Méthodologies de simulation moléculaire pour le calcul de grandeurs thermodynamiques d'association : théorie et application à différents types de complexes

Aziz Ghoufi

▶ **To cite this version:**

Aziz Ghoufi. Méthodologies de simulation moléculaire pour le calcul de grandeurs thermodynamiques d'association : théorie et application à différents types de complexes. Chimie analytique. Université Blaise Pascal - Clermont-Ferrand II, 2006. Français. ⟨NNT : 2006CLF21655⟩. ⟨tel-00691202⟩

HAL Id: tel-00691202

https://theses.hal.science/tel-00691202v1

Submitted on 25 Apr 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

HAL Authorization

Numéro d'ordre : D.U. **1655**

UNIVERSITE BLAISE PASCAL

(U.F.R. de Recherche Scientifique et Technique)

ECOLE DOCTORALE DES SCIENCES FONDAMENTALES

N° : 487

THESE

présentée pour obtenir le grade de

DOCTEUR D'UNIVERSITE

(Spécialité : Chimie Physique)

PAR

Ghoufi Aziz

Diplômé d'études approfondies

Méthodologies de simulation moléculaire pour le calcul de grandeurs thermodynamiques d'association : théorie et application à différents types de complexes.

Soutenue publiquement le 23 Juin 2006, devant la commission d'examen

Rapporteur :	A. Boutin (CR)
Rapporteur :	F. Calvo (CR)
Directeur :	P. Malfreyt (Pr)
Examinateur :	A. Pádua (Pr)
Examinateur :	A. Pimpinelli (Pr)
Président du jury :	P. Ungerer (Pr)

à Nadia, pour son amour, sa patience et son soutien inconditionnel,
et Lily Meg pour avoir agréablement perturbée la rédaction de cette thèse.

Remerciements

Je tiens à remercier Patrice Malfreyt, directeur de cette thèse, pour la confiance et la liberté qu'il m'a accordée et pour la collaboration fructueuse que nous avons eue, tant sur un plan scientifique, dans notre effort d'améliorer certaines théorie, que sur le plan humain. Sa disponibilité, la clarté de ces explications, sa patience et sa détermination ont été pour moi un exemple et m'ont permis de travailler dans les meilleurs conditions possibles.

J'exprime ma gratitude à Vladimir Mayer, directeur du laboratoire, pour m'avoir acceuilli au sein du laboratoire de thermodynamique des solutions et des polymères (UMR CNRS 6003) à l'université Blaise Pascal de Clermont Ferrand.

Je suis très reconnaissant envers Anne Boutin, Florent Calvo, Agilio Pádua, Alberto Pimpinelli et Phillipe Ungerer qui m'ont fait l'honneur d'examiner et de corriger ce travail avec attention.

Je remercie également toute l'équipe du laboratoire qui ont rendu mon séjours très agréable. Je remercie particuliérement Florent Goujon et Christine Bonal, pour les nombreuses conversations stimulantes que nous avons eues. Je remercie aussi Jean Pierre Morel et Nicole Morel Desrosiers pour leurs aides diverses dans ce travail. Enfin, je tiens à remercier Johnny, Julien et Johan pour les nombreuses conversations plus ou moins scientifiques que nous avons eues.

Un grand merci à Nadia et Hakim, dont le soutien et la présence ont été indispensable pour mener ce travail à terme. Je témoigne enfin toute ma reconaissance à toute celles et tous ceux que je n'ai pas cité ici pour leur soutien à différents niveaux.

Liste des acronymes et des symboles

Acronyme	**Signification**
FEP	*Free Energy Perturbation*
FEP^m	méthode FEP modifiée
TI	*Thermodynamics Integration*
FDTI	*Finite Difference Thermodynamics Integration*
TP	*Temperature Perturbation*
PMF	*Potential of Mean Force*
DM	Dynamique Moléculaire
MC	Monte Carlo
DFT	*Density Functional Theory*
cdm	centre de masse
FCS	Force Contrainte utilisant l'algorithme SHAKE
FCFS	Force Contrainte utilisant l'algorithme FAST SHAKE
D.L.	Développement Limité

Abréviation	**Signification**
Cal^{4-}	calix[4]arène *p*-sulfonaté
Alc^s	cas des alcools seules en solution
Alc^c	cas des alcools en présence du calixarène
Dio^s	cas des diols seules en solution
Dio^c	cas des diols en présence du calixarène
In_d	intégration de distribution de distance

Symbole	Signification
G, H, S, K	enthalpie libre, enthalpie, entropie et constante d'équilibre
Q_{NVT}	fonction de partition dans l'ensemble NVT
Q_{NpT}	fonction de partition dans l'ensemble NpT
$\mathcal{H}$	hamiltonien
$\mathcal{L}$	lagrangien
$\mathcal{U}$	énergie potentielle
$\mathcal{K}$	énergie cinétique
$\mathbf{r_i}$	vecteur position de l'atome i
$\mathbf{v_i}$	vecteur vitesse de l'atome i
$\mathbf{p_i}$	vecteur impulsion de l'atome i
$\mathbf{F_i}$	vecteur force agissant sur l'atome i
Δt	pas d'intégration
$\mathcal{W}$	potentiel de force moyenne
P	pression
T	température
V	volume
β	température réciproque : $\frac{1}{kT}$
k_B	constante de Boltzmann
$\mathcal{N}$	nombre d'Avogadro
$d\tau$	volume infinitésimal
r_g	rayon de giration
$\mathcal{A}$	coefficient d'asphéricité
γ_i	moment d'inertie selon les axes x($i=1$), y($i=2$), z($i=3$)
r_c	rayon de coupure
λ	paramètre de couplage
E^{MM}	énergie calculée en dynamique moléculaire
E^{QM}	énergie calculée en mécanique quantique
E^{Fit}	fonction d'ajustement
$\mathcal{W}^{MM}$	PMF calculé en dynamique moléculaire
$\mathcal{W}^{Corr}$	PMF de correction
$\mathcal{W}^{Tot}$	PMF total
r_{min}	position du puits de potentiel
$< ... >$	moyenne d'ensemble d'une propriété
k_1,k_2, k_3	vecteur unitaire dans l'espace de Fourier
α	paramètre de convergence dans la sommation d'Ewald
d_l	borne limite d'intégration
d_c	distance contrainte

Table des matières

Remerciements **5**

Liste des acronymes et des symboles **7**

1 Introduction **13**

1.1 Contexte scientifique et description du système 13

1.2 Thermodynamique et mécanique statistique 14

1.3 Objectifs 17

I Méthodologie et applications **21**

2 Méthode de dynamique moléculaire **23**

2.1 Introduction 23

2.2 Mécanique analytique : lagrangien et hamiltonien 29

2.3 Résolution des équations canoniques du mouvement 33

2.4 Modélisation des interactions intermoléculaires 36

2.5 Mise en œuvre d'une simulation moléculaire 40

2.6 Protocole de simulation 43

3 Relation «micro-macro» : association entre le calixarène et les alcools linéaires **49**

3.1 Introduction 50

3.2 Protocole 50

3.3 Complexation avec les alcools 52

3.4 Pourquoi le méthanol n'est-il pas complexé ? 62

3.5 Complexation avec les diols 63

3.6 Résumé 68

4 Calculs de $\Delta\Delta G$, $\Delta\Delta H$, $\Delta\Delta S$: application à des complexes de sphère externe 69

4.1 Introduction 69

4.2 Rappels 73

4.3 *Free Energy Perturbation* : FEP 80

4.4 *Thermodynamics Integration* : TI 87

4.5 *Temperature Perturbation* : TP 90

4.6 Nouvelle méthodologie 92

4.7 Validation 97

4.8 Calculs de grandeurs d'hydratation et d'association 102

4.9 Résumé 111

5 Potentiel de force moyenne et calcul de $\Delta_r G^o$, $\Delta_r H^o$ et $\Delta_r S^o$: application à des complexes d'insertion et de sphère externe 113

5.1 Introduction et objectifs 113

5.2 Théorie 113

5.3 Potentiel de force moyenne de la paire d'ions $[Ca^{2+} \cdots SO_4^{2-}]$ 128

5.4 Potentiel de force moyenne pour différents types de complexes du calixarène . 138

5.5 Résumé 148

6 Méthodologie pour la prise en compte des interactions cations$-\pi$ 151

6.1 Introduction et problématique 151

6.2 Méthodologie 152

6.3 Mise en œuvre 164

6.4 Résumé 172

7 Conclusions et perspectives 173

7.1 Conclusions 173

7.2 Perspectives 176

II Théorie **177**

8 Mécanique statistique **181**

8.2 Description d'un système 181
8.3 Facteur de Boltzmann et fonction de partition 182
8.4 Notion d'ensemble statistique 187
8.5 Fonction de partition dans l'ensemble NVT 188
8.6 Fonction de partition dans l'ensemble NpT 190
8.7 Fonctions d'état 193

9 Simulation numérique de type dynamique moléculaire **195**

9.1 Déterminisme et imprédictibilité 195
9.2 Algorithme de Berendsen 201
9.3 Sommation d'Ewald 213
9.4 Algorithme de contrainte 223
9.5 Méthode de non-équilibre 235

10 Rappels sur la théorie de la DFT **239**

Bibliographie **245**

Index **251**

Chapitre 1

Introduction

1.1 Contexte scientifique et description du système

Un des axes de recherche de la thématique «Energétique et Simulation d'Edifices Supramoléculaires et de Systèmes Interfaciaux» du Laboratoire de Thermodynamique des Solutions a pour objectif la caractérisation thermodynamique des processus d'interaction entre récepteurs et substrats [1]. La recherche de ligands, capables de reconnaître sélectivement des molécules cibles en formant un complexe supramoléculaire conduit à de nombreuses applications. Les calixarènes sont des composés capables de former des complexes avec des molécules organiques ou inorganiques en solution. Les calixarènes sont en général très peu solubles dans l'eau ; l'introduction en position *para* (voir figure (2.2), section (2.6.1)) de groupements sulfonatés a permis d'obtenir des calixarènes hydrosolubles, particulièrement intéressants dans différents domaines tels que la reconnaissance spécifique de molécules biologiquement actives et le mimétisme d'enzymes. Dans le but d'identifier les facteurs qui contrôlent les processus d'association des calixarènes avec des biomolécules, les propriétés de complexation avec les acides aminés lysine et arginine ont été étudiées par microcalorimétrie. Il a été mis en évidence le rôle joué par les différents groupes fonctionnels : chaîne alkyle, groupements ammonium, hydroxyle et carboxylate. Selon le substrat complexé, différents types d'interactions peuvent être impliqués (ioniques, hydrophobes, van der Waals, $\pi - \pi$, cation$-\pi$, liaisons hydrogène). Au niveau fondamental, il est important de bien appréhender ces différents types d'interactions.

Dans ce but, des mesures microcalorimétriques ont été réalisées sur la complexation du calix[4]arène *p*-sulfonaté avec des substrats modèles alkylammonium, tétraalkylammonium, alcools aliphatiques, alkyldiammoniums, alkylguanidiniums, carboxylates et aminocarboxylates

en solution aqueuse [2, 3, 4]. Concernant l'interaction entre des espèces fortement chargées (interactions de type purement ioniques), l'association entre ce macrocycle et différents ions inorganiques (alcalino-terreux et lanthanides) a également été étudiée [1].

Il a été mis en évidence des comportements thermodynamiques très différents pour les ions organiques et inorganiques. Alors que les enthalpies libres standard de complexation sont du même ordre de grandeur pour ces deux types d'ions, les enthalpies et entropies de complexation diffèrent complètement. La complexation des ions inorganiques (complexe de sphère externe) est en effet fortement contrôlée par l'entropie ($\Delta_r H^o > 0$ et $T\Delta_r S^o \gg 0$) alors que la complexation des ions organiques (complexe d'insertion) est contrôlée par l'enthalpie ($\Delta_r H^o \ll 0$ et $T\Delta_r S^o <$ ou > 0). Les ions inorganiques sont impliqués dans la formation de paires d'ions avec les groupements sulfonate du calixarène, alors que pour les ions organiques, des groupements alkyles pénètrent dans la cavité du calixarène.

La caractérisation thermodynamique de ces types de complexes est essentielle pour identifier les facteurs qui contrôlent les processus de complexation. La thermodynamique fournit une grandeur macroscopique mais ne renseigne pas directement sur la structure du complexe. Une étude par simulation moléculaire a alors débuté sur ces complexes de façon à obtenir une interprétation moléculaire du processus de complexation. En effet, la simulation moléculaire par l'intermédiare de la mécanique statistique, permet d'avoir une vision microscopique du système et de calculer les grandeurs thermodynamiques.

1.2 Thermodynamique et mécanique statistique

La thermodynamique est l'étude de la relation entre les différentes propriétés du système à l'équilibre. C'est, avant tout, une science expérimentale, qui est très précieuse d'un point de vue pratique dans de nombreux domaines tels que la biologie, la chimie, la géologie, la physique et les sciences de l'ingénieur. Une des applications les plus importantes de la thermodynamique concerne l'analyse d'équilibres chimiques où elle permet de déterminer les conditions de température et de pression optimisant une réaction chimique. Les propriétés les plus importantes de la thermodynamique sont probablement la variation d'enthalpie libre ($\Delta_r G^o$), d'enthalpie

($\Delta_r H^o$) et d'entropie ($\Delta_r S^o$). Ces trois grandeurs permettent de caractériser l'état du système. La thermodynamique fut développée dans les années 1800, lorsque la théorie atomique de la matière ne faisait pas l'unanimité. Ainsi, les lois de la thermodynamique ne reposent sur aucune théorie atomique ou moléculaire, elles en sont toutes indépendantes. C'est, peut-être, en ce point que résident la force et la faiblesse de la thermodynamique, puisque ces lois sont invariantes par rapport à l'évolution de la théorie de la structure atomique ou moléculaire. Par contre, cette thermodynamique dite classique n'offre qu'une vision restreinte de la structure moléculaire.

Avec le développement de la théorie atomique et moléculaire au début des années 1900, on attribua à la thermodynamique une interprétation et une base moléculaire. Ainsi, est née la thermodynamique statistique, permettant de lier les moyennes de grandeurs moléculaires à des grandeurs thermodynamiques macroscopiques. La mécanique statistique vise précisement à décrire les systèmes macroscopiques. Le passage du microscopique au macroscopique, qui met en jeu un nombre incommensurable, le nombre d'Avogadro ($6.02\ 10^{23}\ \mathrm{mol}^{-1}$), ne peut être réalisé que par le biais de méthodes probabilistes. En effet, il est inconcevable d'écrire les 10^{23} équations de mouvement et encore plus de les résoudre. Il en résulte que les grandeurs observées ne sont en fait que des valeurs moyennes présentant des variations aléatoires (fluctuations statistiques). C'est Boltzmann qui a initié ces travaux en donnant une interprétation statistique à l'entropie thermodynamique par la relation $S = k \ln \Omega$, où Ω est le nombre d'états microscopiques. On peut démontrer que Ω est égal à la distribution la plus probable. Si le système est contraint en énergie et en nombre de particules, alors l'utilisation de la méthode des multiplicateurs de Lagrange et d'une statistique appropriée (Maxwell-Boltzmann, Fermi-Dirac,...) permet de déterminer la distribution la plus probable. Cette démarche permet ainsi de déterminer le nombre de systèmes dans un état j d'énergie $E_j(N, V)$. La probabilité qu'un système choisi aléatoirement soit dans l'état $E_j(N, V)$ est proportionnelle à la fonction de partition du système. Cette fonction de partition est une notion fondamentale en mécanique statistique, puisqu'elle va permettre de déterminer l'ensemble des fonctions thermodynamiques. En effet, à partir des relations $S = k \ln \Omega$ et $U = \sum_i N_i \epsilon_i$ (N_i et ϵ_i sont respectivement le nombre de particules et l'énergie du système i), il est possible de déter-

miner les fonctions thermodynamiques telles que A,G,U,H,P,V ou K[1]. Comme les valeurs observées expérimentalement sont des grandeurs moyennes en mécanique statistique, il faut donc pouvoir échantillonner de manière conséquente l'ensemble des états constituant le système.

La simulation numérique permet justement cet échantillonnage. Les moyennes définies en mécanique statistique sont purement statiques et sont définies sur l'ensemble des états accessibles au système. Or, cette moyenne ne correspond pas à la moyenne temporelle qu'on réalise en simulation de type dynamique moléculaire. En effet, dans ce dernier cas, on réalise des simulations durant un intervalle de temps fini et on détermine la moyenne sur ces simulations. C'est Gibbs qui a proposé de remplacer le calcul de la moyenne temporelle par celui de la moyenne d'ensemble. En effet, la moyenne temporelle s'écrit de la manière suivante :

$$\bar{A}(\mathbf{r}) = \frac{1}{P}\sum_{i=1}^{P} \lim_{t\to\infty} \frac{1}{t-t_i}\int_{t_i}^{t} A(\mathbf{r},t)dt'$$

où i est un indice qui parcourt un ensemble de conditions initiales et t_i le temps initial correspondant à l'indice i. Si l'ensemble des conditions initiales est compatible avec l'ensemble statistique approprié, on a alors :

$$\bar{A}(\mathbf{r}) = \lim_{t_1\to\infty} \frac{1}{t-t_1}\int_{t_i}^{t} < A(\mathbf{r},t) > dt'$$

Comme la moyenne d'ensemble ne dépend pas des conditions initiales, on peut écrire que :

$$\bar{A}(\mathbf{r}) =< A(\mathbf{r}) >$$

Cette égalité traduit le principe ergodique. Ainsi, en conjugant les techniques de simulation numérique et la mécanique statistique, il est possible de déterminer de nombreuses propriétés thermodynamiques. La simulation numérique peut être considérée comme la troisième méthodologie scientifique après la théorie et l'expérience. La simulation permet, par le biais de la

[1] A=énergie de Helmholtz, G=enthalpie libre, U=énergie interne, H=enthalpie, P=pression, V=volume, K=constante d'équilibre

mécanique statistique et/ou de la thermodynamique, de calculer des grandeurs directement comparables aux résultats expérimentaux. Cette confrontation permet de valider les modèles et les algorithmes utilisés. Une fois validées, les simulations permettent de donner une interprétation microscopique des grandeurs macroscopiques calculées. Il apparaît donc essentiel de mettre au point des méthodologies sûres et efficaces permettant l'évaluation correcte de ces grandeurs. En outre, ces techniques de simulation peuvent, dans certains cas, prendre le relais sur les mesures expérimentales (toxicité des produits utilisés, temps de mesure très longs, coût financier important, domaines de pression et de température trop élevés, etc.).Dans ce cas, on attribue alors un rôle prédictif à la simulation moléculaire.

1.3 Objectifs

Le premier travail que nous avons mené, dans le cadre d'un stage de DEA (A.Mendès) [5] sur le complexe formé entre le calix[4]arène *p*-sulfonaté et l'ion La^{3+} a montré que le complexe est à sphère de coordination externe c'est-à-dire un complexe où le ligand et le cation sont séparés par plusieurs molécules d'eau. La présence du calixarène ne perturbe que la seconde couche d'hydratation de La^{3+}. Disposant des valeurs expérimentales d'enthalpie libre standard de complexation, nous avons utilisé un cycle thermodynamique approprié afin de vérifier si les paramètres et les méthodologies utilisés dans ce travail respectent la sélectivité. Le calcul des différences d'enthalpie libre standard de complexation effectué à partir de deux formalismes de perturbation différents avec les cations La^{3+} et Eu^{3+} d'une part, et La^{3+} et Yb^{3+} d'autre part, donne un excellent accord avec les grandeurs obtenues expérimentalement par microcalorimétrie.

Enfin, les simulations ont montré une désolvatation partielle de la cavité du calixarène dans le complexe avec l'ion La^{3+} et une désolvatation totale avec les ions tétraalkylammoniums [5]. Nous avons également effectué des calculs de perturbation thermodynamique afin de calculer la variation d'enthalpie libre de complexation entre le complexe formé avec l'ion tétraméthylammonium (Me_4N^+) et celui formé avec l'ion méthylammonium ($MeNH_3^+$). Cette étude est très intéressante au niveau de la méthodologie des calculs de perturbation et d'intégration thermodynamique [6]. En effet, ce type de calcul nécessite de perturber le cation

Me_4N^+ en $MeNH_3^+$ seul en solution aqueuse et dans le complexe en solution aqueuse. Ces calculs de perturbation ont été réalisés en décomposant la perturbation en trois étapes : une étape pour les perturbations intramoléculaires (liaisons), une deuxième étape pour la perturbation des paramètres Lennard-Jones et une troisième pour la transformation des charges. Nous avons développé des algorithmes appropriés et montré l'importance de chacune de ces transformations en calculant leur contribution dans la variation d'enthalpie libre totale. Nous avons montré que la variation d'enthalpie libre de complexation calculée est en accord avec la valeur expérimentale, avec une déviation maximale inférieure à 13 %. C'est un excellent résultat compte tenu des différentes contributions à prendre en compte dans ce type de calcul. Nous montrons également que les contributions intramoléculaires sur le calcul de l'enthalpie libre de complexation ne peuvent pas être négligées et surtout qu'elles ne s'annulent pas forcément dans le cycle thermodynamique. Nous avons donné une interprétation moléculaire des différentes contributions et expliqué pourquoi la contribution due à la transformation des charges passe d'une valeur négative pour la mutation de l'ion seul à une valeur positive pour la transformation du cation dans le complexe. La présence des groupements sulfonatés réorganise la seconde couche d'hydratation du cation, ce qui a pour effet de positionner ces molécules d'eau avec les atomes d'hydrogène pointant vers ceux du cation méthylammonium avec une augmentation linéaire de l'énergie électrostatique cation-eau durant le processus de transformation des charges.

J'ai opté pour une présentation du mémoire en deux parties afin de dégager dans la première partie les principaux résultats avec les méthodologies correspondantes : dans la deuxième partie, sont décrits les différents algorithmes et les différentes techniques qui sont à la base de ce travail. Dans cette partie, nous revenons en détail sur la démonstration des formules utilisées par ces algorithmes.

Le deuxième chapitre de cette thèse est consacré à la présentation de la technique de la simulation moléculaire de type dynamique moléculaire et au protocole de simulation. Le *troisième* chapitre concerne la relation structure-énergie pour des complexes du calixarène avec des alcools linéaires. Dans le chapitre 4, nous discutons les méthodes permettant de déterminer $\Delta\Delta G$ et $\Delta\Delta S$. Nous présenterons une nouvelle méthode permettant un calcul plus

précis de $\Delta\Delta S$. Enfin, nous testerons l'ensemble de ces méthodes sur le calcul de propriétés d'hydratation et d'association. Le chapitre 5 se rapporte au calcul des grandeurs absolues à partir de la détermination du potentiel de force moyenne. Nous commencerons par établir les formules permettant de déterminer le potentiel de force moyenne et le grandeurs absolues. Nous étudierons l'influence de certaines approximations sur l'allure du profil du potentiel de force moyenne et sur le calcul de ΔG et ΔS. Nous appliquerons ces calculs à des complexes d'insertion et de sphère externe. Le chapitre 6 traite de la méthodologie mise au point pour tenir compte des interactions cation-π pour la détermination de potentiel de force moyenne. Cette méthodologie sera ensuite appliquée à la détermination des grandeurs absolues d'association entre la calixarène et l'ion césium. Le chapitre 7 regroupera les conclusions de ce travail

La deuxième partie débute avec le chapitre 8 où nous faisons des rappels de mécanique statistique. Le chapitre 9 revient sur certains algorithmes de dynamique moléculaire, notamment sur l'algorithme SHAKE qui constitue la base du calcul du potentiel de force moyenne par la méthode de la force contrainte. Enfin, le chapitre 10 est un bref rappel de la théorie de la DFT. Ce chapitre traite des origines de la fonctionnelle utilisée, B3LYP.

Première partie

Méthodologie et applications

Sommaire - I

2 **Méthode de dynamique moléculaire** 23

3 **Relation «micro-macro» : association entre le calixarène et les alcools linéaires** 49

5 **Calcul de $\Delta\Delta G$, $\Delta\Delta H$, $\Delta\Delta S$: application à des complexes de sphère externe** 69

4 **Potentiel de force moyenne et calcul de $\Delta_r G^o$, $\Delta_r H^o$ et $\Delta_r S^o$: application à des complexes d'insertion et de sphère externe** 113

6 **Méthodologie pour la prise en compte des interactions cations$-\pi$** . . 151

7 **Conclusions** 173

Chapitre 2

Méthode de dynamique moléculaire

2.1 Introduction

Tout système naturel ou synthétique est composé de molécules constituées d'atomes dont la structure interne est faite de noyaux et d'électrons. Le comportement de ces entités est en principe décrit par la mécanique quantique relativiste présentée à partir des postulats suivants :

1 - **Hypothèse de L. De Broglie (1923)** : à toute particule de masse m et de vitesse v, il est associé un phénomène ondulatoire dont la longueur d'onde λ est donnée par la relation $\lambda = \frac{h}{p}$ où p est la quantité de mouvement telle que $p = mv$.

2 - **Postulat de M. Born (1926)** : la description complète du comportement d'un électron peut se faire par une fonction mathématique ψ appelée fonction d'onde définie par l'équation de Schrödinger. Le carré de cette fonction $\psi^2_{x,y,z}$ au point x,y,z représente la probabilité de trouver l'électron en ce point.

3 - **Relation de Heisenberg** : ce principe marque la limite de validité des concepts macroscopiques, de particules et d'ondes. En effet, elle stipule que la détermination simultanée sur un même système, de deux grandeurs A et B incompatibles, est affectée d'une incertitude intrinsèque telle que le produit des écarts-types $\Delta A \Delta B$ ne peut en aucune façon être inférieur à une limite finie qui dépend de la nature de A et B.

4 - **Postulat de Schrödinger** : l'équation de Schrödinger relie la fonction ψ permettant d'accéder à la probabilité de présence de l'électron et aux valeurs quantifiées de l'énergie qu'il peut avoir (équation (2.1)). Ainsi, la mécanique quantique est essentiellement basée sur la résolution de cette équation.

$$\begin{aligned} \hat{\mathcal{H}}\psi &= E\psi \qquad (2.1) \\ \text{tel que } \hat{\mathcal{H}} &= -\sum_{i=1}^{N}\frac{1}{2}\nabla_i^2 - \sum_{K=1}^{M}\frac{1}{2}\frac{\nabla_K^2}{M_K} - \sum_{i=1}^{N}\sum_{A=1}^{M}\frac{Z_A}{r_{iA}} + \sum_{i=1}^{N}\sum_{j>i}^{N}\frac{1}{r_{ij}} + \sum_{A=1}^{M}\sum_{B>A}^{M}\frac{Z_A Z_A}{r_{AB}} \end{aligned}$$

Dans cette équation, $\hat{\mathcal{H}}$ est l'opérateur hamiltonien, ψ est une fonction d'onde, E l'énergie propre du système. Le premier terme de l'expression de l'hamiltonien correspond à l'énergie cinétique des électrons, le deuxième terme est l'énergie cinétique des noyaux. Dans les deux cas, l'opérateur laplacien correspond à l'opérateur énergie cinétique où M_K est la masse du noyau K. Le troisième terme correspond aux interactions entre les noyaux et les électrons où Z_A est le numéro atomique du noyau A et r_{iA} la distance séparant le noyau de l'électron. Le quatrième terme est relatif aux interactions entre électrons, r_{ij} est la distance entre deux électrons. Le dernier terme correspond aux interactions entre noyaux. Dans cette expression, M est le nombre de noyaux et N le nombre d'électrons. Ces postulats sont justifiés parce que le développement logique du formalisme mathématique qu'ils définissent et mettent en œuvre conduit à des résultats de calculs cohérents avec l'observation. Ainsi, se dessine le premier objectif de tout calcul et de toute théorie : la confrontation avec l'expérience. Dans cette optique, la mécanique quantique a très vite atteint ses limites notamment en terme de nombre de calculs, augmentant de manière exponentielle avec la taille du système, réduisant ainsi le champ d'investigation de cette discipline à celui de systèmes de petite taille. Par ailleurs, cette mécanique permet seulement l'étude et le calcul des propriétés statiques et non dynamiques à l'origine de nombreuses propriétés macroscopiques. Finalement, il a fallu mettre en œuvre des méthodes permettant l'accès à ces grandeurs. Le point de départ est évidemment la mécanique quantique. En effet, les approximations suivantes et successives ont permis de réduire considérablement le nombre de calculs :

1 - Les vitesses des particules sont faibles par rapport à celle de la lumière, ce qui nous permet de sortir du cadre relativiste.

2 - L'approximation de Born-Oppenheimer selon laquelle le mouvement des électrons est beaucoup plus rapide que celui des noyaux.

L'hamiltonien de l'équation (2.1) peut être écrit comme :

$$\hat{\mathcal{H}} = \mathbf{T}_N + \mathbf{H}^{'} \tag{2.2}$$

$\mathbf{T}_N$ repésente l'opérateur énergie cinétique des noyaux et $\mathbf{H}^{'}$ les termes restants de l'équation (2.1). Soit $\psi_T(..., \mathbf{r}_i, ..., \mathbf{r}_k, ...)$ la fonction d'onde du système électron(i)-noyau(k)

$$\hat{\mathcal{H}}\psi_T(..., \mathbf{r}_i, ..., \mathbf{r}_k, ...) = \mathcal{E}\psi_T(..., \mathbf{r}_i, ..., \mathbf{r}_k, ...) \tag{2.3}$$

Comme le suggère le principe de superposition des états, il est possible de développer ψ_T dans une base orthogonale de fonctions propres ψ_j de $\mathbf{H}^{'}$ et de fonctions propres $\Theta_j(..., \mathbf{r}_k, ...)$ de l'hamiltonien nucléaire. Les propriétés sur les opérateurs nous permettent alors d'écrire :

$$\psi_T = \sum_{j=1}^{\infty} \Theta_j(..., \mathbf{r}_k, ...)\psi_j(..., \mathbf{r}_i, ..., \mathbf{r}_k, ...) \tag{2.4}$$

où les $\psi_j(..., \mathbf{r}_i, ..., \mathbf{r}_k, ...)$ vérifient l'équation :

$$\mathcal{H}^{'}\psi_j(..., \mathbf{r}_i, ..., \mathbf{r}_k, ...) = U_j\psi_j(..., \mathbf{r}_i, ..., \mathbf{r}_k, ...) \tag{2.5}$$

Dans ces conditions, l'équation (2.3) peut s'écrire :

$$(\mathbf{T}_N + \mathbf{H}^{'}) \sum_{j=1}^{\infty} \Theta_j(..., \mathbf{r}_k, ...)\psi_j(..., \mathbf{r}_i, ..., \mathbf{r}_k, ...) = \mathcal{E} \sum_{j=1}^{\infty} \Theta_j(..., \mathbf{r}_k, ...)\psi_j(..., \mathbf{r}_i, ..., \mathbf{r}_k, ...)$$

Le premier terme de cette expression peut se développer sous la forme :

$$-\sum_K \frac{1}{2M_K} \sum_{j=1}^{\infty} [\psi_j \boldsymbol{\Delta}_K \Theta_j + 2\nabla_K \Theta_j \cdot \nabla_K \psi_j + \Theta_j \boldsymbol{\Delta}_K \psi_j] + \sum_j U_j \Theta_j \psi_j$$

Le premier terme correspond à l'opérateur $\mathbf{T}_N$ tandis que le second correspond à $\mathbf{H}^{'}$. Le premier terme s'obtient en dérivant deux fois la somme de fonctions propres ($d^2ab = bd^2a + ad^2b + 2dadb$). Le deuxième terme est obtenu à partir de la relation (2.5). Si l'on multiplie l'équation précédente par la fonction ψ_p^* et si l'on intègre par rapport aux variables électroniques r_i, on obtient la relation suivante :

$$\mathbf{T}_N\Theta_p - \sum_K \frac{1}{2M_K}\left[\sum_j \langle\psi_p|\nabla_K|\psi_j\rangle .\nabla_K\Theta_j + \sum_j \Theta_j \langle\psi_p|\mathbf{T}_N|\psi_j\rangle\right] + U_p\Theta_p = \mathcal{E}\Theta_p$$

L'approximation de ***Born-Oppenheimer*** consiste à considérer que la masse des noyaux est infiniment plus grande que celle des électrons. Par conséquent, les noyaux se déplacent avec une faible amplitude. Ainsi, toute intégrale électronique contenant le terme cinétique devient négligeable.

$$\langle\psi_p|\nabla_K|\psi_j\rangle = 0.0 \quad \text{et} \quad \langle\psi_p|\mathbf{T}_N|\psi_j\rangle = 0.0$$

Le système d'équations prend alors une forme simple :

$$\begin{aligned} & \mathbf{T}_N\Theta_p + U_p\Theta_p = \mathcal{E}\Theta_p \\ \rightarrow \quad & \mathcal{E}\Theta_p = \mathcal{H}^{BO}\Theta_p \ \text{ tel que } \ \mathcal{H}^{BO} = \mathbf{T}_N + U_p \end{aligned} \tag{2.6}$$

L'utilisation de cet hamiltonien permet de diminuer le nombre de calculs à effectuer puisque les termes électroniques disparaissent. Finalement, l'approximation de Born-Oppenheimer permet de considérer tout système par la position $\mathbf{r}_K$ et l'impulsion $\mathbf{p}_K$ des atomes le constituant. Si l'on considère ces atomes comme étant des points matériels[1] définis par ($\mathbf{r}_K$, $\mathbf{p}_K$), on peut alors appliquer les principes de la physique newtonnienne. En effet, la dynamique du point matériel, dont les lois physiques sur lesquelles elle s'appuie, ont été énoncées partiellement par G. Galilei en 1632 et complètement par I. Newton en 1687 dans les célèbre *Principia Mathematica*, permet de décrire l'espace des phases d'un système ($\mathbf{r}_K$, $\mathbf{p}_K$). Cette description est réalisable grâce notamment à la deuxième loi de Newton (loi fondamentale de la dynamique) :

[1] Un point matériel est défini *cinématiquement* par un seul vecteur vitesse et un seul vecteur accélération et *dynamiquement* par un scalaire appelé masse inerte qui est constant (au cours du temps) et invariant (par changement de référentiel).

Enoncé historique : Le changement de mouvement est proportionnel à la force imprimée et s'effectue suivant la droite par laquelle cette force est imprimée.

Enoncé actuel : Le mouvement d'un *point matériel* A, dans un référentiel galiléen $\mathcal{R}$, de masse m soumis à plusieurs forces dont la somme est $\sum \mathbf{F}$, satisfait à la relation :

$$\frac{d\mathbf{p}}{dt} = \sum \mathbf{F} \quad \text{soit} \quad m\mathbf{a}_{A/\mathcal{R}} = \sum \mathbf{F} \tag{2.7}$$

Ainsi, le potentiel donné par l'équation (2.6) et la deuxième loi de Newton permettent, en théorie, de définir le mouvement d'un atome dans un référentiel galiléen. En effet, si les forces exercées sur l'atome sont conservatives, c'est-à-dire dérivant d'une énergie potentielle, on a alors :

$$m\mathbf{a}_{A/\mathcal{R}} = \sum \mathbf{F} = -\sum \nabla U_p = \frac{d\mathbf{p}_{A/\mathcal{R}}}{dt} = m\frac{d^2\mathbf{r}_{A/\mathcal{R}}}{dt^2} \tag{2.8}$$

Cette relation montre de manière explicite que la connaissance du potentiel et des conditions initiales permet de décrire l'espace des trajectoires de l'atome puisque deux intégrations successives permettent de déterminer la vitesse et la position de cet atome. On peut, ainsi, presque déterminer la trajectoire d'un atome «sur le papier». Néanmoins, étant donnée l'avancée extraordinaire de l'informatique ces cinquantes dernières années, il est possible d'étudier des systèmes moléculaires de plus grande taille (10^3 à 10^6 atomes). Ce type de calcul porte le nom de *simulation numérique de type dynamique moléculaire*.

Finalement, ces méthodes prennent leur source théorique dès le $17^{\text{ème}}$ siècle avec Newton. Mais, ce n'est qu'au début de 1950, lorsque les premiers ordinateurs ont pu être utilisés pour un usage civil, que prennent réellement naissance ces méthodes de simulation[2]. La seule détermination de l'ensemble ($\mathbf{r}^N$, $\mathbf{p}^N$) ne permet pas à elle seule d'accéder aux propriétés mécaniques, thermodynamiques, etc. Il faut, en effet, coupler la dynamique de Newton à la mécanique statistique de L. Boltzmann (le chapitre (8) revient sur les bases de cette mécanique). Ce couplage permet finalement d'avoir une vue microscopique et macroscopique d'un système donné. La recherche et l'étude de cette relation «*micro-macro*» sont devenues un élément essentiel pour la compréhension de tout phénomène physico-chimique. Ces méthodes se situent à une échelle spatiale de l'ordre du nanomètre et une échelle de temps allant de

[2]En particulier à Los Alamos, la machine MANIAC est devenue opérationnelle en 1952.

la picoseconde à la nanoseconde comme le montre la figure (2.1). Tout système se situant sur des échelles de temps et d'espace différentes implique l'utilisation de méthodes basées sur de nouvelles approximations. La figure (2.1) donne un aperçu de ces méthodes et des limites d'espace et de temps pour leurs applications.

