3668		3650	3574 ^c
	ν_2	1595	1664	1589	1610 ^c
		1569		1562	1593 ^c
	ν_3	4075	3883	4051	3730 ^c
		3993	3781	3975	3708 ^c
$(\text{H}_2\text{O})_3/\text{Ar}$	ν_1	3628	3575	3607	3514 ^d
		3608	3528		
	ν_2	1596	1615	1594	1601 ^d
		1578		1576	
	ν_3	1570			
		3974	3863	3958	3699 ^d
$(\text{H}_2\text{O})_4/\text{Ar}$	ν_1	3694			
		3554	3495	3530	3372 ^e
		3541			
	ν_2	1597	1600	1598	
		1586		1582	
		1579			
	ν_3	3961	3794	3955	3694,8 ^c
		3955			
		3948			
		3650	3514	3643	
$(\text{H}_2\text{O})_{6(bag)}/\text{Ar}$	ν_1	3591	3450	3602	
		3582	3390	3551	
		3556		3503	
		3530		3437	
		3491		3345	
		3487			
		3438			
		1649	1670	1628	
	ν_2	1610	1609	1619	
		1600	1587	1580	
		1593			
		1586			
		1561			
		3967	3837	3971	
		3955	3769	3943	
		3941		3920	
	ν_3	3925		3735	
		3919		3720	
		3823			
		3818			

nos simulations restent en accord avec celles obtenues dans la référence [418] (voir tableau 5.1).

Comparaison des spectres MD et TRPMD-GLE et impact des NQEs.

Concentrons nous à présent sur l'impact des NQEs en comparant les spectres obtenus au travers de simulations MD et TRPMD-GLE. Tout d'abord, contrairement à la simulation MD et aux résultats expérimentaux, aucun mode de rotation de H_2O dans la matrice d'argon n'est observé en TRPMD-GLE. En revanche, deux bandes constituées d'un massif de pics d'absorption associées à des mouvements de libration se situent respectivement entre ~ 600 et $\sim 800 \text{ cm}^{-1}$ d'une part et ~ 1100 et 1500 cm^{-1} d'autre part. Un autre effet de l'inclusion des fluctuations quantiques des noyaux en TRPMD-GLE est un décalage vers le rouge systématique des modes d'élongation symétriques ν_1 et asymétriques ν_3 ainsi qu'un décalage vers le bleu systématique du mode de cisaillement $\text{H}-\text{O}-\text{H}$ ν_2 . Pour ce dernier, l'intensité du signal diminue lorsque la taille de l'agrégat augmente. L'élargissement des bandes en TRPMD-GLE et la grande multiplicité de pics qui les constitue complique l'attribution d'une position précise pour les modes de vibrations en TRPMD-GLE. Les positions des bandes en TRPMD-GLE reportées dans le tableau sont donc à considérer avec précaution. Pour pouvoir réaliser une analyse qualitative de l'impact des NQEs nous avons reporté les valeurs correspondant aux pics de plus haute intensité mais il n'est cependant pas possible de quantifier avec précision les décalages induit par les NQEs. Ceci étant dit, les positions approximatives des modes ν_1 et ν_3 issues des simulations TRPMD-GLE sont considérablement plus proches des valeurs expérimentales que celles obtenues en MD. Dans le cas de $\text{H}_2\text{O}/\text{Ar}$, le décalage vers le rouge du mode ν_1 induit par les NQEs est trop important (3813 cm^{-1} en MD, $\sim 3505 \text{ cm}^{-1}$ en TRPMD-GLE et $3638,5 \text{ cm}^{-1}$ ou $3639,2 \text{ cm}^{-1}$ expérimentalement).

5.3.2 Comparaison des spectres IR de $\text{C}_{24}\text{H}_{12}/\text{Ar}$ et $(\text{C}_{24}\text{H}_{12})(\text{H}_2\text{O})_\sigma/\text{Ar}$ en MD et TRPMD-GLE

Dans cette partie, nous discuterons des spectres d'absorption IR de $\text{C}_{24}\text{H}_{12}/\text{Ar}$ et $(\text{C}_{24}\text{H}_{12})(\text{H}_2\text{O})_\sigma/\text{Ar}$ issus de simulations MD et TRPMD-GLE. Dans un premier temps, les effets de l'interaction avec H_2O sur les modes de vibrations du coronène seront discutés pour les spectres MD. Dans un second temps, la même

5. IMPACT DES EFFETS QUANTIQUES NUCLÉAIRES SUR LES SPECTRES IR DE SYSTÈMES $(\text{H}_2\text{O})_n/\text{Ar}$ ET $(\text{PAH})(\text{H}_2\text{O})/\text{Ar}$

discussion sera réalisée sur les spectres TRPMD-GLE. Enfin, les différences entre les spectres MD et TRPMD-GLE seront mises en évidence afin d'évaluer l'impact des NQEs sur les modes de vibration de $\text{C}_{24}\text{H}_{12}/\text{Ar}$ et $(\text{C}_{24}\text{H}_{12})(\text{H}_2\text{O})_\sigma/\text{Ar}$. Tous ces spectres sont présentés dans la figure 5.4. Les valeurs expérimentales des modes de vibration du coronène seul en matrice d'argon issues des références [415] et [425] sont reportées dans le tableau 5.2. Dans ce même tableau, les valeurs théoriques (harmoniques et dynamiques calculées en MD) de la référence [415] et celles calculées dans cette étude en MD et TRPMD-GLE sont également reportées.

Analyse des spectres MD

L'analyse des spectres d'absorption IR issus de simulations de MD a déjà été réalisée par les auteurs de [415]. Cependant, par soucis de rigueur et pour s'assurer de ne pas introduire un quelconque biais dans la comparaison avec les spectres issus de simulations TRPMD-GLE, l'analyse de l'interaction de $\text{C}_{24}\text{H}_{12}$ avec H_2O en matrice d'argon a été réitérée pour les spectres obtenus dans ce travail. Avant de s'intéresser à l'effet de l'interaction, **attardons nous tout d'abord sur les valeurs des modes de vibration du coronène seul**. Dans le tableau 5.2, on peut voir que **les valeurs théoriques calculées en MD sont proches de celles calculées dans la référence [415]**. Des différences de 8 cm^{-1} maximum sont observées pour les modes d'élongation CH (ν_{CH}). Par ailleurs, les positions absolues des modes d'élongation CC (ν_{CC}) et du mode de déformation CH dans le plan (ν_{CH}) ne sont pas en très bon accord avec les valeurs expérimentales. Cette observation a déjà été faite par les auteurs de [415] qui attribuent ce comportement au couplage de ces modes que la SCC-DFTB ne reproduit pas correctement. Les positions des bandes correspondant aux modes γ_{CH} et ν_{CH} , qui ne sont pas couplés à d'autre modes, sont quant à elles en bon accord avec les valeurs expérimentales. Discutons à présent de **l'effet de l'interaction avec H_2O sur les fréquences de vibration du coronène et de H_2O** . Si l'on regarde la figure 5.4, on s'aperçoit que le mode de déformation CCC (δ_{CCC}) en dehors du plan est décalé vers le bleu de $+3\text{ cm}^{-1}$. Les deux modes de déformation CCH (γ_{CH}) entre 800 et 830 cm^{-1} sont respectivement décalés vers le bleu de $+3\text{ cm}^{-1}$ et $+4\text{ cm}^{-1}$. La bande correspondant à la déformation CH dans le plan (δ_{CH}) est décalée vers le rouge de -10 cm^{-1} . Les trois modes ν_{CC} entre 1600 et 1900 cm^{-1} ne sont pas affectés par l'interaction avec H_2O . En revanche, une bande correspondant au mode de cisaillement de l'eau apparaît à

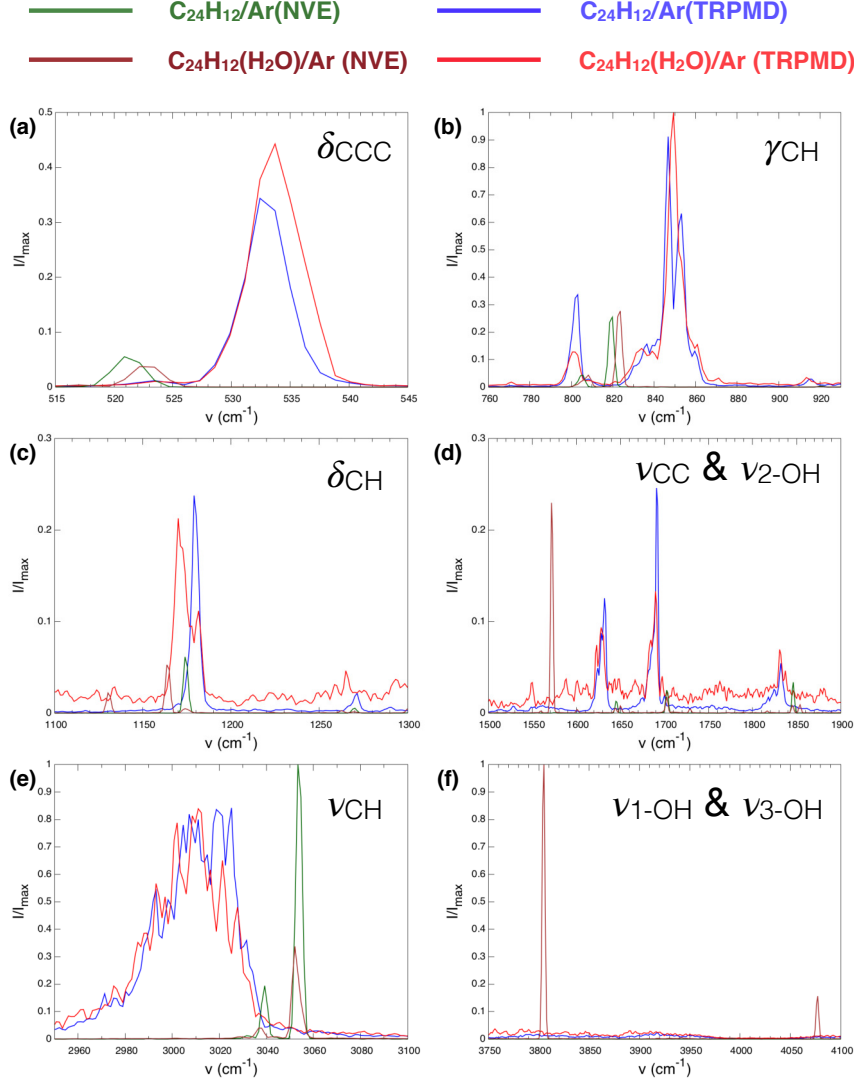


FIGURE 5.4 – Spectres d’absorption IR de $C_{24}H_{12}/Ar$ et $(C_{24}H_{12})(H_2O)_\sigma/Ar$ issus de simulations MD et TRPMD-GLE. Le spectre a été découpé en plusieurs fenêtres correspondant aux modes de vibrations les plus pertinents. Chaque spectre a été normalisé par l’intensité de la bande d’absorption la plus intense. Les spectres en vert et en bleu correspondent respectivement aux spectres de $C_{24}H_{12}/Ar$ issus de simulations MD et TRPMD-GLE tandis que les spectres en marron et en rouge correspondent respectivement aux spectres de $(C_{24}H_{12})(H_2O)_\sigma/Ar$ issus de simulations MD et TRPMD-GLE. Les différentes fenêtres de fréquences d’absorption correspondent aux modes (a) de déformation CCC (δ_{CCC}), (b) de déformation CCH en dehors du plan et dans le plan (γ_{CH}), (c) de déformation CH dans le plan (δ_{CH}), (d) d’élongation CC (ν_{CC}) et de cisaillement HOH (ν_{2-OH}), (e) d’élongation CH (ν_{CH}) et (f) d’élongation OH symétriques (ν_{1-OH}) et asymétriques (ν_{3-OH}).

5. IMPACT DES EFFETS QUANTIQUES NUCLÉAIRES SUR LES SPECTRES IR DE SYSTÈMES $(\text{H}_2\text{O})_n/\text{Ar}$ ET $(\text{PAH})(\text{H}_2\text{O})/\text{Ar}$

1572 cm^{-1} . Dans le spectre de $\text{H}_2\text{O}/\text{Ar}$, ce mode donnait lieu à un massif de pics entre 1562 et 1582 cm^{-1} . L'affinement de la bande d'absorption est dû au fait que la molécule d'eau ne tourne plus lorsqu'elle est complexée avec le coronène. Dans le complexe avec H_2O , les modes ν_{CH} sont décalés de $\sim -2\text{ cm}^{-1}$ vers le rouge. Dans la référence [415], les résultats théoriques et expérimentaux montrent également que ces modes ne sont pas affectés par l'interaction avec H_2O . Expérimentalement, lorsque la quantité d'eau augmente, la bande correspondant aux elongation O–H s'élargit et se décale vers le rouge jusqu'à absorber la bande des modes ν_{CH} . Les positions des modes ν_1 et ν_3 de l'eau ne sont pas modifiées. Ils se situent respectivement à 3806 cm^{-1} et 4077 cm^{-1} dans $\text{H}_2\text{O}/\text{Ar}$ et à 3805 cm^{-1} et 4077 cm^{-1} dans $(\text{C}_{24}\text{H}_{12})(\text{H}_2\text{O})_\sigma/\text{Ar}$.

Pour conclure sur l'effet de l'interaction entre le coronène et le monomère d'eau en position σ , en accord avec les résultats expérimentaux et théoriques présentés dans la référence [415], **les résultats issus de simulations de MD montrent que les positions des bandes correspondant aux modes δ_{CCC} , γ_{CH} , δ_{CH} et ν_{CH} du coronène sont modifiées par la présence de la molécule d'eau. En revanche, ce n'est pas le cas des modes ν_{CC} dont la positions reste inchangé malgré la présence de H_2O .** Dans la référence [415], l'expérience montre que ces bandes sont décalées vers le rouge lorsque la concentration en eau est suffisamment grande pour que des agrégats composés d'au moins trois molécules d'eau commencent à se former. Concernant la bande, à 1572 cm^{-1} (1564 cm^{-1} dans [415]) de $(\text{C}_{24}\text{H}_{12})(\text{H}_2\text{O})_\sigma/\text{Ar}$, les auteurs de [415] supposent qu'elle pourrait correspondre à une bande de très faible intensité apparaissant à 1573 cm^{-1} dans le spectre dynamique de $\text{C}_{24}\text{H}_{12}/\text{Ar}$ ou bien à un très grand décalage vers le rouge d'une bande à 1664 cm^{-1} dans le spectre harmonique. Dans notre cas, nous observons également une bande de très faible intensité à 1574 cm^{-1} dans le spectre de $\text{C}_{24}\text{H}_{12}/\text{Ar}$. Cependant cette bande à 1572 cm^{-1} se situe exactement au centre du massif correspondant au mode de cisaillement ν_2 de $\text{H}_2\text{O}/\text{Ar}$ et il nous semble donc logique de l'attribuer à ce mode là, d'autant plus que la même bande apparaît dans les spectres de $(\text{C}_{20}\text{H}_{10})(\text{H}_2\text{O})/\text{Ar}$ comme nous le verrons plus bas.

Analyse des spectres TRPMD-GLE

Pour bien comprendre quel est l'impact des NQEs sur les spectres d'absorption IR de $\text{C}_{24}\text{H}_{12}/\text{Ar}$ et $(\text{C}_{24}\text{H}_{12})(\text{H}_2\text{O})_\sigma/\text{Ar}$, nous allons dans un premier temps analyser

5.3 Résultats et discussion

TABLE 5.2 – Modes de vibration du coronène seul en matrice d’argon, $C_{24}H_{12}/Ar$. Les valeurs expérimentales des références [415] et [425] ainsi que les valeurs harmoniques et dynamiques (MD) calculées en SCC-DFTB dans la référence [415] sont reportées. Dans la référence [415] les valeurs entre 1500 cm^{-1} et 1900 cm^{-1} ne sont pas reportées car elles sont de trop faibles intensités et occultées par les modes de vibration de l’eau dans cette région du spectre. Nous avons décidé de les reporter dans ce tableau en raison de leur forte intensité en TRPMD-GLE.