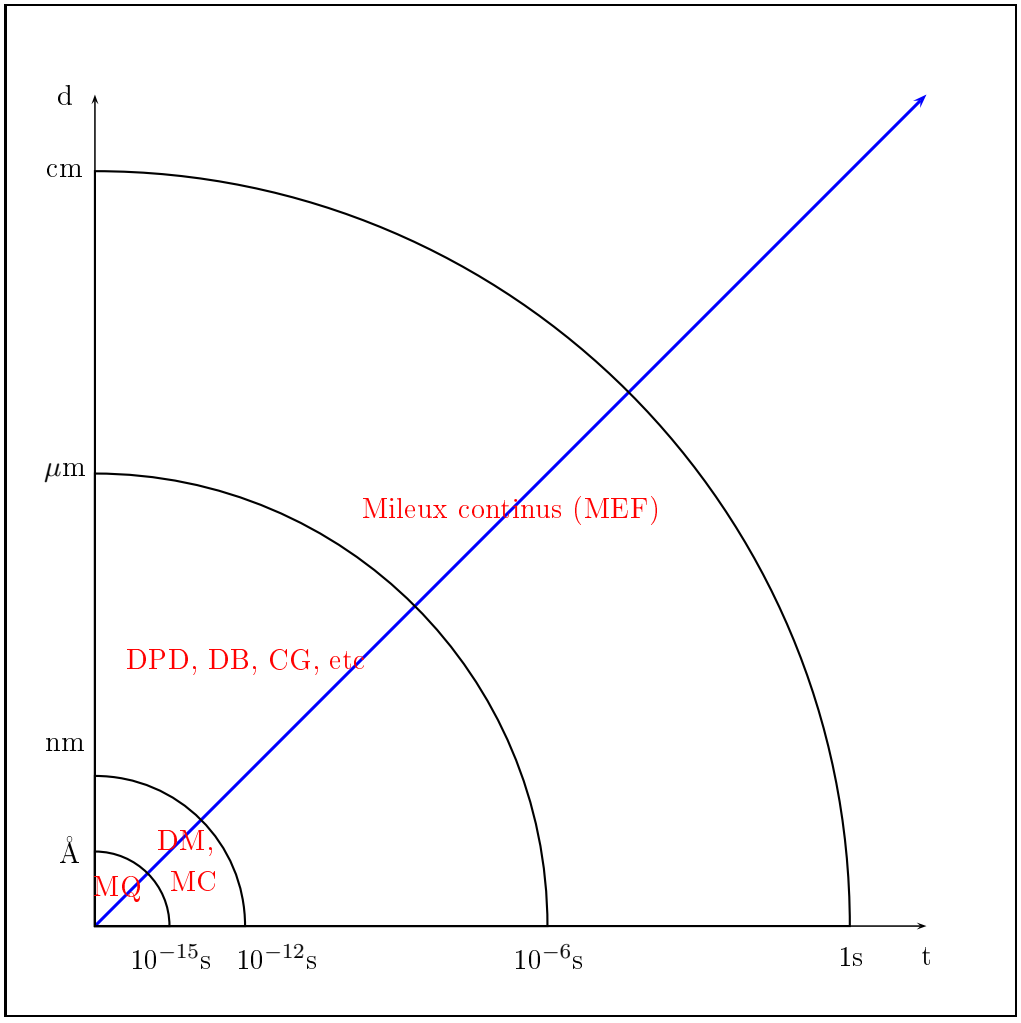

FIG. 2.1 – Méthodes de simulation utilisées en fonction de l'échelle spatiale (d) et temporelle (t). (MQ)=Mécanique quantique, (DM)=Dynamique moléculaire, (MC)=Monte Carlo, (DPD)=Dynamique des particules dissipatives, (DB)=Dynamique brownienne, (MEF)=Méthodes des éléments finis, (CG)=*Coarse graining*. (→) indique le sens d'évolution dont il est question dans la discussion.

Le passage d'une méthode à l'autre dans le sens indiqué sur la figure (2.1) se fait par le biais d'approximations successives. La conséquence est une perte de détail dans l'observa-

tion et l'étude du système. Par contre, on a un gain d'efficacité (temps de calcul plus faible) permettant l'accès à de nouvelles propriétés. La méthode de dynamique moléculaire permet plusieurs types de descriptions permettant de diminuer le temps de calcul. Il y a la description dite explicite où chaque atome est representé par un point matériel. Il existe également une description dite «atome unifié» où c'est un groupe d'atomes que l'on représente par un pseudo-atome. À une échelle plus élevée, c'est une région du fluide qui est décrite de cette manière. À mi-chemin entre les méthodes quantiques et celles de dynamique moléculaire, on trouve des méthodes hybrides : les méthodes de dynamique quantique (*ab initio*, DFT, ...) et celle de dynamique mixte (QM/MM) où les électrons sont prise en compte de manière explicite.

2.2 Mécanique analytique : lagrangien et hamiltonien

Le mouvement d'une particule, et plus généralement d'un système matériel, peut être étudié à partir d'un formalisme développé par les mathématiciens français A. Lagrange et écossais W. Hamilton. Ce formalisme s'appuie essentiellement sur deux fonctions scalaires ayant la dimension d'une énergie : le *lagrangien* $\mathcal{L}$ et l'*hamiltonien* $\mathcal{H}$. Son intérêt principal est qu'il se prête mieux à l'extension de la mécanique et des autres domaines de la physique (optique, physique, quantique, etc.). En outre, il permet de trouver plus facilement les grandeurs physiques qui se conservent (comme on le verra dans la section (9.1.3)).

2.2.1 Lagrangien

Considérons un système matériel $\{\mathbf{r}_i, \dot{\mathbf{r}}_i\}$ en mouvement par rapport à un référentiel galiléen, sous l'action de forces extérieures et intérieures dérivant d'une énergie potentielle totale $\mathcal{E}_p$. On définit le lagrangien du système comme une fonction dépendante des positions et de leurs dérivées par rapport au temps :

$$\mathcal{L}\left(\{\mathbf{r}_i\}, \{\dot{\mathbf{r}}_i\}\right) = \sum_i^N \frac{\dot{\mathbf{r}}_i \dot{\mathbf{r}}_i m_i}{2} - \sum_i^N \sum_{j>i} u(r_{ij}) \quad \text{tel} \quad \text{que } U(\{\mathbf{r}_i\}) = \sum_i^N \sum_{j>i} u(r_{ij})$$

Cette définition n'est valable qu'en mécanique newtonienne. On en déduit alors le moment conjugué $\mathbf{p}_i$ associé à $\mathbf{r}_i$ par la relation suivante :

$$\mathbf{p}_i = \frac{\partial \mathcal{L}\left(\{\mathbf{r}_i\}, \{\dot{\mathbf{r}}_i\}\right)}{\partial \dot{\mathbf{r}}_i} = m\dot{\mathbf{r}}_i$$

Dans certains cas, le moment conjugué est assimilé à la quantité de mouvement. Cela nous permet d'écrire les équations de Lagrange du système soumis uniquement à des forces qui dérivent d'une énergie potentielle :

$$\frac{d\mathbf{p}_i}{dt} - \frac{\partial \mathcal{L}\left(\{\mathbf{r}_i\}, \{\dot{\mathbf{r}}_i\}\right)}{\partial \mathbf{r}_i} = 0$$

2.2.2 Hamiltonien

Définition

Contrairement au lagrangien, l'espace décrit par l'hamiltonien est l'espace des phases constitué de $2N$ variables indépendantes. Dans la théorie d'Hamilton, on s'intéresse plutôt à l'état mécanique du système. On définit l'hamiltonien comme :

$$\mathcal{H}\left(\{\mathbf{r}_i\}, \{\mathbf{p}_i\}\right) = \sum_i^N \dot{\mathbf{r}}_i \mathbf{p}_i - \mathcal{L}\left(\{\mathbf{r}_i\}, \{\dot{\mathbf{r}}_i\}\right) \tag{2.9}$$

Cette transformation, très utilisée en thermodynamique, est connue sous le nom de transformation de Legendre qui permet de passer du lagrangien, fonction associée aux variables $(\mathbf{r}_i, \dot{\mathbf{r}}_i)$ à l'hamilonien, fonction associée aux variables $(\mathbf{r}_i, \mathbf{p}_i)$. On montre que l'hamiltonien d'un système naturel (énergie potentielle ne dépendant pas du temps) n'est autre que son énergie totale. Dans ce contexte de système naturel, il en résulte que les coordonnées d'un point A quelconque du système ne sont fonction que des paramètres de configuration $\{\mathbf{r}_i\}$. Si O est l'origine du référentiel, la vitesse de A notée $\dfrac{d\mathbf{OA}}{dt}$ vaut donc :

$$\mathbf{v}_A = \frac{d\mathbf{OA}}{dt} = \sum_i \left(\frac{\partial \mathbf{OA}}{\partial \mathbf{r}_i}\right) \frac{d\mathbf{r}_i}{dt} = \sum_i \left(\frac{\partial \mathbf{OA}}{\partial \mathbf{r}_i}\right) \dot{\mathbf{r}}_i$$

La forme quadratique de l'énergie cinétique est :

$$\mathcal{E}_k = \frac{1}{2} m_A v_A^2 = \sum_i \sum_j \left(\frac{\partial \mathbf{OA}}{\partial \mathbf{r}_i}\right) \left(\frac{\partial \mathbf{OA}}{\partial \mathbf{r}_j}\right) \dot{\mathbf{r}}_i \dot{\mathbf{r}}_j = \sum_i \sum_j a_{ij} \dot{\mathbf{r}}_i \dot{\mathbf{r}}_j$$

avec $a_{ij} = \left(\frac{\partial \mathbf{OA}}{\partial \mathbf{r}_i}\right)\left(\frac{\partial \mathbf{OA}}{\partial \mathbf{r}_j}\right)$.

L'énergie cinétique ne dépend donc pas explicitement du temps. Comme, $\mathcal{E}_p = \mathcal{E}_p(\{\mathbf{r}_i\})$, il vient :

$$\mathbf{p}_i = \frac{\partial \mathcal{L}\left(\{\mathbf{r}_i\},\{\dot{\mathbf{r}}_i\}\right)}{\partial \dot{\mathbf{r}}_i} = \frac{\partial \mathcal{E}_k\left(\{\dot{\mathbf{r}}_i\}\right)}{\partial \dot{\mathbf{r}}_i} = \sum_j 2a_{ij}\dot{\mathbf{r}}_j \quad \text{et} \quad \sum_i \mathbf{p}_i\dot{\mathbf{r}}_i = 2\sum_i\sum_j a_{ij}\dot{\mathbf{r}}_i\dot{\mathbf{r}}_j = 2\mathcal{E}_k$$

d'où

$$\begin{aligned} \mathcal{H}\left(\{\mathbf{r}_i\},\{\mathbf{p}_i\}\right) &= \sum_i^N \dot{\mathbf{r}}_i\mathbf{p}_i - \mathcal{L}\left(\{\mathbf{r}_i\},\{\dot{\mathbf{r}}_i\}\right) = \mathcal{E}_k + \mathcal{E}_p \\ \text{car} \quad \mathcal{L}\left(\{\mathbf{r}_i\},\{\dot{\mathbf{r}}_i\}\right) &= \mathcal{E}_k - \mathcal{E}_p \end{aligned} \tag{2.10}$$

La relation (2.10) permet de justifier l'utilisation de l'hamiltonien tel qu'on le connaît en dynamique moléculaire, puisqu'elle est équivalente à l'équation (2.6). À partir de ces relations, il est possible de déterminer les équations de mouvement ou équations canoniques.

Équations de mouvement

Pour déterminer ces équations, il suffit d'écrire la différentielle de la fonction d'Hamilton de deux manières différentes, en considérant $\mathcal{H}$ comme fonction de $2N$ variables $(\{\mathbf{r}_i\},\{\mathbf{p}_i\})$ et indépendantes du temps :

$$d\mathcal{H} = \sum_i \left(\frac{\partial \mathcal{H}}{\partial \mathbf{r}_i}d\mathbf{r}_i + \frac{\partial \mathcal{H}}{\partial \mathbf{p}_i}d\mathbf{p}_i\right) \tag{2.11}$$

À présent, différencions $\mathcal{H}$ à partir de la relation (2.9) :

$$\begin{aligned} d\mathcal{H} &= d\sum_i \mathbf{p}_i\dot{\mathbf{r}}_i - d\mathcal{L} = \sum_i (\mathbf{p}_i d\dot{\mathbf{r}}_i + \dot{\mathbf{r}}_i d\mathbf{p}_i) - d\mathcal{L} \\ &= \sum_i (\mathbf{p}_i d\dot{\mathbf{r}}_i + \dot{\mathbf{r}}_i d\mathbf{p}_i) - \sum_i \left(\frac{\partial \mathcal{L}}{\partial \mathbf{r}_i}d\mathbf{r}_i + \frac{\partial \mathcal{H}}{\partial \dot{\mathbf{r}}_i}d\dot{\mathbf{r}}_i\right) \end{aligned}$$

Comme $\mathbf{p}_i = \dfrac{\partial \mathcal{L}}{\partial \dot{\mathbf{r}}_i}$ et $\dot{\mathbf{p}}_i = \dfrac{\partial \mathcal{L}}{\partial \mathbf{r}_i}$, il vient :

$$d\mathcal{H} = \sum_i (\dot{\mathbf{r}}_i d\mathbf{p}_i - \dot{\mathbf{p}}_i d\mathbf{r}_i) \tag{2.12}$$

En identifiant les relations (2.11) et (2.12), on en déduit les équations du premier ordre en $\mathbf{p}_i$ et $\mathbf{r}_i$, appelées équations canoniques du mouvement :

$$\frac{d\mathbf{r}_i}{dt} = \frac{\partial \mathcal{H}}{\partial \mathbf{p}_i} \quad \text{et} \quad \frac{d\mathbf{p}_i}{dt} = -\frac{\partial \mathcal{H}}{\partial \mathbf{r}_i} = -\sum_i \sum_{j>i} \hat{\mathbf{r}}_{ij} u'(r_{ij}) \tag{2.13}$$

$u'(r_{ij})$ n'est autre que la force agissant sur l'atome i. Ces équations associées à l'hamiltonien (équation (2.10)) seront utilisées pour explorer l'espace des phases lors des simulations.

Conservation de l'énergie

Pour un système matériel en mouvement par rapport à un référentiel galiléen, l'hamiltonien $\mathcal{H}$ ne dépend pas explicitement du temps. On en déduit :

$$\frac{d\mathcal{H}}{dt} = \frac{\partial \mathcal{H}}{\partial t} = -\frac{\partial \mathcal{L}}{\partial t} = 0 \quad \text{d'où} \quad \mathcal{H} = \text{constante}$$

La *conservation de l'énergie* est donc reliée à l'invariance par rapport au temps de l'hamiltonien et du lagrangien.

Conservation des moments conjugués

Lorsque l'une des variables $\mathbf{r}_i$ n'apparaît pas dans l'hamiltonien, le lagrangien ou l'énergie potentielle, on dit que que cette variable est cyclique ou cachée. D'après les équations de Lagrange, le moment conjugué $\mathbf{p}_i$ associé est alors une constante du mouvement :

$$\frac{\partial \mathcal{L}}{\partial \mathbf{r}_i} = \frac{d\mathbf{p}_i}{dt} = 0 \quad \text{d'où} \quad \mathbf{p}_i = \text{constante}$$

Ayant déterminé les équations canoniques du mouvement et la forme de l'hamiltonien, il ne reste plus qu'à résoudre ces équations.

2.3 Résolution des équations canoniques du mouvement

La méthode la plus simple pour résoudre les équations de mouvement données par l'équation (2.13) est de combiner un développement en série de Taylor et des différences finies.

2.3.1 Différences finies

Avant de discuter les algorithmes, il est nécessaire de rappeler les lois d'invariance des phénomènes physiques. Ces dernières impliquent des lois de conservations qui jouent un rôle fondamental en mécanique. Ainsi, un des critères de validation d'un algorithme de dynamique moléculaire est le respect inconditionnel de ces lois d'invariance qui sont :

1- Conservation de l'énergie
2- conservation des moments conjugués[3]
3- respect du renversement du temps[4] (condition de symétrie contenue dans les équations de Newton).
4- conservation du volume dans l'espace des phases (on dit alors que la simulation est symplectique).

Il existe plusieurs algorithmes permettant de résoudre ces équations, mais tous sont basés sur des développements en séries de Taylor (discrétisation en temps). En effet, dans ce cas, on évalue les dérivées $\frac{d\mathbf{r}_i}{dt}$ et $\frac{d\mathbf{p}_i}{dt}$ par des différences finies du deuxième et du premier ordre à deux points à droite. On résoud, ensuite, pas à pas ces équations en choisissant un pas d'intégration Δt. Comme pour toutes les méthodes de résolution numérique d'équations différentielles, plus Δt sera petit et plus la solution sera précise et exacte. En outre, Δt doit être inférieur au plus petit temps de relaxation.

$$\mathbf{r}_{n+1} = \mathbf{r}_n + \mathbf{v}_n \Delta t + \frac{\mathbf{f}_n}{2m}(\Delta t)^2 \tag{2.14}$$

$$\mathbf{v}_{n+1} = \mathbf{v}_n + \frac{\mathbf{f}_n}{m}\Delta t \tag{2.15}$$

[3] $\mathbf{v}_1 + \mathbf{v}_2 = \mathbf{v}_1^{'} + \mathbf{v}_2^{'}$ où $\mathbf{v}_i^{'}$ indique la vitesse de la particule i au temps suivant.
[4] Cela revient à dire que si, à l'instant t, on inverse le cours du temps, la trajectoire de la dynamique revient sur ses pas.

Cette solution est considérée comme déficiente puisqu'elle ne respecte pas le renversement du temps. En effet, en changeant Δt en $-\Delta t$ et $n+1$ en $n-1$, les équations précédentes sont différentes. Cette solution, appelée solution d'Euler, peut être améliorée si on considère des dérivées premières et secondes en trois points, comme le montrent les deux relations de l'équation (2.16) :

$$\begin{aligned} \mathbf{r}_{n+1} &= 2\mathbf{r}_n - \mathbf{r}_{n-1} + \frac{\mathbf{f}_n}{2m}(\Delta t)^2 \\ \mathbf{v}_{n+1} &= \frac{\mathbf{r}_{n+1} - \mathbf{r}_{n-1}}{2\Delta t} \end{aligned} \tag{2.16}$$

L'amélioration réside dans le type de variation de l'erreur en Δt. En effet, dans un développement en deux points, l'erreur varie linéairement en Δt c'est-à-dire $\mathcal{O}(\Delta t)$. Le développement en trois points permet une variation d'ordre 4 ($\mathcal{O}(\Delta t^4)$) et d'ordre 2 ($\mathcal{O}(\Delta t^2)$) de l'erreur pour les positions et les vitesses, respectivement. Les positions sont obtenues en développant à l'ordre 4 $\mathbf{r}_{n+1}$ et $\mathbf{r}_{n-1}$ et en les additionnant, tandis que pour les vitesses on les développe à l'ordre 2 et on les soustrait. On pourrait aussi développer en quatre points, abaissant ainsi l'erreur mais augmentant le nombre de calculs. En effet, 99 % du temps de calcul en dynamique moléculaire est consacré aux calcul des forces puisque l'on doit évaluer $\frac{n(n-1)}{2}$ interactions. Avec la méthode en quatre points, il faut calculer quatre fois ces forces soit $\dfrac{4n(n-1)}{2}$ contre $\dfrac{2n(n-1)}{2}$ avec un développement en trois points (Eq.(2.17)).

$$\begin{aligned} f^{'}(x) &= \frac{-f(x+2h) + 8f(x+h) - 8f(x-h) + f(x-2h)}{12h} + \mathcal{O}(h^4) \\ f^{''}(x) &= \frac{-f(x+2h) + 16f(x+h) - 30f(x) + 16f(x-h) - f(x-2h)}{12h^2} + \mathcal{O}(h^4) \end{aligned} \tag{2.17}$$

Finalement, ce choix n'est probablement pas le meilleur même si on peut augmenter le pas d'intégration. En outre, on observe une dérive énergétique aux temps longs. Le développement en trois points semble être un bon compromis entre le temps de calcul, la minimisation de l'erreur et une dérive énergétique faible aux temps longs. Ces relations bien qu'existantes depuis le 17$^{\text{ème}}$ siècle (développées par le mathématicien anglais B. Taylor (1685-1731)), portent le nom d'*Algorithme de Verlet* initialement conçu par Verlet [7] et développé par Störmer [8]. Cet algorithme, contrairement à celui d'Euler, est réversible puisque $\mathbf{r}_{n+1}$ et $\mathbf{r}_{n-1}$ ont un rôle

symétrique. Les principaux inconvénients de cet algorithme sont :

1- La détermination des positions sans la prise en compte des vitesses

2- une faible précision sur la détermination des vitesses.

Pour remédier à ces déficiences, R.W Hockney [9] a proposé l'algorithme de *Verlet Leap Frog* basé sur le fait que les vitesses sont calculées pour des intervalles de temps demi-entiers et les positions sont obtenues pour les intervalles de temps entiers. Si on définit les vitesses pour les temps $(t + \Delta t/2)$ et $(t - \Delta t/2)$ par des taux d'accroissements :

$$\begin{aligned} \mathbf{v}(t + \frac{\Delta t}{2}) &= \frac{\mathbf{r}(t + \Delta t) - \mathbf{r}(t)}{\Delta t} \\ \mathbf{v}(t - \frac{\Delta t}{2}) &= \frac{\mathbf{r}(t) - \mathbf{r}(t - \Delta t)}{\Delta t} \end{aligned}$$

on obtient alors les positions aux temps $(t + \Delta t)$ et $(t - \Delta t)$:

$$\begin{aligned} \mathbf{r}(t + \Delta t) &= \mathbf{r}(t) + \mathbf{v}(t + \frac{\Delta t}{2})\Delta t \\ \mathbf{r}(t - \Delta t) &= \mathbf{r}(t) - \mathbf{v}(t - \frac{\Delta t}{2})\Delta t \end{aligned}$$

En utilisant l'équation principale de l'algorithme de Verlet, on a alors :

$$\begin{aligned} \mathbf{v}(t + \frac{\Delta t}{2}) &= \mathbf{v}(t - \frac{\Delta t}{2}) + \frac{\mathbf{f}(\Delta t)}{2m}\Delta t + \mathcal{O}((\Delta t)^4) \\ \mathbf{v}(t) &= \frac{1}{2}(\mathbf{v}(t + \frac{\Delta t}{2}) + \mathbf{v}(t - \frac{\Delta t}{2})) \end{aligned} \tag{2.18}$$

L'avantage de cette méthode est la prise en compte des vitesses dans le calcul des positions et l'amélioration de la précision du calcul des vitesses. L'inconvénient majeur de cette méthode est l'absence de simultanéité dans le calcul des positions et des vitesses ainsi que l'approximation faite sur le calcul des vitesses à l'instant t. Néanmoins, cette méthode est probablement la méthode la plus employée actuellement. C'est d'ailleurs celle qu'on utilisera dans les simulations afin de respecter l'algorithme original de Berendsen [10] permettant de maintenir une pression et une température constantes. Pour résoudre ces problèmes, Swope *et al.*

[11] ont proposé une méthode appelée *Velocity Verlet*, permettant de calculer les vitesses et les positions en même temps. Les équations utilisées (données dans l'ordre d'éxécution dans l'algorithme) sont les suivantes :

$$\begin{aligned}
\mathbf{r}(t+\Delta t) &= \mathbf{r}(t) + \mathbf{v}(t)\Delta t + \frac{\mathbf{f}(\Delta t)}{2m}(\Delta t)^2 \\
\mathbf{v}(t+\frac{\Delta t}{2}) &= \mathbf{v}(t) + \frac{\mathbf{f}(\Delta t)}{2m}\Delta t \\
\mathbf{f}(t+\Delta t) &= \mathbf{f}(\mathbf{r}(t+\Delta t)) \\
\mathbf{v}(t+\Delta t) &= \mathbf{v}(t+\frac{\Delta t}{2}) + \frac{\mathbf{f}(t+\Delta t)}{2m}\Delta t
\end{aligned}$$

Cet algorithme permet donc de calculer en même temps les vitesses, les positions ainsi que les accélérations. En outre, il est stable aux temps longs (faibles dérives d'énergie). L'inconvénient de cette méthode est le calcul des forces à (t) et $(t+\Delta t)$ à la même itération. Ainsi, le nombre d'interactions à calculer est deux fois plus grand que celui calculé dans l'algorithme de *Verlet Leap Frog* soit $\frac{n(n-1)}{2}$ calculs supplémentaires. Finalement, le compromis idéal entre temps et précision des calculs est l'algorithme de *Verlet Leap Frog*, ce qui explique son abondante utilisation. Cependant, pour des simulations de type mésoscopique, il est préférable d'utiliser la méthode *Velocity Verlet* et donc d'opter pour la précision car le pas d'intégration est supérieur à celui utilisé dans les méthodes de simulation moléculaire. Le choix de l'algorithme de propagation des équations de mouvement dépend du type de simulation et de la taille du système. Évidemment, les algorithmes de *Verlet Leap Frog* et *Velocity Verlet* sont symplectiques, respectent le renversement du temps et conservent l'énergie mécanique. Il existe d'autres algorithmes tout aussi efficaces comme celui de *Gear predictor-corrector* développé par Gear [12] qui consiste à prédire les positions, les vitesses et l'accéleration au pas suivant sans invoquer les équations de mouvement. Ensuite, vient la phase de correction où l'on calcule la différence entre l'accéleration prédite et celle corrigée (c'est-à-dire les forces) et on utilise cette différence pour corriger les positions et les vitesses.

2.4 Modélisation des interactions intermoléculaires

Nous allons à présent voir comment sont calculées les forces permettant de parcourir l'espace des phases. Les systèmes que nous simulons sont des systèmes isolés, ce qui implique que le

système est conservatif. Ainsi, les forces dérivent d'un potentiel (gradient de potentiel) :

$$\mathbf{f}_{ij} = -\nabla(U(\mathbf{r}_{ij})) \tag{2.19}$$

Il suffit donc de se donner un potentiel, d'en prendre le gradient et d'intégrer les équations de mouvement pour obtenir la trajectoire. En général, on distingue deux types d'interactions : les interactions *intramoléculaires* propres aux interactions entre atomes d'une même molécule et les interactions *intermoléculaires* entre atomes de molécules différentes. Chaque potentiel est composé d'un jeu de paramètres déterminés par des méthodes de mécanique quantique ou par ajustement sur des grandeurs thermodynamiques (enthalpie de vaporisation, d'hydratation,...). L'ensemble des paramètres ajustables et des expressions analytiques des courbes d'ajustement forme un champ de forces. Il existe des champs de forces généralistes tels que le champ de forces OPLS développé par Jorgensen et *al.* [13] ou AMBER de Cornell et *al.* [14], très largement utilisés pour des simulations physico-chimiques.

2.4.1 Potentiel intramoléculaire

Le potentiel intramoléculaire permet de rendre compte de la flexibilité des molécules. Celui ci peut être décomposé en quatre termes :

$$U^{intra} = U^{liaisons} + U^{pliages} + U^{torsions} + U^{1-4}$$

Ils décrivent l'élongation des liaisons, la déformation des angles, les torsions pour les angles dièdres périodiques et impropres et les interactions entre atomes séparés par plus de trois liaisons. Les différents potentiels prennent les formes suivantes :

$$\begin{aligned}
U^{liaisons} &= \sum_{liaisons} k_b(r - r_o)^2 \\
U^{pliages} &= \sum_{pliages} k_\theta(\theta - \theta_o)^2 \\
U^{torsions} &= \sum_{torsions} k_\phi[1 + cos(n\phi + \delta)] \\
U^{1-4} &= \sum_{i=1}^{N-1}\sum_{j>i}^{N} 4\epsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right]
\end{aligned}$$

r_o, θ_o sont la distance et l'angle d'équilibre tandis que k_b, k_θ, k_ϕ sont les constantes de force et δ le déphasage de l'angle dièdre. Les angles impropres qui permettent de maintenir une structure fixe (un cycle benzénique plan par exemple) ont la même forme que les potentiels de dièdre. Nous avons modélisé les interactions $1-4$ par un potentiel de type Lennard jones dont les paramètres seront explicités dans le paragraphe suivant.

2.4.2 Potentiel intermoléculaire

Les interactions non liées sont principalement dues aux interactions de *van der Waals* et aux interactions électrostatiques. Il est nécessaire que ce potentiel d'interaction contienne à la fois un terme répulsif à courte distance pour tenir compte de l'impossibilité de recouvrement des atomes et une partie attractive à moyenne et longue distance. Ces différentes interactions peuvent être décrites à l'aide du potentiel donné par l'équation (2.20).

$$U^{tot} = U^{LJ} + U^{ele} = \sum_{i=1}^{N-1} \sum_{j>i}^{N} 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right] + {U_{ij}}^{ele} \tag{2.20}$$

Cette équation montre que le potentiel choisi est un potentiel de paires. Le premier terme correspond au potentiel de Lennard-Jones, lequel est composé de deux termes : un terme répulsif ($\frac{1}{r_{ij}^{12}}$) et un terme attractif ($\frac{1}{r_{ij}^{6}}$), relatif aux interactions de London (interactions dipôle induit-dipôle induit en $\frac{1}{r^6}$) et Debye (interactions dipôle permanent-dipôle induit en $\frac{1}{r^6}$) et Keesom (interactions dipôle induit-dipôle induit en $\frac{1}{r^6}$ et plus importante que les deux autres). La constante ϵ_{ii} fixe l'échelle d'énergie et la constante σ_{ii} le diamètre de l'atome. Chaque atome i se voit ainsi attribuer un jeu de paramètres $\{\sigma_{ii}, \epsilon_{ii}\}$, déterminé par des procédures d'ajustement. Les interactions mixtes sont évaluées ensuite par des lois de mélange de composition. La plus utilisée est celle de Lorentz et Berthelot :

$$\epsilon_{ij} = \sqrt{\epsilon_{ii}\epsilon_{jj}} \tag{2.21}$$

$$\sigma_{ij} = \frac{\sigma_{ii} + \sigma_{jj}}{2} \tag{2.22}$$

Il existe d'autre règles de mélange, plus sophistiquées pour le calcul de propriétés spécifiques comme celle de Kong [15] ou Waldman-Kagler [16]. Le deuxième terme correspondant à l'énergie électrostatique ${U_{ij}}^{ele}$ est décrit dans la section suivante.

Interactions électrostatiques

Contrairement aux interactions de van der Waals, ces interactions sont à longue portée ($\frac{1}{r}$). elles sont régies par des forces coulombiennes entre charges. L'expression de l'énergie entre deux charges q_1 et q_2 est donnée ci dessous :

$$U^c = -\sum_{i=1}^{N-1}\sum_{j>i}^{N}\frac{q_i q_j}{4\pi\epsilon_o r_{ij}} \tag{2.23}$$

où ϵ_o est la permittivité du vide $\epsilon_o = 8.854.10^{-12}C^2J^{-1}mol^{-1}$. Soit un système de N particules dans une boîte cubique de longueur L avec conditions aux limites périodiques (voir section (2.5.1)), possédant des charges telles que la charge totale du système soit nulle. Comme on le verra dans la section (2.5.1), les conditions aux limites périodiques permettent de simuler un réseau inifini en répliquant une infinité d'images de la boîte centrale. Si $\mathbf{r}_i$ est le vecteur position de l'atome i, ce même atome sera repéré dans les boîtes images par la coordonnée $\mathbf{r}_i + \mathbf{n}L$ où $\mathbf{n}$ est un vecteur de translation. L'énergie totale du système vaut alors :

$$U^c = -\frac{1}{2}\sum_{i=1}^{N}\sum_{j,\mathbf{n}}^{*}\frac{q_i q_j}{|r_{ij} + \mathbf{n}L|} \tag{2.24}$$

L'étoile indique que la somme est faite sur toutes les boîtes et sur toutes les particules, hormis $j = i$ quand $\mathbf{n} = \vec{0}$. Pour des potentiels à courte portée (décroissance plus rapide que $\frac{1}{r^3}$), on fait l'approximation que l'interaction entre deux particules i et j est en très bon accord avec celle calculée entre l'atome i et l'image de l'atome j la plus proche de i (convention d'image minimum). Pour des potentiels à longue portée, cette approximation n'est pas suffisante car l'énergie d'interaction décroît trop lentement pour que l'on puisse se limiter à la première image. Dans ce cas, il faut tenir compte de l'ensemble des boîtes. Pour remédier à cette difficulté, la somme d'Ewald consiste à séparer le potentiel coulombien en deux termes, un terme à courte portée et un terme à longue portée calculé dans l'espace réciproque (*cf* section (9.3) pour la démonstration complète). Les expressions de l'énergie et de la force correspondante sont données par les équations suivantes :

$$
\begin{aligned}
U^C &= \sum_{\mathbf{k}\neq 0} \frac{2\pi}{k^2 V} \exp\left(-\frac{k^2}{4\alpha^2}\right) \left| \sum_{j=1}^{N} q_j \exp\left(i\mathbf{k}(\mathbf{r}_j)\right) \right|^2 + \sum_{i=1}^{N-1} \sum_{l=i+1}^{N} q_i q_l \frac{erfc(\alpha \left|\mathbf{r}_i - \mathbf{r}_l\right|)}{\left|\mathbf{r}_i - \mathbf{r}_l\right|} \\
&- \sum_{j=1}^{N} \left[q_j^2 \frac{\alpha}{\sqrt{\pi}} + \sum_{i=1,\neq j}^{N_m} q_i q_j \frac{erf(\alpha \left|\mathbf{r}_i - \mathbf{r}_j\right|)}{2\left|\mathbf{r}_i - \mathbf{r}_j\right|} \right] \qquad (2.25)
\end{aligned}
$$

$$
\begin{aligned}
\mathbf{f}^C &= -q_i \sum_{\mathbf{k}\neq 0} \exp\left(i\mathbf{k}(\mathbf{r}_i)\right) \frac{4\pi \mathbf{k} i}{k^2 V} \exp\left(-\frac{k^2}{4\alpha^2}\right) \left| \sum_{j=1}^{N} q_j \exp\left(i\mathbf{k}(\mathbf{r}_j)\right) \right| \\
&+ \sum_{i=1}^{N-1} \sum_{l=i+1}^{N} q_i q_l \frac{\mathbf{r}_{il}}{r_{il}^3} \left[erfc(\alpha \left|\mathbf{r}_{il}\right|) + \frac{2\alpha \left|\mathbf{r}_{il}\right|}{\sqrt{\pi}} \exp -(\alpha^2 \left|\mathbf{r}_{il}\right|^2) \right] \qquad (2.26) \\
&- \sum_{j=1}^{N} \sum_{i=1,\neq j}^{N_m} q_i q_j \frac{\mathbf{r}_{ij}}{r_{ij}^3} \left[erf(\alpha \left|\mathbf{r}_{ij}\right|) - \frac{\alpha \left|\mathbf{r}_{ij}\right|}{2\sqrt{\pi}} \exp -(\alpha^2 \left|\mathbf{r}_{ij}\right|^2) \right]
\end{aligned}
$$

La démonstration de ces relations est réalisée dans la deuxième partie de cette thèse.

2.5 Mise en œuvre d'une simulation moléculaire

2.5.1 Conditions périodiques aux limites

Dans les simulations de taille finie, le problème des effets de bords est traité par les conditions périodiques aux limites. Cette méthode permet de simuler un réseau infini autour du système initial appelé alors boîte centrale. En effet, cette méthode consiste à répliquer à travers l'espace la cellule initiale, afin d'obtenir un réseau infini et donc de s'affranchir des problèmes liés aux effets de bords. Ainsi, lorsqu'une molécule sort par une face de la cellule centrale, une de ses images rentre par la face opposée. Finalement, on continue de traiter uniquement la boîte centrale. Le problème qui peut exister est l'interaction d'un atome de la boîte centrale avec sa propre image. Comme les cellules sont répliquées par translation d'un vecteur $\mathbf{n}L$, il faut alors calculer les interactions pour des distances inférieures à L. Ainsi, l'énergie totale est évaluée en calculant les interactions dans la cellule centrale et dans le réseau infini :

$$
U_{tot} = \frac{1}{2} \sum_{\mathbf{n}}^{'} \sum_{i=1}^{N} \sum_{j=1}^{N} U(|\mathbf{r}_{ij}| + \mathbf{n}) \qquad (2.27)
$$

Le $'$ indique que pour la cellule centrale ($\mathbf{n} = 0$) i est différent de j. $\mathbf{n} = (n_x L, n_y L, n_z L)$

2.5.2 Convention d'image minimum

L'énergie d'interaction d'une particule i correspond à l'interaction de cette particule avec les $N-1$ particules de la cellule centrale. Il faut, en outre, tenir compte des interactions entre i et les autres particules l des cellules répliquées (images). Ainsi, on se retrouve avec un nombre important d'interactions à calculer. Il faut, donc, trouver une méthode permettant de réduire ce nombre. Pour cela, on définit une région de même forme et de même dimension que la cellule centrale dans laquelle la particule i est située au centre. Ainsi, la particule i interagit avec les $N-1$ particules de cette nouvelle cellule, c'est-à-dire les $N-1$ images périodiques les plus proches de i. La sommation est donc réduite au calcul de $(N(N-1)/2)$ distances. Il est possible de diminuer encore le temps de calcul en considérant un rayon de coupure (r_c). En effet, on considère seulement les interactions inférieures à r_c (au-delà de r_c, les interactions sont supposées nulles). Ce rayon de coupure est choisi inférieur à une demie longueur de boîte afin qu'il n'y ait pas plus d'une image de chaque particule prise en compte. Comme on néglige les interactions pour lesquelles $r > r_c$, on introduit inévitablement une erreur sur l'énergie totale qui n'est pas négligeable. Il est possible d'estimer ce terme correctif : pour cela, nous devons déterminer l'expression de l'énergie totale.

$$E = kT^2 \left(\frac{\partial \ln Q_{NpT}}{\partial T} \right)_{N,p}$$

$$\text{tel que } Q_{NpT} = CT^{3N/2} \int\int \exp\left(\frac{-\left[U(\mathbf{r}^N) + pV\right]}{kT} \right) d\mathbf{r}^N dV$$

C est une constante fonction de la longeur d'onde de De Broglie. Si on suppose que le terme pV est constant on obtient alors :

$$E = \frac{3NkT}{2} + \left\langle U(\mathbf{r}^N) \right\rangle + pV$$

Il faut à présent exprimer le deuxième terme en fonction de la distribution radiale, puisque celle-ci, à longue distance, est égale à 1 (pas d'ordre à longue distance). Pour cela, on reprend l'expression de $\left\langle U(\mathbf{r}^N) \right\rangle$. En effet, on suppose que $U(\mathbf{r}^N)$ correspond à la somme de $N(N-1)/2$ termes. Ainsi, si on utilise $u(\mathbf{r}_{12})$ comme potentiel de réference, on a :

$$\left\langle U(\mathbf{r}^N)\right\rangle = \frac{N(N-1)}{2}\int\int u(\mathbf{r}_{12})\left[\frac{\int\int \exp\left(\frac{-\left[U(\mathbf{r}^N)+pV\right]}{kT}\right)d\mathbf{r}^{N-2}dV}{\int\int \exp\left(\frac{-\left[U(\mathbf{r}^N)+pV\right]}{kT}\right)d\mathbf{r}^{N}dV}\right]$$

Étant donnée l'expression de la densité de probabilité (équation (2.28)), on obtient la relation (2.29).

$$\rho^2(\mathbf{r}_{12}) = N(N-1)\frac{\int\int \exp\left(\frac{-\left[U(\mathbf{r}^N)+pV\right]}{kT}\right)d\mathbf{r}^{N-2}dV}{\int\int \exp\left(\frac{-\left[U(\mathbf{r}^N)+pV\right]}{kT}\right)d\mathbf{r}^{N}dV} \tag{2.28}$$

$$\left\langle U(\mathbf{r}^N)\right\rangle = \frac{1}{2}\int\int\int u(\mathbf{r}_{12})\rho^2(\mathbf{r}_{12})dV d\mathbf{r}_1 d\mathbf{r}_2 \tag{2.29}$$

Si on généralise, on obtient la relation (2.30)[5] [17]. À partir de l'équation (2.30), nous obtenons l'expression de l'énergie E (équation (2.31))

$$\left\langle U(\mathbf{r}^N)\right\rangle = \frac{N^2}{2V}\int\int 4\pi r^2 u(r)g(r)dV dr \tag{2.30}$$

$$E = \frac{3NkT}{2} + \frac{N^2}{2V}\int\int 4\pi r^2 u(r)g(r)dV dr + pV \tag{2.31}$$

Le terme correctif étant déterminé au-delà de r_c, nous avons $g(r) = 1$, et donc :

$$E^{corr} = \frac{N^2}{2V}\int\int_{r_c}^{\infty} 4\pi r^2 u(r)dV dr + (pV)^{corr}$$

Remarque : le terme $(pV)^{corr}$ peut aussi être déterminé de la même manière. On montre facilement [17] que :

[5] La distribution radiale peut être exprimée de la manière suivante :

$$g^n = \frac{V^n N!}{N^n(N-n)!}\frac{\int\int \exp\left(\frac{-\left[U(\mathbf{r}^N)+pV\right]}{kT}\right)d\mathbf{r}^{N-n}dV}{\int\int \exp\left(\frac{-\left[U(\mathbf{r}^N)+pV\right]}{kT}\right)d\mathbf{r}^{N}dV}$$

.

$$pV = NkT - \frac{2N^2}{3V}\int_0^\infty \pi r^2 r \frac{du(r)}{dr} g(r) dr = NkT - \frac{2N^2}{3V}\int_0^\infty \pi r^2 w(r) g(r) dr$$

où $w(r) = -\frac{du}{dr}$ est le viriel. Le terme correctif est alors donné par l'équation (2.32). Cela nous permet d'écrire l'énergie totale corrigée (équation (2.33)).

$$(pV)^{corr} = -\frac{2N^2}{3V}\int_{r_c}^\infty \pi r^2 w(r) dr \tag{2.32}$$

$$E^{corr} = \frac{N^2}{2V}\int_{r_c}^\infty 4\pi r^2 u(r) dr - \frac{2N^2}{3V}\int_{r_c}^\infty \pi r^2 w(r) dr \tag{2.33}$$

Ces termes correctifs doivent surtout être pris en considération dans des ensembles où le volume subit de grosses fluctuations (NpT, μVT) [18].