Mode de vibration	Exp. (cm^{-1})		Calc. (cm^{-1})			
	Ref. [415]	Ref. [425]	Ref. [415]	Ce travail		
			Harm.	MD	MD	TRPMD-GLE
Élongation CH (ν_{CH} , e_1u)	3106,7	3105,9				3025
	3099,9	3099,0				3018
	3077,3	3077,2				3011
	3067,3	3066,8				3007
	3051,8	3051,4	3053	3049	3053	3005
	3034,5	3034,3	3038	3037	3039	2998
	3029,3	3029,2		3024	3032	2993
						2980
Élongation CC (ν_{CC}) + déformation CH dans le plan (δ_{CH} , e_1u)					1846	1832
					1703	1691
					1645	1632
	1318,3	1318,3	1455	1575		
	1316,6	1317,4	1268	1465	1456	1371
				1268	1270	1271
Déformation CH dans le plan (δ_{CH} , e_1u)	1137,0	1137,0	1172	1174	1174	1179
Déformation CCH en dehors du plan (γ_{CH} , a_2u)	857,6	857,0	818	819	820	847-853
Déformation CCH dans le plan (γ_{CH} , e_1u)	842,7	846,1	802	805	805	802
Déformation CCH en dehors du plan (γ_{CH} , e_2u)	771,5	771,6	771	771	772	
Déformation CCC en dehors du plan (δ_{CCC} , a_2u)	549,9	550,5	519	521	521	532

les spectres TRPMD-GLE comme nous venons de le faire pour les spectres MD, c’est à dire discuter à partir des simulations TRPMD-GLE de l’effet de l’interaction avec H_2O sur les fréquences de vibration du coronène et de H_2O . Les différences entre les spectres MD et TRPMD-GLE seront discutées dans la section suivante.

Le mode de déformation CCC (a_2u) en dehors du plan est décalé vers le bleu de $+2\text{ cm}^{-1}$. Le mode de déformation CCH (e_2u) en dehors du plan apparaît à 772 cm^{-1} dans $(C_{24}H_{12})(H_2O)_\sigma$ mais est indistinguable du bruit de fond dans le spectre du coronène. Les deux modes de déformation CCH (e_1u et a_2u) entre 800 cm^{-1} et 880 cm^{-1} ne sont pas affectés par l’interaction avec H_2O . La présence de deux bandes dans le spectre du coronène est un artefact de la moyenne des dix spectres. Expérimentalement, ce mode est décalé vers le bleu lorsque suffisamment d’eau est ajouté pour des agrégats composés de plusieurs molécules puissent se former. La bande correspondant à la déformation CH dans le plan (e_1u) est décalée vers le rouge de -9 cm^{-1} . Les deux premiers modes ν_{CC} entre 1600 cm^{-1} et 1900 cm^{-1} sont

5. IMPACT DES EFFETS QUANTIQUES NUCLÉAIRES SUR LES SPECTRES IR DE SYSTÈMES $(\text{H}_2\text{O})_n/\text{Ar}$ ET $(\text{PAH})(\text{H}_2\text{O})/\text{Ar}$

décalés vers le rouge de -4 cm^{-1} et -1 cm^{-1} . La position du troisième mode n'est pas modifiée par la présence de H_2O . Cependant il n'est pas réellement possible de d'interpréter le décalage pour les deux premiers pics car l'amplitude du décalage est de l'ordre de la précision sur la convergence des spectres. Concernant le mode de cisaillement ν_2 de l'eau aucune bande n'est distinguable. En revanche, un épaulement diffus et de faible intensité entre 1500 cm^{-1} et 2000 cm^{-1} apparaît. Il convient de rappeler que pour H_2O seul en matrice d'argon, le mode de cisaillement donne lieu à un massif de pics entre 1550 cm^{-1} et 1750 cm^{-1} . Les modes ν_{CH} donnent également un massif de pics entre 2980 cm^{-1} et 3040 cm^{-1} pour $\text{C}_{24}\text{H}_{12}/\text{Ar}$ et $(\text{C}_{24}\text{H}_{12})(\text{H}_2\text{O})_\sigma/\text{Ar}$. Il semble y avoir un léger décalage vers le rouge du massif qui reste cependant difficile à quantifier car il n'est pas possible de distinguer un pic d'absorption maximum précis. Enfin, les modes ν_1 et ν_3 de l'eau donnent deux légers épaulements de très faible intensité étendus entre $\sim 3200 \text{ cm}^{-1}$ et $\sim 4000 \text{ cm}^{-1}$ et $\sim 4100 \text{ cm}^{-1}$ et $\sim 4900 \text{ cm}^{-1}$.

Comparaison des spectres MD et TRPMD-GLE et impact des NQEs

En regardant le tableau 5.2 et les spectres tracés en figure 5.4, on voit que l'inclusion des fluctuations quantiques des noyaux en TRPMD-GLE induit un décalage des modes de vibration de $\text{C}_{24}\text{H}_{12}/\text{Ar}$ et du complexe $(\text{C}_{24}\text{H}_{12})(\text{H}_2\text{O})_\sigma/\text{Ar}$. Ainsi, le mode de déformation CCC (δ_{CCC}) en dehors du plan est décalé vers le bleu de $+11 \text{ cm}^{-1}$ pour $\text{C}_{24}\text{H}_{12}/\text{Ar}$ et $(\text{C}_{24}\text{H}_{12})(\text{H}_2\text{O})_\sigma/\text{Ar}$. Comme dit précédemment, le mode de déformation CCH (γ_{CH}) en dehors du plan (e_2u), à 772 cm^{-1} dans $(\text{C}_{24}\text{H}_{12})(\text{H}_2\text{O})_\sigma/\text{Ar}$ mais est indistinguable du bruit de fond pour $\text{C}_{24}\text{H}_{12}/\text{Ar}$. Pour les deux modes de déformation CCH (γ_{CH}) entre 800 cm^{-1} et 880 cm^{-1} , l'attribution des bandes d'absorption est compliquée. Il semble que le mode de déformation CCH (γ_{CH} , a_2u) soit décalé vers $\sim 850 \text{ cm}^{-1}$ mais le mode de déformation CCH (γ_{CH} , e_1u) pourrait correspondre à la bande à 802 cm^{-1} auquel cas les NQEs auraient pour effet d'écarter ces deux bandes d'absorption ($11\text{-}15 \text{ cm}^{-1}$ expérimentalement, 16 cm^{-1} en harmonique, $14\text{-}15 \text{ cm}^{-1}$ en MD et $\sim 30 \text{ cm}^{-1}$ en TRPMD-GLE). Cependant il est possible que le mode de déformation CCH (γ_{CH} , e_1u) corresponde à l'épaulement sur la gauche de la bande (γ_{CH} , a_2u) entre 830 cm^{-1} et 840 cm^{-1} . Dans ce cas l'écart entre les deux bandes subirait le même décalage vers le bleu dû aux fluctuations quantiques des noyaux et l'écart entre les bandes serait également de l'ordre de $\sim 10\text{-}15 \text{ cm}^{-1}$. Les modes de déformation CH dans le plan (δ_{CH}) sont décalées

vers le bleu de $+5\text{ cm}^{-1}$ pour $\text{C}_{24}\text{H}_{12}/\text{Ar}$ comme pour $(\text{C}_{24}\text{H}_{12})(\text{H}_2\text{O})_\sigma/\text{Ar}$. Les fluctuations quantiques des noyaux décalent la bande de déformation CH dans le plan (δ_{CH}) initialement à 1270 cm^{-1} en MD vers le bleu de $+1\text{ cm}^{-1}$ pour $\text{C}_{24}\text{H}_{12}/\text{Ar}$ et $+3\text{ cm}^{-1}$ pour $(\text{C}_{24}\text{H}_{12})(\text{H}_2\text{O})_\sigma/\text{Ar}$. Les trois bandes entre 1600 cm^{-1} et 1900 cm^{-1} sont plus fortes en intensité pour $\text{C}_{24}\text{H}_{12}/\text{Ar}$ et $(\text{C}_{24}\text{H}_{12})(\text{H}_2\text{O})_\sigma/\text{Ar}$ et sont décalées vers le rouge de -13 cm^{-1} , -11 cm^{-1} et 13 cm^{-1} pour $\text{C}_{24}\text{H}_{12}/\text{Ar}$ et -20 cm^{-1} , -13 cm^{-1} et -14 cm^{-1} pour $(\text{C}_{24}\text{H}_{12})(\text{H}_2\text{O})_\sigma/\text{Ar}$. Concernant les deux bandes des modes ν_{CH} , les fluctuations quantiques des noyaux en TRPMD-GLE induisent un élargissement du signal et donnent un massif de pics entre 2980 cm^{-1} et 3040 cm^{-1} pour les deux systèmes. Dans [415], un massif similaire est observé en MD et expérimentalement. Globalement le massif est décalé vers le rouge de $\sim 35\text{ cm}^{-1}$ par rapport aux signaux obtenus en MD. Enfin, comme mentionné plus haut, aucune bande d'absorption correspondant aux modes ν_1 et ν_3 de l'eau est observable en TRPMD-GLE.

Conclusion

Les tendances relatives obtenues en MD sont similaires en TRPMD-GLE. En effet, les décalages induit par l'interaction de $\text{C}_{24}\text{H}_{12}$ avec une molécule d'eau en σ modifie qualitativement et quantitativement les positions des bandes d'absorption de $\text{C}_{24}\text{H}_{12}$ en MD et en TRPMD-GLE de la même façon. **L'inclusion des effets thermiques en MD semble donc suffisante pour rendre compte de ces décalages.** De plus la TRPMD-GLE engendre un signal bruité qui complique l'analyse des spectres et l'attribution des signaux. Enfin, les modes d'élongation ν_1 et ν_3 de l'eau dans $(\text{C}_{24}\text{H}_{12})(\text{H}_2\text{O})_\sigma/\text{Ar}$ sont fortement atténués en TRPMD-GLE.

5.3.3 Comparaison des spectres IR de $\text{C}_{20}\text{H}_{10}/\text{Ar}$ et $(\text{C}_{20}\text{H}_{10})(\text{H}_2\text{O})/\text{Ar}$ en MD et TRPMD-GLE

Les spectres MD et TRPMD-GLE de $\text{H}_2\text{O}/\text{Ar}$, $\text{C}_{20}\text{H}_{10}/\text{Ar}$ et des complexes $(\text{C}_{20}\text{H}_{10})(\text{H}_2\text{O})/\text{Ar}_{ext,int,side}$ sont représentés sur la figure 5.5. Les spectres expérimentaux réalisés par l'équipe de Joëlle Mascetti (ISM, Bordeaux) sont représentés sur la figure 5.6. Enfin, les modes de vibrations déterminés expérimentalement et calculées en SCC-DFTB à partir de simulations MD et TRPMD-GLE sont reportés

5. IMPACT DES EFFETS QUANTIQUES NUCLÉAIRES SUR LES SPECTRES IR DE SYSTÈMES $(\text{H}_2\text{O})_n/\text{Ar}$ ET $(\text{PAH})(\text{H}_2\text{O})/\text{Ar}$

dans le tableau 5.3. Dans un premier temps nous discuterons de l'impact des NQEs sur la description théorique des modes de vibrations de $\text{C}_{20}\text{H}_{10}/\text{Ar}$ en comparant les spectres MD et TRPMD-GLE. Le spectre expérimental nous servira de référence pour cette discussion. Ensuite, l'effet de l'interaction du corannulène avec une molécule d'eau sur les modes de vibrations de $\text{C}_{20}\text{H}_{10}$ sera commenté à partir de l'analyse des spectres MD et TRPMD-GLE. Là aussi, les résultats seront comparés aux valeurs expérimentales.

Nous verrons que contrairement au coronène, l'interaction avec une molécule d'eau n'a pas d'effet notable sur la positions des pics d'absorption IR du corannulène. Par ailleurs, comme pour le coronène, les effets thermiques décrit par la MD sont suffisant pour décrire les modes de vibrations en accord avec les résultats expérimentaux d'autant que les spectres issus de simulations TRPMD-GLE sont difficiles à interpréter en raison de l'élargissement des pics et du bruit.

Impact des NQEs sur les modes de vibration de $\text{C}_{20}\text{H}_{10}/\text{Ar}$.

Pour les bandes d'absorption jusqu'à 1600 cm^{-1} correspondant aux modes γ_{CH} et δ_{CH} , les valeurs issues de simulations TRPMD-GLE sont en meilleur accord avec les valeurs expérimentales que celles issues de simulation MD. Le massif de pics d'absorption se situant à 1600 cm^{-1} correspondant à un mode de couplage $\delta_{CC} + \nu_{CC}$ présent dans les spectres théoriques (voir figure 5.5) et expérimentaux (voir figure 5.6) n'a pas été reporté dans le tableau par les expérimentateurs car il se superpose au mode de cisaillement ν_2 de l'eau ce qui complique l'attribution des bandes dans cette région du spectre. Le spectre MD fait apparaître deux petites bandes de faible intensité à 1620 cm^{-1} et 1647 cm^{-1} . Dans le spectre TRPMD-GLE, ces bandes sont de plus forte intensité et se situent à 1606 cm^{-1} et 1633 cm^{-1} . Étonnamment, la bande à 1633 cm^{-1} coïncide exactement avec une bande du massif correspondant au mode ν_2 de $\text{H}_2\text{O}/\text{Ar}$. Concernant les modes d'élongations ν_{CH} , un massif entre $\sim 3030\text{ cm}^{-1}$ et $\sim 3070\text{ cm}^{-1}$ est observé expérimentalement. En MD, ce massif est décalé vers le rouge à $\sim 2955\text{-}3040\text{ cm}^{-1}$. La prise en compte des fluctuations quantiques des noyaux en TRPMD-GLE n'améliore pas la position de ce massif. Au contraire elle l'élargit et le décale vers le rouge à $\sim 2935\text{-}3005\text{ cm}^{-1}$. Enfin, les modes observés expérimentalement entre 3600 cm^{-1} et 3080 cm^{-1} correspondent à des traces d'eau présentes dans la matrice d'argon. Ces modes n'apparaissent ni dans les spectres MD ni dans les spectres TRPMD-GLE.

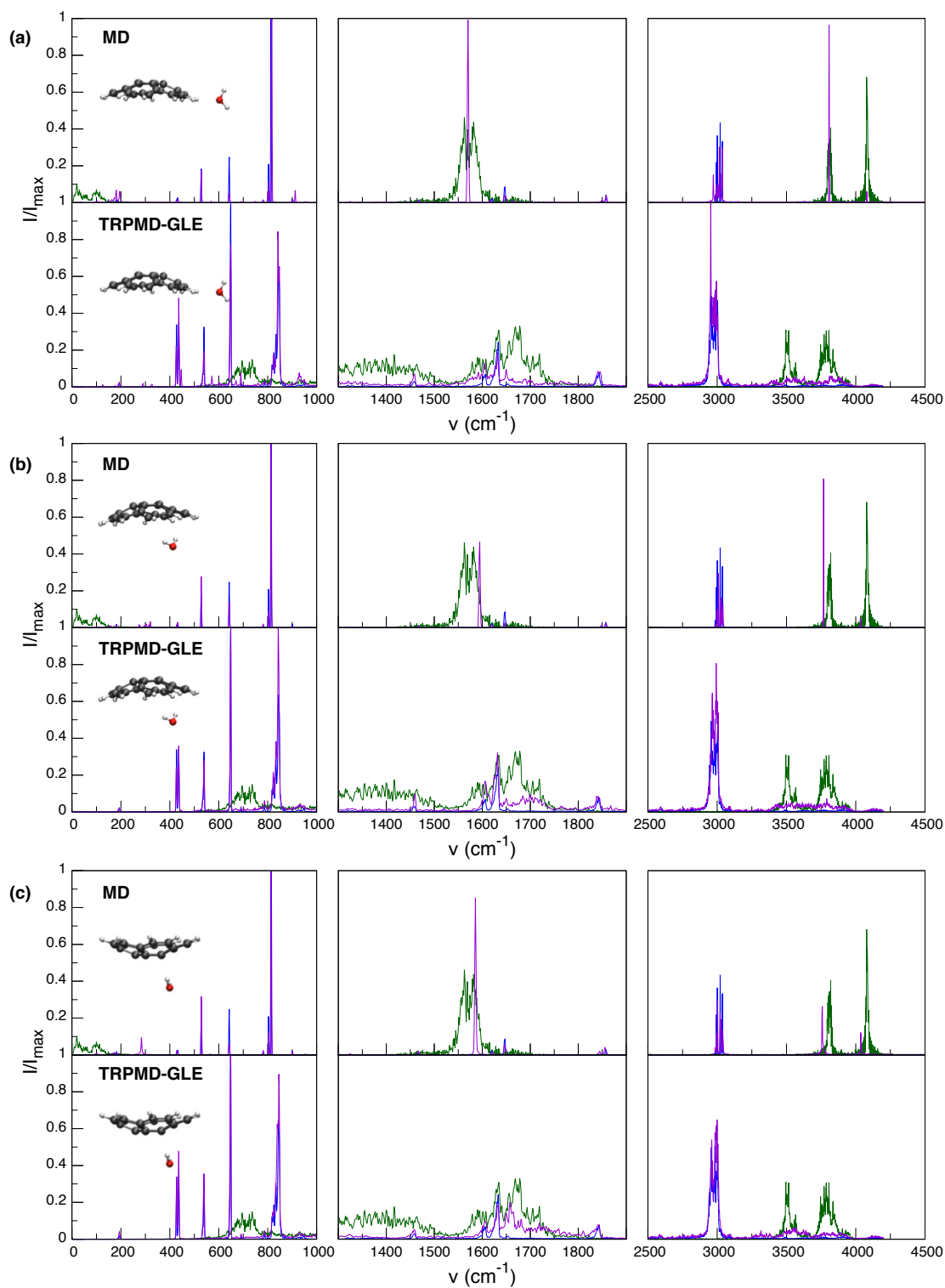


FIGURE 5.5 – Spectre d'absorption IR théoriques de $\text{H}_2\text{O}/\text{Ar}$ (vert), $\text{C}_{20}\text{H}_{10}/\text{Ar}$ (bleu) et $(\text{C}_{20}\text{H}_{10})(\text{H}_2\text{O})/\text{Ar}$ (violet) à 10 K. Le spectre de $\text{H}_2\text{O}/\text{Ar}$ a été redimensionné pour faciliter la visualisation.

5. IMPACT DES EFFETS QUANTIQUES NUCLÉAIRES SUR LES SPECTRES IR DE SYSTÈMES $(\text{H}_2\text{O})_n/\text{Ar}$ ET $(\text{PAH})(\text{H}_2\text{O})/\text{Ar}$

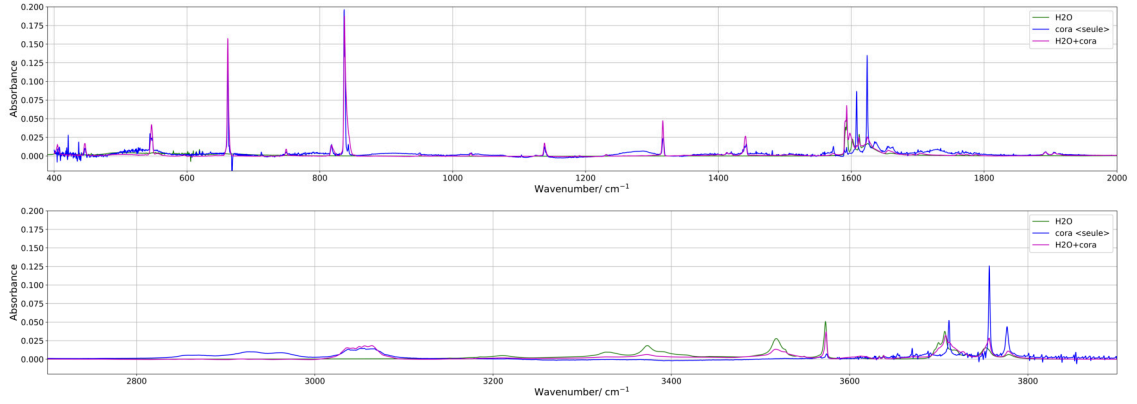


FIGURE 5.6 – Spectre d’absorption IR expérimental de $\text{H}_2\text{O}/\text{Ar}$ (vert), $\text{C}_{20}\text{H}_{10}/\text{Ar}$ (bleu) et des complexes (a) *side*, (b) *int*, (c) *ext* de $(\text{C}_{20}\text{H}_{10})(\text{H}_2\text{O})/\text{Ar}$ (violet) à 10 K. Le spectre de $\text{C}_{20}\text{H}_{10}/\text{Ar}$ présente des bandes correspondant aux modes de l’eau en raison de la présence de traces d’eau dans la matrice. Le spectre de $(\text{C}_{20}\text{H}_{10})(\text{H}_2\text{O})/\text{Ar}$ ne correspond pas au spectre de $\text{C}_{20}\text{H}_{10}$ en interaction avec une seule molécule d’eau comme dans nos simulations mais à $\text{C}_{20}\text{H}_{10}$ en interaction avec plusieurs agrégats de différentes tailles.

Impact de l’interaction avec H_2O sur les modes de vibrations de $\text{C}_{20}\text{H}_{10}/\text{Ar}$.

Expérimentalement, le plus grand décalage observé entre le spectre de $\text{C}_{20}\text{H}_{10}/\text{Ar}$ et $(\text{C}_{20}\text{H}_{10})(\text{H}_2\text{O})/\text{Ar}$ est un décalage vers le bleu de $+3\text{ cm}^{-1}$ pour la bande γ_{CH} initialement à 816 cm^{-1} pour $\text{C}_{20}\text{H}_{10}/\text{Ar}$. Cette absence de décalage des positions des bandes d’absorption est globalement reproduite par les simulations MD. On observe toutefois un léger décalage vers le bleu pour les modes d’élongation ν_{CH} . En TRPMD-GLE, l’élargissement du massif correspondant aux modes ν_{CH} et le bruitage qui multiplie les pics ne permet pas de dire si l’interaction avec H_2O induit un décalage de la positions des bandes d’absorption. Toutefois, si les positions des pics de plus haute intensité sont modifiées par l’interaction avec H_2O , la position moyenne du massif reste la même. L’interaction avec le corannulène dans la matrice d’argon affecte les modes de vibrations de H_2O de façon totalement différente en MD et en TRPMD-GLE. Nous discuterons d’abord de l’effet de l’interaction sur les modes de vibrations de l’eau en MD. Dans ce cas, l’interaction avec le corannulène amincit les bandes d’absorption de l’eau. Cela s’explique par le fait que la molécule d’eau est beaucoup moins mobile lorsqu’elle interagit avec le corannulène que lorsqu’elle est seule en matrice. Les positions des bandes sont également mo-

5.3 Résultats et discussion

TABLE 5.3 – Modes de vibration théoriques issus de simulations de MD et TRPMD-GLE et expérimentaux du corannulène seul et en présence d’eau en matrice d’argon. Théoriquement, seul les complexes avec le monomère d’eau dans les conformations *int*, *ext* et *side* (voir figure 5.2) sont considérées dans cette étude. Les valeurs expérimentales nous ont été transmises par l’équipe de Joëlle Mascetti (ISM, Bordeaux).

Mode de vibration	C ₂₀ H ₁₀ /Ar			(C ₂₀ H ₁₀)(H ₂ O)/Ar						
	Exp.	BOMD	TRPMD	Exp.	<i>int</i>	<i>ext</i>	<i>side</i>	<i>int</i>	<i>ext</i>	<i>side</i>
δ_{CCC}	548	529	540	548	529	529	529	540	539	539
	663	643	649	663	643	643	643	649	648	649
γ_{CH}	750	781	786	752	781	781	781	786	788	788
	816	803	817-825	819	804	804	807	822	825	824
	838	813	834-844	838	815	815	817	834-843	845	842
δ_{CH}	1142	1121	1124	1142	1121	1123	1121	1124	1125	
	1231	1179	1182	1233	1176	1178	1174	1179	1179	1182
					1271	1273	1191			
						1303	1273			
$\delta_{CH} + \nu_{CC}$	1319	1324	1320-1328	1319	1324	1324	1324			
	1418			1416						
	1443	1468	1459	1443	1465	1467	1468	1460	1459	1459
		1620	1606	1510	1496		1496			
	1572	1647	1633	1572	1595	1586	1569	1606	1606	1604
						1622	1617	1632	1633	1632
						1646	1646	1700	1650	1650
						1663	1663		1658	
						1677	1678			
	1892	1858	1844	1892	1849	1845	1850	1839	1810	1841
ν_{CH}	3034	2985	2957-2965	3034	2996-3001	2988	2971	2965	2961	2953
	3053	2994	2979-2984	3053	3008	2996	2992	2993	2991	2970-2976
	3067	3001	2996-3003	3067	3025	3003	3007		3000	2985
		3021			3033	3021-3026	3016			2994
		3038			3042	3041	3032-3037			3003
ν_{1-OH} et ν_{3-OH}				3575			3627		~3316	
				3709					~3554	
				3755	3768	3759	3812		~3828	
					4034	4036	4080			

difiées en MD. Le mode de cisaillement ν_2 est décalé vers le bleu de $+25 \text{ cm}^{-1}$ et $+16 \text{ cm}^{-1}$ pour les systèmes *int* et *ext* tandis qu’il n’est pas affecté dans le cas du système *side*. Les modes d’élongation symétriques (ν_1) et asymétriques (ν_3) sont également affectés. La bande d’absorption correspondant au mode ν_1 est décalée vers le rouge de -45 cm^{-1} , -54 cm^{-1} et -1 cm^{-1} pour les systèmes *int*, *ext* et *side*. La bande correspondant au mode ν_3 est décalé vers le rouge de -46 cm^{-1} , -44 cm^{-1} pour les systèmes *int* et *ext* mais il n’est pas affecté dans le cas du système *side*. À présent, si l’on se concentre sur les spectres TRPMD-GLE, l’effet de l’interaction avec le corannulène sur les modes de vibrations de l’eau est totalement différent. Les modes de librations de H₂O/Ar ont complètement disparus. Comme dans le cas

5. IMPACT DES EFFETS QUANTIQUES NUCLÉAIRES SUR LES SPECTRES IR DE SYSTÈMES $(\text{H}_2\text{O})_n/\text{Ar}$ ET $(\text{PAH})(\text{H}_2\text{O})/\text{Ar}$

du coronène, cela s'explique par la moindre mobilité de la molécule d'eau lorsqu'elle interagit avec le corannulène. Pour le mode de cisaillement ν_2 , l'élargissement des bandes complique l'analyse tant pour $\text{H}_2\text{O}/\text{Ar}$ que pour $(\text{C}_{24}\text{H}_{12})(\text{H}_2\text{O})/\text{Ar}$. De plus ce mode de vibration se superpose aux bandes des modes $\delta_{CH} + \nu_{CC}$ du corannulène à 1606 cm^{-1} et 1633 cm^{-1} . Pour les bandes des modes ν_1 et ν_3 , on observe le même phénomène que dans le cas du coronène. Les deux modes se fondent en un massif diffus et de très faible intensité entre $\sim 3375\text{ cm}^{-1}$ et $\sim 4000\text{ cm}^{-1}$ pour les systèmes *int* et *side* et entre $\sim 3250\text{ cm}^{-1}$ et $\sim 4000\text{ cm}^{-1}$ pour le système *ext*.

5.4 Conclusions et perspectives

Dans cette étude, les spectres d'absorption IR des systèmes $(\text{H}_2\text{O})_{1-4,6(bag)}/\text{Ar}$, $\text{C}_{24}\text{H}_{12}/\text{Ar}$, $(\text{C}_{24}\text{H}_{12})(\text{H}_2\text{O})/\text{Ar}$, $\text{C}_{20}\text{H}_{10}/\text{Ar}$ et $(\text{C}_{20}\text{H}_{10})(\text{H}_2\text{O})/\text{Ar}$ ont été calculés à partir de simulations de MD et TRPMD-GLE. Le calcul de l'énergie et des forces a été réalisé avec un potentiel SCC-DFTB implémenté dans le code deMonNano tandis que l'intégration des équations du mouvement a été réalisée avec le code i-PI. Ainsi, l'effet de l'interaction de deux PAHs (le coronène ($\text{C}_{24}\text{H}_{12}$) et le corannulène ($\text{C}_{20}\text{H}_{10}$)) avec une molécule d'eau sur les modes de vibrations des PAHs a pu être évalué et nos résultats ont systématiquement été comparés aux résultats expérimentaux. Par ailleurs, l'impact des NQEs sur la description des modes de vibrations de ces différents systèmes a également été étudié en comparant les spectres issus de simulations TRPMD-GLE à ceux issus de simulations MD. Selon le type de système étudié, la prise en compte des fluctuations quantiques des noyaux en TRPMD-GLE ne modifie pas les modes de vibrations de la même façon. Pour les **agrégats d'eau purs**, comme attendu, on observe un **décalage vers le rouge et un élargissement systématique des modes d'élongations O–H en TRPMD-GLE**. Cet élargissement des signaux empêche toute comparaison quantitative des spectres TRPMD-GLE aux spectres expérimentaux. Toutefois, **les positions approximatives des bandes d'absorption en TRPMD-GLE sont clairement plus proches des valeurs expérimentales que celles obtenues en MD**. Pour l'étude de l'interaction du coronène avec une molécule d'eau en matrice d'argon, les effets thermiques simulés en MD sont suffisant pour reproduire les observations expérimentales et l'inclusion des NQEs en

5.4 Conclusions et perspectives

TRPMD-GLE n’apporte pas d’information supplémentaire. Au contraire, les bandes sont plus larges et bruitées en TRPMD-GLE, ce qui complique l’attribution des bandes. De plus, la prise en compte des fluctuations quantiques des noyaux induit une atténuation des modes d’élongation ν_1 et ν_3 de l’eau qui n’est observée ni expérimentalement, ni en MD. Dans le cas du corannulène seul, les positions absolues des modes γ_{CH} et δ_{CH} en dessous de 1600 cm^{-1} obtenus en TRPMD-GLE sont plus proches des valeurs expérimentales que celles obtenues en MD. Cependant, comme pour le coronène, les bandes d’absorption sont élargies en TRPMD-GLE et les modes d’élongation ν_1 et ν_3 de l’eau sont atténués. Ainsi, la TRPMD-GLE ne semble pas être adaptée à l’étude des spectres d’absorption IR de ces systèmes. On retrouve dans la littérature ce constat mitigé concernant l’étude des spectres de vibration en TRPMD-GLE. En 2019, Q. Yu et J. Browman [426] ont constaté que la TRPMD-GLE permettait de décrire le décalage vers le rouge des modes d’élongation O–H dans les agrégats $(\text{H}_2\text{O})_3\text{H}^+$ et $(\text{H}_2\text{O})_4\text{H}^+$ à 100 K mais que l’élargissement des bandes d’absorption empêchait toute comparaison quantitative avec le spectre expérimental. Yair Litman *et al.* [427, 428] ont étudié le spectre d’absorption IR d’une molécule de porphycène en TRPMD-GLE et ont pu reproduire qualitativement le spectre expérimental en faisant apparaître une bande correspondant à une liaison N–H \cdots H qui n’apparaissait pas en MD. Plus récemment en 2021, V. S. Sandeep Inakollu et Haibo Yu ont étudié les modes de vibration de la sérine déprotonée en MD et TRPMD-GLE à 298 K et 100 K. Dans cette étude, les spectres MD comme TRPMD-GLE reproduisent relativement bien le spectre expérimental mais la prise en compte des NQEs en TRPMD-GLE n’apporte pas d’information supplémentaire.

Conclusions générales et perspectives

Conclusions générales et perspectives

Conclusions générales

Cette thèse a porté sur l'**exploration de paysages énergétiques complexes** et sur la prise en compte des **effets quantiques des noyaux** en dynamique moléculaire pour permettre l'étude de propriétés d'**agrégats moléculaires**, principalement des **agrégats d'eau**. Pour de tels systèmes, le calcul de propriétés dépendants de la température nécessite la prise en compte des fluctuations thermiques. Pour cela, il est nécessaire d'explorer les surfaces d'énergie potentielle de manière la plus ergodique possible afin d'échantillonner avec précision l'ensemble des configurations accessibles à une température donnée. Pour des systèmes caractérisés par de nombreux degrés de liberté comme les agrégats d'eau, l'exploration ergodique de la PES par des méthodes standards comme la dynamique moléculaire peut s'avérer difficile voire impossible. Des méthodes d'échantillonnage accéléré telles que la méthode *Parallel-Tempering Molecular Dynamics* permettent d'accélérer cette exploration. Par ailleurs, une description précise des agrégats requiert la prise en compte des effets quantiques des noyaux. Pour cela, l'approche *Path-Integral Molecular Dynamics* basée sur le formalisme des intégrales de chemin permet de traiter les NQEs en dynamique moléculaire tout en restant dans un formalisme classique.