2.6 Protocole de simulation

Le calix[4]arène-*p* sulfonaté (symobolysé par Cal^{4-}) a été modélisé en utilisant la version tout atome du champ de forces AMBER [14]. Les différents termes du potentiel ont été donnés dans la section (2.4). Les interactions électrostatiques sont calculées en utilisant la sommation d'Ewald [18, 19]. Les interactions de dispersion-répulsion sont calculées avec un modèle Lennard-Jones, dans lequel nous avons utilisé les règles de mélange de Lorentz et Berthelot. Dans le champ de forces AMBER, les interactions non liées entre atomes séparés par exactement trois liaisons sont réduites par un facteur 0.5 [14]. Les contraintes sont maintenues en utilisant l'algorithme SHAKE [20]. Les cycles benzéniques du calixarène sont maintenus plans grâce à six angles de torsions impropres. L'eau a été modélisée en utilisant le modèle TIP3P [21]. Les paramètres liés aux groupements sulfonate sont tirés de l'article suivant [22]. Nous simulons le calixarène à pH=2. À ce pH, il a été montré que les quatre groupements hydroxyde du calixarène sont protonés [23, 24, 25]. Les algorithmes choisis pour simuler les ensembles NVT et NpT sont ceux de Berendsen [10]. Les constantes de couplage sont de 0.1 ps pour la température et 0.5 ps pour la pression. Nous combinons les techniques *Multiple Time Steps* [26] et la liste de Verlet pour diminuer le temps simulation. Pour cela, nous choisissons un rayon de Verlet de 14 Å et un rayon de calcul explicite des forces de 8 Å . Ce calcul est réalisé

tout les 10 fs avec un pas d'intégration de 2 fs.

Les simulations concernant les complexes sont réalisées dans une boîte cubique d'une longueur de 30 Å avec 900 molécules d'eau. Le rayon de coupure est de 12 Å. Afin de conserver l'électroneutralité de la solution, des contre-ions sont rajoutés. Dans le cas de complexes, il s'agit des ions sodium. Le nombre de contre-ions varie évidemment en fonction de la charge portée par le substrat étudié. Nous vérifions que les distance $[Na^+ \cdots Na^+]$ et $[Na^+ \cdots \mathrm{Cal}^{4-}]$ sont plus grandes que le rayon de coupure. Le tableau (2.1) résume l'ensemble des paramètres utilisés. Pour l'ensemble des simulations, l'équilibration a duré 500 ps et l'acquisition 1 ns. Les propriétés structurales sont calculées sur 50 000 configurations durant la phase d'acquisition. Les configurations sont générées en utilisant la version parallèle de DL_POLY_MD [27]. Les charges atomiques sont évaluées par un calul quantique HF sur un base 6-31G** en utilisant la procédure CHELPG [28]. Dans cette méthode, les charges atomiques sont ajustées sur des grilles de points distribuées réguliérement dans un cube pour reproduire le potentiel électrostatique moléculaire.

Or, le champ de forces est établi à partir de calcul quantique HF sur un base 6-31G*. Nous vérifions que les deux bases donnent des résulats équivalents en ce qui concerne le calcul de charges. Les déviations sont de l'ordre de 0.07 e. Ces calculs ont été effectués avec le logiciel GAMESS [28].

	Complexe	Cation seul
k_1,k_2,k_3	8,8,8	7,7,7
α	0.2651	0.2651
σ	10^{-6}	10^{-6}
r_c / Å	12.0	12.0
r_v / Å	14.0	14.0
r_{mts} / Å	8.0	8.0
N_{mts}	5	5
T/K	298.15	298.15
p/bar	1.013	1.013
τ_T/ps	0.5	0.5
τ_p/ps	0.1	0.1
$\Delta T/ps$	2.0	2.0
N_{H_20}	900	512
L / Å	30.0	24.0
CI	Na^+	Cl^-

TAB. 2.1 – Paramètres utilisés dans nos simulations. (k_i) vecteurs réciproques pour la sommation d'Ewald, (α) paramètre de convergence pour la sommation d'Ewald, (σ) tolérance pour la sommation d'Ewald, (r_c) rayon de coupure, (r_v) rayon de Verlet, (r_{mts}) distance en-dessous de laquelle on effectue un calcul explicite des forces, (N_{mts}) nombre de pas de simulation nécessaire pour calculer explicitement des forces, (T) température, (p) pression, (τ_T) constante de couplage pour la température selon l'algorithme de Berendsen, (τ_p) constante de couplage pour la pression selon l'algorithme de Berendsen, (Δt) pas d'intégration, (N_{H_20}) est le nombre de molécules d'eau utilisé dans les simulations, (L) est la longueur de la boîte cubique, (CI) est le contre ion utilisé.

2.6.1 Description du calixarène

Le calixarène est formé de quatre unités benzéniques sulfonatées et hydroxylées en position *para* formant une structure en forme de vase représentée sur la figure (2.2). Il est possible de montrer que la cavité du macrocycle peut être représentée par une sphère en calculant le coefficient d'asphéricité $\mathcal{A}$ [29, 30] du macrocycle dont l'expression est donnée par l'équation (2.34). Si $\mathcal{A} \to 0$, la forme est sphérique tandis que si $\mathcal{A} \to 1$ la forme est allongée. Concernant le macrocycle, nous obtenons une valeur de $\mathcal{A} = 0.08$ indiquant une forme quasi-sphérique. En outre, nous attribuons à cette sphère un rayon égal au rayon de giration [31] (équation (2.35)). La figure (2.3) permet de visualiser la forme sphérique et le rayon de giration.

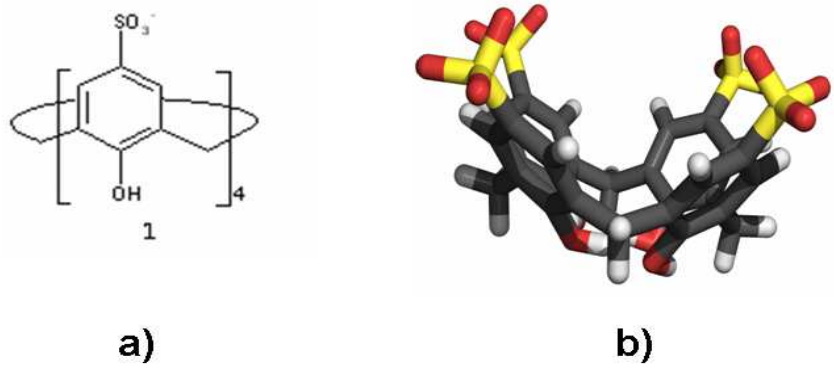

FIG. 2.2 – Structure du calixarène.

$$\mathcal{A} = \frac{1}{2}\frac{\left\langle \sum_{i>j} (\gamma_i - \gamma_j)^2 \right\rangle}{\left\langle \sum_{i=1}^{3} \gamma_i^2 \right\rangle} \tag{2.34}$$

$$r_g^k = \frac{\sum_{i=1}^{N} m_i.({r_i}^k - {r_{cm}}^k)^2}{\sum_i^N m_i} \tag{2.35}$$

$$r_g = \sqrt{\sum_{k=1}^{3} r_g^k} \tag{2.36}$$

Les γ_i sont les composantes x $(i = 1)$, y $(i = 2)$ et z $(i = 3)$ du moment d'inertie γ. Dans la relation (2.35), k a la même signification que (i) dans l'expression de γ_i. m_i représente la masse de l'atome i constituant le macrocycle. r_{cm} est la position du centre de masse du macrocycle. r_i est la position de l'atome i constituant le macrocycle. La figure (2.3) montre une sphère de rayon le rayon de giration $r_g$occupant la cavité du calixarène.

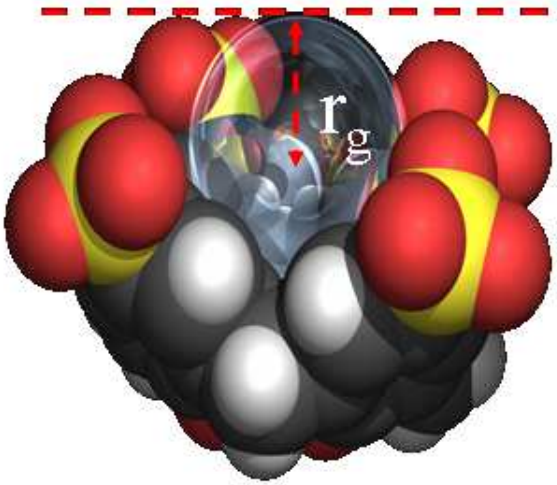

FIG. 2.3 – Visualisation de la cavité sphérique et du rayon de giration du calixarène.

Chapitre 3

Relation «micro-macro» : association entre le calixarène et les alcools linéaires

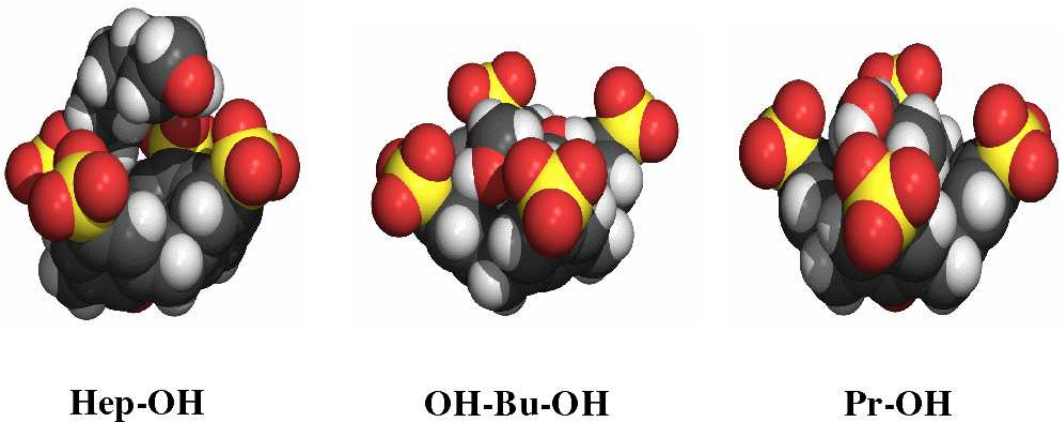

FIG. 3.1 – Conformations de complexes du calixarène avec trois substrats hydroxylés : l'heptanol (Hep-OH), le 1,4-butanediol (OH-Bu-OH) et le propanol (OH-Pr). Hep=heptyl,Bu=butyl, Pr=propyl .

3.1 Introduction

Dans ce chapitre, nous étudions la complexation de molécules d'alcools à 25^oC. Notre objectif est de donner une description microscopique de ces associations en relation avec les résultats expérimentaux. La caractérisation structurale et énergétique des complexes formés avec les alcools et les diols nécessite le calcul des contributions énergétiques (électrostatique et de dispersion-répulsion) entre le substrat et le macrocycle et l'analyse de la désolvatation des espèces en solution.

3.2 Protocole

Les caractéristiques principales des simulations effectuées sont indiquées dans la section (2.6). Nous avons utilisé le champ de forces developpé par Ota et Brünger [32] pour modéliser les alcools et les diols. La forme de ce modèle est celui donné dans la section (2.6). La différence majeure avec le champ de forces AMBER est la modélisation des angles de torsion. En effet, dans ce modèle une décomposition en *cosinus* a été préférée afin de permettre une plus grande flexibilité aux molécules d'alcool.

L'ensemble des simulations est réalisé avec une configuration de départ où le macrocycle et la molécule d'alcool sont très proches. Néanmoins, pour être sûr que la conformation de départ ne correspond pas à un puits local, nous avons testé deux configurations de départ où la distance séparant le macrocycle du substrat est supérieure à 12 Å. En général, un temps de 15 ps est nécessaire pour que cette distance atteigne la distance d'équilibre caractéristique des complexes d'insertion (2 à 5 Å). La figure (3.2) montre dans les deux cas une insertion du substrat dans la cavité du macrocycle. La figure (3.3) est une aide visuelle pour représenter la cavité et les points caractéristiques nécessaires à notre étude.

Les charges atomiques sur les molécules d'alcools sont évaluées par un calcul quantique HF sur une base $6-31G^{**}$ en utilisant la procédure CHELPG [28]. Ces calculs ont été effectués avec le logiciel GAMESS [28].

Pour plus de clarté, nous employons la notation suivante :

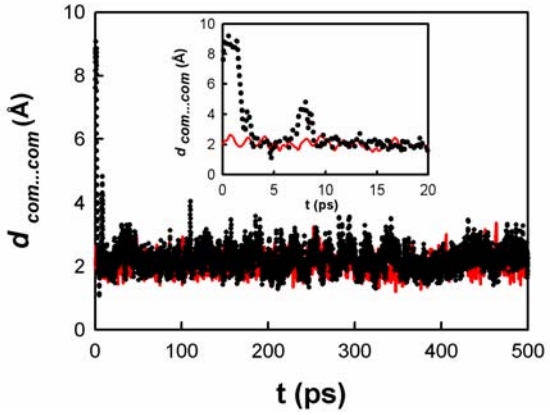

FIG. 3.2 – Distance entre le centre de masse du calixarène et celui du propanol en fonction du temps pour deux configurations initiales différentes. (···) correspond à la situation où les deux molécules sont éloignées de 10 Å. (—) représente la situation où les deux molecules sont proches l'une de l'autre.

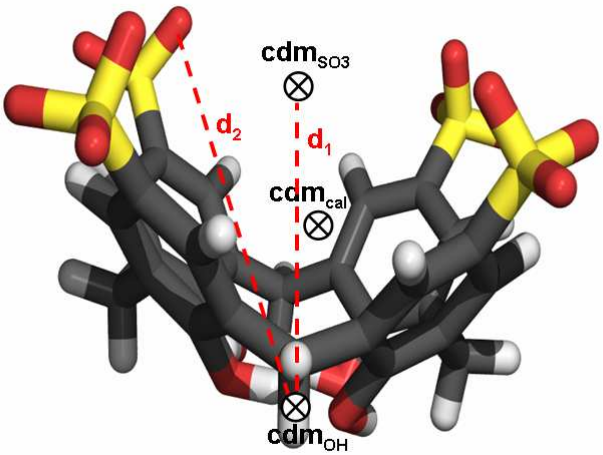

FIG. 3.3 – Représentation du macrocycle et des points caractéristiques dont on se servira dans notre étude structurale. cdm_{SO_3} représente le centre de masse des groupements sulfonate. $cdm_{Cal^{4-}}$ est le centre de masse du calixarène. cdm_{OH} indique le centre de masse des quatre groupes hydroxyde du calixarène. d_1 et d_2 sont des distances caractéristiques que l'on utilisera pour notre étude. d_1 est la distance entre cdm_{SO_3} et cdm$_{OH}$. d_2 est la distance entre cdm_{OH} et le centre de masse d'un groupement sulfonate. d_2 représente la distance limite d'insertion.

- Alcool seul en solution : Alc^s, alcool en présence du macrocycle : Alc^c
- diol seul en solution : Dio^s, diol en présence du macrocycle : Dio^c.

3.3 Complexation avec les alcools

3.3.1 Étude structurale et énergétique

La définition du rayon de giration évoquée dans la section (2.6.1) a été particulèrement efficace pour démontrer la complexation des cations alkylammonium [5]. Néanmoins, pour l'étude de molécules linéaires telles que les alcools, il est nécessaire de définir de nouvelles distances (figure (3.3)) pour vérifier que la molécule échantillonne convenablement la région au dessus des groupements sulfonate. En effet, le rayon de giration est localisé au niveau des groupements sulfonate. On montrera, par la suite, que les alcools ne s'insèrent pas profondément dans la cavité, mais sont situés au dessus des groupements sulfonate. Les distances d_1 et d_2 (figure (3.3)) permettent de caractériser convenablement ce type d'insertion. Pour les diols et les alcools, ces distances sont respectivement égales à 6.0 et 7.5 Å.

Les distances entre le centre de masse du calixarène et celui d'un alcool (de l'éthanol à l'heptanol) sont représentées sur la figure (3.4.a). Nous observons que cette distance varie de 2.0 ± 0.3 Å à 4.2 ± 0.7 Å. En outre, elle augmente avec la longueur de la chaîne alkyle. Le fait que cette distance soit toujours plus faible que le rayon de giration du calixarène (r_g=6 Å) indique que la chaîne alkyle est partiellement insérée dans la cavité du macrocycle.

Afin de décrire correctement le processus d'insertion, nous avons tracé sur la figure (3.4.b), l'intégration des distributions de distances entre le centre de masse des quatre groupements hydroxyde et celui du groupe hydroxyde de l'alcool. Cette figure montre que plus la longueur de la chaîne alkyle augmente, plus le groupe OH s'éloigne du haut de la cavité. Le groupe OH de l'éthanol et du propanol peut être considéré comme complètement inséré dans la cavité. Ce même groupement n'est que partiellement inséré pour les alcools à chaîne alkyle plus longue. À partir du propanol, les distances sont toutes supérieures à d_2 indiquant clairement que le groupement OH est à l'extérieur de la cavité, c'est-à-dire au-dessus des groupements sulfonate. Sur la figure (3.4.c), nous avons représenté les courbes d'intégration de distances des cations La^{3+} et Me_4N^+ en plus de celles de la molécule d'heptanol. Cette figure a pour objectif

de mettre en évidence les différentes types de complexes (d'insertion ou de sphère externe). Pour tous les alcools, nous avons évalué le nombre d'atomes insérés dans la cavité en prenant comme limite supérieure la distance d_2 de 7.5 Å. Ce nombre varie presque linéairement de l'éthanol (9 atomes présents dans la cavité) à l'heptanol (14.4 atomes en moyenne). Le nombre d'atomes insérés dans la cavité peut être mis en relation avec la variation d'enthalpie d'association obtenue expérimentalement par microcalorimétrie. Ainsi, nous avons représenté sur la figure (3.4.d) la variation d'enthalpie et le nombre d'atomes insérés dans la cavité en fonction de la longueur de la chaîne alkyle. Cette courbe montre la corrélation entre $\Delta_r H^o$ et n (nombre de d'atomes de la molécule). Cette relation a déjà été mise en évidence avec les cations tétraalkylammonium [5]. De plus, cette figure montre la présence d'un plateau d'insertion à partir du pentanol. Cela suggère un éloignement maximal de la chaîne pentyle par rapport à la cavité.

Afin de corroborer cette corrélation entre le nombre d'atomes insérés et $\Delta_r H$, on peut déterminer l'énergie intermoléculaire d'association. Cette dernière décrit les interactions électrostatiques et de dispersion-répulsion. Dans le tableau (3.1), nous avons représenté les parties répulsive et attractive du potentiel Lennard-Jones, ainsi que le terme électrostatique. Ce tableau montre que le terme dispersif est le terme prédominant par rapport aux termes répulsif et électrostatique. En outre, ce terme dispersif suit une évolution similaire à celle du nombre d'atomes insérés dans la cavité et à celle de $\Delta_r H$ en fonction de la longueur de la chaîne alkyle. Contrairement aux alcools, le cation La^{3+} donne un complexe de sphère externe dont les interactions prédominantes sont d'origine électrostatique. Ce calcul d' énergie est donc en accord avec l'expérience [4]. F. Perret et *al* [4] associent aux complexes d'insertion un contrôle enthalpique lié aux interactions de type dispersion-répulsion. Ils associent aussi aux complexes à sphère externe un contrôle entropique lié aux interactions électrostatiques.

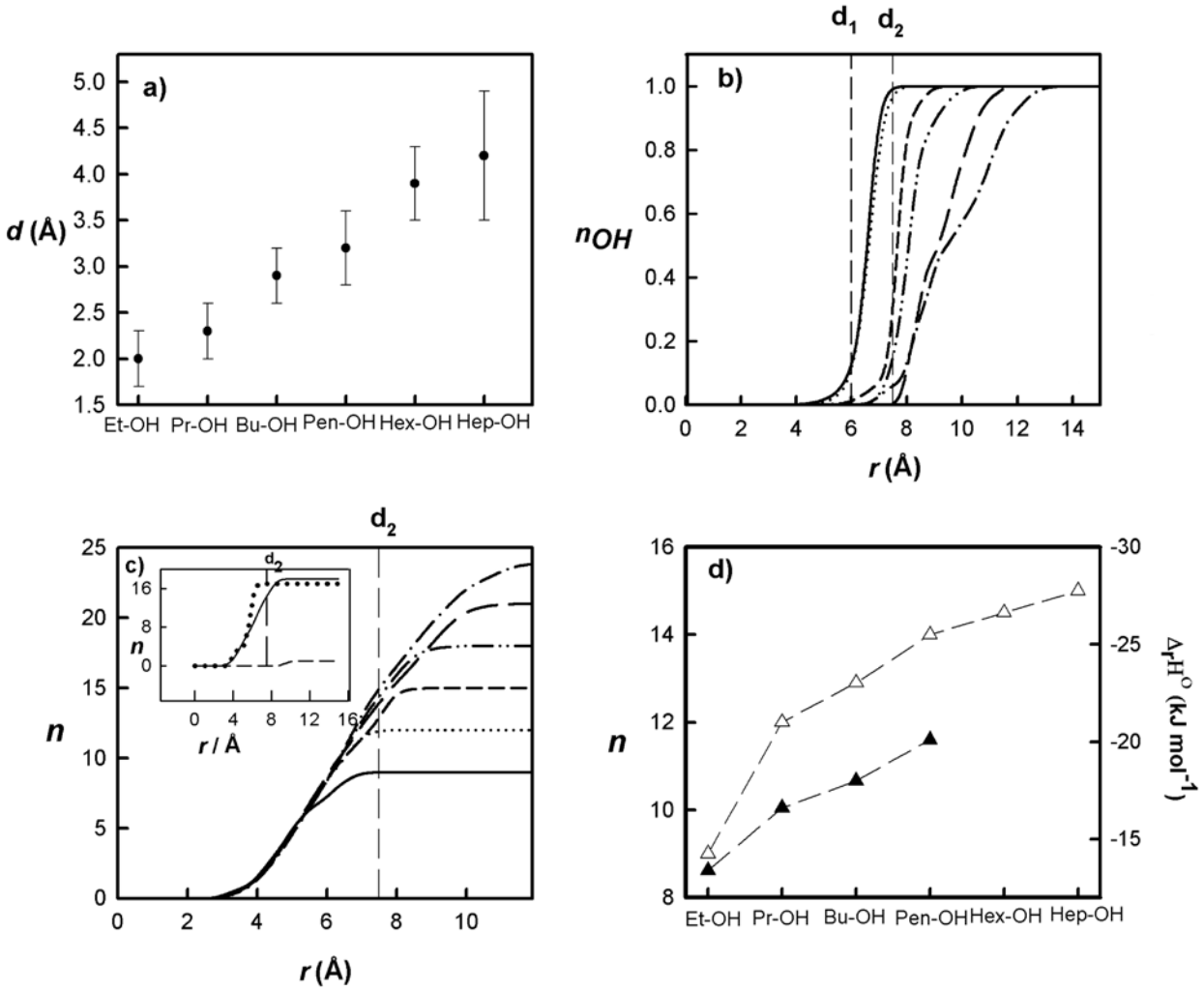

FIG. 3.4 – a) Moyenne de la distance entre le centre de masse du macrocycle et le centre de masse de d'alcool (de Et-OH à Hep-OH). b) Intégration de la distribution de distance entre le point cdm_{OH} (*cf* figure (3.3)) et le groupement OH de la molécule d'alcool (de Et-OH à Hep-OH). (—) Et-OH, $(\cdots)$ Pr-OH, (- - -) Bu-OH, $(-\cdot\cdot-)$ Pen-OH, (— — —) Hex-OH, $(-\cdot-)$ Hep-OH. c) Intégration de la distribution de distance entre chaque atome de la chaîne alkyle de la molécule d'alcool (de Et-OH à Hep-OH) et le point cdm_{OH}. On utilisera la même légende que la figure (3.4.b). La deuxième figure correspond à l'intégration de la distribution de distance (définie précédement dans (a) et (b)) relative aux ions La^{3+} $(---)$, Me_4N^+ (....) et à la molécule d'heptanol (- - -). d) (▲) Variation d'enthalpie de complexation obtenue par microcalorimétrie représentée sur l'axe de droite. (△) Nombre d'atomes insérés dans la cavité représenté sur l'axe de gauche.

	$E_{dispersive}$ (kJ mol^{-1})	$E_{repulsive}$ (kJ mol^{-1})	$E_{lectrostatique}$ (kJ mol^{-1})
Et-OH	-65.2 $_{12}$	30.5_{19}	-3.6_{4}
Pr-OH	-76.9 $_{11}$	35.9_{11}	-5.4_{15}
Bu-OH	-91.9 $_{13}$	50.6_{16}	-6.4_{21}
Pen-OH	-98.2 $_{12}$	52.2_{11}	-6.9_{28}
Hex-OH	-101.5_{14}	55.2_{17}	-7.4_{13}
Hep-OH	-104.9_{18}	55.6_{11}	-6.5_{22}
La^{3+}	-12.1 $_{18}$	64.1_{11}	-479.1_{87}

TAB. 3.1 – Interactions de dispersion-répulsion (partie dispersive et répulsive) et électrostatique entre le calixarène et l'alcool. A titre de comparaison, nous avons ajouté les interactions relatives à la complexation du cation La^{3+}. Le nombre -5.4 $_{15}$ indique -5.4 $\pm$1.5.

3.3.2 Étude de la désolvatation

La solvatation de la cavité du macrocycle peut être étudiée par l'intermédiaire de la distribution radiale entre le centre de masse du calixarène et l'atome d'oxygène de l'eau. La figure (3.5.a) indique ces distributions pour les différents alcools étudiés (de Et-OH à Hep-OH) et pour le calixarène seul en solution. À titre de comparaison, nous avons reporté cette distribution radiale pour le calixarène en présence de l'ion tetraméthylammonium. L'intégration de la distribution radiale entre le centre de masse du calixarène et l'eau jusqu'au rayon de giration indique le nombre de molécules d'eau à l'intérieur de la cavité. Ce nombre varie entre 4.6 et 5.7 (désolvatation partielle) en moyenne pour la série d'alcools étudiée. Il est de 8.3 pour le macrocycle seul en solution et 0.0 (désolvatation totale) en présence de l'ion Me_4N^+.

Afin d'étudier en détail l'influence du groupement hydroxyde, nous avons calculé les distributions radiales de l'atome d'oxygène du groupe OH de l'alcool avec l'atome d'oxygène de l'eau dans le cas Alc^c. Ces distributions sont représentées sur la figure (3.5.b). Cette figure montre que la position de la première sphère de solvatation est indépendante de la longueur de la chaîne alkyle. On en déduit que la présence du calixarène n'affecte pas cette première sphère d'hydratation. Ce fait est corroboré par l'intégration des distributions radiales qui donne le

nombre de molécules d'eau à l'intérieur des couches de solvatation. Ce nombre est reporté dans la figure (3.5.c) pour chaque alcool et pour les cas Alc^c et Alc^s. Ce nombre est identique dans les deux cas et pour chaque alcool. Contrairement à la première couche de solvatation, la seconde couche est très perturbée par la présence du complexe formé par rapport au cas Alc^s. En outre, la forme de la courbe pour le deuxième pic varie avec la longueur de la chaîne alkyle. De l'éthanol jusqu'au pentanol, en présence du macrocycle, le pic relatif à la seconde couche de solvatation du groupe OH est petit et étroit. Il est mieux défini dans le cas de l'hexanol et de l'heptanol. Cette forme se rapproche de celle obtenue pour le cas Alc^s. On en déduit que la présence du calixarène n'affecte pas la deuxième couche d'hydratation dans le cas de l'hexanol et du pentanol. Nous vérifions ces résultas en calculant le nombre de molécules d'eau dans la seconde couche de solvatation (figure (3.5.c)). Cette figure montre que le nombre de molécules d'eau dans la seconde couche de solvatation est différent entre les cas Alc^s et Alc^c. À partir de l'hexanol, cette différence semble s'atténuer. On en déduit donc que plus l'alcool est loin du macrocycle et plus il tend à retrouver la solvatation du cas Alc^s[1].

De plus, nous montrons que l'évolution du nombre de molécules d'eau dans la seconde couche d'hydratation du groupement OH en fonction de la longueur de la chaîne alkyle est différente pour les cas Alc^s et Alc^c. En effet, nous observons une augmentation linéaire du nombre de molécules d'eau dans la seconde couche du groupement OH (n_w^2) en fonction du nombre de carbones de la chaîne alkyle pour le cas Alc^s. Ce résultat peut être expliqué si l'on calcule le nombre d'atomes de carbone qui ont pénétré dans la seconde sphère de solvatation du groupement OH (n_c^2). Ainsi, nous trouvons une corrélation entre n_c^2 et n_w^2 (figure (3.5.d)). En effet, nous montrons que n_c^2 diminue lorsque la longueur de la chaîne alkyle augmente. On peut supposer que cette disparition de matière hydrophobe dans la seconde couche est alors compensée par l'arrivée de molécules d'eau. On peut expliquer ce phénomène par une linéarisation progressive de la chaîne. Cette hypothèse est peu probable étant donnée la nature hydrophobe de la chaîne alkyle. Par contre, on peut imaginer un aspect partiellement incurvé de la chaîne lié aux interactions hydrophobes. Cette courbure est limitée par la présence du groupe OH (hydrophile) qui évite toute circularité de la molécule. Dans le complexe, n_w^2 diminue faiblement en passant de l'éthanol au pentanol avec un comportement atypique pour

[1] quand le pic de la seconde couche n'est pas très prononcé en présence du complexe, nous choisissons comme valeur limite d'intégration la distance de limite de la seconde couche pour l'alcool seul en solution

l'hexanol et l'heptanol (figure (3.5.c)). Pour essayer de comprendre cela, nous avons tracé sur la figure (3.5.c) le nombre d'atomes d'oxygène des groupements sulfonate localisés dans la seconde sphère de solvatation du groupement OH de l'alcool. Le nombre total d'atomes d'oxygène correspondant aux atomes d'oxygène des groupements SO_3^- et aux molécules d'eau est représenté par la courbe en rouge sur la figure (3.5.c). Celui-ci est en parfait accord avec le nombre d'atomes d'oxygène de la seconde sphère de solvatation dans le cas Alc^s. Les atomes d'oxygène du groupement SO_3^- remplacent ceux des molécules d'eau de la second sphère, ce qui permet de reconstruire une sphère de solvatation avec un nombre de coordination approprié (identique au cas Alc^s). On assiste alors à un phénomène de désolvatation-oxygénation (pseudo-solvatation). Nous voyons sur cette même figure que l'hexanol et l'heptanol tendent à retrouver une solvatation identique au cas Alc^s. Ce résultat s'explique par une augmentation de la longueur de la chaîne alkyle. Cela implique un éloignement du groupement OH par rapport aux groupements sulfonate. Dans ce cas, la seconde sphère de solvatation de OH est constituée seulement d'eau puisque l'alcool est très éloigné des groupements sulfonate. Cette situation ne permet pas aux atomes d'oxygène des groupements SO_3^- de s'introduire dans la seconde sphère d'hydratation du groupement OH de l'alcool.

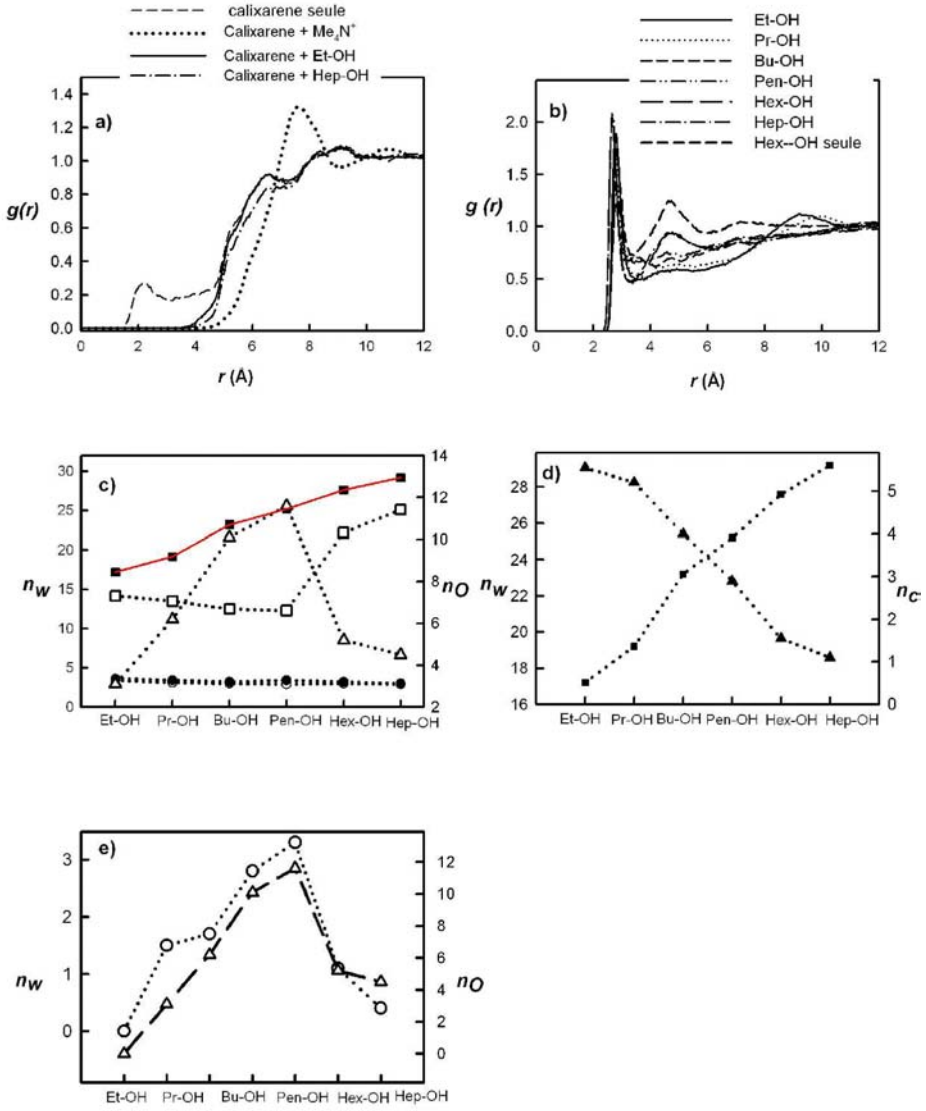

FIG. 3.5 – a) Distribution radiale entre le centre de masse du calixarène et les atomes d'oxygène des molécules ($g_{[cdm_{Cal^{4-}}-O(H_2O)]}$) d'eau (de Et-OH à Hep-OH) b) $g_{[OH_{Al}-O(H_2O)]}$ de Et-OH à Hep-OH. OH_{Al} représente le groupement hydroxyde de l'alcool. c) Nombre de molécules d'eau présentes dans la première couche de solvatation du groupement OH de l'alcool seul en solution (•) et de l'alcool en présence du complexe (○). (■) Nombre de molécules d'eau présentes dans la seconde sphère d'hydratation pour l'alcool seul en solution, et en présence du calixarène (□). (△) Nombre d'atomes d'oxygène appartenant aux groupements sulfonate dans la deuxième couche de solvatation du groupement OH de l'alcool représenté sur l'axe de droite. (—) Nombre total d'atomes d'oxygène (eau + sulfonate) dans la seconde sphère de solvatation du groupement OH de l'alcool. d) (■) Nombre de molécules d'eau présentes dans la seconde couche de solvatation du groupe OH dans le cas des alcools seuls en solution, représenté sur l'axe de gauche. (▲) Nombre d'atomes de carbone de la chaîne alkyle dans la seconde sphère de solvatation du groupement OH, représenté sur l'axe de droite. e) (△) Nombre d'atomes d'oxygène des groupements sulfonate dans la seconde sphère de solvatation du groupe OH, représenté sur l'axe de droite (○) Nombre de ponts de liaisons hydrogène entre le groupe SO_3^- et l'alcool, représenté sur l'axe de gauche.

3.3.3 Formation de ponts de liaisons hydrogène

Afin de compléter cette analyse concernant la solvatation des espèces en solution, nous calculons (figure (3.5.e)) le nombre de ponts de liaisons hydrogène entre les molécules d'eau, les groupements sulfonate et le groupement OH de l'alcool. Les molécules d'eau, dont il est question, sont celles se trouvant à la fois dans la seconde sphère de solvatation du groupe OH et dans la première couche des groupes sulfonate. Nous avons utilisé un critère de distance basé sur un calcul quantique pour définir la liaison hydrogène [6, 33]. Pour caractériser de façon énergétique ces ponts, nous avons calculé l'énergie totale entre le groupement SO_3^- et les molécules d'eau se trouvant dans sa première couche de solvatation. Pour le macrocycle seul en solution, la contribution énergétique est de -131 $\pm$ 6 kJ mol^{-1}. Pour les complexes formés avec les alcools, cette contribution est de l'ordre de -139 $\pm$ 4 kJ mol^{-1} pour les molécules d'eau ne participant pas aux ponts de liaisons hydrogène. Elle est de -240 $\pm$ 6 kJ mol^{-1} à -270 $\pm$ 3 kJ mol^{-1} pour les molécules d'eau impliquées dans des ponts de liaisons hydrogène. Ce calcul montre la nécessité de considérer la solvatation du substrat et du macrocycle par l'intermédiaire de la création d'un pont de liaisons hydrogène énergétiquement favorable. Ce pont est donc enthalpiquement favorable et entropiquement défavorable puisque l'on perd des degrés de liberté en pontant de la sorte le groupe SO_3^- et OH par l'intermédiaire de molécules d'eau. Ces résultats sont en bon accord avec l'expérience [4].

Les figures (3.6.a) et (3.6.b) montrent l'intégration des distributions de distance entre le centre de masse des quatre groupements sulfonate et le centre de masse du groupe hydroxyde pour les molécules d'éthanol et d'heptanol. La figure (3.6.a) montre que l'éthanol est équidistant des quatre groupes sulfonate puisqu'on obtient quatre courbes pratiquement identiques. En outre, la position du plateau d'insertion indique que 100 % des atomes constituant la molécule se trouvent dans la cavité. Cette conclusion est possible si on prend comme limite supérieure de cavité le rayon de giration qui est de l'ordre de 6 Å. L'heptanol présente un comportement différent puisque deux des quatre courbes présentent un seul plateau tandis que les deux autres montrent deux plateaux à 55 % et à 45%. La présence de ces deux paliers indique que le groupement OH se trouve dans deux positions préférentielles diamétralement opposées. Pour représenter cette situation, on pourra s'aider du schéma de la figure (3.7). Celui-ci montre qu'une structure possible permettant d'expliquer ces courbes est celle indi-

quée sur la figure. Quelle que soit la position du groupement OH $p1$ ou $p2$ les distances d_1 et d_2 restent identiques (de même pour d_3 et d_4), ce qui implique un unique palier. Par contre, d_5 et d_8 (de même pour d_6 et d_7) sont différentes selon la position du groupe OH, ce qui implique ainsi deux paliers dans l'hypothèse d'un mouvement du groupement OH.

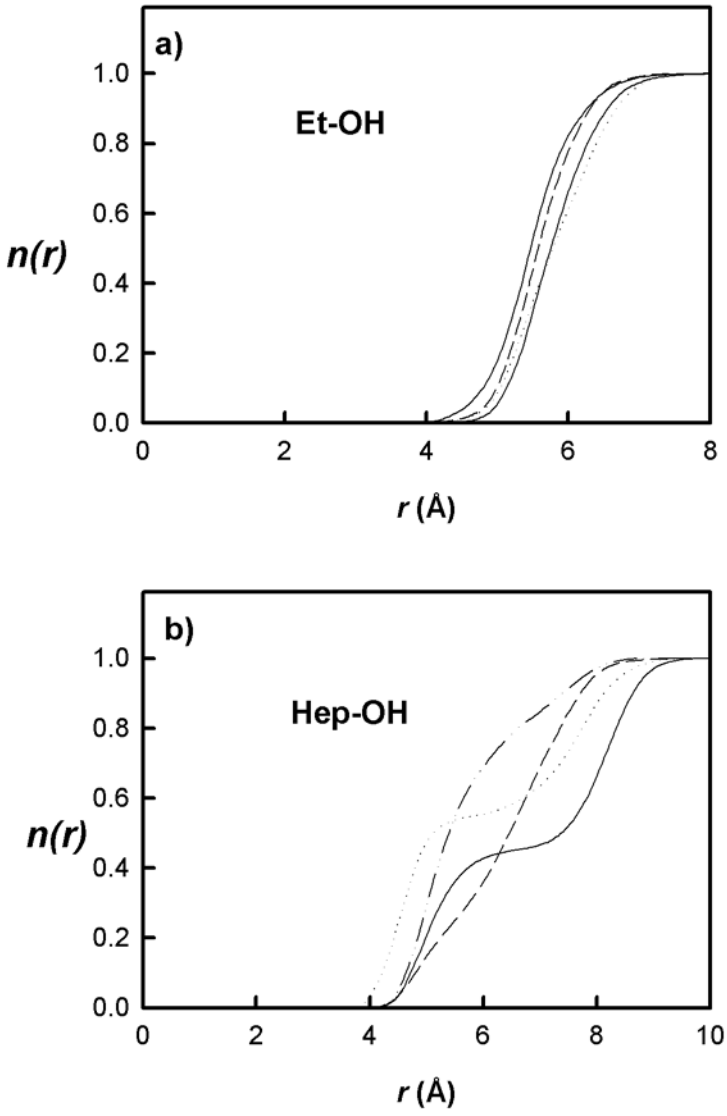

FIG. 3.6 – a) Intégration de la distribution de distance entre le centre de masse du groupement OH de l'alcool et le centre de masse de chacun des quatre groupements sulfonate pour l'éthanol (a) et l'heptanol (b).

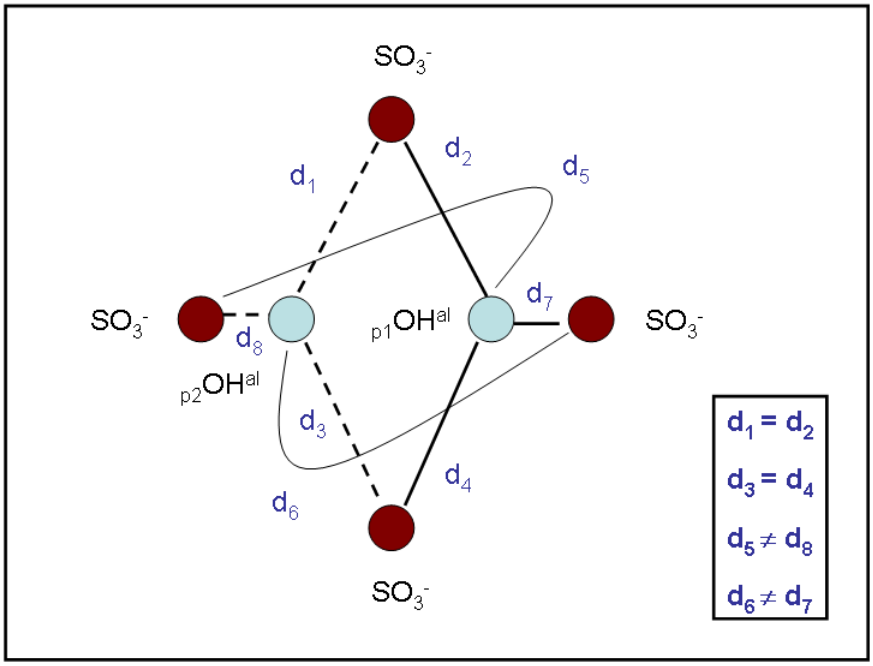

FIG. 3.7 – Schéma permettant de visualiser le phénomène microscopique d'échange de positions du groupe OH de l'alcool. Les distances d_i représentent les différentes distances [SO_3^--OH], $p1$ et $p2$ sont les deux positions préferentiellement échantillonnées par le groupe OH. (•) indique la position des groupements SO_3^-. (•) indique les deux positions que peut prendre le groupement OH.

3.4 Pourquoi le méthanol n'est-il pas complexé ?

Le méthanol est traité à part puisque c'est la seule molécule de la série qui n'est pas complexée en présence du calixarène[4]. Afin de vérifier et d'expliquer cette absence d'association, nous avons réalisé une simulation du méthanol en présence du macrocycle. Cette absence de complexation ne peut être imputée à la petite taille du méthanol, puisque nous avons montré que des molécules de même taille formaient des complexes d'insertion (cas de l'ion méthylammonium [6]).

Afin de simuler convenablement cette molécule et de s'assurer de l'absence de complexation, nous avons construit deux configurations initiales qui différent par la distance entre les centres de masse du calixarène et de la molécule de méthanol ($d_{[cdm_{Me-OH}\cdots cdm_{Cal^{4-}}]}$). Nous prenons deux distances caractéristiques de 2 Å et 12 Å. Afin de donner un interprétation microscopique correcte, nous avons réalisé une simulation où la distance $[cdm_{Me-OH}\cdots cdm_{Cal^{4-}}]$ est contrainte à 1.5 Å. Nous avons choisi cette distance afin de simuler un pseudo-complexe d'insertion. Le méthanol est inséré dans la cavité avec le groupe hydroxyde orienté en direction des groupements sulfonate. Les deux simulations réalisées pour des configurations initiales où les distances $d_{cdm_{Me-OH}\cdots cdm_{Cal^{4-}}}$ sont de 2 Å et 12 Å conduisent au même résultat, c'est-à-dire une absence de complexation. Dans le cas de la simulation à distance contrainte, un calcul d'énergie entre le méthanol et le macrocycle permet de constater une interaction faible (-2.4 $\pm$ 0.4 kJ mol^{-1}). L'énergie est certes faible mais négative, on peut donc envisager une complexation (la détermination du potentiel de force moyenne de l'association est réalisée dans le chapitre (5)). Ainsi, il doit exister un phénomène plus énergetique que l'association ne peut compenser. La figure (3.8) montre la distribution radiale entre le groupe OH et les atomes d'oxygène de l'eau pour les cas Alc^s et Alc^c : cette figure montre que lorsque le méthanol est dans la cavité, le groupement OH perd sa première couche de solvatation (en comparaison au cas Alc^s), c'est-à-dire trois molécules d'eau. Cette solvatation donne une contribution enthalpique favorable puisque l'énergie d'interaction [OH$\cdots$H$_2$O] est de -241 $\pm$ 6 kJ mol^{-1}. En outre, la seconde couche d'hydratation perd 20 molécules d'eau, soit une perte de -19 kJ mol^{-1} par molécule d'eau. Ces calculs énergétiques montrent que l'énergie d'association de -2.4 $\pm$ 0.4 kJ mol^{-1} ne permet pas de compenser le coût énergétique de la désolvatation de la première et la deuxième couche de solvatation (241+19 kJ mol^{-1}). Cette

étude a donc permis de corroborer l'absence de complexation établie par microcalorimétrie et d'expliquer l'origine de ce phénomène.

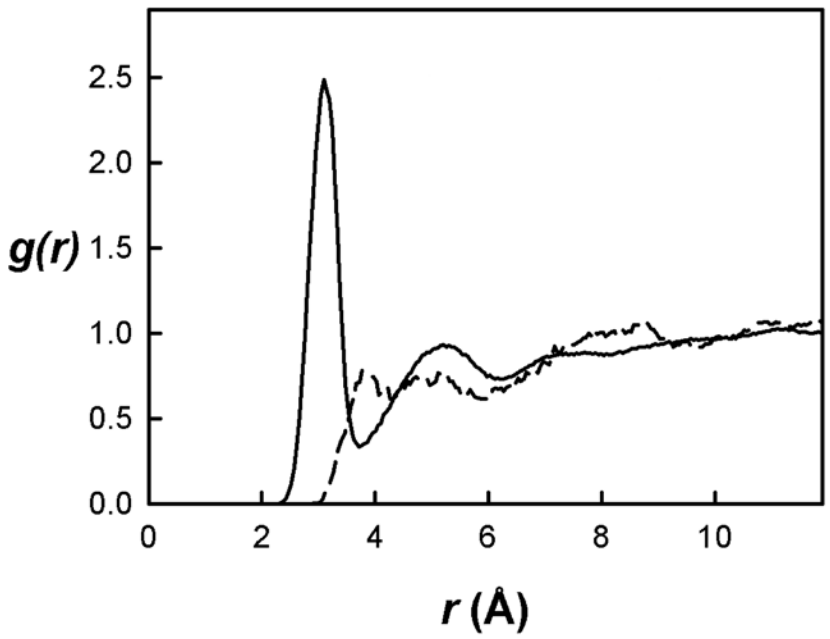

FIG. 3.8 – Distributions radiales entre le groupement OH du méthanol et les atomes d'oxygène de l'eau pour le méthanol seul (—) en solution et en présence du complexe (– – –).

3.5 Complexation avec les diols

Cette section est consacrée à l'étude de la complexation de la molécule 1,4-butanediol (OH-Bu-OH) et de la molécule 1,5-pentanediol (OH-Pen-OH) par le calixarène. Les mesures microcalorimétriques [4] montrent des comportements thermodynamiques identiques ($\Delta_r H^o$, $\Delta_r S^o$) pour OH-Pen-OH et Pen-OH, tandis que $\Delta_r H^o$ et $\Delta_r S^o$ augmentent lorsqu'on passe de OH-Bu-OH à Bu-OH. Un des objectifs ici sera de donner une interprétation microscopique à ce comportement thermodynamique.

Les figures (3.9.a) et (3.9.b) montrent l'intégration des distances entre les groupes OH des

diols et le centre de masse cdm_{OH} du macrocycle (figure (3.3)). Dans le cas des diols, on observe une insertion plus profonde dans la cavité par rapport aux alcools de taille correspondante. Dans le cas du 1,4-butanediol, les deux groupements OH peuvent être considérés comme complètement insérés, à condition de prendre comme limite supérieure la distance $d_2 = 7.5$ Å représentée sur la figure (3.3). On observe dans ce cas des taux d'insertion de OH de l'ordre de 100 % et de 94 % alors que ce taux est de l'ordre de 31 % pour le butanol. Ces taux sont de 63 % et 72 % pour le 1,5-pentanediol et de 15 % pour le pentanol. Cette courbe montre aussi que les diols à chaîne courte s'insèrent plus profondément dans la cavité. La figure (3.9.a) suggère que l'un des groupements OH du 1,4-butanediol (OH_1) est plus inséré que le deuxième (OH_2). La courbe d'intégration de (OH_2) se prolonge aux longues distances de la même manière que le groupe OH du butanol. Dans le cas du diol OH-Pen-OH, on observe deux plateaux. L'un à 50 % qui correspond à une distance de 6.5 Å, et le deuxième qui correspond à 50 % à une distance de 9.0 Å. La similitude des courbes d'insertion pour les deux groupes OH du OH-Pen-OH, ainsi que l'existence de deux paliers indiquent que chaque groupe OH échantillonne deux distances privilégiées. Cela met en évidence un mouvement pendulaire des deux groupements OH dans la cavité.

Le tableau (3.2) montre que la distance moyenne entre le centre de masse du calixarène et celui du OH-Bu-OH est plus faible que celle calculée pour OH-Pen-OH. Cela indique, une fois de plus, que la chaîne butyle est plus insérée que la chaîne pentyle. En théorie, cela devrait correspondre à une enthalpie plus négative (dans le cas d'un complexe d'insertion). Ce n'est pas le cas, puisque [$\Delta_r H$(OH-Bu-OH) -Δ_rH(OH-Pen-OH)] =14.9 kJ mol^{-1}. Afin de comprendre ce paradoxe, nous avons calculé le nombre de molécules dans la première et la deuxième couche de solvatation des deux groupes hydroxyde. Nous avons aussi calculé l'énergie [$OH\cdots H_2O$] pour les cas Dio^s et Dio^c. Les résultats sont rassemblés dans le tableau (3.2). Dans le cas du 1,5-pentanediol, nous observons que la première couche de solvatation des groupements OH n'est pas perturbée par la complexation, alors que la seconde couche perd au moins 13 molécules d'eau. Un comportement similaire est observé dans le cas du groupe OH pour le pentanol. Par contre, il existe une différence majeure dans le cas du 1,4-butanediol. En effet, les deux groupes OH ont des couches de solvatation différentes. Ce résultat est corroboré par les distributions radiales de la figure (3.9.c). En effet, les distributions radiales $g_{O(OH)-0(H_2O)}$ pour les deux groupes OH du butanediol sont différentes. L'intégration de ces courbes montre

que l'un des groupes hydroxyde perd deux molécules d'eau par rapport au cas Dio^s, contrairement au 1,5-pentanediol et aux autres alcools où la première couche restait inchangée en comparaison aux cas Alc^s. Pour plus de clarté, appelons OH_1 le groupe dont la première couche reste intacte et OH_2 celui dont la première couche est désolvatée. La deuxième couche de solvatation de OH_2 est elle aussi plus désolvatée que OH_1 (tableau (3.2)). Cette désolvatation partielle apporte une contribution énergétique positive comme le montre le tableau (3.2), ce qui suggère une enthalpie défavorable. Ce résultat serait alors en bon accord avec les résultats expérimentaux. Ainsi, le fait que l'entropie et l'enthalpie d'association augmentent lorsqu'on passe du 1,5-pentanediol au 1,4-butanediol peut être expliqué par une désolvatation accrue de l'un des deux groupements hydroxyde de OH-Bu-OH. Pour terminer, on remarquera que malgré des grandeurs thermodynamiques semblables, les molécules OH-Pen et OH-Pen-OH présentent des structures d'hydratation différentes. En effet, la deuxième couche des groupements OH du diol se désolvate davantage que celle de l'alcool correspondant, ce qui implique une contribution enthalpique défavorable. Par conséquent, il existe un phénonomène susceptible de compenser cette perte d'énergie. Cette compensation est due au fait que le nombre de ponts de liaisons hydrogène entre les groupements SO_3^- et OH a été multiplié par deux. Or, nous avons montré précédemment que ces liaisons hydrogène sont énergétiquement très favorables. On remarque une variation d'enthalpie identique pour les molécules Pen-OH et OH-Pen-OH ; cela s'explique par le fait que pour OH-Pen-OH, il existe une compensation énergétique entre les processus de désolvatation (qui est défavorable au niveau enthalpique) et de création de ponts de liaisons hydrogène (favorable au niveau enthalpique).

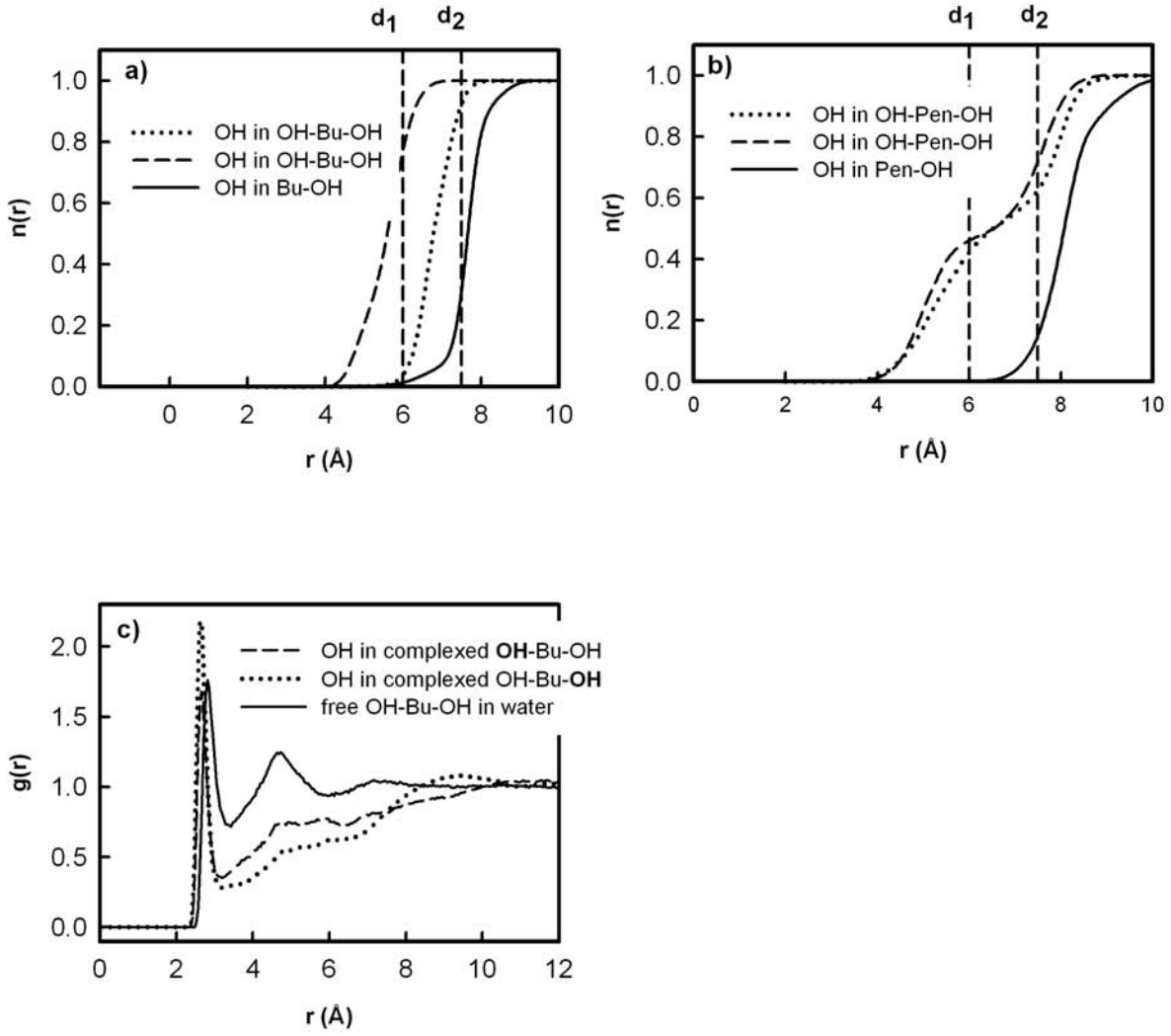

FIG. 3.9 – Intégration de la distribution de distance (In_d : Abréviation utilisée pour représenter les intégration de distribution de distances) entre l'atome d'oxygène du groupement hydroxyde des diols et le centre de masse cdm_{OH} [In_d] (défini sur la figure (3.2)).In_d a été calculée pour les complexes impliquant le diol OH-Bu-OH (a) et le diol OH-Pen-OH (b). a) (—) In_d pour l'un des groupements OH dans le cas de la molécule OH-Bu-OH, (– – –) In_d pour le second groupement OH dans le cas de la molécule OH-Bu-OH , (...) In_d dans le cas de la molécule Bu-OH. b) (—) In_d pour l'un des groupements OH dans le cas de la molécule OH-Pen-OH, (– – –) In_d pour le second groupement OH dans le cas de la molécule OH-Pen-OH , (...) In_d dans le cas de la molécule Pen-OH. c) Distribution radiale entre l'atome d'hydrogène du groupement hydroxyde de la molécule OH-Bu-OH et les atomes d'oxygène des molécules d'eau (—) pour le premier groupement OH dans le cas du substrat en présence du complexe, pour le deuxième groupement OH dans le cas du substrat en présence du complexe (...) et dans le cas du substrat seul en solution (—).

	OH-Pen-OH^{libre}		OH-Pen-OH^{lié}		OH-Bu-OH^{libre}		OH-Bu-OH^{lié}	
$\Delta_r H^o$			-19.3				-4.4	
$T\Delta_r S^o$			-11.1				-2.8	
$< d_{cdm..cdm} >$			5.2_{13}				2.9_3	
N_{ins}			14.6				15.7	
	OH_1	**OH_2**	**OH_1**	**OH_2**	**OH_1**	**OH_2**	**OH_1**	**OH_2**
$N(H_2O)_1$	3.1	3.3	3.1	3.5	3.2	3.0	3.1	1.2
$E(OH-H_2O)_1$	-240_6	-235_5	-244_6	-242_7	-239_5	-238_5	-243_5	-108_4
$N(H_2O)_2$	26.2	26.8	12.3	12.6	24.2	23.8	14.8	8.6
$E(OH-H_2O)_2$	-22_1	-23_1	-7_1	-7_1	-20_1	-20_1	-7_1	-5_1

TAB. 3.2 – Variations d'enthalpie et d'entropie de complexation (kJ mol^{-1}) de OH-Pen-OH et OH-Bu-OH par le calixarène dans l'eau à 25^{o}C [4]. $< d_{cdm..cdm} >$ est la distance moyenne entre le centre de masse du macrocycle et le centre de masse de la molécule d'alcool. N_{ins} est le nombre d'atomes insérés dans la cavité. Les exposants ^{libre} et ^{lié} correspondent aux cas où le diol est seul en solution et ou il est complexé. $N(H_2O)_1$ et $N(H_2O)_2$ sont les nombres de molécules d'eau, respectivement dans la première et la seconde couche d'hydratation du groupement OH. $E(OH-H_2O)_1$ et $E(OH-H_2O)_2$ représentent l'énergie entre le groupement OH et une molécule d'eau, respectivement pour la première et deuxième couche de solvatation. Le nombre -244_6 indique -244 $\pm$6.

3.6 Résumé

L'étude qui a été menée dans ce chapitre a permis de donner un exemple du lien «micro-macro». En effet, nous avons donné une vue énergétique et structurale de l'association de molécules d'alcool et de diols linéaires avec le calixarène. Nous avons ainsi utilisé les simulations de type dynamique moléculaire pour montrer que les complexes formés avec ce type de substrat sont des complexes d'insertion gouvernés essentiellement par les interactions de dispersion-répulsion entre les cycles benzéniques du calixarène et la chaîne alkyle. En outre, nous avons montré qu'il existe une relation linéaire entre le nombre d'atomes insérés et l'enthalpie d'association, de l'éthanol au pentanol. Par ailleurs, nous avons montré que pour la série propanol-heptanol, le groupement OH est situé à l'extérieur de la cavité. En effet, le coût énergétique de la désolvatation engendré par une insertion du groupement OH ne peut pas être compensé par les interactions de dispersion-répulsion. Nous avons ainsi montré que le contrôle enthalpique pour ces complexes est dû aux interactions de dispersion-répulsion d'une part, et d'autre part à la formation de ponts de liaisons hydrogène formés entre les groupements sulfonate du macrocycle, l'eau et le groupement hydroxyde de l'alcool.

Par ailleurs, nous avons clairement établi les causes de l'absence de complexation du méthanol par le calixarène. Nous avons montré que le coût énergétique de la désolvatation de la première et la deuxième couche, engendré par une insertion du substrat ne peut être compensé par les interactions entre le macrocycle et le méthanol.

Nous avons aussi étudié la complexation du 1,5-pentanediol et du 1,4-butanediol. Le 1,5-pentanediol et le pentanol ont le même comportement thermodynamique ($\Delta_r H^o$ et $\Delta_r S^o$ sont très proches). Néanmoins, les complexes possèdent des structures de solvatation différentes. C'est par le biais d'une compensation énergétique relative à des phénomènes de désolvatation-création et de création de ponts liaisons hydrogène que OH-Pen et OH-Pen-OH donnent des résultats microcalorimétriques très proches. Lorsqu'on passe du 1,4-butanediol au butanol, les changements entropiques ($\Delta_r S^o(Bu-OH) < \Delta_r S^o(OH-Bu-OH)$) et enthalpiques ($\Delta_r H^o(Bu-OH) < \Delta_r H^o(OH-Bu-OH)$) sont essentiellement liés à la désolvatation partielle de la première et deuxième couche d'hydratation d'un des deux groupements OH.

Chapitre 4

Calculs de $\Delta\Delta G$, $\Delta\Delta H$, $\Delta\Delta S$: application à des complexes de sphère externe

L'objectif de cette partie est la détermination des grandeurs thermodynamiques à partir de la mécanique statistique. En effet, il est possible d'écrire les princpales fonctions thermodynamiques à partir de la fonction de partition. Cette fonction de partition est une fonction de normalisation de la probabilité d'accéder à un état microscopique donné. L'expression de cette fonction dépend de l'ensemble statistique dans lequel on travail. Étant donnée les grandeurs auxquelles nous avons accés au laboratoire (G et H), nous avons choisi de travailler dans l'ensemble NpT.

4.1 Introduction

D'une manière générale, il est possible de décrire l'état microscopique d'un système par l'opérateur hamiltonien total qui est exprimé en fonction des vecteurs positions $\mathbf{r}$ et des vecteurs impulsions $\mathbf{p}$ de chacun des atomes du système.

$$\mathcal{H}(\mathbf{r}^N, \mathbf{p}^N) = \sum_{i=1}^{N} \frac{\mathbf{p}_i^2}{2m_i} + \mathcal{U}\left(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, ..., \mathbf{r}_N\right) \tag{4.1}$$

Dans cette expression, m_i représente la masse de l'atome i, $\mathcal{U}(\mathbf{r}^N)$ est l'énergie potentielle décrivant les interactions entre les atomes. Le premier terme représente l'énergie cinétique totale. Dans l'ensemble statistique NpT, la fonction de partition du système est donnée par l'équation (4.2).

$$Q_{NpT} = \frac{1}{h^{3N} V^o N!} \int \exp\left(-\frac{pV}{kT}\right) dV \int\int \exp\left(\frac{-\mathcal{H}(\mathbf{r}^N, \mathbf{p}^N)}{kT}\right) d\mathbf{r}^N d\mathbf{p}^N \tag{4.2}$$

Dans le cas d'une séparation de l'opérateur Hamilton en deux termes, un terme d'énergie cinétique et un terme d'énergie potentielle, nous pouvons réécrire cette fonction de partition en utilisant la longueur d'onde de De Broglie (la démonstration est effectuée dans le chapitre (8).

$$Q_{NpT} = \frac{1}{N!}\,\frac{1}{\Lambda^{3N}V^o}\int \exp\left(-\frac{pV}{kT}\right)dV \int \exp\left(\frac{-U(\mathbf{r}^N)}{kT}\right)d\mathbf{r}^N \tag{4.3}$$

On reconnaît dans cette expression les variables classiques décrivant le système : le volume V, la température T , la constant de Boltzmann k, la constante de Planck h et la pression p. V^o n'est autre qu'un facteur d'homogénéisation permettant d'obtenir une fonction de partition adimensionnelle. Λ est la longueur d'onde de De Broglie dont l'expression est :

$$\Lambda = \left(\frac{h^2}{2\pi mkT}\right)^{\frac{1}{2}} \tag{4.4}$$

Cette longueur d'onde de De Broglie correspond à une fonction de partition translationnelle. Les parties rotationnelle, nucléaire , électronique et vibrationelle sont prises en compte dans le terme lié aux fluctuations des positions. La mécanique statistique nous permet d'accéder aux autres grandeurs thermodynamiques par l'utilisation des relations suivantes :

$$\begin{aligned}
G_{N,p,T} &= -kT\ln Q_{NpT} && (4.5)\\
S_{N,p,T} &= \left(-\frac{\partial G(N,p,T)}{\partial T}\right)_{N,p} && (4.6)\\
&= k\ln Q_{NpT} + kT\left(\frac{\partial \ln Q_{NpT}}{\partial T}\right)_{N,p} &&\\
H_{N,p,T} &= G_{NpT} + TS_{NpT} && (4.7)\\
&= kT^2\left(\frac{\partial \ln Q_{NpT}}{\partial T}\right)_{N,p} && (4.8)
\end{aligned}$$

L'un des objectifs de la mécanique statistique est la détermination et la compréhension de grandeurs macroscopiques à partir de propriétés microscopiques du système. La simulation numérique est un outil de choix permettant de générer des états microscopiques, c'est-à-dire le N-uplet $(\mathbf{r},\mathbf{p})^N$ à partir duquel il est possible de déterminer les propriétés thermodynamiques et mécaniques [17, 18, 34, 35] (pression, température, énergie libre, etc.).

Lorsque l'on s'intéresse à la stabilité et à la selectivité d'un processus la propriété thermodynamique la plus calculée est probablement l'enthalpie libre [5, 6, 36]. En effet, celle-ci renseigne sur la stabilité énergétique du système. Ainsi, si on confronte les résultats expérimentaux à ceux issus d'un calcul par simulation, on peut alors valider d'un point de vue énergétique un modèle ou une théorie. On comprend dès lors la nécessité de développer une méthodologie robuste et efficace permettant ce type de calcul. Plusieurs méthodes ont vu le jour depuis ces dix dernières années :

1- *Umbrella sampling [37]*
2- *reversible scaling [38]*
3- *theory of perturbation [39]*
4- *nonequilibrium free energy [40]*
5- *overlapping distribution [41]*
6- *adiabatic switching [42].*

Les méthodes les plus utilisées sont probablement les méthodes de perturbation et d'intégration [39, 43, 44] pour leur simplicité et leur efficacité. Ces méthodes ont donné de très bons résultats dans de nombreux cas, tels que le calcul de solubilité [49], de grandeurs thermodynamiques d'association [6, 5, 36], d'hydratation [45, 46, 47, 48], etc.

La détermination de la variation d'enthalpie libre permet de déterminer la stabilité d'un système par son signe et d'évaluer la sélectivité de récepteurs vis à vis de plusieurs substrats. Mais, en aucun cas, elle ne donne une vision microscopique du système. De plus, elle ne permet pas de discriminer deux processus possédant la même variation d'enthalpie libre (ΔG). Il faut, dans ce cas, déterminer la variation *d'enthalpie* et *d'entropie* du système. La détermination de ces deux grandeurs thermodynamiques est fondamentale en chimie supramoléculaire si on veut identifier la force motrice d'une association entre un substrat et un récepteur. Le calcul de ces fonctions thermodynamiques est plus difficile et a fait l'objet de beaucoup moins d'études. La détermination de ces grandeurs a été réalisée dans l'ensemble NVT par l'utilisation de méthodes de perturbation (FEP) [44, 39], d'intégration (TI) [44, 50, 51] et de perturbation de température (TP) [52]. Les erreurs sur le calcul de $\Delta\Delta S$ sont plus importantes que celles concernant $\Delta\Delta G$. La source de ces erreurs est attribuée au fait qu'on prend en compte l'hamiltonien total dans le calcul et non plus des variations de celui-ci. Cela implique un échantillonnage de l'espace des phases conséquent afin de réduire les imprécisions sur ce type de calcul [53].

Notre objectif est la détermination précise de ces grandeurs thermodynamiques dans le cas de propriétés d'hydratation et d'association. À cet effet, nous ferons un bref historique sur les méthodes de perturbation, puis nous établirons les formules opérationnelles permettant l'évaluation de $\Delta\Delta_r S$ et $\Delta\Delta_r G$. Enfin, nous testerons ces formules opérationnelles sur le calcul de propriétés d'hydratation et de complexation.

4.2 Rappels

4.2.1 Théorie de la perturbation : Zwanzig (1954) [39]

La théorie de la perturbation développée par Zwanzig en 1954 permet d'obtenir la variation d'énergie libre entre un système de référence et ce même système à l'état perturbé. Soit $A^{(0)}$ le système de référence, $A^{(1)}$ le système perturbé et P la perturbation ($A^{(0)} \xrightarrow{P} A^{(1)}$). Quelle que soit la nature de cette perturbation, l'hamiltonien du système perturbé peut s'écrire $\mathcal{H}^{(1)} = \mathcal{H}^{(0)} + \delta\mathcal{H}$. Les fonctions de partition de l'état de référence et de l'état perturbé s'écrivent :

$$Q^{(0)}_{NpT} = \int \exp\left(-\beta p V\right) dV \int\int \exp\left(-\beta \mathcal{H}^{(0)}(\mathbf{r}^N, \mathbf{p}^N)\right) d\mathbf{r}^N d\mathbf{p}^N \tag{4.9}$$

$$Q^{(1)}_{NpT} = \int \exp\left(-\beta p V\right) dV \int\int \exp\left(-\beta (\mathcal{H}^{(0)}(\mathbf{r}^N, \mathbf{p}^N) + \delta\mathcal{H}(\mathbf{r}^N, \mathbf{p}^N))\right) d\mathbf{r}^N d\mathbf{p}^N \tag{4.10}$$

Il faut bien comprendre que cet état perturbé est obtenu à partir de l'état de référence. Ainsi, l'espace des phases parcouru dans létat perturbé est le même que celui parouru dans l'état de référence. Seul l'hamiltonien change, les autres variables restent identiques dans les deux états. On a donc : $\mathbf{r}^{(1)} = \mathbf{r}^{(0)}, \mathbf{p}^{(1)} = \mathbf{p}^{(0)}, V^{(1)} = V^{(0)}$. Si on multiplie l'équation (4.10) par $Q^{(0)}_{NpT}/Q^{(0)}_{NpT}$, on obtient :

$$\begin{aligned} Q^{(1)}_{NpT} \;=\; & \int \exp\left(-\beta p V\right) dV \int\int \exp\left(-\beta \mathcal{H}^{(0)}(\mathbf{r}^N, \mathbf{p}^N)\right) d\mathbf{r}^N d\mathbf{p}^N \cdot \\ & \frac{\int \exp\left(-\beta p V\right) dV \int\int \exp\left(-\beta (\mathcal{H}^{(0)}(\mathbf{r}^N, \mathbf{p}^N) + \delta\mathcal{H}(\mathbf{r}^N, \mathbf{p}^N))\right) d\mathbf{r}^N d\mathbf{p}^N}{\int \exp\left(-\beta p V\right) dV \int\int \exp\left(-\beta \mathcal{H}^{(0)}(\mathbf{r}^N, \mathbf{p}^N)\right) d\mathbf{r}^N d\mathbf{p}^N} \end{aligned}$$

En multipliant le numérateur par $\exp\left(-\beta\mathcal{H}^{(0)}(\mathbf{r}^N, \mathbf{p}^N)\right) \exp\left(\beta\mathcal{H}^{(0)}(\mathbf{r}^N, \mathbf{p}^N)\right)$, nous obtenons l'équation suivante :

$$\begin{aligned} Q^{(1)}_{NpT} \quad &= \quad Q^{(0)}_{NpT} \left\langle \exp\left(-\beta\delta\mathcal{H}(\mathbf{r}^N, \mathbf{p}^N)\right)\right\rangle_{(0)} \\ &\text{avec} \quad \delta\mathcal{H}(\mathbf{r}^N, \mathbf{p}^N) = \mathcal{H}^{(1)}(\mathbf{r}^N, \mathbf{p}^N) - \mathcal{H}^{(0)}(\mathbf{r}^N, \mathbf{p}^N) \end{aligned} \tag{4.11}$$

Ainsi, la variation d'énergie libre entre ces deux états vaut :

$$\Delta G_{0\to 1} = G^{(1)} - G^{(0)} = -kT \ln \left[\frac{Q^{(1)}_{NpT}}{Q^{(0)}_{NpT}} \right] = -kT \ln \left[\left\langle \exp\left(-\beta \delta \mathcal{H}(\mathbf{r}^N, \mathbf{p}^N)\right) \right\rangle_{(0)} \right] \quad (4.12)$$

Si l'état final que l'on veut atteindre est trop différent de l'état de référence, alors l'espace des phases parcouru dans l'état perturbé qui est celui de l'état de référence ne correspondra pas à celui de l'état final. Dans ces conditions, l'équation (4.10) n'est plus valide. En effet, l'approximation qui est faite dans cette théorie est de postuler l'équivalence des espaces des phases parcouru dans l'état de référence et dans l'état final pour calculer les moyennes d'ensemble.