Pour pouvoir appliquer ces méthodes à l'étude d'agrégats moléculaires, **nous avons réalisé plusieurs implémentations originales dans deux programmes de chimie théorique** : deMonNano et deMon2k. Ces implémentations nous ont permis d'étudier la **transition de phase dans les agrégats $(\text{H}_2\text{O})_6$ et $(\text{H}_2\text{O})_8$** a des niveaux de théorie jamais atteints précédemment. Nous avons également pu étudier l'impact des NQEs sur les propriétés physico-chimiques de plusieurs systèmes moléculaires. Ainsi nous avons étudié l'impact des NQEs sur les **déplacements**

CONCLUSIONS GÉNÉRALES ET PERSPECTIVES

chimiques ^{13}C de carboxylates linéaires, sur la distribution d'isomères à basse température dans les agrégats $(\text{H}_2\text{O})_6$ et $(\text{H}_2\text{O})_8$ et sur les **spectres d'absorption infrarouge** dans des agrégats d'eau et des complexes PAH/ H_2O solvatés par une matrice d'argon.

Ces travaux ont été présentés en 4 chapitres dont les principaux résultats sont les suivants :

- **Implémentation de l'algorithme PTMD dans deMon2k et application à la transition de phase dans les agrégats $(\text{H}_2\text{O})_6$ et $(\text{H}_2\text{O})_8$.** Une part importante de ce travail de thèse a porté sur l'implémentation de l'algorithme PTMD dans le code deMon2k. À notre connaissance, c'est la première fois que cette méthode est implémentée nativement dans un code de DFT utilisant une combinaison linéaire d'orbitales gaussiennes pour la résolution des équations de Kohn-Sham. Pour pouvoir efficacement paralléliser l'algorithme à l'aide de bibliothèques MPI sans avoir à modifier la structure parallèle de deMon2k dédiée au calcul de la structure électronique, **nous avons développé une routine qui permet de créer et de manipuler des groupes et sous-groupes de processus MPI.** En plus de la version standard de l'algorithme PTMD réalisant des **échanges de répliques de manière synchrone**, nous avons également implémenté une version réalisant des **échanges asynchrones**. Cette implémentation a été appliquée à l'étude de la transition de phase dans les agrégats $(\text{H}_2\text{O})_6$ et $(\text{H}_2\text{O})_8$. À notre connaissance, **c'est la première fois que de telles études sont réalisées à ce niveau de théorie.** Afin d'évaluer l'influence du potentiel sur la transition de phase de l'**hexamère d'eau $(\text{H}_2\text{O})_6$** , nous avons comparé les résultats obtenus en ADFT avec la fonctionnelle BLYP et en SCC-DFTB. Si les deux méthodes attestent de la coexistence de nombreux isomères à basse températures, les distributions de ces isomères en fonction de la température sont différentes pour ces deux potentiels. De plus, les courbes de capacité calorifique sont très différentes. **La courbe issue de simulations SCC-DFTB présente un pic bien défini suggérant une transition de phase du premier ordre tandis qu'en DFT, avec la fonctionnelle BLYP, la courbe de capacité calorifique suggère une transition de phase plus continue et donc plus proche d'une transition de phase du second ordre.** Ces différences dans

les courbes de capacité calorifique entre les deux méthodes ont été rationalisées en terme d'énergie de liaison des isomères de basse énergie. Concernant l'**octamère d'eau (H₂O)₈**, l'étude de la distribution d'isomère et de la courbe de capacité calorifique a été réalisée en SCC-DFTB ainsi qu'en DFT avec les fonctionnelles BLYP et PBE0. La température de transition de phase obtenue au niveau SCC-DFTB est beaucoup plus faible qu'en DFT, que ce soit avec la fonctionnelle BLYP ou PBE0. Finalement, les résultats sur (H₂O)₆ obtenus à l'aide des implémentations synchrones et asynchrones ont été comparés. Les résultats démontrent que pour (H₂O)₆, **notre implémentation synchrone est caractérisée par une bonne efficacité parallèle**. Cela rend cette implémentation tout à fait utilisable sur des supercalculateurs.

- **Étude de l'impact des NQEs sur les déplacements chimiques ¹³C de carboxylates linéaires.** Pour réaliser cette étude, nous avons **implémenté une interface entre le programme deMon2k et un programme dédié à la dynamique moléculaire dans le formalisme PIMD, i-PI**. Nous avons ensuite réalisé de nombreuses simulations afin de valider l'utilisation d'un potentiel QM/MM pour décrire des carboxylates linéaires en interaction avec un agrégat d'eau. Les carboxylates ont été traités par un potentiel ADFT avec la fonctionnelle BLYP et l'agrégat composé de 34 molécules d'eau a été traité par un champ de force. Le type de thermostat et le nombre de répliques à inclure dans le polymère cyclique afin de converger les NQEs a également été testé et validé. Les résultats montrent que les fluctuations quantiques des atomes d'hydrogènes des carboxylates autour de leur position d'équilibre ont pour effet de systématiquement diminuer les déplacements chimiques ¹³C. De plus, **la prise en compte des NQEs améliore systématiquement la comparaison avec les données expérimentales** lorsque l'on considère les déplacements chimiques ¹³C relatifs au sein d'une même molécule.
- **Étude de l'impact des NQEs sur les distributions d'isomères de (H₂O)₆ et (H₂O)₈.** Pour réaliser cette étude, nous avons implémenté une interface entre les codes deMonNano et i-PI. Nous avons ainsi pu réaliser des simulations PTPIMD à partir desquelles nous avons calculé les distributions d'isomères de (H₂O)₆ et (H₂O)₈ en fonction de la température et en tenant compte des NQEs. Les études de ce type sont extrêmement rares dans la littérature. **Ce travail et la démonstration que de telles simulations**

CONCLUSIONS GÉNÉRALES ET PERSPECTIVES

peuvent être convergées correctement grâce à l'utilisation d'un potentiel SCC-DFTB ouvrent ainsi la voie à de futures études originales. Nos simulations démontrent que pour $(\text{H}_2\text{O})_6$, l'isomérisation des isomères *prism* en *book* débute 20 K plus bas lorsque l'on prend en compte les NQEs et les distributions d'isomères se retrouvent fortement perturbées sur une large gamme de températures. Dans le cas de $(\text{H}_2\text{O})_8$, les NQEs n'ont qu'un très léger impact sur les distributions d'isomères. Cela est attribué à la grande stabilité de l'isomère de plus basse énergie comparé aux autres isomères.

- **Étude de l'impact des NQEs sur les spectres d'absorption IR d'agrégats d'eau, de PAHs et de complexes PAH/eau solvatés par une matrice d'argon.** Ces simulations ont également été réalisées à l'aide de l'interface entre les codes deMonNano et i-PI. Nous avons repris dans ce chapitre les travaux réalisés précédemment concernant l'étude théorique de spectres d'absorption IR d'agrégats d'eau, de coronène $\text{C}_{24}\text{H}_{12}$ et de complexes $(\text{C}_{24}\text{H}_{12})(\text{H}_2\text{O})$ solvatés dans des matrices d'argon pour tenter d'y inclure les NQEs. Nous avons également étudié le corannulène, $\text{C}_{20}\text{H}_{10}$, en matrice d'argon ainsi que son complexe avec une molécule d'eau dans trois conformations différentes. Les résultats montrent que la **prise en compte des NQEs a un impact important sur les modes d'élongation O–H** dans le cas des agrégats d'eau en matrice. Les bandes d'absorptions correspondant à ces modes sont fortement décalées vers le rouge. Dans le cas des PAH et des complexes PAH/eau, la prise en compte des NQEs apporte peu d'information supplémentaire et l'élargissement des bandes d'absorption complique l'attribution des bandes. De plus, dans les spectres d'absorption des complexes PAH/eau issus de simulations incluant les NQEs, les modes de vibrations O–H sont fortement amortis.

Perspectives

Ce travail de thèse offre de nombreuses perspectives pour des travaux futurs en particulier grâce aux **implémentations originales** qui ont été réalisées. L'implémentation que nous avons proposée de l'algorithme PTMD dans le code deMon2k a en effet pour vocation à être utilisée par la communauté des chimistes théoriciens

depuis son intégration officielle à la version 6.0.8 du code. L'absence de code DFT intégrant cet algorithme, la puissance des supercalculateurs actuels, et la large gamme d'applications potentielles nous laissent penser qu'il sera largement utilisé dans les années à venir.

Il est important de noter que l'implémentation de la structure MPI en deux couches que nous avons réalisée pourra à l'avenir **servir à paralléliser de nombreuses autres fonctionnalités du code deMon2k**. Le calcul des fréquences de vibration harmoniques, le calcul des couplage spin-spin ainsi que l'implémentation native de l'algorithme PIMD dans le code deMon2k sont autant d'exemples qui pourraient à l'avenir tirer parti de notre travail. Plus personnellement, s'il m'était donné l'occasion de continuer à travailler sur la physico-chimie des agrégats, l'étude de deux types de systèmes chimiques pourrait être envisagée grâce à cette approche : des **agrégats d'aluminium** et des **agrégats d'eau protonés et déprotonés**. Dans le premier cas, il s'agit d'une étude que j'ai déjà démarrée en collaboration avec Andreas M. Köster sur la transition de phase dans les agrégats Al_{27}^+ et Al_{28}^+ . Dans le second cas, les agrégats d'eau protonés et déprotonés sont très mal décrits au niveau SCC-DFTB et leur description au niveau DFT avec fonctionnelle hybride pourrait permettre de revisiter une grande variété de propriétés de ces systèmes à un niveau de précision jamais atteint jusqu'alors.

Les interfaces réalisées entre le code i-PI et les codes deMon2k et deMonNano pourront également être mises à profit dans les années à venir. Nous travaillons actuellement à **l'intégration officielle de l'interface deMon2k/i-PI dans le code deMon2k**, là encore, pour la mettre à disposition de la communauté d'utilisateurs la plus large possible. Au vu de l'importance des effets quantiques nucléaires sur les systèmes contenant des atomes d'hydrogène, il y a fort à parier que cette interface pourra donner lieu à de nombreuses applications. Finalement, l'interface deMonNano/i-PI pourra être utilisée pour l'étude de nouveaux systèmes au niveau SCC-DFTB. Une direction envisagée serait d'utiliser cette interface pour reparamétriser les potentiels SCC-DFTB O–H à l'aide de données expérimentales sur l'eau liquide mais en prenant en compte les effets quantiques nucléaires lors des comparaisons entre théorie et expérience. Ceci permettrait de **développer des paramètres SCC-DFTB spécifiques aux systèmes aqueux sans engendrer un double comptage des effets quantiques nucléaires**.

Annexes

Annexes

- Script de soumission de job i-PI/deMonNano sur un nœud avec socket UNIX.
- Scripts de soumission de job i-PI/deMonNano sur plusieurs nœuds (ou postes) avec une socket INET

ANNEXES

```
#####
# Script a executer sur un noeud de calcul #
#####

#!/bin/bash
#SBATCH -J JOBNAME                # Nom du job slurm
#SBATCH -N 1                      # Nb noeuds
#SBATCH -t WALLTIME              # Temps horloge
#SBATCH -p PARTITION             # Partition
#SBATCH --cpus-per-task=CPUPERTASK # Nb coeurs par instance deMonNano
#SBATCH --ntasks-per-node=NTASKPERNODE # 1+N avec N nb instances deMonNano

# Creation du dossier IPI
mkdir IPI

# Creation du dossier pour N instances de deMonNano
mkdir DEMON

# Copie de la geometrie initiale dans IPI
cp init.xyz IPI

# Copie de l'input i-PI dans IPI
cp input.xml IPI

# Exection de i-PI
cd IPI
PATHTOIPI/i-pi input.xml </dev/null &>log &
cd ../

# Sleep pour s'assurer que i-PI ait eu le temps d'ouvrir la socket
sleep 10

# Boucle sur les N instances de deMonNano
for i in {1..NTASKPERNODE-1} ; do

    # Creation du dossier de l'instance i de deMonNano
    mkdir DEMON/${i}

    # Copie de l'input deMonNano dans i
    cp deMon.inp DEMON/${i}

    cd DEMON/${i}

    if [ ${i} -lt NTASKPERNODE ] ; then
        # Execution de deMonNano
        PATHTODEMON/deMon.x &
    else
        # Execution de deMonNano
        PATHTODEMON/deMon.x
    fi

    cd ../../

done

wait
```



```
#####
# Script a executer sur un poste de travail local #
# ou sur le noeud principal d'un centre de calcul #
#####

# Creation du dossier IPI
mkdir IPI

# Copie de l'input i-PI dans IPI
cp ../input.xml IPI

cd IPI

# Attribue l'adresse IP du noeud a ADRESS
sed -i "s/ADRESS/$HOSTNAME/g" input.xml

# Execution de i-PI
nohup PATHTOIPI/i-pi input.xml </dev/null &>log &

cd ..

#####
# Script a executer sur un ou plusieurs noeuds de calcul #
#####

#!/bin/bash
#SBATCH -J JOBNAME                # Nom du job slurm
#SBATCH -N 1                      # Nb noeuds
#SBATCH -t WALLTIME               # Temps horloge
#SBATCH -p PARTITION              # Partition
#SBATCH --cpus-per-task=CPUPERTASK # Nb coeurs par instance deMonNano
#SBATCH --ntasks-per-node=NTASKPERNODE # N instances deMonNano

# Cherche l'adresse IP du noeud de calcul
# A adapter en fonction du centre de calcul
myaddress=$(ifconfig | grep -m 3 inet | awk ' END {print $2} ')

# Creation du tunnel ssh permanent
ssh -f -N -L 10000:IPADRESSIPI:65535 -2 ${myaddress}

# Creation du dossier pour N instances de deMonNano
mkdir DEMON
cd DEMON

# Boucle sur les N instances de deMonNano
for i in {1..NTASKPERNODE}; do

    # Creation du dossier de l'instance i de deMonNano
    mkdir ${i}

    # Copie de l'input deMonNano dans i
    cp deMon.inp ${i}

    # Attribue l'adresse IP du noeud qui execute IPI a ADRESS
    sed -i "s/ADRESS/IPADRESSIPI/g" ${i}/deMon.inp

    cd ${i}

    if [ ${i} -lt NTASKPERNODE ] ; then
        # Execution de deMonNano
        PATHTODEMON/deMon.x &
    else
        # Execution de deMonNano
        PATHTODEMON/deMon.x
    fi

    cd ../

done

cd ../

wait
```

Ce type de calcul s'exécute en deux temps. Dans un premier il faut exécuter le script de soumission du job i-PI sur le nœud principal (frontale de connexion) ou sur un poste local. Puis dans un deuxième temps, il faut modifier la valeur de IPADRESSIPI dans le script de soumission du code client (ici deMonNano) et exécuter le script de soumission du code client sur un (ou plusieurs) nœud(s) de calcul.

Bibliographie

Bibliographie

- [1] D. Kashchiev. *Nucleation*. Elsevier, **2000**.
- [2] F. A. Cotton. *Inorg. Chem.*, **1964**, 1217–1220.
- [3] W. Knight, K. Clemenger, W. A. de Heer, W. A. Saunders, M. Chou, M. L. Cohen. *Phys. Rev. Lett.*, **1984**, 2141.
- [4] G. Ganteför, M. Gausa, K.-H. Meiwes-Broer, H. O. Lutz. *Z. Phys. D*, **1988**, 253–261.
- [5] B. K. Teo, X. Shi, H. Zhang. *J. Am. Chem. Soc.*, **1992**, 2743–2745.
- [6] H. Buchenau, E. Knuth, J. Northby, J. Toennies, C. Winkler. *J. Chem. Phys.*, **1990**, 6875–6889.
- [7] J. Farges, M. De Feraudy, B. Raoult, G. Torchet. *J. Chem. Phys.*, **1983**, 5067–5080.
- [8] J. Farges, M. De Feraudy, B. Raoult, G. Torchet. *J. Chem. Phys.*, **1986**, 3491–3501.
- [9] J. Stapelfeldt, J. Wörmer, T. Möller. *Phys. Rev. Lett.*, **1989**, 98.
- [10] O. Echt, K. Sattler, E. Recknagel. *Phys. Rev. Lett.*, **1981**, 1121.
- [11] O. Cheshnovsky, S. Yang, C. Pettiette, M. Craycraft, Y. Liu, R. Smalley. *Chemical Phys. Lett.*, **1987**, 119–124.
- [12] E. A. Rohlfing, D. M. Cox, A. Kaldor. *J. Chem. Phys.*, **1984**, 3322–3330.
- [13] P. J. Dyson, J. S. McIndoe. CRC Press, **2018**.
- [14] U. Bodensieck, A. Meister, G. Meister, G. Rheinwald, H. Stoeckli-Evans, G. Siiss-Fink. *Chimia*, **1993**, 189–191.
- [15] F. Cotton, J. Mague. *Inorg. Chem.*, **1964**, 1402–1407.
- [16] A. W. Maverick, J. S. Najdzonek, D. MacKenzie, D. G. Nocera, H. B. Gray. *J. Am. Chem. Soc.*, **1983**, 1878–1882.
- [17] J. Plešek. *Chem. Rev.*, **1992**, 269–278.
- [18] H. He, C. Tyson, S. Bobev. *Crystals*, **2011**, 87–98.
- [19] J. Sangster. *J. Ph. Equilibria Diffus.*, **2006**, 190.

BIBLIOGRAPHIE

- [20] S. Stegmaier, T. F. Fässler. *J. Am. Chem. Soc.*, **2011**, 19758–19768.
- [21] N. Toshima, T. Yonezawa, K. Kushihashi. *J. Chem. Soc. Faraday Tran*, **1993**, 2537–2543.
- [22] K. Liu, J. D. Cruzan, R. J. Saykally. *Science*, **1996**, 929–933.
- [23] V. Stert, W. Radloff, C. Schulz, I. Hertel. *Eur. Phys. J. D*, **1999**, 97–106.
- [24] G. Torchet, H. Bouchier, J. Farges, M. De Feraudy, B. Raoult. *J. Chem. Phys.*, **1984**, 2137–2143.
- [25] M. J. Rusyniak, Y. M. Ibrahim, D. L. Wright, S. N. Khanna, M. S. El-Shall. *J. Am. Chem. Soc.*, **2003**, 12001–12013.
- [26] J. Sinfelt. *J. Cat.*, **1973**, 308–315.
- [27] D. K. Boehme, H. Schwarz. *Angew. Chem. Int. Ed.*, **2005**, 2336–2354.
- [28] J. H. Kwak, L. Kovarik, J. Szanyi. *ACS Catal.*, **2013**, 2449–2455.
- [29] X.-L. Ma, J.-C. Liu, H. Xiao, J. Li. *J. Am. Chem. Soc.*, **2018**, 46–49.
- [30] K. Bhattacharyya, S. Mukherjee. *Bull. Chem. Soc. Jap.*, **2018**, 447–454.
- [31] G. Tang, J. He, J. Liu, X. Yan, K. Fan. *Exploration*, **2021**, 75–89.
- [32] J. K. Gregory, D. C. Clary, K. Liu, M. G. Brown, R. J. Saykally. *Science*, **1997**, 814–817.
- [33] F. N. Keutsch, R. J. Saykally. *Proc. Natl. Acad. Sci. USA*, **2001**, 10533–10540.
- [34] T. Kaneko, T. Akimoto, K. Yasuoka, A. Mitsutake, X. C. Zeng. *J. Chem. Theory Comput.*, **2011**, 3083–3087.
- [35] K. Liu, M. G. Brown, C. Carter, R. J. Saykally, J. K. Gregory, D. C. Clary. *Nature*, **1996**, 501–503.
- [36] K. Liu, M. G. Brown, R. J. Saykally. *J. Phys. Chem. A*, **1997**, 8995–9010.
- [37] K. Nauta, R. E. Miller. *Science*, **2000**, 293–295.
- [38] C. J. Burnham, S. S. Xantheas, M. A. Miller, B. E. Applegate, R. E. Miller. *J. Chem. Phys.*, **2002**, 1109–1122.
- [39] E. G. Diken, W. H. Robertson, M. A. Johnson. *J. Phys. Chem. A*, **2003**, 64–68.
- [40] C. Steinbach, P. Andersson, M. Melzer, J. K. Kazimirski, U. Buck, V. Buch. *Phys. Chem. Chem. Phys.*, **2004**, 3320–3324.
- [41] P. Sennikov, S. Ignatov, O. Schrems. *ChemPhysChem*, **2005**, 392–412.
- [42] V. Vaida. *J. Chem. Phys.*, **2011**, 020901.
- [43] M. Fárník, J. Fedor, J. Kočíšek, J. Lengyel, E. Pluhařová, V. Poterya, A. Pysanenko. *Phys. Chem. Chem. Phys.*, **2021**, 3195–3213.

-
- [44] C. Kolb, J. Jayne, D. Worsnop, M. Molina, R. Meads, A. Viggiano. *J. Am. Chem. Soc.*, **1994**, 10314–10315.
- [45] W. Duley. *Astrophys. J.*, **1996**, L57.
- [46] A. Rimola, M. Sodupe, P. Ugliengo. *Astrophys. J.*, **2012**, 24.
- [47] E. F. Van Dishoeck, E. Herbst, D. A. Neufeld. *Chem. Rev.*, **2013**, 9043–9085.
- [48] A. Rimola, D. Skouteris, N. Balucani, C. Ceccarelli, J. Enrique-Romero, V. Taquet, P. Ugliengo. *ACS Earth Space Chem.*, **2018**, 720–734.
- [49] I. Braud. Ph.D. thesis, Université Paul-Sabatier, **2016**.
- [50] S. Zamith, P. Feiden, P. Labastie, J. M. L’Hermite. *J. Chem. Phys.*, **2010**, 154305.
- [51] S. Zamith, P. Feiden, P. Labastie, J. M. L’Hermite. *Phys. Rev. Lett.*, **2014**.
- [52] S. Zamith, P. Feiden, P. Labastie, J. M. L’Hermite. *Phys. Rev. Lett.*, **2010**, 103401.
- [53] S. Zamith, P. Labastie, J. M. L’Hermite. *J. Chem. Phys.*, **2012**, 214301.
- [54] S. Zamith, P. Labastie, J. M. L’Hermite. *J. Chem. Phys.*, **2013**, 034304.
- [55] I. Braud, J. Boulon, S. Zamith, J. M. L’Hermite. *J. Phys. Chem. A*, **2015**, 6017–6023.
- [56] G. Bruni. Ph.D. thesis, Université Claude Bernard - Lyon I, **2010**.
- [57] S. Legendre. Ph.D. thesis, Université de Caen, Basse-Normandie, **2006**.
- [58] R. Maisonnay. Ph.D. thesis, Université de Caen, Basse-Normandie, **2011**.
- [59] K. Korchagina, A. Simon, M. Rapacioli, F. Spiegelman, J. M. L’Hermite, I. Braud, S. Zamith, J. Cuny. *Phys. Chem. Chem. Phys.*, **2017**, 27288–27298.
- [60] I. Braud, S. Zamith, J. Cuny, L. Zheng, J.-M. L’Hermite. *J. Chem. Phys.*, **2019**, 014303.
- [61] S. Zamith, M.-C. Ji, J.-M. L’Hermite, C. Joblin, L. Dontot, M. Rapacioli, F. Spiegelman. *J. Chem. Phys.*, **2019**, 194303.
- [62] L. Dontot, F. Spiegelman, S. Zamith, M. Rapacioli. *Eur. Phys. J. D*, **2020**, 1–9.
- [63] S. Zamith, J.-M. L’hermite, L. Dontot, L. Zheng, M. Rapacioli, F. Spiegelman, C. Joblin. *J. Chem. Phys.*, **2020**, 054311.
- [64] L. Zheng, J. Cuny, S. Zamith, J.-M. L’Hermite, M. Rapacioli. *Phys. Chem. Chem. Phys.*, **2021**.
- [65] L. Zheng, S. Zamith, M. Rapacioli. *Theor. Chem. Acc.*, **2021**, 1–14.
- [66] K. F. Jensen, D. G. Truhlar. ACS Publications, **1987**.
- [67] J. Almlöf, D. G. Truhlar, T. P. Lybrand. *Interdiscip. Sci. Rev.*, **1990**, 252–263.
- [68] C. P. Sosa, J. Ochterski, J. Carpenter, M. J. Frisch. *J. Comput. Chem.*, **1998**, 1053–1063.

BIBLIOGRAPHIE

- [69] E. Bohm, A. Bhatele, L. V. Kale, M. E. Tuckerman, S. Kumar, J. A. Gunnels, G. J. Martyna. *IBM J. Res. Dev.*, **2008**, 159–175.
- [70] A. Laganà. Springer Science & Business Media, **2012**.
- [71] M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, E. Lindah. *Softwarex*, **2015**, 19–25.
- [72] M. Elstner, D. Porezag, G. Jungnickel, J. Elsner, M. Haugk, T. Frauenheim. *Phys. Rev. B Condens. Matter Mater. Phys.*, **1998**, 7260–7268.
- [73] E. J. Baerends, D. Ellis, P. Ros. *Chem. Phys.*, **1973**, 41–51.
- [74] H. Sambe, R. H. Felton. *J. Chem. Phys.*, **1975**, 1122–1126.
- [75] A. C. Van Duin, S. Dasgupta, F. Lorant, W. A. Goddard. *J. Phys. Chem. A*, **2001**, 9396–9409.
- [76] S. Kirkpatrick, C. D. Gelatt, M. P. Vecchi. *science*, **1983**, 671–680.
- [77] P. J. Van Laarhoven, E. H. Aarts. In *Simulated annealing : Theory and applications*. Springer, **1987**. 7–15.
- [78] J. H. Holland, et al. MIT press, **1992**.
- [79] D. B. Hibbert. *Chemometr. Intell. Lab.*, **1993**, 277–293.
- [80] K. D. Harris, R. L. Johnston, S. Habershon. In *Applications of Evolutionary Computation in Chemistry*. Springer, **2004**. 55–94.
- [81] P. Sibani, J. C. Schön, P. Salamon, J.-O. Andersson. *EPL (EuroPhys. Lett.)*, **1993**, 479.
- [82] J. C. Schön, M. Jansen. *Angew. Chem. Int. Ed. in English*, **1996**, 1286–1304.
- [83] T. Hansson, C. Oostenbrink, W. van Gunsteren. *Curr. Opin. Struct. Biol.*, **2002**, 190–196.
- [84] B. L. Hammond, W. A. Lester, P. J. Reynolds. World Scientific, **1994**.
- [85] B. Cheng, E. A. Engel, J. Behler, C. Dellago, M. Ceriotti. *Proc. Natl. Acad. Sci. USA*, **2019**, 1110–1115.
- [86] A. Eltareb, G. E. Lopez, N. Giovambattista. *Phys. Chem. Chem. Phys.*, **2021**, 6914–6928.
- [87] R. Ramírez, C. Herrero. *Phys. Rev. B*, **2011**, 064130.
- [88] C. Vega, M. M. Conde, C. McBride, J. L. F. Abascal, E. G. Noya, R. Ramírez, L. M. Sesé. *J. Chem. Phys.*, **2010**, 046101.
- [89] M. Ceriotti, J. Cuny, M. Parrinello, D. E. Manolopoulos. *Proc. Natl. Acad. Sci. USA*, **2013**, 15591–15596.

- [90] M. H. Beck, A. Jäckle, G. A. Worth, H.-D. Meyer. *Phys. Rep.*, **2000**, 1–105.
- [91] R. P. Feynman, A. R. Hibbs. McGraw-Hill, **1965**.
- [92] L. B. Skinner, C. Benmore, J. C. Neufeind, J. B. Parise. *J. Chem. Phys.*, **2014**, 214507.
- [93] A. Soper. *Chem. Phys.*, **2000**, 121–137.
- [94] J. L. Abascal, C. Vega. *J. Chem. Phys.*, **2005**, 234505.
- [95] H. Pathak, A. Späh, N. Esmaeildoost, J. A. Sellberg, K. H. Kim, F. Perakis, K. Amann-Winkel, M. Ladd-Parada, J. Koliyadu, T. J. Lane, et al. *Proc. Natl. Acad. Sci. USA*, **2021**.
- [96] C. Angell, W. Sichina, M. Oguni. *J. Phys. Chem.*, **1982**, 998–1002.
- [97] *deMonNano*. <http://demon-nano.ups-tlse.fr/>.
- [98] G. Geudtner, P. Calaminici, J. Carmona-Espíndola, J. M. del Campo, V. D. Domínguez-Soria, R. F. Moreno, G. U. Gamboa, A. Goursot, A. M. Köster, J. U. Revelles, T. Mineva, J. M. Vásquez-Pérez, A. Vela, B. Zúñiga-Gutierrez, D. R. Salahub. *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, **2012**, 548–555.
- [99] V. Kapil, M. Rossi, O. Marsalek, R. Petraglia, Y. Litman, T. Spura, B. Cheng, A. Cuzzocrea, R. H. Meißner, D. M. Wilkins, B. A. Helfrecht, P. Juda, S. P. Bienvenue, W. Fang, J. Kessler, I. Poltavsky, S. Vandenbrande, J. Wieme, C. Corminboeuf, T. D. Kühne, D. E. Manolopoulos, T. E. Markland, J. O. Richardson, A. Tkatchenko, G. A. Tribello, V. Van Speybroeck, M. Ceriotti. *Comput. Phys. Commun.*, **2019**, 214–223.
- [100] E. Schrödinger. *Ann. Phys. (Berl.)*, **1926**, 437–490.
- [101] M. Born, R. Oppenheimer. *Ann. Phys. (Berl.)*, **1927**, 457–484.
- [102] D. R. Hartree. In *Mathematical Proceedings of the Cambridge Philosophical Society*, 1. Cambridge university press, 89–110.
- [103] J. C. Slater. *Phys. Rev.*, **1929**, 1293.
- [104] V. Fock. *Z. Phys.*, **1930**, 126–148.
- [105] C. C. J. Roothaan. *Rev. Mod. Phys.*, **1951**, 69.
- [106] C. Møller, M. S. Plesset. *Phys. Rev.*, **1934**, 618.
- [107] R. J. Bartlett, D. M. Silver. *Int. J. Quantum Chem.*, **1976**, 185–186.
- [108] J. Binkley, J. Pople. *Int. J. Quantum Chem.*, **1975**, 229–236.
- [109] J. A. Pople, J. S. Binkley, R. Seeger. *Int. J. Quantum Chem.*, **1976**, 1–19.
- [110] R. J. Bartlett, I. Shavitt. *Chemical Phys. Lett.*, **1977**, 190–198.

BIBLIOGRAPHIE

- [111] R. Krishnan, J. A. Pople. *Int. J. Quantum Chem.*, **1978**, 91–100.
- [112] R. Krishnan, M. Frisch, J. Pople. *J. Chem. Phys.*, **1980**, 4244–4245.
- [113] S. A. Kucharski, R. J. Bartlett. *Adv. Quantum Chem.*, **1986**, 281–344.
- [114] S. A. Kucharski, J. Noga, R. J. Bartlett. *J. Chem. Phys.*, **1989**, 7282–7290.
- [115] K. Raghavachari, J. A. Pople, E. S. Replogle, M. Head-Gordon. *J. Phys. Chem.*, **1990**, 5579–5586.
- [116] Z. He, D. Cremer. *Int. J. Quantum Chem.*, **1996**, 15–29.
- [117] Z. He, D. Cremer. *Int. J. Quantum Chem.*, **1996**, 71–95.
- [118] Z. He, D. Cremer. *Int. J. Quantum Chem.*, **1996**, 57–69.
- [119] D. Cremer, Z. He. *J. Phys. Chem.*, **1996**, 6173–6188.
- [120] D. Cremer, Z. He. In *Conceptual Perspectives in Quantum Chemistry*. Springer, **1997**. 239–318.
- [121] J. Čížek. *J. Chem. Phys.*, **1966**, 4256–4266.
- [122] O. Sinanoğlu. *J. Chem. Phys.*, **1962**, 706–717.
- [123] R. J. Bartlett. *J. Phys. Chem.*, **1989**, 1697–1708.
- [124] R. J. Bartlett. *Modern Electronic Structure Theory : Part II*, **1995**, 1047–1131.
- [125] B. O. Roos, P. R. Taylor, P. E. Sigbahn. *Chem. Phys.*, **1980**, 157–173.
- [126] K. Andersson, P. A. Malmqvist, B. O. Roos, A. J. Sadlej, K. Wolinski. *J. Phys. Chem.*, **1990**, 5483–5488.
- [127] K. Andersson, P.-Å. Malmqvist, B. O. Roos. *J. Chem. Phys.*, **1992**, 1218–1226.
- [128] C. Angeli, R. Cimiraglia, S. Evangelisti, T. Leininger, J.-P. Malrieu. *J. Chem. Phys.*, **2001**, 10252–10264.
- [129] C. Angeli, R. Cimiraglia, J.-P. Malrieu. *Chemical Phys. Lett.*, **2001**, 297–305.
- [130] C. Angeli, R. Cimiraglia, J.-P. Malrieu. *J. Chem. Phys.*, **2002**, 9138–9153.
- [131] L. H. Thomas. In *Mathematical proceedings of the Cambridge philosophical society*. Cambridge University Press, 542–548.
- [132] E. Fermi. *Rend. Accad. Naz. Lincei*, **1927**, 602–607.
- [133] P. Hohenberg, W. Kohn. *Phys. Rev.*, **1964**, B864.
- [134] W. Kohn, S. L. J. *phys. Rev*, **1965**, A1133.
- [135] P. A. Dirac. In *Mathematical proceedings of the Cambridge philosophical society*. Cambridge University Press, 376–385.
- [136] A. D. Becke. *Phys. Rev. A*, **1988**, 3098–3100.