4.2.2 Paramètre de couplage : Kirkwood (1935) [54]

Cette méthode permet de déterminer les variations d'enthalpie libre entre deux systèmes chimiques ou physiques A et B différents. Si on considère B comme une perturbation de A, on peut écrire l'équation suivante :

$$\begin{aligned} \Delta G_{A\to B} &= G^{(B)} - G^{(A)} = \quad (4.13) \\ &= -kT \ln \left[\frac{Q^{(B)}_{NpT}}{Q^{(A)}_{NpT}} \right] = -kT \ln \left[\left\langle \exp\left(-\beta(\mathcal{H}^{(B)}(\mathbf{r}^N, \mathbf{p}^N) - \mathcal{H}^{(A)}(\mathbf{r}^N, \mathbf{p}^N)\right) \right\rangle_{(A)} \right] \end{aligned}$$

Ici, N est le nombre de particules constituant le système. Si l'espace des phases de B est différent de A alors cette relation ne peut pas donner de bons résultats car les relations : $\mathbf{r}^{(B)} = \mathbf{r}^{(A)}, \mathbf{p}^{(B)} = \mathbf{p}^{(A)}, V^{(B)} = V^{(A)}$ ne sont plus valides. Afin de contourner ce problème, il est d'usage de décomposer le chemin thermodynamique entre A et B en sous états couplés à A et B par l'intermédiaire du paramètre de couplage λ variant entre 0 et 1. Soit Nt le nombre total de subdivisions, on identifie chaque subdivision par i. $(i = Nt)$ est l'état correspondant à B et $(i = 0)$ celui de l'état A (*cf* figure (4.1)). Si on considére la théorie de Zwanzig, la relation (4.13) donne l'équation (4.14).

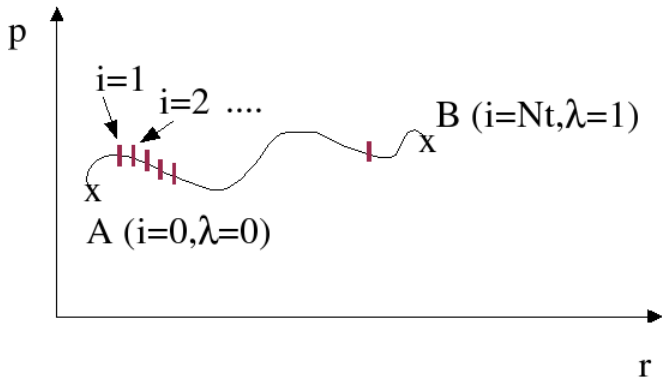

FIG. 4.1 – Schéma permettant de visualiser la décomposition du chemin thermodynamique entre les états A et B dans l'espace des phases. λ est le paramètre de couplage, i permet d'identifier chaque micro-état. Nt est le nombre total de subdivisions.

$$\begin{aligned}
\Delta G_{A\to B} &= G^{(B)} - G^{(A)} = \left(G^{(Nt)} - G^{(Nt-1)}\right) + \left((G^{(Nt-1)} - G^{(Nt-2}\right) .... + \left((G^{(1)} - G^{(0)}\right) \\
&= -kT \ln \left[\left(\frac{Q_{NpT}^{(Nt)}}{Q_{NpT}^{(Nt-1)}} \right)_{Nt-1} \left(\frac{Q_{NpT}^{(Nt-1)}}{Q_{NpT}^{(Nt-2)}} \right)_{Nt-2} .... \left(\frac{Q_{NpT}^{(1)}}{Q_{NpT}^{(0)}} \right)_{0} \right] \\
&= -kT \ln \left[R^{(Nt-1)} R^{(Nt-2)} .... R^{(0)} \right] = \sum_i \Delta G^{(i)} \text{ tel que } \Delta G^{(i)} = -kT \ln R^{(i)} \quad (4.14)
\end{aligned}$$

$()_{Nt-1}$ indique que le rapport est évalué à l'état $Nt-1$. On considère que chaque fonction au numérateur est une perturbation (au sens de Zwanzig) de la fonction qui se trouve au dénominateur. Cette relation postule que les espaces des phases parcourues dans les états i et $i-1$ pour le calcul de ΔG à l'état $i-1$ sont identiques. C'est en cela que réside le point faible de cette théorie. En effet, dans l'état $i-1$, l'hamiltonien de l'état i vaut $\left[\mathcal{H}(\mathbf{r}^{i-1}, \mathbf{p}^{i-1})\right]_{i-1}^{i}$. Or, l'hamiltonien de l'état i dans l'état i vaut $\left[\mathcal{H}(\mathbf{r}^{i}, \mathbf{p}^{i})\right]_{i}^{i}$. L'indice indique l'état où est effectué le calcul, l'exposant indique l'état auquel correspond l'hamiltonien. Pour que ces deux hamiltoniens soient identiques, il faut que l'amplitude de la perturbation soit très faible. Cela implique un nombre de sous-états très importants. Il faudrait être capable de calculer l'hamiltonien suivant $\left[\mathcal{H}(\mathbf{r}^{i}, \mathbf{p}^{i})\right]_{i-1}^{i}$. On sort alors du cadre théorique de la théorie de Zwanzig : on considère alors que, les espaces des phases parcourus dans l' état perturbé et dans l'état de référence sont différents. Dans ces conditions, on suppose que la perturbation appliquée à l'état de référence $i-1$ affecte les variables dépendant des paramètres perturbés (variables

cachées ou non cachées (section (2.3.2)). Si les paramètres d'interaction sont perturbés alors l'énergie potentielle et les forces exercées sur chaque atome sont aussi perturbées .

Pour que les deux premières équations de la relation (4.14) soient équivalentes, il est nécessaire que chaque rapport R^i ait une fonction de partition en commun. Or, l'application de la méthode de perturbation de Zwanzig ne permet pas cette condition, puisque les espaces des phases parcourus sont différents pour chaque rapport R^i : c'est le problème d'orthogonalité. Afin d'améliorer ceci, nous décrivons dans la section (4.6) une nouvelle approche permettant d'évaluer plus précisement l'espace des phases parcouru dans'état perturbé à partir de celui de l'état de référence.

4.2.3 Approche alternative du concept de perturbation

Il existe une autre manière d'appréhender la méthode de perturbation. En effet, on justifie l'équivalence des espaces des phases parcourus entre les deux états par l'existence d'un recouvrement des enveloppes des espaces des phases échantillonnés entre l'état de référence et l'état perturbé (*cf* schéma de la figure (4.2)). La figure (4.2) montre qu'il existe une zone de recouvrement entre les deux espaces des phases parcourus pour un micro-état donné. Néanmoins, les points d'intégration des domaines intégrables ne sont pas les mêmes. Les méthodes de perturbation suggèrent l'équivalence des points d'intégration pour les deux états. Cette équivalence est justifiée par un recouvrement entre les domaines d'intégration. Il n'est pas possible de considérer les points de l'état perturbé puisqu'il serait alors impossible d'écrire les moyennes d'ensemble dans la mesure où les volumes infinitésimaux sont différents. La figure (4.2) montre que, malgré une zone d'intégration en commnun pour les deux états, les points d'intégration sont différents. Finalement, il faudrait déterminer un nouveau jeu de coordonnées et de moments conjugués de l'état perturbé à partir de l'état de référence en respectant l'égalité entre les volumes infinitésimaux. Cela permettrait ainsi d'obtenir une vision plus juste de l'état perturbé en termes énergétiques et structuraux. Même si cette structure n'est pas équivalente à celle de l'état perturbé, elle est plus réaliste que l'état de référence. Dans la section (4.6) nous proposons un calcul permettant de déterminer un nouveau jeu de coordonnées et de moments conjugués de l'état perturbé à partir de l'intégration des équations du mouvement. Pour valider notre calcul, nous démontrons que le nouvel espace engendré permet d'écrire les moyennes d'ensemble. Nous devons donc montrer l'équivalence des volumes

infinitésimaux.

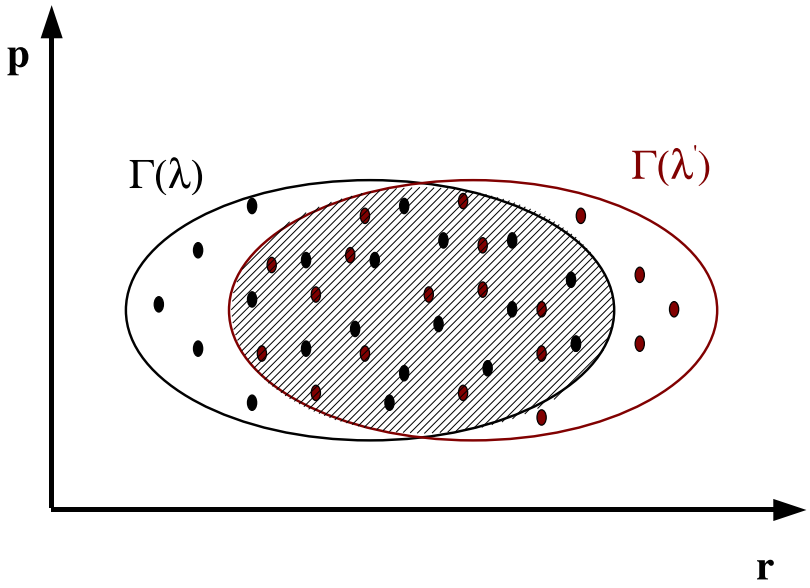

FIG. 4.2 – Recouvrement des espaces des phases entre l'état de référence et l'état perturbé. $\Gamma(\lambda)$ est l'espace des phases de l'état de référence et $\Gamma(\lambda^{'})$ représente l'espace des phases parcourus de l'état perturbé. Les cercles indiquent les points d'intégration relatifs aux positions atomiques.

Il est possible de montrer que si $(r,p) \neq (r^{'},p^{'})$ alors $d\tau = d\tau^{'}$ en utilisant une tranformation canonique. En effet, supposons que les couples $(r^{'},p^{'})$ et (r,p) sont liés par la relation $r^{'} = r + \mathcal{F}(\ddot{r})$. Soit d$\tau$ l'élément de volume dans l'espace des phases tel que $d\tau = dr_1...dr_N dp_1...dp_N$. On montre dans la section (9.1) que le volume $d\tau$ est invariant par rapport à une transformation canonique particulière (développement en série de Taylor). Dans notre cas, la transformation canonique est la transformation de (r,p) en $(r^{'},p^{'})$

$$r^{'}(r_1,...,r_N;t)\ ,\ p^{'}(p_1,...,p_N;t) \tag{4.15}$$

telle que les équations d'Hamilton gardent la même forme avec ces nouvelles variables. Soit $\mathcal{H}^{'}(r_1^{'},...,r_N^{'},p_1^{'},...,p_N^{'};t)$ la fonction de Hamilton exprimée avec les variables $(r^{'},p^{'})$. Dans le cas d'une transformation canonique, les équations de mouvement sont conservées [66]. Ce résultat se traduit par la relation (4.16).

$$\dot{r_i^{'}} = \frac{\partial \mathcal{H}^{'}}{\partial p_i}\ ,\ \dot{p_i^{'}} = \frac{\partial \mathcal{H}^{'}}{\partial r_i} \tag{4.16}$$

L'invariance du volume par rapport aux transformations canoniques s'exprime par l'équation (4.17). C'est cette relation que nous devons démontrer pour la transformation canonique envisagée.

$$\int dr_1...dr_N dp_1...dp_N = \int dr_1^{'}...,dr_N^{'}dp_1^{'}...dp_N^{'} \tag{4.17}$$

En effet, avec le changement de variables décrit par la relation (4.15), on a :

$$\int dr_1...dr_N dp_1...dp_N = \int |J|\, dr_1^{'}...,dr_N^{'}dp_1^{'}...dp_N^{'} \tag{4.18}$$

où $|J|$ est le jacobien de la transformation à l'ordre un. Or, ce jacobien est égal à :

$$J = \begin{vmatrix} \frac{\partial r_1^{'}}{\partial r_1} & ... & ... & \frac{\partial r_1^{'}}{\partial r_N} & ... & ... & \frac{\partial r_1^{'}}{\partial p_N} \\ ... & ... & ... & ... & ... & ... & ... \\ \frac{\partial r_N^{'}}{\partial r_1} & ... & ... & \frac{\partial r_N^{'}}{\partial r_N} & ... & ... & \frac{\partial r_N^{'}}{\partial p_N} \\ \frac{\partial p_1^{'}}{\partial r_1} & ... & ... & \frac{\partial p_1^{'}}{\partial r_N} & ... & ... & \frac{\partial p_1^{'}}{\partial p_N} \\ ... & ... & ... & ... & ... & ... & ... \\ \frac{\partial p_N^{'}}{\partial r_1} & ... & ... & \frac{\partial p_N^{'}}{\partial r_N} & ... & ... & \frac{\partial p_N^{'}}{\partial p_N} \end{vmatrix}$$

Celui-ci peut aussi s'écrire de la manière suivante :

$$J = \sum_i \sum_j \left[\frac{\partial r^{'}{}_i}{\partial r_j} \frac{\partial p^{'}{}_i}{\partial p_j} - \frac{\partial r^{'}{}_i}{\partial p_j} \frac{\partial p^{'}{}_i}{\partial r_j} \right] \tag{4.19}$$

Le changement de variables effectué correspond à une transformation canonique puisque les équations de mouvement sont conservées (hypothèse de départ) : $\dot{r}^{'}{}^i = \frac{\partial \mathcal{H}^{'}}{\partial p^{'}{}_i}$, $\dot{p}^{'}{}_i = -\frac{\partial \mathcal{H}^{'}}{\partial r^{'}{}_i}$. Dans ces conditions, les crochets de Poisson sont les mêmes que ceux exprimés par les variables initiales, c'est-à-dire :

$$\{r^{'}{}_i, r^{'}{}_j\} = 0,\ \{p^{'}{}_i, p^{'}{}_j\} = 0,\ \{r^{'}{}_i, p^{'}{}_j\} = \delta_{ij} \tag{4.20}$$

Nous obtenons alors la relation (4.21) :

$$\frac{\partial r^{'}{}_i}{\partial r_j}\frac{\partial p^{'}{}_i}{\partial p_j} = 1 \; et \; \frac{\partial r^{'}{}_i}{\partial p_j}\frac{\partial p^{'}{}_i}{\partial r_j} = 0 \rightarrow J = 1 \tag{4.21}$$

En effet, $r^{'}{}_i$ ne dépend pas de p_i et $p^{'}{}_i$ ne dépend pas de r_i (équation (4.15)). En outre, on montre que $\dfrac{\partial r^{'}{}_i}{\partial r_j} = 1$ puisque $r^{'}{}_i = r_i + \Delta t^2 \mathcal{F}$ (cette relation sera démontrée dans la section (4.6)) où $\mathcal{F}$ est une fonctionnelle qui dépend de la dérivée seconde de r_i par rapport au temps. Comme nous avons supposé un développement au premier ordre, on peut négliger les termes du deuxième ordre. Nous montrons de manière explicite l'invariance du volume par rapport à la tranformation canonique envisagée. On en déduit donc l'existence du couple $(r^{'}{}_i, p^{'}{}_i) \neq (r_i, p_i)$ tel $d\tau^{'} = d\tau$. Le choix d'un développement d'ordre un permet de travailler avec un jacobien de forme simple. Dans ce cas, seules les dérivées premières sont considérées. Celles-ci peuvent être déterminées par les équations canoniques.

Une méthode susceptible d'abaisser davantage ce problème d'orthogonalité serait le couplage du paramètre de Kirkwood avec le temps. Dans ce cas, cela nous permettrait de subdiviser davantage l'espace des phases parcouru sans perte de temps de calcul. Néanmoins, comme l'hamiltonien est différent pour chaque pas de perturbation, la fonction de partition de chaque micro-état est différente. Dans ces conditions, la dynamique que l'on engendre n'est plus à l'équilibre mais se trouve dans un état hors équilibre puisque l'énergie totale du système n'est pas conservée ($\Leftrightarrow \frac{d\mathcal{H}}{dt} \neq 0$). Les formules opérationnelles permettant de calculer ΔG et ΔS pour ce formalisme hors-équilibre ont été démontrées dans la section (9.5).

4.3 *Free Energy Perturbation* : FEP

4.3.1 Calcul de ΔG

Avant de commencer, il est nécessaire d'écrire l'expression rigoureuse de la variation d'enthalpie libre (équation (4.22)). À partir de la relation (4.22), nous effectuerons des approximations successives permettant d'accéder aux expressions les plus couramment utilisées dans la littérature. Cette démarche a pour objectif d'évaluer la nature et l'expression de chacune des approximations. La suite des développements sera réalisée en considérant la fonction de partition dans l'ensemble NpT. Commençons par écrire la différence d'enthalpie libre entre les états λ et $\lambda^{'}$.

$$\Delta G_\lambda = G_{\lambda'} - G_\lambda = -kT \ln \frac{\int\int\int \exp\left(-\frac{(\mathcal{H}(\mathbf{r}^N,\mathbf{p}^N)+pV)_{\lambda'}}{kT}\right) d\mathbf{r}^N_{\lambda'} d\mathbf{p}^N_{\lambda'} dV_{\lambda'}}{\int\int\int \exp\left(-\frac{(\mathcal{H}(\mathbf{r}^N,\mathbf{p}^N)+pV)_{\lambda}}{kT}\right) d\mathbf{r}^N_{\lambda} d\mathbf{p}^N_{\lambda} dV_{\lambda}} \tag{4.22}$$

Comme le montre cette expression, nous avons considéré la fonction de partition à $\lambda^{'}$ et λ. Cela implique de considérer que l'espace des phases parcourus dans les deux états sont différents. Il est évident que cette expression est inexploitable puisqu'il est impossible de déterminer une moyenne d'ensemble, laquelle se calcule dans un seul et unique espace des phases comme le montre la relation suivante :

$$< X >_{NpT} = \frac{\int\int\int X(\mathbf{r},\mathbf{p}) \exp\left(\frac{-\left(\mathcal{H}(\mathbf{r}^N,\mathbf{p}^N)+pV\right)}{kT}\right) d\mathbf{r}^N d\mathbf{p}^N dV}{\int\int\int \exp\left(\frac{-(\mathcal{H}(\mathbf{r}^N,\mathbf{p}^N)+pV)}{kT}\right) d\mathbf{r}^N d\mathbf{p}^N dV} \tag{4.23}$$

Néanmoins, si on applique la théorie de Zwanzig c'est-à-dire que l'on considère l'état perturbé comme une perturbation de l'état de référence, on peut écrire l'équation suivante :

$$\Delta G_\lambda = G_{\lambda'} - G_\lambda = -kT \ln \frac{\int\int\int \exp\left(-\frac{(\mathcal{H}_{\lambda'}(\mathbf{r}^N,\mathbf{p}^N)+pV)}{kT}\right) d\mathbf{r}^N_{\lambda} d\mathbf{p}^N_{\lambda} dV_{\lambda}}{\int\int\int \exp\left(-\frac{(\mathcal{H}_{\lambda}(\mathbf{r}^N,\mathbf{p}^N)+pV)}{kT}\right) d\mathbf{r}^N_{\lambda} d\mathbf{p}^N_{\lambda} dV_{\lambda}}$$

Comme on l'a vu dans la section (4.2) il existe une alternative permettant de postuler l'invariance du volume infinitésimal entre les deux états tout en considérant des N-uplet $(\mathbf{r},\mathbf{p})^N$ différents entre les deux états. Cela est possible si et seulement si les couples $(\mathbf{r},\mathbf{p})^N$ à l'état perturbé respectent les équations canoniques du mouvement et s'ils sont calculés à partir de l'état de référence. Il est possible de montrer l'existence d'un couple $(\mathbf{r},\mathbf{p})^N$ à l'état perturbé différent de celui de l'état de référence vérifiant l'invariance du volume infinitésimal. Pour cela, il suffit de postuler cette invariance ($d\tau = d\tau'$) et de développer cette expression :

$$\begin{aligned}
d\tau = d\tau' \quad &\Leftrightarrow \quad d\mathbf{r}^N_\lambda d\mathbf{p}^N_\lambda = d\mathbf{r}^N_{\lambda'} d\mathbf{p}^N_{\lambda'} \\
&\Leftrightarrow \quad d(\mathbf{r}^N_\lambda \mathbf{p}^N_\lambda)\left[\frac{\mathbf{r}^N_\lambda}{d\mathbf{r}^N_\lambda} + \frac{\mathbf{p}^N_\lambda}{d\mathbf{p}^N_\lambda}\right]^{-1} = d(\mathbf{r}^N_{\lambda'} d\mathbf{p}^N_{\lambda'})\left[\frac{\mathbf{r}^N_{\lambda'}}{d\mathbf{r}^N_{\lambda'}} + \frac{\mathbf{p}^N_{\lambda'}}{d\mathbf{p}^N_{\lambda'}}\right]^{-1} \\
&\Leftrightarrow \quad d(\mathbf{r}^N_\lambda \mathbf{p}^N_\lambda)\left[\frac{d\ln(\mathbf{r}^N_\lambda \mathbf{p}^N_\lambda)}{d\ln(\mathbf{r}^N_\lambda) d\ln(\mathbf{p}^N_\lambda)}\right]^{-1} = d(\mathbf{r}^N_{\lambda'} \mathbf{p}^N_{\lambda'})\left[\frac{d\ln(\mathbf{r}^N_{\lambda'} \mathbf{p}^N_{\lambda'})}{d\ln(\mathbf{r}^N_{\lambda'}) d\ln(\mathbf{p}^N_{\lambda'})}\right]^{-1}
\end{aligned} \tag{4.24}$$

La solution la plus simple de cette équation est :

$$\mathbf{r}_{\lambda}^{N} = \mathbf{r}_{\lambda'}^{N} \quad , \quad \mathbf{p}_{\lambda}^{N} = \mathbf{p}_{\lambda'}^{N} \tag{4.25}$$

Il existe une autre solution moins simple qui respecte l'hypothèse de départ $d\tau = d\tau'$:

$$\mathbf{r}_{\lambda}^{N}\mathbf{p}_{\lambda}^{N} = \mathbf{r}_{\lambda'}^{N}\mathbf{p}_{\lambda'}^{N} \text{ avec } \mathbf{r}_{\lambda}^{N} \neq \mathbf{r}_{\lambda'}^{N} \text{ et } \mathbf{p}_{\lambda}^{N} \neq \mathbf{p}_{\lambda'}^{N} \tag{4.26}$$

Cette équation nous permet donc de réécrire la relation (4.22), en considérant que les volumes dans les deux états sont identiques.

$$\Delta G_{\lambda} = -kT \ln \frac{\int\int\int \exp\left(-\frac{(\mathcal{H}_{\lambda'}(\mathbf{r}^N,\mathbf{p}^N)+pV)}{kT}\right) d\mathbf{r}^N d\mathbf{p}^N dV}{\int\int\int \exp\left(-\frac{(\mathcal{H}_{\lambda}(\mathbf{r}^N,\mathbf{p}^N)+pV)}{kT}\right) d\mathbf{r}^N d\mathbf{p}^N dV} \tag{4.27}$$

Ce calcul montre l'existence des couples $(\mathbf{r},\mathbf{p})_{\lambda'}^{N}$ différents des couples $(\mathbf{r},\mathbf{p})_{\lambda}^{N}$ si et seulement si les équations canoniques sont respectées dans l'état perturbé et si le N-uplet $(\mathbf{r},\mathbf{p})_{\lambda'}^{N}$ de l'état perturbé est fonction du N-uplet $(\mathbf{r},\mathbf{p})_{\lambda}^{N}$ de l'état de référence. En outre, il faut que ces relations vérifient l'égalité $d\tau = d\tau'$. Il nous faut, à présent, définir une méthode permettant de déterminer ce nouveau jeu de paramètres.

Avant de discuter la détermination de ces couples $(\mathbf{r},\mathbf{p})_{\lambda'}^{N}$, nous allons développer les calculs menant aux formules opérationnelles de perturbations. Pour cela, nous commencerons par considérer l'hamiltonien total. Ces calculs ont pour point de départ la relation (4.27). La détermination de la moyenne d'ensemble nécessite de définir les variables suivantes :

$$\begin{aligned} E^{\mathcal{H}} &= \exp\left(-\left[\frac{\mathcal{H}_{\lambda}(\mathbf{r}^N,\mathbf{p}^N)+pV}{kT}\right]\right)\cdot\exp\left(\frac{\mathcal{H}_{\lambda}(\mathbf{r}^N,\mathbf{p}^N)+pV}{kT}\right) \\ I^{\mathcal{H}} &= \frac{\int\int\int \exp\left(-\left[\frac{\mathcal{H}_{\lambda}(\mathbf{r}^N,\mathbf{p}^N)+pV}{kT}\right]\right) d\mathbf{r}^N d\mathbf{p}^N dV}{\int\int\int \exp\left(-\left[\frac{\mathcal{H}_{\lambda}(\mathbf{r}^N,\mathbf{p}^N)+pV}{kT}\right]\right) d\mathbf{r}^N d\mathbf{p}^N dV} \end{aligned}$$

L'obtention de la moyenne d'ensemble se fait en multipliant le numérateur de la relation (4.27) par $\mathrm{I}^{\mathcal{H}}$:

$$\begin{aligned}\Delta G_\lambda &= \\ &- kT\ln\frac{\int\int\int \exp\left(\frac{\mathcal{H}_{\lambda'}(\mathbf{r}^N,\mathbf{p}^N)+pV}{kT}-(\frac{\mathcal{H}_\lambda(\mathbf{r}^N,\mathbf{p}^N)+pV}{kT})\right)\cdot\exp\left(\frac{\mathcal{H}_\lambda(\mathbf{r}^N,\mathbf{p}^N)+pV}{kT}\right)d\mathbf{r}^N d\mathbf{p}^N dV}{\int\int\int \exp\left(-\frac{\mathcal{H}_\lambda(\mathbf{r}^N,\mathbf{p}^N)+pV}{kT}\right)d\mathbf{r}^N d\mathbf{p}^N dV}\end{aligned}$$

Ce qui nous permet d'écrire après simplification :

$$\Delta G_\lambda = -kT\ln\left\langle\exp\left(-\left[\frac{\mathcal{H}_{\lambda'}(\mathbf{r}^N,\mathbf{p}^N)-\mathcal{H}_\lambda(\mathbf{r}^N,\mathbf{p}^N)}{kT}\right]\right)\right\rangle_\lambda \tag{4.28}$$

L'indice sur l'hamiltonien indique que l'on tient compte des paramètres externes qui ont été perturbés, en gardant un espace des phases parcourus identique pour les deux états. Si on décompose l'hamiltonien total en un terme d'énergie cinétique et d'énergie potentielle, nous obtenons la relation (4.29).

$$\begin{aligned}\Delta G_\lambda &= \\ &- kT\ln\left\langle\exp\left(-\left[\frac{\mathcal{U}_{\lambda'}(\mathbf{r}^N)-\mathcal{U}_\lambda(\mathbf{r}^N)}{kT}\right]-\frac{v_\lambda^2}{2kT}\left[m_{\lambda'}-m_\lambda\right]\right)\right\rangle_\lambda\end{aligned} \tag{4.29}$$

Si on utilise la longueur d'onde de De Broglie pour évaluer le terme cinétique nous obtenons :

$$\Delta G_\lambda = -kT\ln\left[\left(\frac{m_{\lambda'}}{m_\lambda}\right)^{\frac{3}{2}}\right]\frac{\int\int\exp\left(-\left[\frac{\mathcal{U}_{\lambda'}(\mathbf{r}^N)+pV}{kT}\right]\right)d\mathbf{r}^N dV}{\int\int\exp\left(-\left[\frac{\mathcal{U}_\lambda(\mathbf{r}^N)+pV}{kT}\right]\right)d\mathbf{r}^N dV} \tag{4.30}$$

Définissons les termes suivants :

$$\begin{aligned}E^{\mathcal{U}} &= \exp\left(-\left[\frac{\mathcal{U}_\lambda(\mathbf{r}^N,\mathbf{p}^N)+pV}{kT}\right]\right)\cdot\exp\left(\frac{\mathcal{U}_\lambda(\mathbf{r}^N,\mathbf{p}^N)+pV}{kT}\right) \\ I^{\mathcal{U}} &= \frac{\int\int\int\exp\left(-\left[\frac{\mathcal{U}_\lambda(\mathbf{r}^N,\mathbf{p}^N)+pV}{kT}\right]\right)d\mathbf{r}^N d\mathbf{p}^N dV}{\int\int\int\exp\left(-\left[\frac{\mathcal{U}_\lambda(\mathbf{r}^N,\mathbf{p}^N)+pV}{kT}\right]\right)d\mathbf{r}^N d\mathbf{p}^N dV}\end{aligned}$$

En réitérant le même type de calcul effectué avec l'hamiltonien total (multiplication par E^U), on a :

$$\Delta G_\lambda = -kT \ln \left(\left[\frac{m_{\lambda'}}{m_\lambda} \right]^{\frac{3}{2}} \left\langle \exp \left(- \left[\frac{\mathcal{U}_{\lambda'}(\mathbf{r}^N) - \mathcal{U}_\lambda(\mathbf{r}^N)}{kT} \right] \right) \right\rangle_\lambda \right) \tag{4.31}$$

En théorie les relations (4.28) et (4.31) sont identiques, on s'attend donc à ce que les deux calculs fournissent les mêmes résultats : nous verrons par la suite que ce n'est pas le cas. Si on néglige le terme massique, les équations (4.28) et (4.31) donnent le résultat connu :

$$\Delta G_\lambda = -kT \ln \left\langle \exp \left(- \left[\frac{\mathcal{U}_{\lambda'}(\mathbf{r}^N) - \mathcal{U}_\lambda(\mathbf{r}^N)}{kT} \right] \right) \right\rangle_\lambda \tag{4.32}$$

4.3.2 Calcul de ΔS

L'entropie est donnée par la relation suivante :

$$\begin{aligned} S(N,p,T) &= k \ln Q_{NpT} + kT \left(\frac{\partial \ln Q_{NpT}}{\partial T} \right)_{N,p} \\ &= k \ln Q_{NpT} + kT \frac{1}{Q_{NpT}} \left(\frac{\partial Q_{NpT}}{\partial T} \right)_{N,p} \end{aligned}$$

La variation d'entropie entre les deux états λ et λ' vaut :

$$\begin{aligned} \Delta S_\lambda &= k \ln \frac{[Q_{NpT}]_{\lambda'}}{[Q_{NpT}]_\lambda} \\ &+ kT \left(\frac{1}{[Q_{NpT}]_{\lambda'}} \frac{\partial [Q_{NpT}]_{\lambda'}}{\partial T} - \frac{1}{[Q_{NpT}]_\lambda} \frac{\partial [Q_{NpT}]_\lambda}{\partial T} \right) \end{aligned}$$

Nous avons trois termes à déterminer. Le premier correspond à la variation d'enthalpie libre précédemment calculée.

Terme à λ :

$$\begin{aligned} \frac{\partial [Q_{NpT}]_\lambda}{\partial T} &= \frac{1}{N! \Lambda^{3N} V^o} \int \int \left(\frac{\mathcal{U}_\lambda(\mathbf{r}^N) + pV}{kT^2} \right) \exp \left(- \left[\frac{\mathcal{U}_\lambda(\mathbf{r}^N) + pV}{kT} \right] \right) d\mathbf{r}^N dV \\ &+ \frac{1}{N! \Lambda^{3N} V^o} \frac{3N}{2T} \int \int \exp \left(- \left[\frac{\mathcal{U}_\lambda(\mathbf{r}^N) + pV}{kT} \right] \right) d\mathbf{r}^N dV \end{aligned}$$

Le terme $\left(\frac{\partial P}{\partial T}\right)_P$ est nul puisque l'ensemble considéré est l'ensemble NpT. Le terme $\left(\frac{\partial V}{\partial T}\right)_P$ est nul car le volume instantané ne dépend pas de T contrairement à $< V >$. L'expression $\frac{3N}{2T}$ est issue du développement suivant :

$$\begin{aligned}\frac{\partial}{\partial T}\left(\frac{1}{\Lambda^{3N}}\right) &= \left(\frac{2mk\pi}{h^2}\right)^{\frac{3N}{2}} \cdot \frac{3N}{2} \cdot (T)^{\frac{3N}{2}-1} \\ &= \frac{3N}{2T}\left(\frac{1}{\Lambda^{3N}}\right)\end{aligned}$$

La suite des calculs donne :

$$\begin{aligned}\frac{1}{[Q_{NpT}]_\lambda}\frac{\partial [Q_{NpT}]_\lambda}{\partial T} &= \frac{\int\int \left(\frac{\mathcal{U}_\lambda(\mathbf{r}^N)+pV}{kT^2}\right) \exp\left(-\left[\frac{\mathcal{U}_\lambda(\mathbf{r}^N)+pV}{kT}\right]\right) d\mathbf{r}^N dV}{\int\int \exp\left(-\left[\frac{\mathcal{U}_\lambda(\mathbf{r}^N)+pV}{kT}\right]\right) d\mathbf{r}^N dV} \\ &+ \frac{3N}{2T}\end{aligned}$$

$$\frac{1}{[Q_{NpT}]_\lambda}\frac{\partial [Q_{NpT}]_\lambda}{\partial T} = \left\langle \frac{\mathcal{U}_\lambda(\mathbf{r}^N)+pV}{kT^2} \right\rangle_\lambda + \frac{3N}{2T}$$

Terme à λ' :

$$\frac{1}{[Q_{NpT}]_{\lambda'}}\frac{\partial [Q_{NpT}]_{\lambda'}}{\partial T} = \frac{\int\int \left(\frac{\mathcal{U}_{\lambda'}(\mathbf{r}^N)+pV}{kT^2}\right) \exp\left(-\left[\frac{\mathcal{U}_{\lambda'}(\mathbf{r}^N)+pV}{kT}\right]\right) d\mathbf{r}^N dV}{\int\int \exp\left(-\left[\frac{\mathcal{U}_{\lambda'}(\mathbf{r}^N)+pV}{kT}\right]\right) d\mathbf{r}^N dV} + \frac{3N}{2T}$$

Multiplions ce terme par $\frac{E^U}{E^U}$:

$$\begin{aligned}&\frac{1}{[Q_{NpT}]_{\lambda'}}\frac{\partial [Q_{NpT}]_{\lambda'}}{\partial T} = \\ &\frac{\int\int \left(\frac{\mathcal{U}_{\lambda'}(\mathbf{r}^N)+pV}{kT^2}\right) \exp\left(-\left[\frac{\mathcal{U}_{\lambda'}(\mathbf{r}^N)-\mathcal{U}_\lambda(\mathbf{r}^N)}{kT}\right]\right) \exp\left(-\left[\frac{\mathcal{U}_\lambda(\mathbf{r}^N)+pV}{kT}\right]\right) d\mathbf{r}^N dV}{\int\int \exp\left(-\left[\frac{\mathcal{U}_{\lambda'}(\mathbf{r}^N)-\mathcal{U}_\lambda(\mathbf{r}^N)}{kT}\right]\right) \exp\left(-\left[\frac{\mathcal{U}_\lambda(\mathbf{r}^N)+pV}{kT}\right]\right) d\mathbf{r}^N dV} + \frac{3N}{2T}\end{aligned}$$

La multiplication par I^U nous permet d'obtenir :

$$\left[\frac{\int\int\left(\frac{\mathcal{U}_{\lambda'}(\mathbf{r}^N)+pV}{kT^2}\right)\exp\left(-\left[\frac{\mathcal{U}_{\lambda'}(\mathbf{r}^N)-\mathcal{U}_{\lambda}(\mathbf{r}^N)}{kT}\right]\right)d\mathbf{r}^N dV}{\int\int\exp\left(-\left[\frac{\mathcal{U}_{\lambda'}(\mathbf{r}^N,)-\mathcal{U}_{\lambda}(\mathbf{r}^N)}{kT}\right]\right)\exp\left(-\left[\frac{\mathcal{U}_{\lambda}(\mathbf{r}^N)+pV}{kT}\right]\right)d\mathbf{r}^N dV}\right.$$
$$\left.\cdot\exp\left(-\left[\frac{\mathcal{U}_{\lambda}(\mathbf{r}^N)+pV}{kT}\right]\right)d\mathbf{r}^N dV\right]\cdot\frac{\int\int\exp\left(-\left[\frac{\mathcal{U}_{\lambda}(\mathbf{r}^N,\lambda)+pV}{kT}\right]\right)d\mathbf{r}^N dV}{\int\int\exp\left(-\left[\frac{\mathcal{U}_{\lambda}(\mathbf{r}^N)+pV}{kT}\right]\right)d\mathbf{r}^N dV}+\frac{3N}{kT}$$

Cette relation se simplifie en l'équation suivante :

$$\frac{1}{[Q_{NpT}]_{\lambda'}}\frac{\partial\left[Q_{NpT}\right]_{\lambda'}}{\partial T}=\frac{\left\langle\left(\frac{\mathcal{U}_{\lambda'}(\mathbf{r}^N)+pV}{kT^2}\right)\exp\left(-\left[\frac{\mathcal{U}_{\lambda'}(\mathbf{r}^N)-\mathcal{U}_{\lambda}(\mathbf{r}^N)}{kT}\right]\right)\right\rangle_\lambda}{\left\langle\exp\left(-\left[\frac{\mathcal{U}_{\lambda'}(\mathbf{r}^N)-\mathcal{U}_{\lambda}(\mathbf{r}^N)}{kT}\right]\right)\right\rangle_\lambda}+\frac{3N}{2T}$$