- [137] R. Colle, O. Salvetti. *Theor. Chem. Acc.*, **1975**, 329–334.
- [138] R. Colle, O. Salvetti. *J. Chem. Phys.*, **1983**, 1404–1407.
- [139] C. Lee, W. Yang, R. G. Parr. *Phys. Rev. B*, **1988**, 785–789.
- [140] J. P. Perdew, M. Ernzerhof, K. Burke. *J. Chem. Phys.*, **1996**, 9982–9985.
- [141] C. Adamo, V. Barone. *J. Chem. Phys.*, **1999**, 6158–6170.
- [142] M. Rapacioli, F. Spiegelman, D. Talbi, T. Mineva, A. Goursot, T. Heine, G. Seifert. *J. Chem. Phys.*, **2009**, 194303.
- [143] W. Jorgensen, J. Tirado-Rives. *J. Am. Chem. Soc.*, **1988**, 1657–1666.
- [144] W. L. Jorgensen, D. S. Maxwell, J. Tirado-Rives. *J. Am. Chem. Soc.*, **1996**, 11225–11236.
- [145] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein. *J. Chem. Phys.*, **1983**, 926–935.
- [146] H. M. Senn, W. Thiel. *Angew. Chem. Int. Ed.*, **2009**, 1198–1229.
- [147] A. Warshel, M. Levitt. *J. Mol. Biol.*, **1976**, 227–249.
- [148] M. J. Field, P. A. Bash, M. Karplus. *J. Comput. Chem.*, **1990**, 700–733.
- [149] G. A. Tribello, M. Ceriotti, M. Parrinello. *Proc. Natl. Acad. Sci. USA*, **2012**, 5196–5201.
- [150] B. J. Alder, T. E. Wainwright. *J. Chem. Phys.*, **1957**, 1208–1209.
- [151] S. Charati, S. Stern. *Macromolecules*, **1998**, 5529–5535.
- [152] I.-C. Yeh, G. Hummer. *J. Phys. Chem. B*, **2004**, 15873–15879.
- [153] G. Pranami, M. H. Lamm. *J. Chem. Theory Comput.*, **2015**, 4586–4592.
- [154] M. Mondello, G. S. Grest. *J. Chem. Phys.*, **1997**, 9327–9336.
- [155] B. Hess. *J. Chem. Phys.*, **2002**, 209–217.
- [156] Y. Zhang, A. Otani, E. J. Maginn. *J. Chem. Theory Comput.*, **2015**, 3537–3546.
- [157] S. Aragon, R. Pecora. *J. Chem. Phys.*, **1976**, 1791–1803.
- [158] B. Guillot. *J. Chem. Phys.*, **1991**, 1543–1551.
- [159] R. Vuilleumier, D. Borgis. *J. Chem. Phys.*, **1999**, 4251–4266.
- [160] M.-P. Gageot, M. Sprik. *J. Phys. Chem. B*, **2003**, 10344–10358.
- [161] S. Corcelli, C. Lawrence, J. Skinner. *J. Chem. Phys.*, **2004**, 8107–8117.
- [162] J. D. Honeycutt, H. C. Andersen. *J. Phys. Chem.*, **1987**, 4950–4963.
- [163] C. Briant, J. Burton. *J. Chem. Phys.*, **1975**, 2045–2058.

BIBLIOGRAPHIE

- [164] J. P. Postma, H. J. Berendsen, J. R. Haak. In *Faraday Symposia of the Chemical Society*. Royal Society of Chemistry, 55–67.
- [165] W. Sekkal, B. Bouhafs, H. Aourag, M. Certier. *J. Phys. Condens. Matter*, **1998**, 4975.
- [166] S.-T. Lin, M. Blanco, W. A. Goddard III. *J. Chem. Phys.*, **2003**, 11792–11805.
- [167] M. Parrinello, A. Rahman. *Phys. Rev. Lett.*, **1980**, 1196.
- [168] P. Ballone, W. Andreoni, R. Car, M. Parrinello. *Phys. Rev. Lett.*, **1988**, 271.
- [169] M. Karplus, J. A. McCammon. *Nat. Struct. Biol.*, **2002**, 646–652.
- [170] G. Gutierrez, B. Johansson. *Phys. Rev. B*, **2002**, 104202.
- [171] M. Levitt. *J. Mol. Biol.*, **1983**, 723–764.
- [172] W. C. Swope, J. W. Pitera, F. Suits. *J. Phys. Chem. B*, **2004**, 6571–6581.
- [173] J. D. Chodera, W. C. Swope, J. W. Pitera, K. A. Dill. *Multiscale Model. Simul.*, **2006**, 1214–1226.
- [174] Y. Sugita, Y. Okamoto. *Chemical Phys. Lett.*, **1999**, 141–151.
- [175] R. Car, M. Parrinello. *Phys. Rev. Lett.*, **1985**, 2471–2474.
- [176] J. T. Batina. *AIAA J.*, **1991**, 327–333.
- [177] J. C. Butcher. John Wiley & Sons, **2016**.
- [178] C. Runge. *Math. Ann.*, **1895**, 167–178.
- [179] W. Kutta. *Z. Math. Phys.*, **1901**, 435–453.
- [180] J. C. Butcher. *Appl. Numer. Math.*, **1996**, 247–260.
- [181] C. W. Gear. Technical report, Argonne National Lab., Ill., **1966**.
- [182] K. Diethelm, N. J. Ford, A. D. Freed. *Nonlinear Dyn.*, **2002**, 3–22.
- [183] L. Verlet. *Phys. Rev.*, **1967**, 98.
- [184] F. Fraige, P. Langston. *Adv. Powder Technol.*, **2004**, 227–245.
- [185] W. C. Swope, H. C. Andersen, P. H. Berens, K. R. Wilson. *J. Chem. Phys.*, **1982**, 637–649.
- [186] H. J. Berendsen, J. v. Postma, W. F. van Gunsteren, A. DiNola, J. R. Haak. *J. Chem. Phys.*, **1984**, 3684–3690.
- [187] S. Nosé. *J. Chem. Phys.*, **1984**, 511–519.
- [188] W. G. Hoover. *Phys. Rev. A*, **1985**, 1695–1697.
- [189] F. Legoll, M. Luskin, R. Moeckel. *Arch. Ration. Mech. Anal.*, **2007**, 449–463.
- [190] G. J. Martyna, M. L. Klein, M. Tuckerman. *J. Chem. Phys.*, **1992**, 2635–2643.

-
- [191] L. F. L. Oliveira, J. Cuny, M. Morinière, L. Dontot, A. Simon, F. Spiegelman, M. Rapacioli. *Phys. Chem. Chem. Phys.*, **2015**, 17079–17089.
- [192] A. Cheng, K. M. Merz. *J. Phys. Chem.*, **1996**, 1927–1937.
- [193] S. Jang, Y. Pak, S. Shin. *J. Chem. Phys.*, **2002**, 4782–4786.
- [194] S. K. Sahoo, N. N. Nair. *Front. Chem*, **2018**, 275.
- [195] M. Salonen, I. Napari, H. Vehkamäki. *Mol. Simul.*, **2007**, 245–251.
- [196] S. Urata, J. Irisawa, A. Takada, W. Shinoda, S. Tsuzuki, M. Mikami. *J. Phys. Chem. B*, **2005**, 4269–4278.
- [197] J. O. Daldrop, M. Saita, M. Heyden, V. A. Lorenz-Fonfria, J. Heberle, R. R. Netz. *Nat. Commun.*, **2018**, 1–7.
- [198] M. Baer, G. Mathias, I.-F. W. Kuo, D. J. Tobias, C. J. Mundy, D. Marx. *ChemPhysChem*, **2008**, 2703–2707.
- [199] D. J. Tobias, G. J. Martyna, M. L. Klein. *J. Phys. Chem.*, **1993**, 12959–12966.
- [200] J. A. Morrone, R. Car. *Phys. Rev. Lett.*, **2008**, 017801.
- [201] T. Schneider, E. Stoll. *Phys. Rev. B*, **1978**, 1302.
- [202] G. Bussi, D. Donadio, M. Parrinello. *J. Chem. Phys.*, **2007**, 014101.
- [203] M. Ceriotti, G. Bussi, M. Parrinello. *Phys. Rev. Lett.*, **2009**, 020601.
- [204] M. Ceriotti, G. Bussi, M. Parrinello. *Phys. Rev. Lett.*, **2009**, 030603.
- [205] M. Ceriotti, G. Bussi, M. Parrinello. *J. Chem. Theory Comput.*, **2010**, 1170–1180.
- [206] G. Bussi, M. Parrinello. *Comput. Phys. Commun.*, **2008**, 26–29.
- [207] M. Ceriotti. *GLE4MD*. <http://gle4md.org/>, **2014**.
- [208] M. Ceriotti, D. E. Manolopoulos. *Phys. Rev. Lett.*, **2012**, 100604.
- [209] Y. Badyal, D. Price, M.-L. Saboungi, D. Haefner, S. Shastri. *J. Chem. Phys.*, **2002**, 10833–10837.
- [210] F. Paesani, W. Zhang, D. A. Case, T. E. Cheatham III, G. A. Voth. *J. Chem. Phys.*, **2006**, 184507.
- [211] F. Paesani, S. Iuchi, G. A. Voth. *J. Chem. Phys.*, **2007**, 074506.
- [212] T. E. Markland, S. Habershon, D. E. Manolopoulos. *J. Chem. Phys.*, **2008**, 194506.
- [213] S. Habershon, T. E. Markland, D. E. Manolopoulos. *J. Chem. Phys.*, **2009**, 024501.
- [214] A. Zeidler, P. S. Salmon, H. E. Fischer, J. C. Neufeind, J. M. Simonson, T. E. Markland. *J. Phys. Condens. Matter*, **2012**, 284126.
- [215] T. Spura, C. John, S. Habershon, T. D. Kühne. *Mol. Phys.*, **2015**, 808–822.

BIBLIOGRAPHIE

- [216] M. Riera, E. Lambros, T. T. Nguyen, A. W. Götz, F. Paesani. *Chem. Sci.*, **2019**, 8211–8218.
- [217] D. Berta, D. Ferenc, I. Bakó, Á. Madarász. *J. Chem. Theory Comput.*, **2020**, 3316–3334.
- [218] I. Bakó, Á. Madarász, L. Pusztai. *J. Mol. Liq.*, **2021**, 115192.
- [219] D. Chandler, P. G. Wolynes. *J. Chem. Phys.*, **1981**, 4078–4095.
- [220] S. Habershon, D. E. Manolopoulos, T. E. Markland, T. F. Miller. *Annu. Rev. Phys. Chem.*, **2013**, 387–413. PMID : 23298242.
- [221] J. Cao, G. A. Voth. *J. Chem. Phys.*, **1994**, 6168–6183.
- [222] M. E. Tuckerman, D. Marx, M. L. Klein, M. Parrinello. *J. Chem. Phys.*, **1996**, 5579–5588.
- [223] M. Rossi, M. Ceriotti, D. E. Manolopoulos. *J. Chem. Phys.*, **2014**, 234116.
- [224] M. Rossi, V. Kapil, M. Ceriotti. *J. Chem. Phys.*, **2018**, 102301.
- [225] R. H. Swendsen, J. S. Wang. *Phys. Rev. Lett.*, **1986**, 2607–2609.
- [226] C. GEYER. In *Computing science and statistics : Proceedings of 23rd Symposium on the Interface Interface Foundation, Fairfax Station, 1991*. 156–163.
- [227] B. A. Berg, T. Neuhaus. *Phys. Lett. B*, **1991**, 249–253.
- [228] B. A. Berg, T. Neuhaus. *Phys. Rev. Lett.*, **1992**, 9.
- [229] K. Hukushima, K. Nemoto. *J. Phys. Soc. Jpn.*, **1996**, 1604–1608.
- [230] M. Falcioni, M. W. Deem. *J. Chem. Phys.*, **1999**, 1754–1766.
- [231] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, E. Teller. *J. Chem. Phys.*, **1953**, 1087–1092.
- [232] F. Louisnard, G. Geudtner, A. M. Köster, J. Cuny. *Theor. Chem. Acc.*, **2021**, 95.
- [233] C. Dellago, P. G. Bolhuis, F. S. Csajka, D. Chandler. *J. Chem. Phys.*, **1998**, 1964–1977.
- [234] C. Dellago, P. Bolhuis, P. L. Geissler, et al. *Adv. Chem. Phys.*, **2002**.
- [235] P. G. Bolhuis, D. Chandler, C. Dellago, P. L. Geissler. *Annu. Rev. Phys. Chem.*, **2002**, 291–318.
- [236] M. R. So/rensen, A. F. Voter. *J. Chem. Phys.*, **2000**, 9599–9606.
- [237] F. Montalenti, M. Sørensen, A. Voter. *Phys. Rev. Lett.*, **2001**, 126101.
- [238] R. J. Zamora, B. P. Uberuaga, D. Perez, A. F. Voter. *Annu. Rev. Chem. Biomol. Eng.*, **2016**, 87–110.

- [239] D. Bertsimas, J. Tsitsiklis. *Stat. Sci.*, **1993**, 10–15.
- [240] K. A. Dowsland, J. M. Thompson. In *Handbook of Natural Computing*. Springer Berlin Heidelberg, Berlin, Heidelberg, **2012**. 1623–1655.
- [241] A. Laio, M. Parrinello. *Proc. Natl. Acad. Sci. USA*, **2002**, 12562–12566.
- [242] A. Laio, A. Rodriguez-Forteza, F. L. Gervasio, M. Ceccarelli, M. Parrinello. *J. Phys. Chem. B*, **2005**, 6714–6721.
- [243] B. Ensing, M. De Vivo, Z. Liu, P. Moore, M. L. Klein. *Acc. Chem. Res.*, **2006**, 73–81.
- [244] A. Barducci, M. Bonomi, M. Parrinello. *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, **2011**, 826–843.
- [245] D. Branduardi, G. Bussi, M. Parrinello. *J. Chem. Theory Comput.*, **2012**, 2247–2254.
- [246] P. Tiwary, M. Parrinello. *Phys. Rev. Lett.*, **2013**, 230602.
- [247] G. M. Torrie, J. P. Valleau. *J. Comput. Phys.*, **1977**, 187–199.
- [248] A. Warmflash, P. Bhimalapuram, A. R. Dinner. *J. Chem. Phys.*, **2007**, 114109.
- [249] J. Kästner. *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, **2011**, 932–942.
- [250] X. Yuan, A. N. Cormack. *Comput. Mater. Sci.*, **2002**, 343–360.
- [251] M. Jieli, T. Miyake, M. Aida. *Bull. Chem. Soc. Jap.*, **2007**, 2131–2136.
- [252] M. Jieli, M. Aida. *J. Phys. Chem. A*, **2009**, 1586–1594.
- [253] D. Akase, H. Teramae, M. Aida. *Chemical Phys. Lett.*, **2015**, 51–56.
- [254] M. Aida, D. Akase. *Pure Appl. Chem.*, **2019**, 301–316.
- [255] A. P. Bartók, R. Kondor, G. Csányi. *Phys. Rev. B*, **2013**, 184115.
- [256] P. Labastie, R. L. Whetten. *Phys. Rev. Lett.*, **1990**, 1567.
- [257] R. Poteau, F. Spiegelmann, P. Labastie. *Z. Phys. D*, **1994**, 57–68.
- [258] F. London. *J. phys. radium*, **1937**. 8, 397–409.
- [259] R. McWeeny. *Mol. Phys.*, **1958**. 1, 311–321.
- [260] D. Davies. *J. Chem. Soc. Faraday Trans.*, **1961**. 57, 2081–2088.
- [261] H. Hamerka. *Rev. Mod. Phys.*, **1962**. 34, 87.
- [262] J. Pople. *J. Chem. Phys.*, **1962**. 37, 53–59.
- [263] G. Hall, A. Hardisson. *Proc. Math. Phys. Eng. Sci.*, **1962**. 268, 328–338.
- [264] A. Amos, H. F. Roberts. *J. Chem. Phys.*, **1969**. 50, 2375–2381.
- [265] R. Marinenko, K. F. Heinrich, F. Ruegg. *Micro-homogeneity studies of NBS Standard Reference Materials, NBS Research Materials, and other related samples*. 260-265. Department of Commerce, National Bureau of Standards, **1979**.