Nous obtenons ainsi la relation finale :

$$\begin{aligned}T\Delta S(\lambda)&=kT\ln\left[\frac{m_{\lambda'}}{m_\lambda}\right]^{\frac{3}{2}}\left\langle\exp\left(-\left[\frac{\mathcal{U}_{\lambda'}(\mathbf{r}^N)-\mathcal{U}_{\lambda}(\mathbf{r}^N)}{kT}\right]\right)\right\rangle_\lambda\\&+kT^2\left(\frac{\left\langle\left(\frac{\mathcal{U}_{\lambda'}(\mathbf{r}^N)+pV}{kT^2}\right)\exp\left(-\left[\frac{\mathcal{U}_{\lambda'}(\mathbf{r}^N)-\mathcal{U}_{\lambda}(\mathbf{r}^N)}{kT}\right]\right)\right\rangle_\lambda}{\left\langle\exp\left(-\left[\frac{\mathcal{U}_{\lambda'}(\mathbf{r}^N)-\mathcal{U}_{\lambda}(\mathbf{r}^N)}{kT}\right]\right)\right\rangle_\lambda}\right)\\&-kT^2\left(\left\langle\frac{\mathcal{U}_{\lambda}(\mathbf{r}^N)+pV}{kT^2}\right\rangle_\lambda\right)\end{aligned}\tag{4.33}$$

En considérant l'hamiltonien total, on obtient la relation suivante :

$$\begin{aligned}T\Delta S(\lambda)&=kT\ln\left\langle\exp\left(-\left[\frac{\mathcal{H}_{\lambda'}(\mathbf{r}^N,\mathbf{p}^N)-\mathcal{H}_{\lambda}(\mathbf{r}^N,\mathbf{p}^N)}{kT}\right]\right)\right\rangle_\lambda\\&+kT^2\left(\frac{\left\langle\left(\frac{\mathcal{H}_{\lambda'}(\mathbf{r}^N,\mathbf{p}^N)+pV}{kT^2}\right)\exp\left(-\left[\frac{\mathcal{H}_{\lambda'}(\mathbf{r}^N,\mathbf{p}^N)-\mathcal{H}_{\lambda}(\mathbf{r}^N,\mathbf{p}^N)}{kT}\right]\right)\right\rangle_\lambda}{\left\langle\exp\left(-\left[\frac{\mathcal{H}_{\lambda'}(\mathbf{r}^N,\mathbf{p}^N)-\mathcal{H}_{\lambda}(\mathbf{r}^N,\mathbf{p}^N)}{kT}\right]\right)\right\rangle_\lambda}\right)\\&-kT^2\left(\left\langle\frac{\mathcal{H}_{\lambda}(\mathbf{r}^N,\mathbf{p}^N)+pV}{kT^2}\right\rangle_\lambda\right)\end{aligned}\tag{4.34}$$

4.4 *Thermodynamics Integration* : TI

4.4.1 Calcul de ΔG

Cette méthode détermine la variation d'énergie libre entre deux états en considérant que la fonction d'état G peut s'écrire sous la forme d'une différentielle totale. Ainsi, la variation est obtenue par intégration de la dérivée partielle de G par rapport à λ.

$$\Delta G(\lambda) = \int_0^1 \frac{\partial G}{\partial \lambda} d\lambda = -kT \int_0^1 \frac{1}{Q_{NpT}} \frac{\partial Q_{NpT}}{\partial \lambda} d\lambda \tag{4.35}$$

Pour plus de simplicité dans l'écriture des formules, nous écrirons la longueur d'onde de De Broglie de la manière suivante :

$$\Lambda^{3N} = \left(\frac{C}{m}\right)^{\frac{3N}{2}}$$

La dérivée de la fonction de partition par rapport à λ nous permet d'écrire :

$$\begin{aligned}
\frac{\partial \left[Q_{NpT}\right]}{\partial \lambda} &= \frac{3N}{2} \frac{m_\lambda}{N!C}^{\frac{3N}{2}-1} \frac{\partial \left[m_\lambda\right]}{\partial \lambda} \int\int \exp\left(-\left[\frac{\mathcal{U}_\lambda(\mathbf{r}^N)+pV}{kT}\right]\right) d\mathbf{r}^N dV \\
&- \frac{m_\lambda}{CN!}^{\frac{3N}{2}} \int\int \frac{1}{kT} \frac{\partial \mathcal{U}_\lambda(\mathbf{r}^N)}{\partial \lambda} \exp\left(-\left[\frac{\mathcal{U}_\lambda(\mathbf{r}^N)+pV}{kT}\right]\right) d\mathbf{r}^N dV
\end{aligned}$$

Soit, en divisant cette expression par la fonction de partition :

$$\begin{aligned}
\frac{1}{Q_{NpT}} \frac{\partial \left[Q_{NpT}\right]}{\partial \lambda} &= \frac{3N}{2} m_\lambda^{-1} \frac{\partial \left[m_\lambda\right]}{\partial \lambda} \frac{\int\int \exp\left(-\left[\frac{\mathcal{U}_\lambda(\mathbf{r}^N)+pV}{kT}\right]\right) d\mathbf{r}^N dV}{\int\int \exp\left(-\left[\frac{\mathcal{U}_\lambda(r^N)+pV}{kT}\right]\right) d\mathbf{r}^N dV} \\
&- \frac{\int\int \frac{1}{kT} \frac{\partial \mathcal{U}_\lambda(\mathbf{r}^N)}{\partial \lambda} \exp\left(-\left[\frac{\mathcal{U}_\lambda(\mathbf{r}^N)+pV}{kT}\right]\right) d\mathbf{r}^N dV}{\int\int \exp\left(-\left[\frac{\mathcal{U}_\lambda(r^N)+pV}{kT}\right]\right) d\mathbf{r}^N dV}
\end{aligned} \tag{4.36}$$

Nous obtenons alors :

$$\Delta G(\lambda) = \int_0^1 \left[-\frac{3NkT}{2} \cdot \left\langle \frac{1}{m_\lambda} \cdot \frac{\partial m_\lambda}{\partial \lambda} \right\rangle_\lambda + \left\langle \frac{\partial \mathcal{U}_\lambda(\mathbf{r}^N)}{\partial \lambda} \right\rangle_\lambda \right] d\lambda \tag{4.37}$$

m_λ et $\mathcal{U}_\lambda(\mathbf{r}^N)$ n'indiquent pas dans cette expression un calcul réalisé à λ (comme pour la méthode FEP), mais seulement une dépendance vis-à-vis du paramètre de couplage. Le premier terme peut évidemment s'écrire sans $<>_\lambda$ puisque la masse ne fluctue pas. En considérant l'hamiltonien total, on a :

$$\frac{1}{Q_{NpT}}\frac{\partial\left[Q_{NpT}\right]}{\partial\lambda} = -\frac{\int\int \frac{1}{kT}\left[\frac{\partial\mathcal{U}_\lambda(\mathbf{r}^N)}{\partial\lambda} + \frac{\partial m_\lambda}{\partial\lambda}\sum\mathbf{v_i^2}\right]\exp\left(-\left[\frac{\mathcal{H}_\lambda(\mathbf{r}^N,\mathbf{p}^N)+pV}{kT}\right]\right)d\mathbf{r}^N dV}{\int\int\exp\left(-\left[\frac{\mathcal{H}_\lambda(\mathbf{r^N},\mathbf{p^N})+\mathbf{pV}}{kT}\right]\right)d\mathbf{r}^N dV} \tag{4.38}$$

ce qui nous permet d'obtenir :

$$\Delta G(\lambda) = \int_0^1 \left[\left\langle\left[\frac{\partial\mathcal{U}_\lambda(\mathbf{r}^N)}{\partial\lambda} + \frac{\partial m_\lambda}{\partial\lambda}\sum\mathbf{v_i^2}\right]\right\rangle_\lambda\right]d\lambda \tag{4.39}$$

Si on néglige le rapport de masse, nous obtenons la relation classique :

$$\Delta G(\lambda) = \int_0^1 \left[\left\langle\frac{\partial\mathcal{U}_\lambda(\mathbf{r}^N)}{\partial\lambda}\right\rangle_\lambda\right]d\lambda \tag{4.40}$$

En considérant l'hamiltonien total, on a :

$$\Delta G(\lambda) = \int_0^1 \left[\left\langle\frac{\partial\mathcal{H}_\lambda(\mathbf{r}^N,\mathbf{p}^N)}{\partial\lambda}\right\rangle_\lambda\right]d\lambda \tag{4.41}$$

4.4.2 Calcul de ΔS

De la même manière que pour l'enthalpie libre, on peut écrire :

$$\Delta S(\lambda) = \int dS = \int\frac{\partial S}{\partial\lambda}d\lambda$$

soit :

$$\frac{\partial S}{\partial\lambda} = k\frac{\partial\left[\ln Q_{NpT}\right]}{\partial\lambda} + kT\frac{\partial}{\partial\lambda}\left[\frac{\partial\left[\ln Q_{NpT}\right]}{\partial T}\right]$$

Le premier terme correspond à celui calculé pour l'enthalpie libre. Le deuxième se calcule tout aussi aisément :

$$\frac{\partial}{\partial\lambda}\left[\frac{1}{[Q_{NpT}]_\lambda}\frac{\partial\,[Q_{NpT}]_\lambda}{\partial T}\right] = \frac{\partial}{\partial\lambda}\left[\frac{\int\int\left(\frac{\mathcal{U}_\lambda(\mathbf{r}^N)+pV}{kT^2}\right)\exp\left(-\left[\frac{\mathcal{U}_\lambda \mathbf{r}^N)+pV}{kT}\right]\right)d\mathbf{r}^N dV}{\int\int\exp\left(-\left[\frac{\mathcal{U}_\lambda(\mathbf{r}^N)+pV}{kT}\right]\right)d\mathbf{r}^N dV}\right]$$

$$\begin{aligned}\frac{\partial}{\partial\lambda}\left[\frac{1}{[Q_{NpT}]_\lambda}\frac{\partial\,[Q_{NpT}]_\lambda}{\partial T}\right] &= \left[\left\langle\left(\frac{\mathcal{U}_\lambda(\mathbf{r}^N)+pV}{kT^2}\right)\right\rangle_\lambda\left\langle\frac{1}{kT}\frac{\partial\mathcal{U}_\lambda(\mathbf{r}^N)}{\partial\lambda}\right\rangle\right]_\lambda \\ &- \left[\left\langle\frac{1}{k^2T^3}\frac{\partial\mathcal{U}_\lambda(\mathbf{r}^N)}{\partial\lambda}\left(\mathcal{U}_\lambda(\mathbf{r}^N)+pV\right)\right\rangle_\lambda\right] \\ &+ \left\langle\frac{1}{kT^2}\frac{\partial\mathcal{U}_\lambda(\mathbf{r}^N)}{\partial\lambda}\right\rangle_\lambda\end{aligned}$$

Si on utilise la relation (4.36), on obtient :

$$\frac{\partial\,[\ln Q_{NpT}]}{\partial\lambda} = \frac{3N}{2}\left\langle\frac{1}{m_\lambda}\cdot\frac{\partial m_\lambda}{\partial\lambda}\right\rangle_\lambda - \left\langle\frac{1}{kT}\frac{\partial\mathcal{U}_\lambda(\mathbf{r}^N)}{\partial\lambda}\right\rangle_\lambda$$

On a :

$$\begin{aligned}\frac{\partial S}{\partial\lambda} &= \frac{3kN}{2}\left\langle\frac{1}{m_\lambda}\cdot\frac{\partial m_\lambda}{\partial\lambda}\right\rangle_\lambda + \left[\left\langle\left(\frac{\mathcal{U}_\lambda(\mathbf{r}^N)+pV}{kT^2}\right)\right\rangle_\lambda\left\langle\frac{1}{kT}\frac{\partial\mathcal{U}_\lambda(\mathbf{r}^N)}{\partial\lambda}\right\rangle_\lambda\right][kT] \\ &- \left[\left\langle\frac{1}{k^2T^3}\frac{\partial\mathcal{U}_\lambda(\mathbf{r}^N)}{\partial\lambda}\left(\mathcal{U}_\lambda(\mathbf{r}^N)+pV\right)\right\rangle_\lambda\right][kT]\end{aligned}$$

La variation d'entropie et la variation d'enthalpie ($\Delta H = \Delta G + T\Delta S$) s'écrivent :

$$\begin{aligned}\Delta S(\lambda) &= \int[kT]\left[\left[\left\langle\left(\frac{\mathcal{U}_\lambda(\mathbf{r}^N)+pV}{kT^2}\right)\right\rangle_\lambda\left\langle\frac{1}{kT}\frac{\partial\mathcal{U}_\lambda(\mathbf{r}^N)}{\partial\lambda}\right\rangle_\lambda\right]\right. \\ &- \left.\left[\left\langle\frac{1}{k^2T^3}\frac{\partial\mathcal{U}_\lambda(\mathbf{r}^N)}{\partial\lambda}\left(\mathcal{U}_\lambda(\mathbf{r}^N)+pV\right)\right\rangle_\lambda\right] + \frac{3kN}{2\,[kT]}\left\langle\frac{1}{m_\lambda}\cdot\frac{\partial m_\lambda}{\partial\lambda}\right\rangle_\lambda\right]d\lambda\end{aligned} \tag{4.42}$$

$$\begin{aligned}\Delta H(\lambda) &= \int kT^2\left[\left[\left\langle\left(\frac{\mathcal{U}_\lambda(\mathbf{r}^N)+pV}{kT^2}\right)\right\rangle_\lambda\left\langle\frac{1}{kT}\frac{\partial\mathcal{U}_\lambda(\mathbf{r}^N)}{\partial\lambda}\right\rangle_\lambda\right]\right. \\ &- \left.\left[\left\langle\frac{1}{k^2T^3}\frac{\partial\mathcal{U}_\lambda(\mathbf{r}^N)}{\partial\lambda}\left(\mathcal{U}_\lambda(\mathbf{r}^N)+pV\right)\right\rangle_\lambda\right] + \frac{1}{kT^2}\left\langle\frac{\partial\mathcal{U}_\lambda(\mathbf{r}^N)}{\partial\lambda}\right\rangle_\lambda\right]d\lambda\end{aligned} \tag{4.43}$$

Si l'on considère l'hamiltonien total, on obtient, une relation similaire. Dans les deux cas $\frac{\partial \mathcal{H}_\lambda}{\partial \lambda}$ et $\frac{\partial m_\lambda}{\partial \lambda}$ sont déterminés par une différence finie :

$$\frac{\partial \mathcal{H}_\lambda(\mathbf{r}^N, \mathbf{p}^N)}{\partial \lambda} = \frac{\mathcal{H}(\mathbf{r}^N, \mathbf{p}^N, \lambda') - \mathcal{H}(\mathbf{r}^N, \mathbf{p}^N, \lambda)}{\partial \lambda} \tag{4.44}$$

L'évaluation de l'intégrale peut être réalisée par l'utilisation de diverses méthodes d'intégration numérique (méthode de Gauss, méthode des rectangles,...) ou bien encore par intégration d'une fonction d'ajustement.

La méthode FDTI (*Finite Difference Thermodynamics Integration*) [6] utilise le formalisme de la méthode FEP pour calculer la dérivée partielle de X ($X = G$, S ou H) par rapport à λ (équation (4.45)). Le principal avantage de cette méthode, par rapport aux méthodes FEP et TI est une convergence plus rapide. Nous avons utilisé cette méthode lors d'études précédentes [5, 6].

$$\frac{\partial X}{\partial \lambda} \simeq \frac{\Delta X}{\partial \lambda} \rightarrow \Delta X = \int_0^1 \frac{\partial X}{\partial \lambda} d\lambda = \int_0^1 \frac{\Delta X}{\partial \lambda} d\lambda \tag{4.45}$$

4.5 *Temperature Perturbation* : TP

Cette méthode permet d'évaluer l'entropie et l'enthalpie d'une manière différente. La variation d'énergie libre est déterminée par le méthode FEP. Dans le section (4.3), nous avons vu que la variation d'entropie pouvait s'écrire :

$$\Delta S_\lambda = k \ln \frac{[Q_{NpT}]_{\lambda'}}{[Q_{NpT}]_\lambda} + kT \left(\frac{1}{[Q_{NpT}]_{\lambda'}} \frac{\partial [Q_{NpT}]_{\lambda'}}{\partial T} - \frac{1}{[Q_{NpT}]_\lambda} \frac{\partial [Q_{NpT}]_\lambda}{\partial T} \right)$$

Dans la méthode FEP, on dérive de manière explicite la fonction de partition en fonction de la température, alors qu'ici cette dérivée est déterminée par une différence finie entre les températures $T' = T + \Delta T$ et $T'' = T - \Delta T$, comme le montre l'équation suivante :

$$\frac{\partial \ln [Q_{NpT}]_\lambda}{\partial T} = \frac{\ln \left[Q_{NpT'}\right]_\lambda - \ln \left[Q_{NpT''}\right]_\lambda}{2\Delta T} = \frac{1}{2\Delta T} \ln \frac{\left[Q_{NpT'}\right]_\lambda}{\left[Q_{NpT''}\right]_\lambda}$$

Avec cette nouvelle expression, on a alors :

$$\Delta S_\lambda = k \ln \frac{[Q_{NpT}]_{\lambda'}}{[Q_{NpT}]_\lambda} + \frac{kT}{2\Delta T}\left[\ln\left[\frac{Q_{NpT'}}{Q_{NpT''}}\right]_{\lambda'} - \ln\left[\frac{Q_{NpT'}}{Q_{NpT''}}\right]_{\lambda}\right] \tag{4.46}$$

Nous devons donc évaluer trois termes. Le premier ayant été déterminé dans la section précédente, nous devons calculer les deux termes restants, l'un à λ et l'autre à $\lambda^{'}$.

Terme à $\lambda^{'}$:

Le rapport de fonctions de partition relatif aux deux températures est donné par :

$$\left[\frac{Q_{NpT'}}{Q_{NpT''}}\right]_{\lambda'} = \frac{\Lambda_{T''}^{3N}\int\int \exp\left(-\left[\frac{\mathcal{U}_{\lambda'}(\mathbf{r}^N)+pV}{kT'}\right]\right) d\mathbf{r}^N dV}{\Lambda_{T'}^{3N}\int\int \exp\left(-\left[\frac{\mathcal{U}_{\lambda'}(\mathbf{r}^N)+pV}{kT''}\right]\right) d\mathbf{r}^N dV} \tag{4.47}$$

le premier terme valant :

$$\frac{\Lambda_{T''}^{3N}}{\Lambda_{T'}^{3N}} = \left(\frac{T'}{T''}\right)^{\frac{3N}{2}}$$

Si nous multiplions le dénominateur et le numérateur par E^U et I^U, nous pouvons écrire :

$$\left[\frac{Q_{NpT'}}{Q_{NpT''}}\right]_{\lambda'} = \left(\frac{T'}{T''}\right)^{\frac{3N}{2}} \frac{\left\langle \exp\left(-\left[\frac{\mathcal{U}_{\lambda'}(\mathbf{r}^N)+pV}{kT'} - \left(\frac{\mathcal{U}_{\lambda}(\mathbf{r}^N)+pV}{kT}\right)\right]\right)\right\rangle_\lambda}{\left\langle \exp\left(-\left[\frac{\mathcal{U}_{\lambda'}(\mathbf{r}^N)+pV}{kT''} - \left(\frac{\mathcal{U}_{\lambda}(\mathbf{r}^N)+pV}{kT}\right)\right]\right)\right\rangle_\lambda}$$

Terme à λ :

Pour cet état, la méthode de calcul est identique à la précédente. Le calcul débute de la manière suivante :

$$\left[\frac{Q_{NpT'}}{Q_{NpT''}}\right]_{\lambda} = \frac{\Lambda_{T''}^{3N}\int\int \exp\left(-\left[\frac{\mathcal{U}_{\lambda}(\mathbf{r}^N)+pV}{kT'}\right]\right) d\mathbf{r}^N dV}{\Lambda_{T'}^{3N}\int\int \exp\left(-\left[\frac{\mathcal{U}_{\lambda}(\mathbf{r}^N)+pV}{kT''}\right]\right) d\mathbf{r}^N dV} \tag{4.48}$$

Après simplification, nous obtenons la relation suivante.

$$\left[\frac{Q_{NpT'}}{Q_{NpT''}}\right]_\lambda = \left(\frac{T'}{T''}\right)^{\frac{3N}{2}} \frac{\left\langle \exp\left(-\left[\frac{\mathcal{U}_\lambda(\mathbf{r}^N)+pV}{kT'} - \left(\frac{\mathcal{U}_\lambda(\mathbf{r}^N)+pV}{kT}\right)\right]\right)\right\rangle_\lambda}{\left\langle \exp\left(-\left[\frac{\mathcal{U}_\lambda(\mathbf{r}^N)+pV}{kT''} - \left(\frac{\mathcal{U}_\lambda(\mathbf{r}^N)+pV}{kT}\right)\right]\right)\right\rangle_\lambda}$$

Cette équation se simplifie en :

$$\left[\frac{Q_{NpT'}}{Q_{NpT''}}\right]_\lambda = \left(\frac{T'}{T''}\right)^{\frac{3N}{2}} \frac{\left\langle \exp\left(-\left[\left(\mathcal{U}_\lambda(\mathbf{r}^N)+pV\right)\left[\frac{1}{kT'} - \frac{1}{kT}\right]\right]\right)\right\rangle_\lambda}{\left\langle \exp\left(-\left[\left(\mathcal{U}_\lambda(\mathbf{r}^N)+pV\right)\left[\frac{1}{kT''} - \frac{1}{kT}\right]\right]\right)\right\rangle_\lambda}$$

pour finalement donner l'équation (4.49). On obtient la même relation si on considère l'hamiltonien total, en remplaçant $\mathcal{U}_\lambda(\mathbf{r}^N)$ par $\mathcal{H}_\lambda(\mathbf{r}^N, \mathbf{p}^N)$ et en ôtant le terme massique.

$$\Delta S^o(\lambda) = k \ln\left[\left(\frac{m_{\lambda'}}{m_\lambda}\right)^{\frac{3}{2}}\right] \left\langle \exp\left(-\left[\frac{\mathcal{U}_{\lambda'}(\mathbf{r}^N) - \mathcal{U}_\lambda(\mathbf{r}^N)}{kT}\right]\right)\right\rangle_\lambda \tag{4.49}$$

$$+ \quad \frac{kT}{2\Delta T} \ln\left[\frac{\left\langle \exp\left(-\left[\frac{\mathcal{U}_{\lambda'}(\mathbf{r}^N)+pV}{kT'} - \left(\frac{\mathcal{U}_\lambda(\mathbf{r}^N)+pV}{kT}\right)\right]\right)\right\rangle_\lambda}{\left\langle \exp\left(-\left[\frac{\mathcal{U}_{\lambda'}(\mathbf{r}^N)+pV}{kT''} - \left(\frac{\mathcal{U}_\lambda(\mathbf{r}^N)+pV}{kT}\right)\right]\right)\right\rangle_\lambda}\right. \tag{4.50}$$

$$\left. \cdot \quad \frac{\left\langle \exp\left(-\left[\left(\mathcal{U}_\lambda(\mathbf{r}^N)+pV\right)\left[\frac{1}{kT''} - \frac{1}{kT}\right]\right]\right)\right\rangle_\lambda}{\left\langle \exp\left(-\left[\left(\mathcal{U}_{\lambda'}(\mathbf{r}^N)+pV\right)\left[\frac{1}{kT'} - \frac{1}{kT}\right]\right]\right)\right\rangle_\lambda}\right]$$

4.6 Nouvelle méthodologie

Toutes les méthodes employées jusqu'à présent, ont été démontrées en considérant que les positions et les vitesses à l'état perturbé sont identiques à celles de l'état de référence :

$$\mathbf{r}_\lambda = \mathbf{r}_{\lambda'} \tag{4.51}$$

$$\mathbf{v}_\lambda = \mathbf{v}_{\lambda'} \tag{4.52}$$

Ces considérations sont valides dans le cadre de la théorie de Zwanzig. Pour les raisons citées dans la section (4.2) nous sortons du cadre de cette théorie afin d'homogénéiser la théorie des perturbations avec la notion de paramètres de couplage introduit par Kirkwood. Pour cela, nous postulons la dépendance des couples $(\mathbf{r}, \mathbf{p})^N$ par rapport à λ. En effet, cette dépendance permet de légitimer davantage la décomposition du chemin thermodynamique en de multiples sous-états.

Notre objectif est de déterminer une méthodologie permettant de tenir compte de l'influence des paramètres perturbés sur les variables $\mathbf{r}$ et $\mathbf{p}$. En effet, l'espace des phases parcourus dans les états de référence et perturbé ne pas sont les mêmes. La figure (4.3) représentant $\sum_i \left(\mathbf{r_i(t)} - \mathbf{r'_i(t)}\right)^2$ en fonction du temps, montre une différence faible mais non négligeable dans les trajectoires.

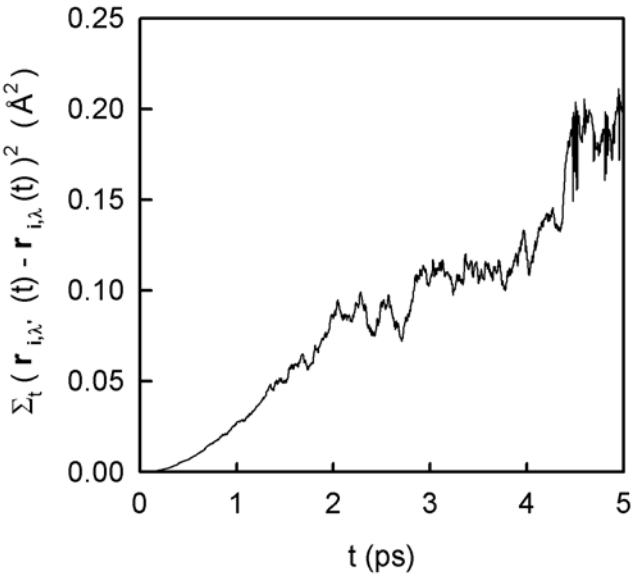

FIG. 4.3 – Trajectoire de la somme du carré de la différence des positions pour l'atome perturbé entre l'état perturbé $(\mathbf{r}^i_{\lambda'}, \mathbf{p}^i_{\lambda'})$ et l'état de référence $(\mathbf{r}^i_{\lambda}, \mathbf{p}^i_{\lambda})$.

Il n'est pas possible de considérer explicitement l'espace des phases de l'état perturbé puisque l'écriture des moyennes d'ensemble deviendrait impossible. Il faut donc déterminer le N-doublet $\left[\mathbf{r}^N, \mathbf{p}^N\right]_{\lambda'}$ à partir du N-doublet $\left[\mathbf{r}^N, \mathbf{p}^N\right]_{\lambda}$. Pour cela, nous utilisons la dynamique moléculaire qui va permettre de lier l'état de référence à l'état perturbé par la variable

temps (t). En effet, on peut écrire que :

$$\mathbf{v}(t-\frac{\Delta t}{2}) = \frac{\mathbf{r}(t)-\mathbf{r}(t-\Delta t)}{\Delta t} \quad (4.53)$$

$$\mathbf{v}(t+\frac{\Delta t}{2}) = \frac{\mathbf{r}(t-\Delta t)-\mathbf{r}(t)}{\Delta t} \quad (4.54)$$

À partir de ces équations, nous pouvons obtenir les nouvelles positions :

$$\mathbf{r}(t+\Delta t) = \mathbf{r}(t)+\mathbf{v}(t+\frac{\Delta t}{2}) \quad (4.55)$$

$$\mathbf{v}(t+\frac{\Delta t}{2}) = \mathbf{v}(t-\frac{\Delta t}{2})+\frac{\mathbf{F}(t)}{m}\Delta t \quad (4.56)$$

$$\Leftrightarrow \mathbf{r}(t+\Delta t) = \mathbf{r}(t)+\mathbf{v}(t-\frac{\Delta t}{2})\Delta t+\frac{\mathbf{F}(t)}{m}\Delta t^2 \quad (4.57)$$

On peut déterminer l'expression de $\mathbf{r}$ pour le pas $(t+n\Delta t)$ par un raisonnement par récurrence :

$$\mathbf{r}^{n\Delta t} = \mathbf{r}^0 + n\mathbf{v}^0\Delta t + \left[n\frac{\mathbf{F}^0}{m} + \sum_{i=1}^{n-1,n>1}\frac{(n-i)\mathbf{F}^{i\Delta t}}{m}\right]\Delta t^2 \quad (4.58)$$

Cette relation entre les positions à $(t^{'}=0)$ et les positions $(t^{'}=t)$ est corroborée par l'équation de Liouville :

$$f\left[\mathbf{r}^N(t),\mathbf{p}^N(t)\right] = \exp\left(iLt\right) f\left[\mathbf{r}^N(0),\mathbf{p}^N(0)\right] \quad (4.59)$$

L'objectif de ce développement est de déterminer le N-doublet $\left(\mathbf{r}^N,\mathbf{p}^N\right)_{\lambda'}$ à partir de $\left(\mathbf{r}^N,\mathbf{p}^N\right)_{\lambda}$. Pour y arriver, nous devons écrire l'équation (4.58) à l'état $\lambda^{'}$ et la comparer à l'état λ. Il est nécessaire de considérer un temps origine identique pour les deux états. Dans ces conditions, il est possible de déterminer une relation liant les coordonnées à l'état λ et à l'état $\lambda^{'}$.

$$\mathbf{r}_{\lambda'}(n\Delta t) = \mathbf{r}_{\lambda}(n\Delta t)+\Delta t^2\left(n\left[\frac{\mathbf{F}_{\lambda'}(0)}{m_{\lambda'}}-\frac{\mathbf{F}_{\lambda}(0)}{m_{\lambda}}\right]+\sum_{i=1}^{n-1,n>1}(n-i)\left[\frac{\mathbf{F}_{\lambda'}(i\Delta t)}{m_{\lambda'}}-\frac{\mathbf{F}_{\lambda}(i\Delta t)}{m_{\lambda}}\right]\right) \quad (4.60)$$

Cette relation peut s'écrire de nouveau :

$$\mathbf{r}_{\lambda'}(n\Delta t) = \mathbf{r}_{\lambda}(n\Delta t) + \Delta t^2 \left(n \left[\frac{\mathbf{F}_{\lambda'}(0)}{m_{\lambda'}} - \frac{\mathbf{F}_{\lambda}(0)}{m_{\lambda}} \right] + \sum_{i=1}^{n-1,n>1} (n-i) \left[\frac{\overset{\circlearrowleft}{\mathbf{F}}_{\lambda'}(i\Delta t)}{m_{\lambda'}} - \frac{\mathbf{F}_{\lambda}(i\Delta t)}{m_{\lambda}} \right] \right) \tag{4.61}$$

Pour calculer les forces s'exerçant sur les atomes à $\lambda^{'}$ avec précision, on doit appliquer une méthode itérative avec un critère de convergence ($\circlearrowleft$). En effet, une fois que l'on calcule $\mathbf{r}_{\lambda'}$, on évalue la force $\mathbf{F}_{\lambda'}$. Ce procédé est alors appliqué jusqu'à convergence du calcul. La figure (4.4) illustre ce procédé de récurrence.

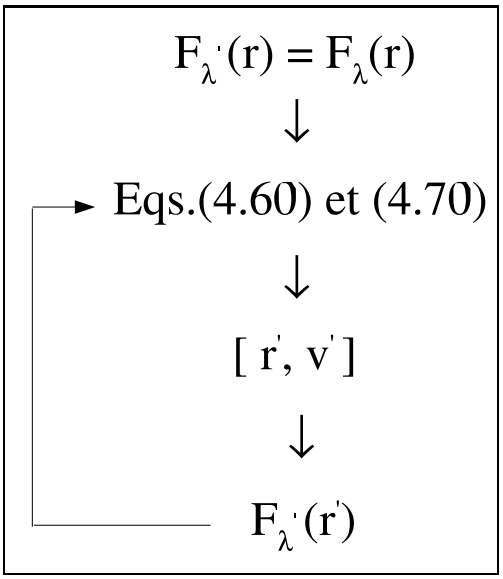

FIG. 4.4 – Schéma représentant le processus itératif pour le calcul de la force $\mathbf{F}_{\lambda'}(\mathbf{r}^N_{\lambda'})$.

Si nous considérons l'hamiltonien total dans le calcul de perturbation, il est nécessaire de pouvoir calculer les vitesses à $\lambda^{'}$ pour évaluer l'énergie cinétique. Nous utilisons alors le même raisonnement que celui qui nous a permis d'évaluer les positions. Ainsi, un développement en série de Taylor nous permet d'écrire :

$$\mathbf{v}(t + \frac{\Delta t}{2}) = \mathbf{v}(t - \frac{\Delta t}{2}) + \frac{\mathbf{F}(t)}{m} \cdot \frac{\Delta t}{2} \tag{4.62}$$

$$\mathbf{v}(t) = \frac{1}{2} \left(\mathbf{v}(t + \frac{\Delta t}{2}) + \mathbf{v}(t - \frac{\Delta t}{2}) \right) \tag{4.63}$$

L'objectif est de lier la vitesse correspondant au temps t à la vitesse correspondant au temps $t = 0$ pour comparer les états λ et $\lambda^{'}$. Cette relation peut être déterminée par un raisonnement de récurrence. Cette approche est propre à la dynamique moléculaire, puisque l'on profite du processus itératif d'intégration des équations de mouvement. Ce raisonnement n'a pas son équivalent avec les simulations de type Monte Carlo basées sur un processus markovnien. Nous présentons ici un cours extrait de la récurrence :

$$\mathbf{v}(0), \mathbf{r}(0) \leftarrow \text{valeurs initiales} \tag{4.64}$$

$$\mathbf{v}(\frac{t}{2}) \leftarrow \mathbf{v}(0), \mathbf{r}(0) \tag{4.65}$$

$$\mathbf{r}(t) \leftarrow \mathbf{r}(0), \mathbf{v}(\frac{t}{2}) \tag{4.66}$$

$$\mathbf{v}(\frac{3t}{2}) \leftarrow \mathbf{v}(\frac{t}{2}), \mathbf{r}(t) \tag{4.67}$$

$$\mathbf{r}(2t) \leftarrow \mathbf{r}(t), \mathbf{v}(\frac{3t}{2}) \tag{4.68}$$

Dans ce schéma, les vitesses à $(t - \frac{\Delta t}{2})$ sont calculées par l'équation (4.62). Il est alors facile de montrer que :

$$\mathbf{v}(n\Delta t) = \mathbf{v}(0) + \frac{\Delta t}{m}\left[\sum_{i=0}^{n-1} \mathbf{F}(i\Delta t) + \frac{\mathbf{F}(n\Delta t)}{2}\right] \tag{4.69}$$

Si on écrit cette relation pour les deux états, on obtient alors par une différence l'équation (4.70) que l'on peut écrire comme l'équation (4.71) si on exprime la récurrence ($\circlearrowleft$).

$$\mathbf{v}_{\lambda'}(n\Delta t) = \mathbf{v}_{\lambda}(n\Delta t) + \Delta t\left[\left(\sum_{i=0}^{n-1,n>1}\left(\frac{\mathbf{F}_{\lambda'}(i\Delta t)}{m_{\lambda'}} - \frac{\mathbf{F}_{\lambda}(i\Delta t)}{m_{\lambda}}\right)\right) + \frac{\mathbf{F}_{\lambda'}(n\Delta t)}{2m_{\lambda'}} - \frac{\mathbf{F}_{\lambda}(n\Delta t)}{2m_{\lambda}}\right] \tag{4.70}$$

$$\mathbf{v}_{\lambda'}(n\Delta t) = \mathbf{v}_{\lambda}(n\Delta t) + \Delta t\left[\left(\sum_{i=0}^{n-1,n>1}\left(\frac{\overset{\circlearrowleft}{\mathbf{F}}_{\lambda'}(i\Delta t)}{m_{\lambda'}} - \frac{\mathbf{F}_{\lambda}(i\Delta t)}{m_{\lambda}}\right)\right) + \frac{\overset{\circlearrowleft}{\mathbf{F}}_{\lambda'}(n\Delta t)}{2m_{\lambda'}} - \frac{\mathbf{F}_{\lambda}(n\Delta t)}{2m_{\lambda}}\right] \tag{4.71}$$

4.7 Validation

Afin de valider la méthode mise au point sur des calculs de grandeurs thermodynamiques, il est nécessaire de s'assurer que l'ensemble $(\mathbf{r}, \mathbf{p})^N$ que l'on a déterminé est correct. Pour cela, nous vérifions les équilibres thermique et mécanique. Nous vérifions, en outre, que la distribution de vitesse, est une distribution de Maxwell-Boltzmann. Enfin, nous vérifions si l'ensemble déterminé est bien une solution de l'équation (4.24).

Équilibres thermique et mécanique

L'algorithme de Berendsen que l'on utilise permet par l'application d'un facteur de mise à l'échelle $\mathcal{F}$ (équation (4.72)) de maintenir la température constante.

$$\mathcal{F} = \left[1 + \frac{\Delta t}{\tau_T}\left\{\frac{T_o}{T\left(t - \frac{\Delta t}{2}\right)} - 1\right\}\right]^{\frac{1}{2}} \quad avec\ \tau_T = \frac{1}{2\gamma} \tag{4.72}$$

γ est le coefficient de friction introduit pour dissiper l'énergie cinétique. Dans l'équation (4.72), nous pouvons considérer le deuxième terme comme un terme correctif. Comme on applique déjà ce facteur aux vitesses de l'état λ, la mise à l'échelle des vitesses à l'état $\lambda^{'}$ est systématique. Cela est possible puisque le deuxième terme de l'équation (4.72) est indépendant de la vitesse. Pour démontrer qu'il n'est pas nécessaire de remettre à l'échelle les vitesses à l'état $\lambda^{'}$, on calcule la température et la pression à l'état $\lambda^{'}$. Ces deux variables sont fonction des vitesses et des positions calculées. Ce calcul nous permettra aussi de vérifier les équilibres thermiques et mécaniques. La figure (4.5) montre la température instantanée calculée à l'état de référence et à l'état perturbé pour la perturbation de l'ion Ca^{2+} en l'ion Ba^{2+} dans l'eau. Cette figure montre que les équilibres thermique et mécanique sont assurés à l'état $\lambda^{'}$. On en déduit donc qu'il n'est pas nécessaire de mettre à l'échelle les vitesses et les positions dans l'état perturbé une fois que cette mise à l'échelle est effectuée sur les vitesses et les positions à l'état de référence.
Remarque : La même étude peut être réalisée sur le facteur de mise à l'échelle concernant les positions pour maintenir la pression constante.