BIBLIOGRAPHIE

- [266] K. Friedrich, G. Seifert, G. Grossmann. *Z. Phys. D*, **1990**. *17*, 45–46.
- [267] R. Ditchfield. *J. Chem. Phys.*, **1972**. *56*, 5688–5691.
- [268] G. Schreckenbach, T. Ziegler. *J. Phys. Chem.*, **1995**. *99*, 606–611.
- [269] *abalone*. <http://www.biomolecular-modeling.com/Abalone/>.
- [270] R. Salomon-Ferrer, D. A. Case, R. C. Walker. *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, **2013**, 198–210.
- [271] S. Plimpton. *J. Comput. Phys.*, **1995**, 1–19.
- [272] A. Buin, R. Comin, A. H. Ip, E. H. Sargent. *J. Phys. Chem. C*, **2015**, 13965–13971.
- [273] T. D. Kühne, M. Iannuzzi, M. Del Ben, V. V. Rybkin, P. Seewald, F. Stein, T. Laino, R. Z. Khaliullin, O. Schütt, F. Schiffmann, et al. *J. Chem. Phys.*, **2020**, 194103.
- [274] M. Feig, J. Karanicolas, C. L. Brooks III. *J. Mol. Graph.*, **2004**, 377–395.
- [275] A. M. Köster, J. U. Reveles, J. M. del Campo. *J. Chem. Phys.*, **2004**, 3417–3424.
- [276] A. De La Lande, A. Alvarez-Ibarra, K. Hasnaoui, F. Cailliez, X. Wu, T. Mineva, J. Cuny, P. Calaminici, L. López-Sosa, G. Geudtner, I. Navizet, C. G. Iriepa, D. R. Salahub, A. M. Köster. *Molecules*, **2019**.
- [277] G. Geudtner, F. Janetzko, A. M. Köster, A. Vela, P. Calaminici. *J. Comput. Chem.*, **2006**, 483–490.
- [278] G. Bussi. *Nuovo Cimento C*, **2009**, 61–65.
- [279] M. Losada, S. Leutwyler. *J. Chem. Phys.*, **2002**. *117*, 2003–2016.
- [280] M. Losada, S. Leutwyler. *J. Chem. Phys.*, **2003**. *119*, 304–312.
- [281] B. Santra, A. Michaelides, M. Fuchs, A. Tkatchenko, C. Filippi, M. Scheffler. *J. Chem. Phys.*, **2008**. *129*, 194111.
- [282] A. K. Kelkkanen, B. I. Lundqvist, J. K. Nørskov. *J. Chem. Phys.*, **2009**, 046102.
- [283] P. L. Silvestrelli. *Chemical Phys. Lett.*, **2009**. *475*, 285–288.
- [284] D. M. Bates, G. S. Tschumper. *J. Phys. Chem. A*, **2009**. *113*, 3555–9.
- [285] R. J. Saykally, D. J. Wales. *Science*, **2012**. *336*, 814–815.
- [286] C. Pérez, M. T. Muckle, D. P. Zaleski, N. A. Seifert, B. Temelso, G. C. Shields, Z. Kisiel, B. H. Pate. *Science*, **2012**, 897–901.
- [287] J. M. Pedulla, K. D. Jordan. *Chem. Phys.*, **1998**, 593–601.
- [288] P. Calaminici, R. Flores-Moreno, A. M. Köster. *CoLe*, **2005**, 164–171.
- [289] P. Calaminici, F. Janetzko, A. M. Köster, R. Mejia-Olvera, B. Zuniga-Gutierrez. *J. Chem. Phys.*, **2007**, 044108.

- [290] N. Godbout, D. R. Salahub, J. Andzelm, E. Wimmer. *Can. J. Chem.*, **1992**, 560–571.
- [291] A. Simon, M. Rapacioli, J. Mascetti, F. Spiegelman. *Phys. Chem. Chem. Phys.*, **2012**, 6771–6786.
- [292] A. Simon, F. Spiegelman. *J. Chem. Phys.*, **2013**, 194309.
- [293] A. Simon, F. Spiegelman. *J. Chem. Phys.*, **2013**.
- [294] A. N. Tharrington, K. D. Jordan. *J. Phys. Chem. A*, **2003**, 7380–7389.
- [295] J. O. Richardson, D. J. Wales, S. C. Althorpe, R. P. McLaughlin, M. R. Viant, O. Shih, R. J. Saykally. *J. Phys. Chem. A*, **2013**, 6960–6966.
- [296] D. J. Wales, M. P. Hodges. *Chem. Phys. Lett.*, **1998**, 65–72.
- [297] J. Kim, D. Majumdar, H. M. Lee, K. S. Kim. *J. Chem. Phys.*, **1999**, 9128–9134.
- [298] P. Nigra, M. A. Carignano, S. Kais. *J. Chem. Phys.*, **2001**, 2621–2628.
- [299] M. Ocasio, G. E. López. *Chemical Phys. Lett.*, **2002**, 168–174.
- [300] S. D. Belair, J. S. Francisco. *Phys. Rev. A*, **2003**, 063206.
- [301] S. S. Xantheas, E. Aprá. *J. Chem. Phys.*, **2004**, 823–828.
- [302] J. Hernández-Rojas, B. S. González, T. James, D. J. Wales. *J. Chem. Phys.*, **2006**, 224302.
- [303] E. Asare, A.-R. Musah, E. Curotto, D. L. Freeman, J. D. Doll. *J. Chem. Phys.*, **2009**, 184508.
- [304] C. M. Neal, A. K. Starace, M. F. Jarrold. *Phys. Rev. B Condens. Matter Mater. Phys.*, **2007**, 054113.
- [305] J. M. Vásquez-Pérez, P. Calaminici, A. M. Köster. *Comput. Theor. Chem.*, **2013**, 229–232.
- [306] C. M. Maupin, B. Aradi, G. A. Voth. *J. Phys. Chem. B*, **2010**, 6922–6931.
- [307] T. H. Choi, K. D. Jordan. *J. Phys. Chem. B*, **2010**, 6932–6936.
- [308] P. Goyal, M. Elstner, Q. Cui. *J. Phys. Chem. B*, **2011**, 6790–6805.
- [309] T. H. Choi, R. Liang, C. M. Maupin, G. A. Voth. *J. Phys. Chem. B*, **2013**, 5165–5179.
- [310] V. S. Bryantsev, M. S. Diallo, A. C. Van Duin, W. A. Goddard III. *J. Chem. Theory Comput.*, **2009**, 1016–1026.
- [311] R. Shi, X. Huang, Y. Su, H. G. Lu, S. D. Li, L. Tang, J. Zhao. *J. Phys. Chem. A*, **2017**, 3117–3127.
- [312] M. Tuckerman, K. Laasonen, M. Sprik, M. Parrinello. *J. Chem. Phys.*, **1995**, 150–161.

BIBLIOGRAPHIE

- [313] M. Tuckerman, K. Laasonen, M. Sprik, M. Parrinello. *J. Phys. Chem.*, **1995**, 5749–5752.
- [314] D. Marx, M. E. Tuckerman, J. Hutter, M. Parrinello. *Nature*, **1999**, 601–604.
- [315] E. G. Stein, L. M. Rice, A. T. Brünger. *J. Magn. Reson.*, **1997**, 154–164.
- [316] F. Brugé, M. Bernasconi, M. Parrinello. *J. Am. Chem. Soc.*, **1999**, 10883–10888.
- [317] D. A. Case. *Acc. Chem. Res.*, **2002**, 325–331.
- [318] L. Banci. *Curr. Opin. Chem. Biol.*, **2003**, 143–149.
- [319] S. Rossano, F. Mauri, C. J. Pickard, I. Farnan. *J. Phys. Chem. B*, **2005**, 7245–7250.
- [320] C. Zazza, G. Mancini, G. Brancato, N. Sanna, V. Barone. *Comput. Theor. Chem.*, **2012**, 53–61.
- [321] M. Dračinský, P. Hodgkinson. *CrystEngComm*, **2013**, 8705–8712.
- [322] D. K. Chakravorty, B. Wang, C. W. Lee, A. J. Guerra, D. P. Giedroc, K. M. Merz. *J. Biomol. NMR.*, **2013**, 125–137.
- [323] R. Pohl, O. Socha, P. Slavíček, M. Šála, P. Hodgkinson, M. Dračinský. *Farad. Disc.*, **2018**, 331–344.
- [324] J. Lobaugh, G. A. Voth. *J. Chem. Phys.*, **1997**, 2400–2410.
- [325] M. C. Böhm, J. Schulte, R. Ramírez. *Chem. Phys. Lett.*, **2000**, 117–124.
- [326] M. C. Böhm, J. Schulte, R. Ramírez. *Int. J. Quantum Chem.*, **2002**, 280–296.
- [327] S. Tang, D. A. Case. *J. Biomol. NMR.*, **2007**, 255–266.
- [328] J. Schulte, R. Ramírez, M. C. Böhm. *Chem. Phys. Lett.*, **2006**, 579–584.
- [329] A. A. Hassanali, J. Cuny, M. Ceriotti, C. J. Pickard, M. Parrinello. *J. Am. Chem. Soc.*, **2012**, 8557–8569.
- [330] M. Dračinský, P. Hodgkinson. *Chem. Eur. J.*, **2014**, 2201–2207.
- [331] K. Sykina, B. Bureau, L. Le Pollès, C. Roiland, M. Deschamps, C. J. Pickard, E. Furet. *Phys. Chem. Chem. Phys.*, **2014**, 17975–17982.
- [332] Y. Mori, Y. Masuda. *Chem. Phys.*, **2015**, 18–29.
- [333] K. Machado, D. Zanghi, V. Sarou-Kanian, S. Cadars, M. Burbano, M. Salanne, C. Bessada. *J. Phys. Chem. C*, **2017**, 10289–10297.
- [334] K. Aidas, A. Møgelhøj, H. Kjær, C. B. Nielsen, K. V. Mikkelsen, K. Ruud, O. Christiansen, J. Kongsted. *J. Phys. Chem. A*, **2007**, 4199–4210.
- [335] N. Özcan, J. Mareš, D. Sundholm, J. Vaara. *Phys. Chem. Chem. Phys.*, **2014**, 22309–22320.

- [336] T. Mineva, Y. Tsoneva, R. Kevorkyants, A. Goursot. *Can. J. Chem.*, **2013**, 529–537.
- [337] Y. Tsoneva, A. Tadjer, T. Mineva. *Int. J. Quantum Chem.*, **2016**, 1419–1426.
- [338] A. Goursot, T. Mineva, J. M. Vásquez-Pérez, P. Calaminici, A. M. Köster, D. R. Salahub. *Phys. Chem. Chem. Phys.*, **2013**, 860–867.
- [339] M. Dracinsky, P. Bour, P. Hodgkinson. *J. Chem. Theory Comput.*, **2016**, 968–973.
- [340] M. Dračinský, J. Vícha, K. Bártová, P. Hodgkinson. *ChemPhysChem*, **2020**, 2075–2083.
- [341] M. Ceriotti, D. E. Manolopoulos, M. Parrinello. *J. Chem. Phys.*, **2011**, 084104.
- [342] M. H. Müser, M. Müller. *J. Chem. Phys.*, **2015**, 174105.
- [343] O. Marsalek, T. E. Markland. *J. Chem. Phys.*, **2016**, 054112.
- [344] F. Briec, H. Dammak, M. Hayoun. *J. Chem. Theory Comput.*, **2016**, 1351–1359.
- [345] V. Kapil, J. VandeVondele, M. Ceriotti. *J. Chem. Phys.*, **2015**.
- [346] A. García, N. Papior, A. Akhtar, E. Artacho, V. Blum, E. Bosoni, P. Brandimarte, M. Brandbyge, J. I. Cerdá, F. Corsetti, et al. *J. Chem. Phys.*, **2020**, 204108.
- [347] P. Giannozzi, O. Baseggio, P. Bonfà, D. Brunato, R. Car, I. Carnimeo, C. Cavazzoni, S. De Gironcoli, P. Delugas, F. Ferrari Ruffino, et al. *J. Chem. Phys.*, **2020**, 154105.
- [348] V. Blum, R. Gehrke, F. Hanke, P. Havu, V. Havu, X. Ren, K. Reuter, M. Scheffler. *Comput. Phys. Commun.*, **2009**, 2175–2196.
- [349] A. Sáenz-Galindo, L. I. López-López, N. Fabiola, A. O. Castañeda-Facio, L. A. Ramírez-Mendoza, K. C. Córdova-Cisneros, D. de Loera-Carrera. In *Carboxylic Acid- Key Role in Life Sciences*. IntechOpen, **2018**.
- [350] M. Dračinský, P. Bour. *J. Chem. Theory Comput.*, **2010**, 288–299.
- [351] J. D. Samaniego-Rojas, L.-I. Hernández-Segura, L. López-Sosa, R. I. Delgado-Venegas, B. Gomez, J.-C. Lambry, A. de la Lande, T. Mineva, J. Alejandro, B. A. Zúñiga-Gutiérrez, et al. In *Multiscale Dynamics Simulations : Nano and Nano-bio Systems in Complex Environments*. Royal Society of Chemistry, **2021**.
- [352] B. I. Dunlap, J. W. D. Connolly, J. R. Sabin. *J. Chem. Phys.*, **1979**, 4993–4999.
- [353] M. Krack, A. M. Köster. *J. Chem. Phys.*, **1998**, 3226–3234.
- [354] B. Zuniga-Gutierrez, G. Geudtner, A. M. Kster. *J. Chem. Phys.*, **2011**, 124108.
- [355] J. Srinivasan, T. E. Cheatham, P. Cieplak, P. A. Kollman, D. A. Case. *J. Am. Chem. Soc.*, **1998**, 9401–9409.
- [356] P. A. Kollman, I. Massova, C. Reyes, B. Kuhn, S. Huo, L. Chong, M. Lee, T. Lee, Y. Duan, W. Wang, et al. *Acc. Chem. Res.*, **2000**, 889–897.