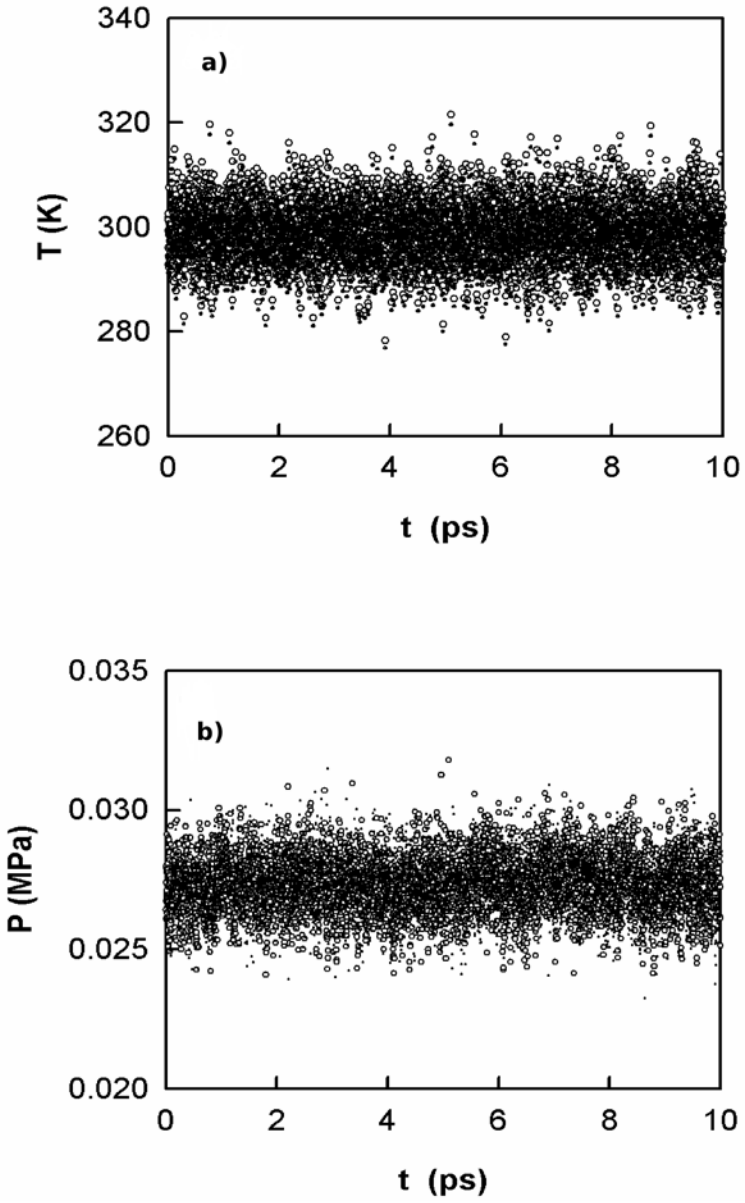

FIG. 4.5 – a) Température en fonction du temps pour les états λ' et λ. (◯) représente la température instantanée dans l'état de référence. (•) représente la température instantanée dans l'état perturbé. b) Pression en fonction du temps pour les états λ' et λ. (◯) représente la pression instantanée dans l'état de référence. (•) représente la pression instantanée dans l'état perturbé.

Distribution des vitesses

On vérifie, en outre, que la distribution des vitesses à l'état $\lambda^{'}$ est gaussienne et qu'elle peut être ajustée par une distribution de Maxwell-Boltzmann (équation (4.73)). Pour cela, on ajuste la distribution des vitesses obtenue à partir de la simulation par l'équation (4.74).

$$\rho\left(\mathbf{v}_{\lambda i}\right) = \left(\frac{m_i}{2\pi kT}\right)^{\frac{1}{2}} \exp - \left(\frac{m_{\lambda i}\mathbf{v}_{\lambda i}^2}{2kT}\right) \quad \text{avec} \quad \lambda = x, y \; or \; z \tag{4.73}$$

$$\rho\left(\mathbf{v}_{ix}\right) = a \exp\left(-\left[b\mathbf{v}_{ix}^2\right]\right) \tag{4.74}$$

La figure (4.6) montre que la distribution des vitesses à l'état $\lambda^{'}$ est gaussienne d'une part et qu'elle est superposable à la distribution de la relation (4.73). Cette étude permet ainsi de vérifier que nos calculs de vitesses à l'état perturbé n'introduisent pas d'erreur.

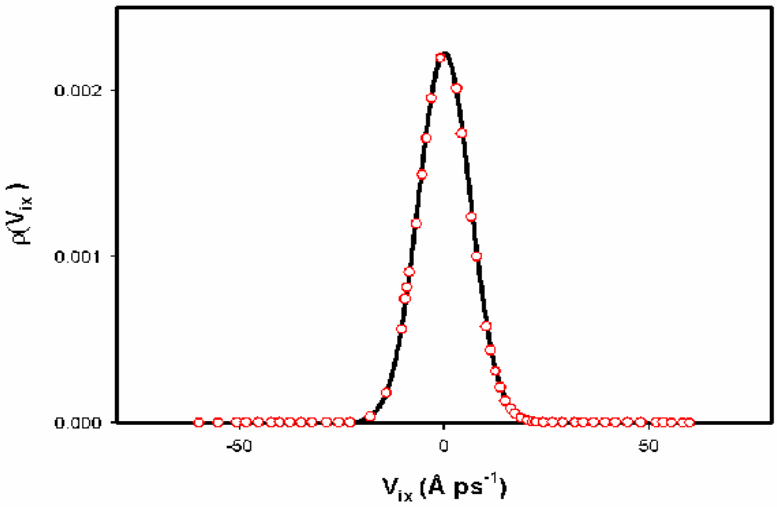

FIG. 4.6 – (—) Distribution de Maxwell-Boltzmann donnée par la relation (4.73). (◯) Distribution des vitesse v_x pour l'atome i dans l'état perturbé, ajustée par l'équation (4.74).

Vérification de $\mathbf{r}_\lambda^N \mathbf{p}_\lambda^N$

Nous devons vérifier la relation (4.26) qui permet en partie de justifier l'existence de ce développement. Pour cela, il suffit de calculer les distributions du produit $\mathbf{r}_\lambda^N \mathbf{p}_\lambda^N$ pour l'état λ et l'état λ'. La figure (4.7) montre que les distributions sont identiques. Par conséquent, la solution envisagée est bien une solution de l'équation (4.24).

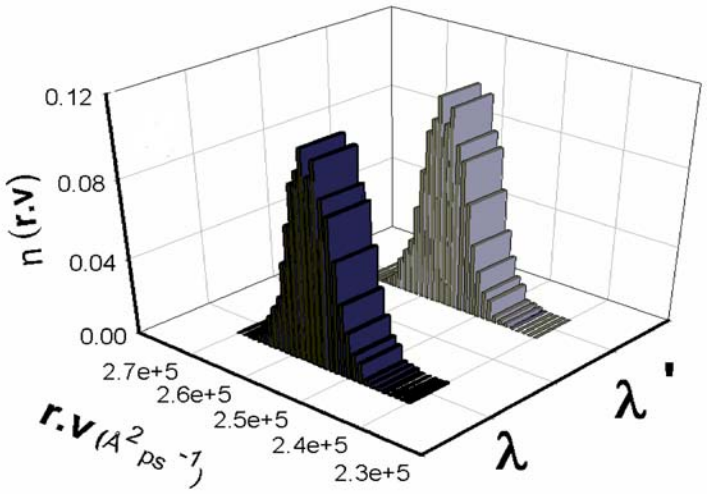

FIG. 4.7 – Distribution du produit $\mathbf{r}_\lambda^N \mathbf{p}_\lambda^N$ pour les états λ et λ'.

Finalement, voici le schéma à suivre pour mener de manière satisfaisante une perturbation :

- Détermination de $\mathbf{F}_{\lambda'}$,
- calcul de $\mathbf{r}_{\lambda'}$ et $\mathbf{v}_{\lambda'}$,
- détermination de $\mathcal{H}_{\lambda'}$,
- application des formules opérationnelles.

Le tableau (4.1) reprend toutes les approximations qui ont été faites, ainsi que la nomenclature utilisée dans la suite du document.

FEP	TI
$\mathcal{M}_I(\mathcal{U}, \mathbf{r}_\lambda, \mathbf{v}_\lambda, \mathbf{r}_{\lambda'} = \mathbf{r}_\lambda, \mathbf{v}_{\lambda'} = \mathbf{v}_\lambda, m_{\lambda'} = m_\lambda)$	
$\Delta G^{[4.32]}$,$\Delta S^{[4.33]}$	$\Delta G^{[4.40]}$,$\Delta S^{[4.42]}$
$\mathcal{M}_{II}(\mathcal{H}, \mathbf{r}_\lambda, \mathbf{v}_\lambda, \mathbf{r}_{\lambda'} = \mathbf{r}_\lambda, \mathbf{v}_{\lambda'} = \mathbf{v}_\lambda, m_{\lambda'} \neq m_\lambda)$	
$\Delta G^{[4.28,4.29]}$,$\Delta S^{[4.34]}$	$\Delta G^{[4.39]}$,$\Delta S^{[4.42]}$
$\mathcal{M}_{III}(\mathcal{U}, \mathbf{r}_\lambda, \mathbf{v}_\lambda, \mathbf{r}_{\lambda'} \neq \mathbf{r}_\lambda, \mathbf{v}_{\lambda'} \neq \mathbf{v}_\lambda, m_{\lambda'} \neq m_\lambda)$	
$\Delta G^{[4.31]}$,$\Delta S^{[4.33]}$	$\Delta G^{[4.37]}$,$\Delta S^{[4.42]}$
$\mathcal{M}_{IV}(\mathcal{H}, \mathbf{r}_\lambda, \mathbf{v}_\lambda, \mathbf{r}_{\lambda'} \neq \mathbf{r}_\lambda, \mathbf{v}_{\lambda'} \neq \mathbf{v}_\lambda, m_{\lambda'} \neq m_\lambda)$	
$\Delta G^{[4.28,4.29]}$,$\Delta S^{[4.34]}$	$\Delta G^{[4.39]}$,$\Delta S^{[4.42]}$

TAB. 4.1 – Résumé des différentes approximations effectuées pour les méthodes FEP et TI. Les formules opérationnelles de ΔG et ΔS sont référencées pour chaque méthodologie.

La méthode développée ici utilise l'intégration des équations de mouvements qui est propre à la dynamique moléculaire. Pour les méthodes Monte Carlo, l'échantillonnage de l'espace des phases est réalisé par un processus stochastique et markovien. L'absence de lien entre les positions de l'état t et celles de l'état initial (absence de mémoire au delà de $t-1$) est un frein essentiel à la détermination de l'espace des phases de l'état perturbé à partir de l'espace des phases de l'état de référence, contrairement à la dynamique moléculaire dont le fonctionnement est basé sur un processus itératif permettant de faire le lien avec l'état initial.

4.8 Calculs de grandeurs d'hydratation et d'association

Notre objectif est de valider la méthodologie développée précédemment à travers l'évaluation de propriétés d'hydratation et de complexation. Pour illustrer nos considérations théoriques nous allons calculer dans un premier temps les propriétés d'hyratation ($\Delta\Delta_h G, \Delta\Delta_h S, \Delta\Delta_h H$) de certains ions inorganiques. Nous appliquerons la méthodologie développée pour le calcul des propriétés d'association ($\Delta\Delta_c G, \Delta\Delta_c S, \Delta\Delta_c H$) à des complexes à sphère externe.

4.8.1 Mise en œuvre

Pour le calcul des propriétés d'hydratation, nous avons perturbé l'ion Ca^{2+} en Ba^{2+}, l'ion Ca^{2+} en Mg^{2+} et l'ion Ca^{2+} en La^{3+}. Chacune des perturbations a été réalisée en décomposant le chemin thermodynamique en 180 sous états. Les caractéristiques des simulations sont données dans la section (2.6). Le temps d'équilibration pour chacune des simulations est de 100 ps alors que le temps d'acquisition est de 100 ps. Le calcul nécessite un temps de simulation total de 36 ns. Les calculs sont réalisés sur 5000 configurations générées durant la phase d'acquisition. Les paramètres des cations (σ, ϵ) ont été déterminés par Aqvist [45]. Les paramètres perturbés sont la masse (m), la charge (q) et les paramètres de Lennard-Jones (σ, ϵ). Pour les atomes i et j, les paramètres sont perturbés de la manière suivante :

$$\begin{aligned}
\epsilon_{ij}(\lambda) &= \lambda\left(\sqrt{(\epsilon_{ii}(1)\epsilon_{jj}}\right) + (1-\lambda)\left(\sqrt{(\epsilon_{ii}(0)\epsilon_{jj})}\right) \\
\sigma_{ij}(\lambda) &= \lambda\left(\frac{\sigma_{ii}(1)+\sigma_{jj}}{2}\right) + (1-\lambda)\left(\frac{\sigma_{ii}(0)+\sigma_{jj}}{2}\right) \\
q_i(\lambda) &= \lambda q_i\,(1) + (1-\lambda) q_i\,(0) \\
m_i(\lambda) &= \lambda m_i\,(1) + (1-\lambda) m_i\,(0)
\end{aligned}$$

λ représente le paramètre de couplage de Kirkwood tel que l'état $\lambda = 0$ correspond à l'état initial et l'état $\lambda = 1$ correspond à l'état final. Enfin, chaque perturbation est réalisée dans les deux sens, direct $(-\Delta\lambda)$ et inverse $(+\Delta\lambda)$.

4.8.2 Grandeurs d'hydratation

Le calcul des propriétés d'hydratation entre deux ions A et B est réalisé par l'intermédiaire du cycle thermodynamique représenté sur la figure (4.8)

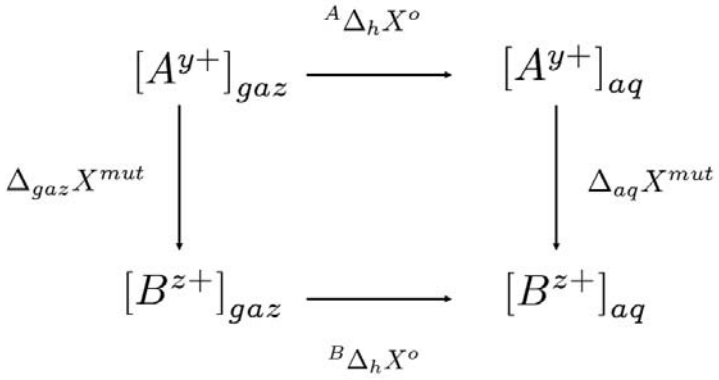

FIG. 4.8 – Cycle thermodynamique permettant d'évaluer la différence de variation d'enthalpie libre, d'enthalpie et d'entropie d'hydratation entre deux ions A^{y+} et B^{z+}. ΔX^{mut} indique une variation de la propriété X pour la perturbation de l'espèce A^{y+} en l'espèce B^{z+}. Cette grandeur est calculée en phase gazeuse et en phase aqueuse.

La relation permettant de lier les grandeurs expérimentales (ΔX^o) et simulées (ΔX^{mut}) est donnée par la relation (4.75)

$$\Delta\Delta_h X^o = {}^{B}\,\Delta_h X^o - {}^{A}\,\Delta_h X^o = \Delta_{aq}X^{mut} - \Delta_{gaz}X^{mut} \tag{4.75}$$

Les relations permettant de calculer la variation d'enthalpie libre par la méthode FEP en prenant en compte les variations de masse sont données par les équations (4.28) , (4.29) et (4.31). Dans l'équation (4.31) le terme «massique» qui apparaît comme une constante, peut être sorti de la moyenne de la façon suivante :

$$\Delta G_\lambda = -\frac{3kT}{2}\ln\left[\frac{m_{\lambda'}}{m_\lambda}\right] - kT\ln\left\langle \exp\left(-\left[\frac{\mathcal{U}_{\lambda'}(\mathbf{r}^N)-\mathcal{U}_\lambda(\mathbf{r}^N)}{kT}\right]\right)\right\rangle_\lambda \tag{4.76}$$

En phase aqueuse, la variation d'enthalpie libre entre deux états peut être calculée par cette relation. En phase gazeuse, seules les parties cinétique et intramoléculaire sont prises en compte. Comme la partie intramoléculaire est nulle, la variation d'enthalpie libre dans cette phase vaut :

$$\Delta G_\lambda = -\frac{3kT}{2}\ln\left[\frac{m_{\lambda'}}{m_\lambda}\right] \tag{4.77}$$

La différence d'enthalpie libre entre les deux phases est égale à :

$$\begin{aligned}\Delta\Delta_h G^o &= {}^B\Delta_h G^o - {}^A\Delta_h G^o = \Delta_{aq}G^{mut} - \Delta_{gaz}G^{mut} \\ &= -\sum_\lambda kT\ln\left\langle\exp\left(-\left[\frac{\mathcal{U}_{\lambda'}(\mathbf{r}^N)-\mathcal{U}_\lambda(\mathbf{r}^N)}{kT}\right]\right)\right\rangle_\lambda\end{aligned} \tag{4.78}$$

Ici l'état A correspond au cas de l'ion A^{y+} et B à celui de l'ion B^{z+}, ce qui revient à considérer l'équation (4.28). Si l'on considère la longueur d'onde de De Broglie, le terme «massique» n'intervient pas. Par contre, si on considère l'hamiltonien total dans la fonction de partition, on ne peut plus négliger ce terme. En effet, le terme cinétique est calculé dans la moyenne. Comme la fonction de partition de l'ion à l'état gazeux est proportionnelle à la longueur d'onde de De Broglie, nous obtenons :

$$\Delta\Delta_h G^o = \frac{3kT}{2}\ln\left[\frac{m_B}{m_A}\right] - kT\sum_\lambda\ln\left\langle\exp\left(-\left[\frac{\mathcal{H}_{\lambda'}(\mathbf{r}^N,\mathbf{p}^N)-\mathcal{H}_\lambda(\mathbf{r}^N,\mathbf{p}^N)}{kT}\right]\right)\right\rangle_\lambda \tag{4.79}$$

$$\Delta\Delta_h G^o = \frac{3kT}{2}\ln\left[\frac{m_B}{m_A}\right] - kT\sum_\lambda\ln\left\langle\exp\left(-\left[\frac{\mathcal{U}_{\lambda'}(\mathbf{r}^N)-\mathcal{U}_\lambda(\mathbf{r}^N)}{kT}\right] - \frac{v_\lambda^2}{2kT}\left[m_{\lambda'}-m_\lambda\right]\right)\right\rangle_\lambda \tag{4.80}$$

Cette relation ne va pas à l'encontre du principe de l'équipartition de l'énergie. En effet, ce dernier lie la moyenne de l'énergie cinétique à la température par la relation $\langle\mathcal{K}\rangle = \frac{3NkT}{2}$. Ainsi, à nombre de particules identiques et pour une même température, on a $\langle\mathcal{K}\rangle_\lambda = \langle\mathcal{K}\rangle_{\lambda'}$. Or, la relation précédente ne tient pas compte de la moyenne de l'énergie cinétique mais de la valeur instantanée. Il est tout à fait possible d'avoir des vitesses instantanées différentes pour les deux états et une moyenne de l'énergie cinétique identique. Si l'on considère la théorie de Zwanzig où l'état perturbé n'est qu'une perturbation de l'état de référence alors nous devons considérer que les vitesses pour les deux états sont identiques. Dans ces conditions, si l'on tient compte de la variation de masse alors on a $\langle\mathcal{K}\rangle_\lambda \neq \langle\mathcal{K}\rangle_{\lambda'}$. Par conséquent, l'équipartition de l'énergie pour l'état perturbé n'est pas respectée. Néanmoins, étant donnée la faible amplitude de la perturbation, on peut faire l'approximation que ce principe est vérifié pour

l'état perturbé. Ces observations sont à reporter dans le calcul de la variation d'entropie. Par ailleurs, nous pouvons faire les mêmes simplifications concernant la méthode d'intégration thermodynamique.

$$\Delta\Delta_h G^o = \int_0^1 \left\langle \frac{\partial \mathcal{U}_\lambda(\mathbf{r}^N)}{\partial \lambda} \right\rangle_\lambda d\lambda \tag{4.81}$$

$$\begin{aligned} \Delta\Delta_h G^o &= \int_0^1 \left[\frac{3NkT}{2} \frac{1}{m_\lambda} \cdot \frac{\partial m_\lambda}{\partial \lambda} + \left\langle \left[\frac{1}{2} \frac{\partial \mathcal{U}_\lambda(\mathbf{r}^N)}{\partial \lambda} + \frac{\partial m_\lambda}{\partial \lambda} \sum \mathbf{v_i^2} \right] \right\rangle_\lambda \right] d\lambda & (4.82) \\ &= \frac{3NkT}{2} \ln \left[\frac{m_A}{m_B} \right] + \int_0^1 \left[\left\langle \left[\frac{1}{2} \frac{\partial \mathcal{U}_\lambda(\mathbf{r}^N)}{\partial \lambda} + \frac{\partial m_\lambda}{\partial \lambda} \sum \mathbf{v_i^2} \right] \right\rangle_\lambda \right] d\lambda & (4.83) \end{aligned}$$

Afin d'étudier les différents cas, nous avons testé ces méthodes sur la perturbation de systèmes simples telle que la perturbation de l'ion Ca^{2+} en l'ion Ba^{2+} et la perturbation de l'ion Ca^{2+} en l'ion Mg^{2+}. Ces exemples ont été choisis pour mesurer l'importance de la prise en compte des masses dans les calcul. Les résultats obtenus sont regroupés dans le tableau (4.2). Ce tableau montre clairement que la seule considération de l'énergie potentielle ($\mathcal{M}_I$) ne permet pas d'évaluer la variation d'entropie et d'enthalpie avec précision. Lorsqu'on considère l'hamiltonien total ($\mathcal{M}_{II}$), on observe de brusques variations par rapport à la méthode $\mathcal{M}_I$. On impute cette différence au fait que les vitesses et les positions à l'état perturbé sont les mêmes qu'à l'état de référence. En effet, lorsqu'on calcule les positions ($\mathcal{M}_{III}$) et les vitesses ($\mathcal{M}_{IV}$) les résultats sont bien plus cohérents avec l'expérience. La méthode $\mathcal{M}_{IV}$ apparaît comme étant une méthodologie sûre pour déterminer la variation d'entropie.

	$\Delta\Delta_h H^o$ (kJ mol^{-1})	$T\Delta\Delta_h S^o$ (kJ mol^{-1})	$\Delta\Delta_h G^o$ (kJ mol^{-1})	$\Delta\Delta_h H^o$ (kJ mol^{-1})	$T\Delta\Delta_h S^o$ (kJ mol^{-1})	$\Delta\Delta_h G^o$ (kJ mol^{-1})
	$Ca^{2+} \rightarrow Ba^{2+}$			**$Ca^{2+} \rightarrow Mg^{2+}$**		
FEP [45]			275			315
Exp. [55]	271	14	257	347	24	323
FEP^I	275_4	2.0_4	273_4	313_4	-5.0_4	318_4
TI^I	275_4	2.0_4	273_4	312_4	-10.0_4	322_5
TP^I	282_4	8.0_4	274_4	349_4	31.0_1	318_4
FEP^{II}	251_4	-15.3_4	266.3_{37}	316_5	-8.0_4	324_5
TI^{II}	255_4	-10.3_4	265.3_4	314_5	-10.0_4	324_5
TP^{II}	257_4	-9.3_4	266.3_4	365_5	41.0_8	324_1
FEP^{III}	269_3	5.7_4	263.3_4	343_4	12.0_4	331_4
TI^{III}	269_4	5.7_4	263.3_4	341_4	11.0_4	330_4
TP^{III}	270.9_3	7.6_4	263.3_4	344_4	23.0_4	321_4
FEP^{IV}	266_3	6.7_4	259.3_4	342_4	11.0_4	331_4
TI^{IV}	266_4	6.7_4	259.3_4	345_4	15.0_4	330_4
TP^{IV}	268_4	8.7_4	259.3_4	344_4	23.0_4	321_4

TAB. 4.2 – Différence de variation d'enthalpie libre, d'enthalpie et d'entropie d'hydratation entre les ions Ca^{2+} et Ba^{2+} et entre les ions Ca^{2+} et Mg^{2+}. Ces grandeurs sont évaluées à partir des différentes méthodologies pour les méthodes FEP, TI et TP. (Exp.) représente les valeurs expérimentales. Nous donnons aussi les valeurs calculées par Aqvist [45] par la méthode FEP. L'ensemble des méthodes est répertorié dans le tableau (4.1). Le nombre 275_4 indique 275 ± 4.

4.8.3 Grandeurs d'association

Le calcul de propriétés d'association se fait par l'utilisation de la méthode $\mathcal{M}_{IV}$. Nous évaluons les différences de variations d'entropie et d'enthalpie libre de complexes à sphère externe ($[\mathrm{Cal}^{4-} \cdots \mathrm{Ca}^{2+}] \rightarrow [\mathrm{Cal}^{4-} \cdots \mathrm{Mg}^{2+}]$ et $[\mathrm{Cal}^{4-} \cdots \mathrm{Ca}^{2+}] \rightarrow [\mathrm{Cal}^{4-} \cdots \mathrm{La}^{3+}]$). Tout comme l'évaluation des propriétés d'hydratation, ce calcul nécessite l'utilisation du cycle thermodynamique représenté sur la figure (4.9). Les résultats concernant ce calcul sont rassemblés dans le tableau (4.3).

Pour la perturbation de l'ion Ca^{2+} en Mg^{2+}, nous avons perturbé les paramètres de Lennard-Jones et la masse de l'ion considéré. Pour la perturbation de l'ion Ca^{2+} en La^{3+}, nous avons perturbé la charge en plus des paramètres précédents. Le protocole de simulation est le même que celui exposé pour le calcul des grandeurs d'hydratation. Ces résultats montrent l'éfficacité de la méthode employée, puisqu'on a seulement 18 % d'erreur sur le calcul de la variation d'entropie. Ces résultats mettent en évidence la nécessité de considérer l'hamiltonien total dans le calcul et d'évaluer le N-doublet $(\mathbf{r}', \mathbf{p}')^N$. La prise en compte de ces grandeurs améliore sensiblement la convergence des résultats comme le montre la figure (4.10).

Nous avons représenté dans le tableau (4.3) les contributions de ΔG, ΔH et ΔS relatives à chaque paramètre perturbé pour l'ion seul en solution et en présence du complexe. La contribution liée aux interactions électrostatiques est de 80%. Ce résultat est en accord avec l'hypothèse selon laquelle les interactions électrostatiques sont prédominantes dans les complexes de sphère externe. Nous observons que la perturbation de l'ion Ca^{2+} en La^{3+} est enthalpiquement favorable ($\Delta\Delta_h H^o < 0$) et entropiquement défavorable ($\Delta\Delta_h S^o < 0$). Ce comportement est dû à l'augmentation du nombre de molécules d'eau dans la première (+2) et la seconde couche (+6) d'hydratation. Nous observons un comportement différent pour la perturbation de l'ion Ca^{2+} en Mg^{2+} puisque nous obtenons une perte de deux molécules d'eau pour la première couche de solvatation et une perte de six molécules d'eau pour la seconde couche. Ce gain de degré de liberté se traduit par une enthalpie défavorable ($\Delta\Delta_h H^o > 0$) et une entropie favorable ($\Delta\Delta_h S^o > 0$). Les mêmes comportements sont observés dans le cas des ions en présence du macrocycle.

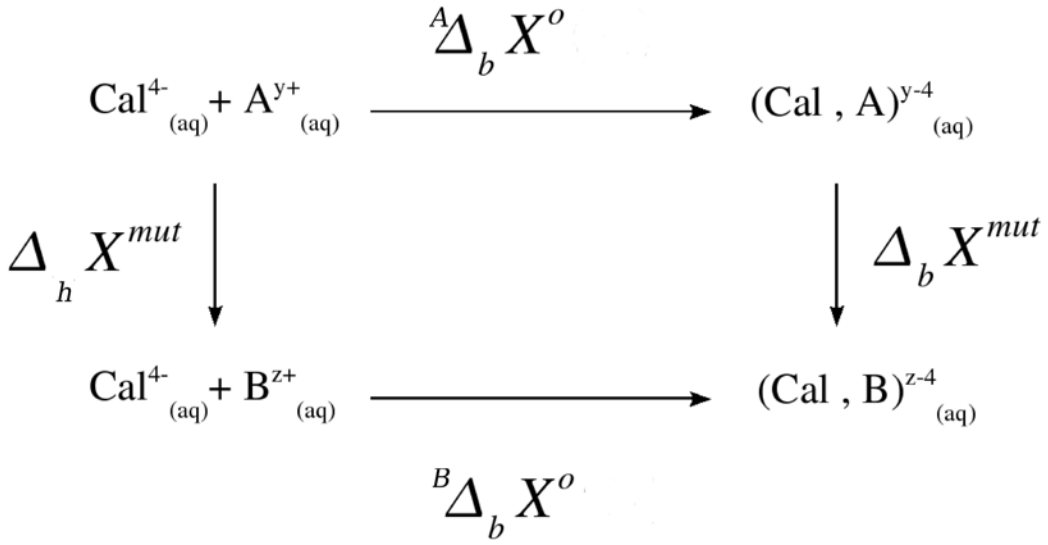

FIG. 4.9 – Cycle thermodynamique permettant d'évaluer la différence de variation d'enthalpie libre, d'enthalpie et d'entropie de complexation des ions A^{y+} et B^{z+} par le calixarène. ΔX^{mut} indique la variation de la propriété X pour la perturbation de l'espèce A^{y+} en l'espèce B^{z+}. On a alors : $\Delta\Delta_b X^o = {}^{B}\Delta_b X^o - {}^{A}\Delta_b X^o = \Delta_b X^{mut} - \Delta_h X^{mut}$.

	$\mathbf{Ca^{2+} \rightarrow La^{3+}}$			$\mathbf{Ca^{2+} \rightarrow Mg^{2+}}$		
	$\Delta\Delta_h H$ (kJ mol^{-1})	$T\Delta\Delta_h S$ (kJ mol^{-1})	$\Delta\Delta_h G$ (kJ mol^{-1})	$\Delta\Delta_h H$ (kJ mol^{-1})	$T\Delta\Delta_h S$ (kJ mol^{-1})	$\Delta\Delta_h G$ (kJ mol^{-1})
LJ.	274_5	8.0_4	266_5	344_4	23.0_4	321_4
Elec.	-1977_6	-71.0_{30}	-1906_6	–	–	–
Tot.	-1703_6	-63.0_{20}	-1640_6	344_4	23.0_4	321_4
Exp.[55]	−1701	-63	-1640	347	24.0	323
	$\Delta_b H^{mut}$ (kJ mol^{-1})	$T\Delta_b S^{mut}$ (kJ mol^{-1})	$\Delta_b G^{mut}$ (kJ mol^{-1})	$\Delta_b H^{mut}$ (kJ mol^{-1})	$T\Delta_b S^{mut}$ (kJ mol^{-1})	$\Delta_b G^{mut}$ (kJ mol^{-1})
LJ.	293_4	3.0_3	290_4	346_5	24.0_3	322_4
Elec.	-1988_4	-54.0_3	-1935_4	–	–	–
Tot.	-1695_4	-51.0_3	-1645_4	346_{50}	24.0_3	322_4
	$\Delta\Delta_b H$ (kJ mol^{-1})	$T\Delta\Delta_b S$ (kJ mol^{-1})	$\Delta\Delta_b G$ (kJ mol^{-1})	$\Delta\Delta_b H$ (kJ mol^{-1})	$T\Delta\Delta_b S$ (kJ mol^{-1})	$\Delta\Delta_b G$ (kJ mol^{-1})
Calc.	-8_4	-13.0_3	5_3	-2_3	-1.0_2	-1_3
Exp [1]	−6	−11	5	−1.7	−1.5	−0.2

TAB. 4.3 – Différence de variation d'enthalpie libre, d'enthalpie et d'entropie de complexation par le calixarène entre les ions Ca^{2+} et Mg^{2+} et entre les ions Ca^{2+} et La^{3+}. Ces grandeurs sont déterminées à partir de la méthodologie $\mathcal{M}_{IV}$ (tableau 4.1) et du formalisme TP. Les valeurs calculées sont comparées aux valeurs expérimentales (Exp.). Nous avons aussi reporté les grandeurs d'hydratation. En outre, nous avons reporté les valeurs relatives aux différentes étapes de la perturbation : une partie liée à la perturbation des charges (Elec.) et une autre liée à la perturbation des paramètres Lennard-Jones (LJ.). Enfin, nous avons représenté la contribution totale (Tot = Elec + LJ). Le nombre 54.0_3 indique 54.0 ± 0.3.

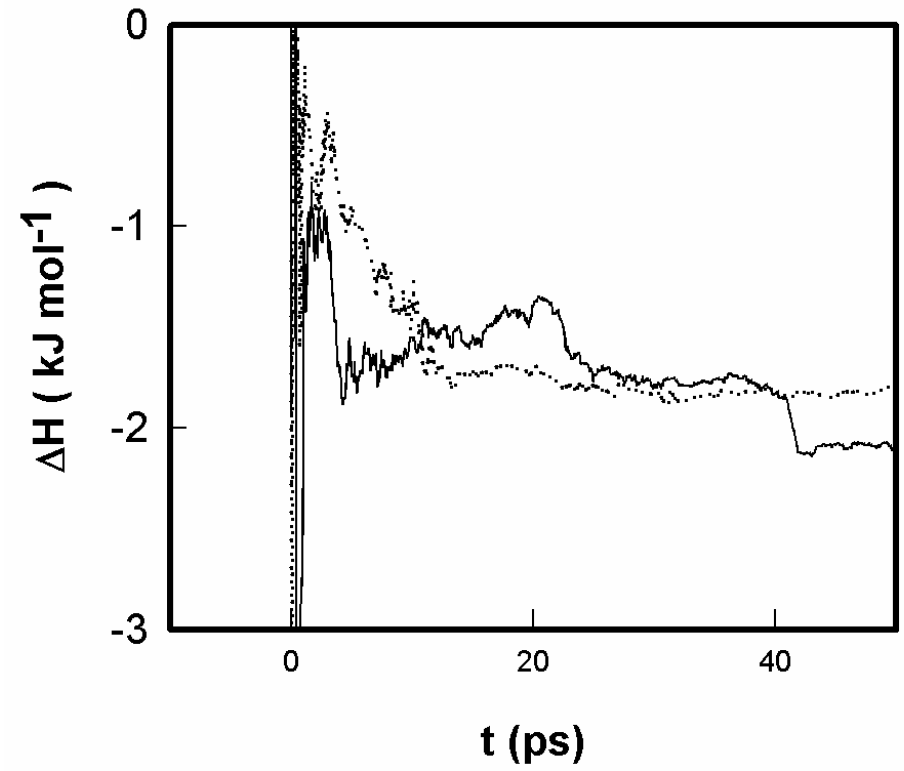

FIG. 4.10 – Variation d'enthalpie instantanée calculée pour la perturbation de Ca^{2+} en Mg^{2+} seul en solution (λ=0.3). (—) est obtenue par les méthodes $\mathcal{M}_I$ et FEP. ($\cdot\ \cdot\ \cdot$) est obtenue par les méthodes $\mathcal{M}_{IV}$ et FEP.

4.9 Résumé

Les méthodes de perturbation et d'intégration ont été très largement utilisées ces dernières années pour le calcul de $\Delta\Delta_r G^o$ contrairement aux calculs de $\Delta\Delta_r S^o$, qui nécessitent une évaluation plus précise de l'hamiltonien total et un échantillonnage conséquent de l'espace des phases.

Notre objectif était d'établir une méthode permettant de calculer la variation d'entropie avec plus de précision. Pour cela, nous avons évalué l'hamiltonien de l'état perturbé en déterminant le N-uplet $(\mathbf{r}, \mathbf{p})_{\lambda'}$: cela a été possible en confrontant les équations du mouvement de chaque atome pour les deux états (référence et perturbé). Il s'agit donc d'une méthode propre à la méthode de dynamique moléculaire. De telles considérations nous ont permis de discuter la théorie de la perturbation de Zwanzig et celle de Kirkwood concernant le paramètre de couplage. En effet, nous avons montré que l'application de la théorie de Zwanzig à chacun des états issus du découplage impliquait des fonctions de partition différentes pour le même état. La méthodologie mise au point n'a pas l'ambition de résoudre ce problème mais plutôt de l'atténuer de telle manière qu'on peut évaluer l'entropie de manière plus précise.

Les résultats obtenus montrent clairement que la méthodologie employée permet d'obtenir des résultats satisfaisants, en accord avec l'expérience sur les grandeurs d'association et d'hydratation. Nous avons été capables, de reproduire des variations d'enthalpie sur des complexes de sphère externe avec des ions inorganiques. En outre, nous avons montré que l'évaluation précise de l'hamiltonien à l'état perturbé permet une amélioration de la convergence du calcul de variation d'entropie. Le point faible de cette méthode est probablement le temps de calcul puisque pour chaque configuration, il faut calculer les forces agissant sur chaque atome du système à l'état perturbé et à l'état de référence. Néanmoins, en parallélisant les codes [65], nous avons diminué le temps de calcul par un facteur 16.

Chapitre 5

Potentiel de force moyenne et calcul de $\Delta_r G^o$, $\Delta_r H^o$ et $\Delta_r S^o$: application à des complexes d'insertion et de sphère externe

5.1 Introduction et objectifs

Ce chapitre a pour objet de comparer les différentes méthodes permettant de calculer le potentiel de force moyenne et de mettre en application ces méthodes pour le calcul de grandeurs thermodynamiques d'association : $\Delta_r S^o$, $\Delta_r H^o$ et $\Delta_r G^o$. Nous étudierons les méthodes de perturbation (FEP) et de force contrainte[79].

Dans un premier temps nous développerons les méthodes permettant de calculer le potentiel de force moyenne. Ensuite, nous comparerons ces méthodes à travers l'étude de la paire d'ions $[Ca^{2+} \cdots SO_4^{2-}]$. Cette comparaison nous permettra de déterminer la méthode la plus efficace pour évaluer les grandeurs thermodynamiques d'association. Dans une troisième partie, nous déterminerons $\Delta_r S^o$, $\Delta_r H^o$ et $\Delta_r G^o$ dans le cas des complexes d'insertion et de sphère externe. Nous comparerons les valeurs calculées aux valeurs expérimentales obtenues au laboratoire par microcalorimétrie. En outre, une étude structurale et énergétique permettra de comprendre la nature des puits de potentiels obtenus et les interactions qui prédominent dans ce type d'association.