BIBLIOGRAPHIE

- [357] N. I. of Advanced Industrial Science, Technology. https://sdb.sdb.aist.go.jp/sdb/s/cgi-bin/cre_index.cgi.
- [358] Y. Wang, J. M. Bowman. *J. Phys. Chem. Lett.*, **2013**, 1104–1108.
- [359] C. J. Tainter, J. L. Skinner. *J. Chem. Phys.*, **2012**, 104304.
- [360] L. X. Dang, T.-M. Chang. *J. Chem. Phys.*, **1997**, 8149–8159.
- [361] Y. Wang, V. Babin, J. M. Bowman, F. Paesani. *J. Am. Chem. Soc.*, **2012**, 11116–11119.
- [362] V. Babin, F. Paesani. *Chemical Phys. Lett.*, **2013**, 1–8.
- [363] E. Belega, K. Tatarenko, D. Trubnikov, E. Cheremukhin. *Russian J. Phys. Chem. B*, **2009**, 404–409.
- [364] J. Jin, W. A. Goddard III. *J. Phys. Chem. C*, **2015**, 2622–2629.
- [365] E. Belega, D. Trubnikov, E. Cheremukhin. *J. Struct. Chem.*, **2015**, 52–57.
- [366] Y. Wang, X. Huang, B. C. Shepler, B. J. Braams, J. M. Bowman. *J. Chem. Phys.*, **2011**, 094509.
- [367] G. S. Fanourgakis, S. S. Xantheas. *J. Chem. Phys.*, **2008**, 074506.
- [368] V. Babin, G. R. Medders, F. Paesani. *J. Phys. Chem. Lett.*, **2012**, 3765–3769.
- [369] G. R. Medders, V. Babin, F. Paesani. *J. Chem. Theory Comput.*, **2014**, 2906–2910.
- [370] V. Vaida, J. E. Headrick. *J. Phys. Chem. A*, **2000**, 5401–5412.
- [371] V. Vaida. *The J. Phys. Chem. A*, **2009**, 5–18.
- [372] H. G. Kjaergaard, T. W. Robinson, D. L. Howard, J. S. Daniel, J. E. Headrick, V. Vaida. *J. Phys. Chem. A*, **2003**, 10680–10686.
- [373] V. Vaida, J. S. Daniel, H. G. Kjaergaard, L. M. Goss, A. F. Tuck. *Q. J. R. Meteorol. Soc.*, **2001**, 1627–1643.
- [374] E. Brodskaya, A. P. Lyubartsev, A. Laaksonen. *J. Chem. Phys.*, **2002**, 7879–7892.
- [375] E. Brodskaya, A. P. Lyubartsev, A. Laaksonen. *J. Phys. Chem. B*, **2002**, 6479–6487.
- [376] X. Li, T. Hede, Y. Tu, C. Leck, H. Ågren. *J. Phys. Chem. Lett.*, **2010**, 769–773.
- [377] S. Höfner, B. Almeida, U. H. Hansmann. *Proteins*, **2007**, 662–669.
- [378] K. A. Korchagina, A. Simon, M. Rapacioli, F. Spiegelman, J. Cuny. *J. Phys. Chem. A*, **2016**, 9089–9100.
- [379] J.-L. Kuo, M. L. Klein. *J. Chem. Phys.*, **2005**, 024516.
- [380] T. Takayanagi, K. Takahashi, A. Kakizaki, M. Shiga, M. Tachikawa. *Chem. Phys.*, **2009**, 196–202.

-
- [381] O. Svoboda, L. Kubelova, P. Slavicek. *J. Phys. Chem. A*, **2013**, 12868–12877.
- [382] S. F. Langley, E. Curotto, D. L. Freeman, J. Doll. *J. Chem. Phys.*, **2007**, 084506.
- [383] K. Sellgren, T. Brooke, R. Smith, T. Geballe. *Astrophys. J.*, **1995**, L69.
- [384] L. Allamandola, D. Hudgins, C. Bauschlicher Jr, S. Langhoff. *Astron. Astrophys.*, **1999**, 659–664.
- [385] S. Thorwirth, P. Theulé, C. A. Gottlieb, M. C. McCarthy, P. Thaddeus. *Proc. Int. Astron. Union.*, **2005**, 455–456.
- [386] J. D. Bregman, P. Temi. *Astrophys. J.*, **2001**, 126.
- [387] J. Chiar, A. Tielens, D. Whittet, W. Schutte, A. Boogert, D. Lutz, E. Van Dishoeck, M. Bernstein. *Astrophys. J.*, **2000**, 749.
- [388] M. De Vries, K. Reihs, H. Wendt, W. Golden, H. Hunziker, R. Fleming, E. Peterson, S. Chang. *Geochim. Cosmochim. Acta*, **1993**, 933–938.
- [389] L. Allamandola, S. Sandford, B. Wopenka. *Science*, **1987**, 56–59.
- [390] J. R. Cronin, S. Chang. In *The chemistry of life's origins*. Springer, **1993**. 209–258.
- [391] G. Moreels, J. Clairemidi, P. Hermine, P. Brechignac, P. Rousselot. *Astron. Astrophys.*, **1994**, 643–656.
- [392] P. Ehrenfreund, S. B. Charnley. *Annu. Rev. Astron. Astrophys.*, **2000**, 427–483.
- [393] Oxford University Press, **2002**.
- [394] J. N. Pitts, D. M. Lokensgard, P. S. Ripley, K. A. Van Cauwenberghe, L. Van Vaeck, S. D. Shaffer, A. J. Thill, W. L. Belser. *Science*, **1980**, 1347–1349.
- [395] M. Callén, M. De la Cruz, J. López, R. Murillo, M. Navarro, A. Mastral. *Chemosphere*, **2008**, 1357–1365.
- [396] L. C. Marr, T. W. Kirchstetter, R. A. Harley, A. H. Miguel, S. V. Hering, S. K. Hammond. *Environ. Sci. Technol.*, **1999**, 3091–3099.
- [397] M. J. Kleeman, J. J. Schauer, G. R. Cass. *Environ. Sci. Technol.*, **1999**, 3516–3523.
- [398] J. Sasaki, S. M. Aschmann, E. S. Kwok, R. Atkinson, J. Arey. *Environ. Sci. Technol.*, **1997**, 3173–3179.
- [399] J. L. Durant, A. L. Lafleur, E. F. Plummer, K. Taghizadeh, W. F. Busby, W. G. Thilly. *Environ. Sci. Technol.*, **1998**, 1894–1906.
- [400] A. Feilberg, T. Nielsen. *Environ. Sci. Technol.*, **2000**, 789–797.
- [401] J. N. Pitts Jr. *Environ. Health Perspect.*, **1983**, 115–140.
- [402] R. Atkinson, J. Arey. *Environ. Health Perspect.*, **1994**, 117–126.

BIBLIOGRAPHIE

- [403] G. A. Umbuzeiro, A. Franco, M. H. Martins, F. Kummrow, L. Carvalho, H. H. Schmeiser, J. Leykauf, M. Stiborova, L. D. Claxton. *Mutat. Res. - Genet. Toxicol. Environ. Mutagen.*, **2008**, 72–80.
- [404] A. Leger, J. Puget. *Astron. Astrophys.*, **1984**, L5–L8.
- [405] L. Allamandola, A. Tielens, J. Barker. *Astrophys. J.l*, **1985**, L25–L28.
- [406] A. G. Tielens. *Annu. Rev. Astron. Astrophys.*, **2008**, 289–337.
- [407] B. A. McGuire, R. A. Loomis, A. M. Burkhardt, K. L. K. Lee, C. N. Shingledecker, S. B. Charnley, I. R. Cooke, M. A. Cordiner, E. Herbst, S. Kalenskii, et al. *Science*, **2021**, 1265–1269.
- [408] M. Bernstein, S. Sandford, A. Mattioda, L. Allamandola. *Astrophys. J.l*, **2007**, 1264.
- [409] J. Bouwman, H. M. Cuppen, A. Bakker, L. J. Allamandola, H. Linnartz. *Astron. Astrophys.*, **2010**, A33.
- [410] J. Bouwman, D. Paardekooper, H. Cuppen, H. Linnartz, L. Allamandola. *Astrophys. J.l*, **2009**, 56.
- [411] M. P. Bernstein, S. A. Sandford, L. J. Allamandola, J. S. Gillette, S. J. Clemett, R. N. Zare. *science*, **1999**, 1135–1138.
- [412] Z. Guennoun, C. Aupetit, J. Mascetti. *Phys. Chem. Chem. Phys.*, **2011**, 7340–7347.
- [413] Z. Guennoun, C. Aupetit, J. Mascetti. *J. Phys. Chem. A*, **2011**, 1844–1852.
- [414] A. De Barros, A. Mattioda, A. Ricca, G. Cruz-Diaz, L. Allamandola. *Astrophys. J.l*, **2017**, 112.
- [415] A. Simon, J. A. Noble, G. Rouaut, A. Moudens, C. Aupetit, C. Iftner, J. Mascetti. *Phys. Chem. Chem. Phys.*, **2017**, 8516–8529.
- [416] R. A. Aziz. *J. Chem. Phys.*, **1993**, 4518–4525.
- [417] C. Iftner, A. Simon, K. Korchagina, M. Rapacioli, F. Spiegelman. *J. Chem. Phys.*, **2014**, 034301.
- [418] A. Simon, C. Iftner, J. Mascetti, F. Spiegelman. *J. Phys. Chem.. A*, **2015**, 2449–67.
- [419] F. Calvo, P. Parneix, N. T. Van-Oanh. *J. Chem. Phys.*, **2010**, 124308.
- [420] X. Michaut, A.-M. Vasserot, L. Abouaf-Marguin. *Vib. Spectrosc.*, **2004**, 83–93.
- [421] J. Perchard. *Chem. Phys.*, **2001**, 217–233.
- [422] J. Ceponkus, P. Uvdal, B. Nelander. *J. Chem. Phys.*, **2010**, 074301.
- [423] J. Ceponkus, G. Karlström, B. Nelander. *J. Phys. Chem. A*, **2005**, 7859–7864.
- [424] J. Ceponkus, P. Uvdal, B. Nelander. *J. Phys. Chem. A*, **2012**, 4842–4850.

BIBLIOGRAPHIE

- [425] A. L. Mattioda, A. Ricca, J. Tucker, C. W. Bauschlicher, L. J. Allamandola. *Astron. J.*, **2009**, 4054–4060.
- [426] Q. Yu, J. M. Bowman. *J. Phys. Chem. A*, **2019**, 1399–1409.
- [427] Y. Litman, J. O. Richardson, T. Kumagai, M. Rossi. *J. Am. Chem. Soc.*, **2019**, 2526–2534.
- [428] Y. Litman, J. Behler, M. Rossi. *Farad. Disc.*, **2019**, 526–546.

Abstract

The objective of this thesis is to study the electronic, structural and thermodynamical properties of different molecular clusters using the methods “Parallel-Tempering Molecular Dynamics” (PTMD) and “Path-Integral Molecular Dynamics” (PIMD). For this purpose, several implementations have been made in the codes deMonNano and deMon2k which are molecular simulation programs based on the “Self-Consistent-Charge Density-Functional based Tight-Binding” (SCC-DFTB) and “Auxiliary Density-Functional Theory” (ADFT) methods respectively. The PTMD allows to efficiently explore the potential energy surface and the PIMD permit to access Nuclear Quantum Effects (NQE) while remaining in a classical formalism.

First of all, the PTMD method has been implemented in the code deMon2K. In order to parallelize the PTMD algorithm without modifying the preexisting MPI (for Message Passing Interface) parallel structure of deMon2k, a routine that permit the creation of many layers of MPI parallelization has been developped. The implementation of the PTMD method has then been validated by studying the phase transition in the water hexamer $(\text{H}_2\text{O})_6$ and the water octamer $(\text{H}_2\text{O})_8$ with an ADFT potential. The phase transition was characterized by the analysis of the heat capacity curve and the relative isomer population as a function of the temperature. The same study has been carried with the deMonNano code and a SCC-DFTB potential. The discrepancies in the phase transition observed between the two potentials have been enlightened by the differences in the binding energies of the low lying isomers.

To be able to run PIMD simulations, the codes deMon2k and deMonNano have been interfaced with the program i-PI. The first application of the combination between these two codes was dedicated to the study of the influence of NQEs on the chemical shielding of carbon-13 in linear carboxylates HCOO^- , $\text{CH}_3(\text{CH}_2)_n\text{COO}^-$ ($n = 0-2, 6$) explicitly solvated by 34 water molecules. This study shows that in PIMD, the quantum delocalization of the protons attached to carbon atoms leads to a systematic down shift of the isotropic shielding of carbon-13 and improve the agreement with experimental results for all systems studied.

The combination of the codes i-PI and deMonNano has been used to study the impact of NQEs on the phase transition of the $(\text{H}_2\text{O})_6$ and $(\text{H}_2\text{O})_8$ clusters. For this purpose, “Parallel-Tempering Path-Integral Molecular Dynamics” (PTPIMD) simulations combining both PTMD and PIMD have been made. The results show that NQEs strongly modify the relative isomer population of $(\text{H}_2\text{O})_6$ at low temperature. However, in the case of $(\text{H}_2\text{O})_8$, NQEs have barely no impact on the relative isomer population.

The combination of the codes i-PI and deMonNano has also been used to study the influence of NQEs on the infrared (IR) absorption spectra of molecular systems solvated by an argon matrix. Among the different systems studied, the inclusion of NQEs improves the agreement between theoretical and experimental spectra for the coronene ($\text{C}_{24}\text{H}_{12}$) and the corannulene ($\text{C}_{20}\text{H}_{10}$). In contrast, for the water clusters $(\text{H}_2\text{O})_n$ ($n = 1, 6$) and for the complexes $\text{H}_2\text{O}-\text{C}_{24}\text{H}_{12}$ and $\text{H}_2\text{O}-\text{C}_{20}\text{H}_{10}$, the IR spectra are very noisy and the method employed seems inappropriate.

Keywords molecular dynamics, water clusters, PTMD, PIMD, nuclear quantum effects, SCC-DFTB, ADFT

Résumé de la thèse

Cette thèse a pour objectif d'étudier les propriétés électroniques, structurales et thermodynamiques de différents agrégats moléculaires à l'aide des méthodes "Parallel-Tempering Molecular Dynamics" (PTMD) et "Path-Integral Molecular Dynamics" (PIMD). Pour cela différentes implémentations ont été réalisées dans les codes deMonNano et deMon2k qui sont des programmes basés sur le calcul de la densité électronique. Le code deMonNano utilise la méthode "Self-Consistent-Charge Density-Functional based Tight-Binding" (SCC-DFTB) et deMon2k la méthode "Auxiliary Density-Functional Theory" (ADFT). Les méthodes PTMD et PIMD permettent respectivement d'explorer efficacement la surface d'énergie potentielle et de prendre en compte les effets quantiques des noyaux (NQE) tout en restant dans un formalisme classique.

Tout d'abord l'approche PTMD a été implémentée dans le code deMon2k. Pour pouvoir paralléliser l'implémentation de la méthode PTMD sans avoir à modifier la structure parallèle MPI (pour Message Passing Interface) préexistante dans le code deMon2k une routine permettant la création de différentes couches de parallélisme MPI a été développée. L'implémentation de la méthode PTMD a ensuite été validée en étudiant la transition de phase dans l'hexamère d'eau (H₂O)₆ et l'octamère d'eau (H₂O)₈. Cette transition de phase a été caractérisée par l'analyse de la courbe de capacité calorifique et la distribution d'isomères en fonction de la température. La même étude a été réalisée en SCC-DFTB avec deMonNano et les différences dans la transition de phase avec les deux potentiels ont été rationalisées en terme d'énergie de liaison des isomères de basse énergie.

Pour pouvoir réaliser des simulations PIMD, les codes deMon2k et deMonNano ont été interfacés avec le programme i-PI. Une première application de la combinaison des codes i-PI et deMon2k a été dédiée à l'étude de l'impact des NQE sur le déplacement chimique du carbone 13 dans des carboxylates linéaires HCOO⁻, CH₃(CH₂)_nCOO⁻ (n = 0-2, 6) explicitement solvatés par 34 molécules d'eau. Ces travaux montrent qu'en PIMD, la délocalisation quantique des protons liés aux atomes de carbone induit un décalage vers le bas de l'écrantage isotrope des carbones 13 systématique et améliore l'accord avec les résultats expérimentaux pour tous les systèmes étudiés.

La combinaison des codes i-PI et deMonNano a été utilisée pour étudier l'impact des NQE sur la transition de phase dans les agrégats (H₂O)₆ et (H₂O)₈. Pour cela, des simulations "Parallel-Tempering Path-Integral Molecular Dynamics" (PTPIMD) combinant les méthodes PTMD et PIMD ont été réalisées. Les résultats montrent que les NQE modifient fortement la distribution d'isomères de (H₂O)₆ à basse température. En revanche dans le cas de (H₂O)₈, les NQE n'ont pratiquement aucun impact sur la distribution d'isomères.

La combinaison des codes i-PI et deMonNano a également été utilisée pour étudier les spectres d'absorption infrarouge (IR) de systèmes moléculaires solvatés par une matrice d'argon en incluant les NQE. Parmi les différents systèmes étudiés, la prise en compte des NQE améliore l'accord avec l'expérience pour les spectres IR théoriques du coronène (C₂₄H₁₂) et du corannulène (C₂₀H₁₀). En revanche pour les agrégats d'eau (H₂O)_n (n = 1, 6) et pour les complexes H₂O-C₂₄H₁₂ et H₂O-C₂₀H₁₀, les spectres IR théoriques sont fortement bruités et la méthode employée ne semble pas adaptée.

Mots-clés : dynamique moléculaire, agrégats aqueux, PTMD, PIMD, effets quantiques des noyaux, SCC-DFTB, ADFT