5.2 Théorie

Le potentiel de force moyenne entre deux particules i et j séparées par une distance r_{ij} est le potentiel lié à la force moyenne exercée par j sur i. En général, ce potentiel est représenté

par $\mathcal{W}^n$. C'est Kirkwood en 1930 [74] qui, à partir de l'expression de la distribution radiale, a déterminé l'expression du potentiel de force moyenne, selon les équations (5.1), (5.2) et (5.3).

$$\rho^{(n)}\left(\mathbf{r}^1,...,\mathbf{r}^n\right) = \frac{N!}{(N-n)!}\frac{\int ... \int \exp\left(-\beta \mathcal{U}^N\right) d\mathbf{r}^{n+1},...,d\mathbf{r}^N}{Z_N} \quad (5.1)$$

$$g^{(n)}\left(\mathbf{r}^1,...,\mathbf{r}^n\right) = \frac{V^n \rho^{(n)}}{N^n} \quad (5.2)$$

$$\mathcal{W}^{(n)}\left(\mathbf{r}^1,...,\mathbf{r}^n\right) = -kT \ln g^{(n)} \quad (5.3)$$

où N est le nombre de particules total dans le volume V. $\rho^{(n)}$ est la probabilité que la particule n soit à r^n dans le volume $\mathrm{d}r^n$, indépendamment du reste de la configuration $(N-n)$. Z_N est l'intégrale de configuration. Le passage de l'équation (5.2) à l'équation (5.3) se fait en dérivant cette première par rapport à ξ qui représente un paramètre de couplage. Ainsi, grâce à ce potentiel de force moyenne, on est capable d'obtenir un profil d'énergie d'interaction entre deux particules en fonction de la distance les séparant. Cette méthode permet ainsi de déterminer le puits de potentiel pour une interaction de paire. En outre, elle est très utile dans l'élaboration de potentiels analytiques [76].

Ces équations ne sont utilisables que si l'on peut échantillonner un ensemble suffisament grand de distances r_{ij}. Ainsi, lorsque l'on se trouve dans le cas simple de la détermination du potentiel de force moyenne entre deux particules dans un milieu constitué de $N-2$ particules identiques, cela ne pose aucun problème. Par contre, lorsque deux entités sont plongées dans un solvant de nature différente, seule la distance d'équilibre sera échantillonnée. Dans ces conditions, le potentiel de force moyenne ne pas peut être calculé.

Il faut donc trouver une méthode permettant d'échantillonner l'ensemble des distances r_{ij}. On peut alors découpler le chemin thermodynamique en plusieurs micro-états k, où l'on fait correspondre à chacun des états une distance fixe r_{ij}^k. Dans ces conditions, il est possible d'échantillonner l'ensemble r_{ij}. Les méthodes basées sur ce concept sont les méthodes de perturbation, d'intégration thermodynamique et la méthode dite de «force contrainte». Il existe une autre méthode basée sur cette idée (*Umbrella Sampling*, développée par Patey

et Valleau [77]), dans laquelle la distance est fixée non pas par une contrainte mais par un potentiel harmonique qui est ajouté à l'énergie configurationnelle. Cette méthode utilise alors l'hamiltonien suivant :

$$\begin{aligned} \mathcal{V}'(\mathbf{r}^N) &= \mathcal{V}(\mathbf{r}^N) + \mathcal{W}(\mathbf{r}^N) \\ \mathcal{W}(\mathbf{r}^N) &= k_{\mathcal{W}}(\mathbf{r}^N - \mathbf{r}_o^N)^2 \end{aligned} \tag{5.4}$$

Ici $\mathcal{W}(\mathbf{r}^N)$ est une fonction de biais. Ce potentiel harmonique est un potentiel biaisé qui implique une distribution de type non-Boltzmann. Ce potentiel biaisé sert à confiner les variations des coordonnées des atomes contraints dans un petit intervalle autour d'une valeur d'équilibre. Les moyennes d'ensemble (équation (5.5)) peuvent être déterminées à partir de la distribution de type non-Boltzmann introduite par Torrie et Valleau [78]. Dans l'équation (5.5), A représente l'énergie de Helmholtz et l'indice $_{\mathcal{W}}$ indique que les moyennes sont réalisées avec la probabilité liée au potentiel biaisé.

$$\langle A \rangle = \frac{\left\langle A(\mathbf{r}^N) \exp\left[\frac{\mathcal{W}(\mathbf{r}^N}{k_B T}\right] \right\rangle_{\mathcal{W}}}{\left\langle \exp\left[\frac{\mathcal{W}(\mathbf{r}^N}{k_B T}\right] \right\rangle_{\mathcal{W}}} \tag{5.5}$$

5.2.1 Perturbation et intégration thermodynamique

10.1.1 Théorie

Il est possible de calculer le potentiel de force moyenne en utilisant les méthodes de perturbation. Dans ce cas, le paramètre perturbé est la distance intermoléculaire séparant les deux particules. Cette distance est maintenue fixe par l'application de l'algorithme SHAKE. L'utilisation de cette méthode suppose que la variation d'enthalpie libre de perturbation est comparable au potentiel de force moyenne. On a alors $\mathcal{W}(\mathbf{r}^N) = \Delta G(\mathbf{r}^N)$. Pour deux particules 1 et 2, on peut écrire $\mathcal{W}(\mathbf{r}^N)$ sous forme d'un rapport de fonctions de partition, dans l'ensemble NpT à partir des équations (5.1) et (5.3) :

$$\mathcal{W}(\mathbf{r}_1, \mathbf{r}_2) = -kT \ln \left[\frac{\int \exp\left(-\frac{pV}{kT}\right) dV \int\int \exp\left(\frac{-\mathcal{H}(\mathbf{r}^N, \mathbf{p}^N)}{kT}\right) d\mathbf{r}^{N-2} d\mathbf{p}^{N-2}}{\int \exp\left(-\frac{pV}{kT}\right) dV \int\int \exp\left(\frac{-\mathcal{H}(\mathbf{r}^N, \mathbf{p}^N)}{kT}\right) d\mathbf{r}^N d\mathbf{p}^N} \right] + Cste \tag{5.6}$$

Le terme au numérateur est la probabilité de trouver la particule 1 dans $d\mathbf{r}_1$ et la particule 2 dans $d\mathbf{r}_2$ indépendamment des $N-2$ molécules restantes. Cette probabilité est obtenue en intégrant la probabilité classique (relation suivante) sur les particules de 3 à N.

$$P^{(N)}(\mathbf{r}^1,...,\mathbf{r}^N) = \frac{\exp\left(\frac{-\mathcal{H}(\mathbf{r}^N,\mathbf{p}^N)}{kT}\right)d\mathbf{r}^1,...,d\mathbf{r}^N}{Z_N}$$

La relation (5.6) peut aussi s'écrire

$$\begin{aligned}
\mathcal{W}(\mathbf{r}^1,\mathbf{r}^2) &= -kT\ln\left[\frac{\int\exp\left(-\frac{pV}{kT}\right)dV\int\int\exp\left(\frac{-\mathcal{H}(\mathbf{r}^N,\mathbf{p}^N)}{kT}\right)d\tau^{'}}{\int\exp\left(-\frac{pV}{kT}\right)dV\int\int\exp\left(\frac{-\mathcal{H}(\mathbf{r}^N,\mathbf{p}^N)}{kT}\right)d\tau}\right] + Cste \\
&= -kT\ln\left[\int\exp\left(-\frac{pV}{kT}\right)dV\int\int\exp\left(\frac{-\mathcal{H}(\mathbf{r}^N,\mathbf{p}^N)}{kT}\right)d\tau^{'}\right] \\
&+ kT\ln\left[\int\exp\left(-\frac{pV}{kT}\right)dV\int\int\exp\left(\frac{-\mathcal{H}(\mathbf{r}^N,\mathbf{p}^N)}{kT}\right)d\tau\right] + Cste
\end{aligned}$$

Dans cette expression $d\tau$ correspond à l'hyperespace total, $d\tau^{'}$ à l'hyperespace non contraint et $d\phi$ l'hyperespace contraint. La relation liant ces trois hyperespaces est $d\tau = d\tau^{'} + d\phi$ [75]. Or, l'enthalpie libre s'écrit dans l'espace non contraint de la manière suivante :

$$G(\tau^{'}) = -kT\ln\left[\int\exp\left(-\frac{pV}{kT}\right)dV\int\int\exp\left(\frac{-\mathcal{H}(\mathbf{r}^N,\mathbf{p}^N)}{kT}\right)d\tau^{'}\right] + Cste$$

Comme on a la relation $G(\tau) = G(\tau^{'}) + G(\phi)$, on peut écrire :

$$G(\phi) = G(\tau) - G(\tau^{'}) = kT\ln\left[\frac{\int\exp\left(-\frac{pV}{kT}\right)dV\int\int\exp\left(\frac{-\mathcal{H}(\mathbf{r}^N,\mathbf{p}^N)}{kT}\right)d\tau^{'}}{\int\exp\left(-\frac{pV}{kT}\right)dV\int\int\exp\left(\frac{-\mathcal{H}(\mathbf{r}^N,\mathbf{p}^N)}{kT}\right)d\tau}\right]$$

En effet pour $G(\tau)$ et $G(\tau^{'})$ la constante est la même. Dans ce cas, on obtient l'égalité suivante :

$$G(\phi) = G(\mathbf{r}^1,\mathbf{r}^2) = -\left[\mathcal{W}(\mathbf{r}^1,\mathbf{r}^2) + Cste\right]$$

Cette relation montre que le potentiel de force moyenne n'est autre que l'enthalpie libre au signe et à une constante près. Les méthodes de perturbation et d'intégration thermodynamique

(équation (5.7)) permettent seulement d'évaluer des différences d'énergie libre absolue entre deux états, mais en aucun cas l'énergie libre du système.

$$\begin{aligned} \Delta G^{FEP} &= \sum_{i}^{i=N} \Delta G = \sum_{\lambda=0}^{\lambda=1} -RT \ln \left\langle \exp -\beta \left[\mathcal{H}(\mathbf{r}^N, \lambda + \delta\lambda) - \mathcal{H}(\mathbf{r}^N, \lambda) \right] \right\rangle_\lambda \\ \Delta G^{TI} &= \int_0^1 \frac{\partial G(\lambda)}{\partial \lambda} d\lambda = \int_0^1 \left\langle \frac{\partial \mathcal{H}^N(\mathbf{r}, \lambda)}{\partial \lambda} \right\rangle_\lambda d\lambda \end{aligned} \tag{5.7}$$

Néanmoins, connaissant la variation d'énergie libre, il est possible d'accéder au profil de potentiel de force moyenne. En effet, il suffit d'avoir un point de référence d'énergie nulle et d'y ajouter progressivement chaque variation intermédiaire liée à chacun des micro-états. Le plus évident est de commencer le calcul lorsque les deux molécules sont suffisament éloignées l'une de l'autre pour ne pas interagir : il est ainsi possible par ces méthodes de perturbation de déterminer le potentiel de force moyenne. Soient deux micro-états consécutifs k et $k-1$ tels que l'état k correspond à l'état d'interaction faible. La variation d'enthalpie libre de perturbation pour ces états s'écrit alors :

$$\begin{aligned} \Delta G^k = G^{k-1} - G^k = G^{k-1} \quad &\Rightarrow \quad G^{k-1} = \Delta G^k \\ \Delta G^{k-1} = G^{k-2} - G^{k-1} \quad &\Rightarrow \quad G^{k-2} = \Delta G^{k-1} + G^{k-1} = \Delta G^{k-1} + \Delta G^k \end{aligned}$$

On montre ainsi que $G^i = \sum_{j=N^t}^{i-1} \Delta_p G^j = -\mathcal{W}$. ($N^t$) est le nombre total de micro-états. L'indice $_p$ indique une grandeur de perturbation.

Le calcul de perturbation se fera par la méthode FEP classique et la méthode FEP^{m}mise au point dans le chapitre (4), dans laquelle on calcule les positions et les vitesses à λ' à partir du N-uplet $(\mathbf{r}^N, \mathbf{p}^N)^\lambda$. Nous pensons que la nécessité d'effectuer ce calcul dans le cas du potentiel de force moyenne est d'autant plus justifiée qu'il existe un déplacement explicite de particules à effectuer. En effet, supposons que l'on déplace deux particules parmi N, il est fort probable qu'on réduise la distance entre les espèces contraintes et une particule de solvant (N-2). On observe alors des dérives énergétiques dues à des recouvrements entre espèces. C'est pour cela qu'il est nécessaire de déterminer l'espace des phases à l'état perturbé par le biais des relations développées dans le chapitre (4). Les formules opérationnelles pour les méthodes FEP et TI, permettant d'évaluer le potentiel de force moyenne sont :

$$\mathcal{W}_{\lambda}^{\mathbf{FEP}} = -kT \ln \left\langle \exp\left(- \left[\frac{\mathcal{H}_{\lambda'}(\mathbf{r}^N, \mathbf{p}^N) - \mathcal{H}_{\lambda}(\mathbf{r}^N, \mathbf{p}^N)}{kT} \right] \right) \right\rangle_{\lambda} \tag{5.8}$$

$$\mathcal{W}_{\lambda}^{\mathbf{TI}} = \int_0^1 \left[\left\langle \frac{\partial \mathcal{H}_{\lambda}(\mathbf{r}^N)}{\partial \lambda} \right\rangle_{\lambda} \right] d\lambda \tag{5.9}$$

Mise en œuvre

Pour utiliser les méthodes de perturbation et d'intégration thermodynamique, nous devons évaluer les coordonnées des atomes qui sont déplacés. Soient deux molécules dont les centres de masse sont indiqués par cm_A et cm_B. On considére les trois états suivants, λ, $\lambda + \Delta\lambda = \lambda^{'}$ et $\lambda - \Delta\lambda = \lambda^{''}$. λ représente le paramètre de couplage introduit dans le chapitre (4). Les exposants $^{'}$ et $^{''}$ indiquent deux états perturbés. Le schéma de la figure (5.1) permet de visualiser ces états.

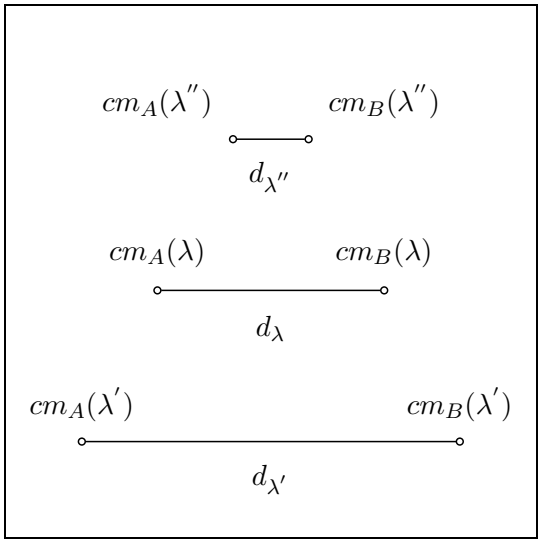

FIG. 5.1 – Schéma permettant de visualiser le déplacement des centres de masse de deux molécules dans le cas d'un élongation ($\lambda^{'}$) de la distance intermoléculaire et dans le cas d'une diminution ($\lambda^{''}$) de cette distance. (λ) désigne le paramètre de couplage pour l'état de référence.

Soit $d_{\Delta\lambda} = d_{\lambda'} - d_{\lambda} = d_{\lambda} - d_{\lambda''}$ la variation de distance entre l'état de référence et l'état perturbé. Si l'amplitude de la perturbation est faible, on peut supposer que la translation se fait selon l'axe $(A-B)^{\lambda}$. Dans ces conditions, on peut calculer les coordonnées des centres de

masse dans les états perturbés, comme le montre le calcul suivant où **r** est le vecteur position du centre de masse.

$$\frac{d_\lambda}{d_{\Delta\lambda}} = \frac{\mathbf{r}_\lambda^{cm_B} - \mathbf{r}_\lambda^{cm_A}}{\mathbf{r}_\lambda^{cm_A} - \mathbf{r}_{\lambda'}^{cm_A}}$$

$$\rightarrow \mathbf{r}_{\lambda'}^{cm_A} = \mathbf{r}_\lambda^{cm_A}. + \frac{d_{\Delta\lambda}}{d_\lambda}\left[\mathbf{r}_\lambda^{cm_B} - \mathbf{r}_\lambda^{cm_A}\right]$$

De la même manière, on a :

$$\mathbf{r}_{\lambda'}^{cm_B} = -\mathbf{r}_\lambda^{cm_B}. + \frac{d_{\Delta\lambda}}{d_\lambda} + \left[\mathbf{r}_\lambda^{cm_B} - \mathbf{r}_\lambda^{cm_A}\right]$$

$$\mathbf{r}_{\lambda''}^{cm_A} = \mathbf{r}_\lambda^{cm_A} + .\frac{d_{\Delta\lambda}}{d_\lambda}\left[-\mathbf{r}_\lambda^{cm_B} + \mathbf{r}_\lambda^{cm_A}\right]$$

$$\mathbf{r}_{\lambda''}^{cm_B} = -\mathbf{r}_\lambda^{cm_B} + \frac{d_{\Delta\lambda}}{d_\lambda}\left[-\mathbf{r}_\lambda^{cm_B} + \mathbf{r}_\lambda^{cm_A}\right]$$

Ainsi, une fois que l'on a calculé les coordonnées des centres de masses des états perturbés, on doit déterminer les coordonnées des atomes constituant les molécules dont on a déplacé les centres de masse. Il ne reste plus qu'à évaluer les interactions entre les atomes perturbés des deux molécules contraintes et les atomes des molécules de solvant. On peut abaisser le nombre d'interactions à calculer si on déplace une seule molécule. Dans ce cas, on a le schéma de la figure 5.2) où $\mathbf{r}_\lambda^{cm_A} = \mathbf{r}_{\lambda'}^{cm_A} = \mathbf{r}_{\lambda''}^{cm_A}$. Ainsi, la position du centre de masse de la molécule que l'on a déplacé, vaut :

$$\mathbf{r}_{\lambda'}^{cm_B} = \mathbf{r}_\lambda^{cm_B} + \mathbf{r}_\lambda^{cm_A}(\frac{d_\lambda}{d_{\lambda'}} - 1)$$

5.2.2 Méthode de la force contrainte

Cette méthode consiste à déterminer la force contrainte et à l'intégrer pour déterminer le potentiel de force moyenne. Pour comprendre l'origine de cette méthode, nous devons dériver

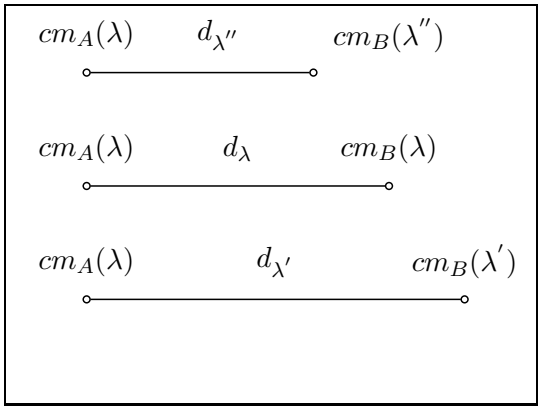

FIG. 5.2 – Cas où un seul centre de masse est déplacé. La légende est la même que celle utilisée sur la figure 5.1.

l'expression du potentiel de force moyenne donnée par les relations (5.1) et (5.3). Le gradient de potentiel de force moyenne ($\mathcal{W}^2$) est donné par la relation (5.10).

$$-\Delta_j \mathcal{W}^{(2)} = \frac{\int ... \int \exp - [\beta U] \, (-\Delta_j U) \, d\mathbf{r}_3 ... d\mathbf{r}_N}{\int ... \int \exp - [\beta U] d\mathbf{r}_3 ... d\mathbf{r}_N} \tag{5.10}$$

L'exposant ^{2} indique que l'on traite deux atomes. Or, $-\Delta_j U$ est la force exercée sur l'atome j par l'ensemble des N atomes constituant le système. Si l'on considère le système contraint, la fonction de partition caractéristique de cet ensemble n'est autre que :

$$Q^c_{NVT} = \int ... \int \exp - [\beta U] d\mathbf{r}_3 ... d\mathbf{r}_N$$

Ainsi, la relation (5.10) s'écrit, en considérant la moyenne d'ensemble dans l'espace contraint :

$$-\Delta_j \mathcal{W}^2 = \langle (-\Delta_j U) \rangle_c = \langle \mathbf{F}_j \rangle_c \tag{5.11}$$

On a alors :

$$\mathcal{W}^{(2)} = \int \langle \mathbf{F}_j \rangle_c \, dr_j \tag{5.12}$$

Finalement, ce gradient de potentiel n'est autre que la moyenne de la force contrainte, d'où le nom *potentiel de force moyenne*. Ainsi, il suffit de déterminer cette force contrainte et de

l'intégrer.

La méthode dite d'Intégration Thermodynamique (TI) fait aussi intervenir cette force contrainte comme le montre J.A.McCammon [79, 80]. Cela implique de découpler l'espace des phases en deux sous-espaces : un espace contraint et un espace non contraint. L'hamiltonien s'écrit :

$$\mathcal{H}(\mathbf{r},\mathbf{p}) = \mathcal{H}(\mathbf{r}^{N_c},\mathbf{p}^{N_c}) + \mathcal{H}(\mathbf{r}^{N-2},\mathbf{p}^{N-2}) \tag{5.13}$$

Ici, N_c représente les molécules ou ions constraints. Ainsi, seul $\mathcal{H}(\mathbf{r}^{N_c},\mathbf{p}^{N_c})$ est fonction de λ. La relation principale d'intégration thermodynamique donnée par l'équation (5.7) peut s'écrire :

$$\mathcal{W}_\lambda = \int_0^1 \left[\left\langle \frac{\partial \mathcal{H}_\lambda(\mathbf{r}^{N_c},\mathbf{p}^{N_c})}{\partial \lambda} \right\rangle_\lambda \right] d\lambda \tag{5.14}$$

Finalement, dans la méthode d'intégration thermodynamique ou dans la méthode de la force contrainte, il faut calculer cette force contrainte. Cela est possible en utilisant la méthode SHAKE [20] ou FAST SHAKE [56].

Procédure SHAKE

La première chose à faire est d'écrire la relation $\dfrac{\partial \mathcal{H}_\lambda(\mathbf{r}^{N_c},\mathbf{p}^{N_c})}{\partial \lambda}$ afin de faire apparaître la force contrainte calculable pendant la simulation. Pour cela, nous devons décomposer ce terme et utiliser les propriétés des dérivées. Dans la suite de la démonstration, nous remplaçons $\mathcal{H}_\lambda(\mathbf{r}^{N_c},\mathbf{p}^{N_c})$ par ${\mathcal{H}^c}_\lambda$ pour plus de souplesse dans les écritures.

$$\begin{aligned}
\left\langle \frac{\partial {\mathcal{H}^c}_\lambda}{\partial \lambda} \right\rangle_\lambda &= \left\langle \frac{\partial {\mathcal{H}^c}_\lambda}{\partial r_{ij}} \cdot \frac{\partial r_{ij}}{\partial \lambda} \right\rangle_\lambda \Leftrightarrow \left\langle \frac{\partial {\mathcal{H}^c}_\lambda}{\partial \lambda} \right\rangle_\lambda = \left\langle \frac{\partial {\mathcal{H}^c}_\lambda}{\partial r_{ij}} \right\rangle_\lambda \cdot \frac{\partial r_{ij}}{\partial \lambda} \\
\left\langle \frac{\partial {\mathcal{H}^c}_\lambda}{\partial \lambda} \right\rangle_\lambda &= \frac{\partial r_{ij}}{\delta \lambda} \left[\left\langle -\frac{\partial {\mathcal{H}^c}_\lambda}{\partial \mathbf{r}_i} \cdot \frac{\partial \mathbf{r}_i}{\partial r_{ij}} + \frac{\partial {\mathcal{H}^c}_\lambda}{\partial \mathbf{r}_j} \cdot \frac{\partial \mathbf{r}_j}{\partial r_{ij}} \right\rangle_\lambda \right]
\end{aligned}$$

Comme la force contrainte est une force conservative, celle-ci dérive automatiquement d'un potentiel. Cela nous permet d'écrire :

$$\begin{aligned}\mathbf{F}_i^c &= -\frac{\partial \mathcal{H}^c{}_\lambda}{\partial \mathbf{r}_i},\ \mathbf{F}_j^c = -\frac{\partial \mathcal{H}^c{}_\lambda}{\partial \mathbf{r}_j} \\ \mathbf{F}_i^c &= -\mathbf{F}_j^c = \mathbf{F}^c\end{aligned}$$

$$\Rightarrow \left\langle \frac{\partial \mathcal{H}^c{}_\lambda}{\partial \lambda} \right\rangle_\lambda = \frac{\partial r_{ij}}{\partial \lambda} \left[\left\langle \mathbf{F}^c \left(\frac{\partial \left(\mathbf{r}_i - \mathbf{r}_j \right)}{\partial r_{ij}} \right) \right\rangle \right] \tag{5.15}$$

Comme :

$$\hat{\mathbf{r}}_{ij} = \frac{\mathbf{r}_{ij}}{r_{ij}} \Rightarrow \frac{\partial \mathbf{r}_{ij}}{\partial r_{ij}} = \hat{\mathbf{r}}_{ij}$$

l'équation (5.15) devient :

$$\left\langle \frac{\partial \mathcal{H}^c{}_\lambda}{\partial \lambda} \right\rangle_\lambda = \frac{\partial r_{ij}}{\partial \lambda} \left\langle \mathbf{F}^c \cdot \hat{\mathbf{r}}_{ij} \right\rangle$$

Ainsi, on a :

$$\mathcal{W} = \int_0^1 \left\langle \frac{\partial \mathcal{H}^c{}_\lambda}{\partial \lambda} \right\rangle_\lambda d\lambda = \int_0^1 \frac{\partial r_{ij}}{\partial \lambda} \left\langle \mathbf{F}^c \cdot \hat{\mathbf{r}}_{ij} \right\rangle_\lambda d\lambda$$

L'objectif est à présent de déterminer cette force contrainte. Cela est possible en utilisant la relation démontrée dans la section (9.4).

$$\begin{aligned}-2 \cdot (\Delta t)^2 \quad & \gamma_1 . \left[\frac{\nabla_j}{m_j} - \frac{\nabla_i}{m_i} \right] . \sigma_1(t) \cdot \left(\mathbf{r}_j^{nc}(t + \Delta t) - \mathbf{r}_i^{nc}(t + \Delta t) \right) + \\ (\Delta t)^4 \qquad & \cdot \gamma_1^2 \left[\left(\frac{\nabla_j}{m_j} - \frac{\nabla_i}{m_i} \right) . \sigma_1(t) \right]^2 = \left(d_{ij}^2 - \left(\mathbf{r}_j^{nc}(t + \Delta t) - \mathbf{r}_i^{nc}(t + \Delta t) \right)^2 \right)\end{aligned}$$

$$\begin{aligned}\Leftrightarrow - \quad 2 \cdot (\Delta t)^2 \quad & \gamma_1 . \left[\frac{\nabla_j}{m_j} - \frac{\nabla_i}{m_i} \right] . \sigma_1(t) \cdot \left(\mathbf{r}_{ji}^{nc}(t + \Delta t) \right) + \\ (\Delta t)^4 \qquad & \cdot \gamma_1^2 \left[\left(\frac{\nabla_j}{m_j} - \frac{\nabla_i}{m_i} \right) . \sigma_1(t) \right]^2 = \left(d_{ji}^2 - \left(\mathbf{r}_{ji}^{nc}(t + \Delta t) \right)^2 \right)\end{aligned}$$

Ici, $\mathbf{r}^{nc}$ représente la distance non contrainte. Comme on a :

$$\begin{aligned}\sigma_1 &= \mathbf{r}_{ji}^2 - d_{ji}^2 \\ \Rightarrow \quad \nabla_j = \frac{\sigma_1}{\mathbf{r}_j} &= 2\mathbf{r}_{ji}\ ,\ \nabla_i = \frac{\sigma_1}{\mathbf{r}_i} = -2\mathbf{r}_{ji}\end{aligned}$$

On peut écrire que :

$$\begin{aligned}\Rightarrow - \quad (\Delta t)^2 \quad & 4\cdot\gamma_1\cdot\ \mathbf{r}_{ji}(t)\cdot\left[\frac{1}{m_j}+\frac{1}{m_i}\right]\cdot\left(\mathbf{r}_{ji}^{nc}(t+\Delta t)\right) + \\ (\Delta t)^4 \quad & \cdot 4\cdot\gamma_1^2\cdot\ \mathbf{r}_{ji}(t)^2\cdot\left[\left(\frac{1}{m_j}+\frac{1}{m_i}\right)\right]^2 = \left(d_{ji}^2 - \left(\mathbf{r}_{ji}^{nc}(t+\Delta t)\right)^2\right)\end{aligned}$$

Le développement de cette expression donne une équation du second ordre :

$$\begin{aligned}\Rightarrow \quad & 4\cdot\gamma_1^2\cdot(\Delta t)^4\cdot\mathbf{r}_{ji}(t)^2\cdot\mu^2 - 4\cdot\gamma_1\cdot(\Delta t)^2\cdot\mathbf{r}_{ji}(t)\cdot\mu\cdot\mathbf{r}_{ji}^{nc}(t+\Delta t) \\ - \quad & \left(d_{ji}^2 - \left(\mathbf{r}_{ji}^{nc}(t+\Delta t)\right)^2\right) = 0 \\ \text{avec} \quad & \mu = \frac{1}{m_j}+\frac{1}{m_i}\end{aligned} \tag{5.16}$$

La solution exacte est :

$$\gamma_1 = -\frac{d_{ji} + \mathbf{r}_{ji}^{nc}(t+\Delta t)}{2\cdot\Delta t^2\cdot\mathbf{r}_{ji}(t)\cdot\mu}$$

Soit σ la contrainte, nous pouvons alors écrire que :

$$\mathbf{F}_i^c(t) = -\sum_{k=1}^{N_c}\gamma_k(t)\cdot\frac{\partial\sigma_k(\mathbf{r}_i(t))}{\partial\mathbf{r}_i(t))}$$

Ici, N_c represente le nombre total de contraintes agissant sur l'atome i. Dans le cas d'une contrainte entre deux atomes i et j, nous obtenons :

$$\mathbf{F}_i^c(t) = -\gamma_1(t)\cdot\frac{\partial\sigma_1(\mathbf{r}_i(t))}{\partial\mathbf{r}_i(t))} = -2\cdot\gamma_1(t)\cdot\mathbf{r}_{ij}(t) \quad avec \quad \sigma_1 = -d_{ij}^2 + r_{ij}^2$$

Le potentiel de force moyenne peut alors s'écrire :

$$\begin{aligned} \mathcal{W}^{FCS} &= \int_0^1 \frac{\partial r_{ij}}{\partial \lambda} \langle 2 \cdot \gamma_1(t) \cdot \mathbf{r}_{ij}(t) \cdot \hat{\mathbf{r}}_{ij} \rangle_\lambda \, d\lambda \\ \Leftrightarrow \quad \mathcal{W} &= \int_0^1 \frac{\partial r_{ij}}{\partial \lambda} \left\langle \frac{d_{ji} - r_{ji}^{nc}(t+\Delta t)}{\Delta t^2 \cdot \mu} \right\rangle_\lambda d\lambda \end{aligned} \tag{5.17}$$

Ici, FCS indique Force Contrainte par la procédure SHAKE. Cette relation met en évidence une dépendance temporelle, ce qui implique que les calculs doivent être menés durant la simulation. Ce calcul peut être simplifié si l'on néglige le terme quadratique [56], ce qui a pour effet de linéariser l'équation. À l'origine, cette approximation a été réalisée pour accroître la vitesse de calcul dans le cas de plusieurs contraintes. Néanmoins, cette approximation s'est généralisé au calcul de potentiel de force moyenne où il n'est pas nécessaire de linéariser le binôme puisqu'on traite une seule constrainte.

Procédure FAST SHAKE

Cette procédure consiste en la linéarisation du polynôme de second ordre précédement obtenu. Cette approximation posée, nous obtenons :

$$4 \cdot \gamma_1 \cdot (\Delta t)^2 \cdot \mathbf{r}_{ji}(t) \cdot \mu \cdot \mathbf{r}_{ji}^{nc}(t+\Delta t) - \left(d_{ji}^2 - \left(\mathbf{r}_{ji}^{nc}(t+\Delta t) \right)^2 \right) = 0 \tag{5.18}$$

La résolution de cette équation donne γ_1 :

$$\gamma_1 = \frac{\left(\mathbf{r}_{ji}^{nc}(t+\Delta t) \right)^2 - d_{ji}^2}{4 \cdot (\Delta t)^2 \cdot \mathbf{r}_{ji}(t) \cdot \mu \cdot \mathbf{r}_{ji}^{nc}(t+\Delta t)}$$

Dans les deux méthodes, les équations de mouvement sont intégrées par le biais d'un cycle itératif pour les contraintes avec un critère de convergence τ de l'ordre de 10^{-6} :

$$\frac{r_{ij}(t+\Delta t) - d_{ij}}{d_{ij}} \leq \tau$$

L'expression finale du potentiel de force moyenne est :

$$\mathcal{W}^{FCFS} = -\int_0^1 \frac{\partial r_{ij}}{\partial \lambda} \left\langle r_{ji}(t) \cdot \frac{\left(\mathbf{r}_{ji}^{nc}(t+\Delta t) \right)^2 - d_{ji}^2}{2 \cdot (\Delta t)^2 \cdot \mathbf{r}_{ji}(t) \cdot \mu \cdot \mathbf{r}_{ji}^{nc}(t+\Delta t)} \right\rangle_\lambda d\lambda \tag{5.19}$$

Ici, $FCFS$ indique Force Contrainte par la procédure FAST SHAKE. Dans les deux cas, nous devons évaluer les positions non contraintes. Cela est réalisé par l'algorithme de Verlet [7] :

$$\mathbf{r}_i(t+\Delta t) = 2\cdot\mathbf{r}_i(t) - \mathbf{r}_i(t-\Delta t) + \dot{\mathbf{r}}_i(t)\cdot\Delta t + \frac{\Delta t^2}{m_i}\cdot\left[\mathbf{F}_i(t) - \sum_{k=1}^{\ell}\gamma_k.\left(\nabla_i.\sigma_k\right)(t)\right]$$

Si on sépare les parties contraintes et non contraintes, on obtient :

$$\mathbf{r}_i(t+\Delta t) = \mathbf{r}_i^{nc}(t+\Delta t) - \frac{\Delta t^2}{m_i}\cdot\left[\sum_{k=1}^{\ell}\gamma_k.\left(\nabla_i.\sigma_k\right)(t)\right]$$

On a alors :

$$\begin{aligned}\mathbf{r}_j(t+\Delta t) &= \mathbf{r}_j^{nc}(t+\Delta t) - 2\cdot\frac{\Delta t^2}{m_j}\cdot\gamma_1.\mathbf{r}_{ji}(t)\\ \mathbf{r}_i(t+\Delta t) &= \mathbf{r}_i^{nc}(t+\Delta t) + 2\cdot\frac{\Delta t^2}{m_i}\cdot\gamma_1.\mathbf{r}_{ji}(t)\end{aligned}$$

À présent que l'on a vu les méthodes permettant de calculer le potentiel de force moyenne, nous allons nous intéresser aux méthodes permettant de calculer les grandeurs absolues.

5.2.3 Évaluation de grandeurs absolues

La détermination des grandeurs absolues ($\Delta_r G^o$,$\Delta_r H^o$ et $\Delta_r S^o$)est basée sur le calcul de la constante d'association proposé par Prue en 1969 [59]. Cette méthode est basée sur la théorie des paires d'ions. On considére que le volume disponible pour un ion B pour former une paire d'ions AB avec un ion A vaut $4\pi r^2\Delta r$, si A et B sont *sphériques* (figure (5.3)).

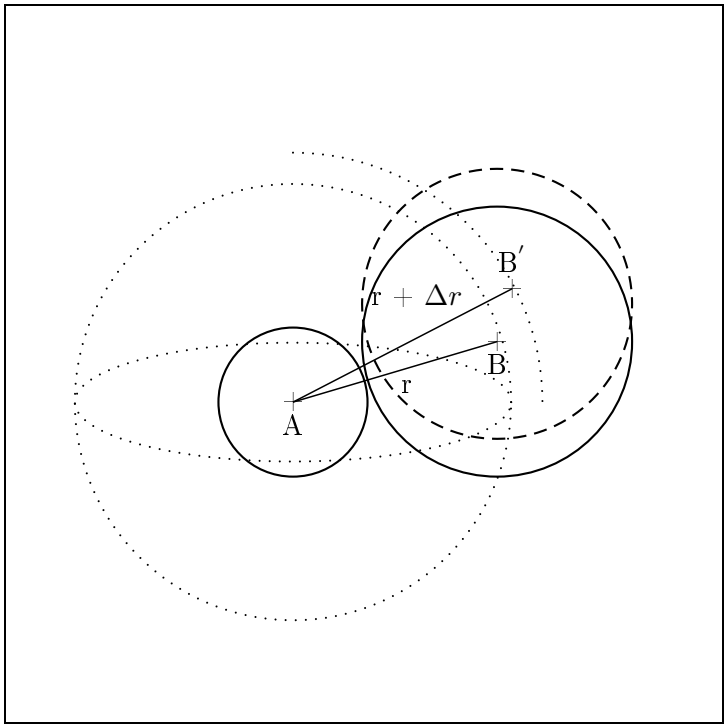

FIG. 5.3 – Définition d'une paire d'ions AB pour le calcul de la constante d'association.

Cette figure nous permet de voir que le volume disponible pour la paire d'ions ainsi formée vaut $4\pi r^2 \Delta r$ si Δr est suffisament petit. Δr représente la variation de distance que peut subir r grâce au mouvement thermique des ions. Le nombre de paires d'ions est égal au nombre d'ions B. Soit $\mathcal{N}.V.C_{AB}$ le nombre de paires d'ions dans le volume V (volume total de la solution) et $\mathcal{N}.V.C_A$ le nombre d'ions A dans le volume V ne participant à aucune paire. Le rapport de ces deux grandeurs est approximativement égal au rapport du volume total occupé par les paires sur le volume total, soit :

$$\frac{\mathcal{N} \cdot V \cdot C_{AB}}{\mathcal{N} \cdot V \cdot C_A} = \frac{4\pi r^2 \Delta r \cdot \mathcal{N}.V.C_B}{V} \tag{5.20}$$

En effet, $4\pi r^2 \Delta r \cdot \mathcal{N}.V.C_B$ représente le volume occupé par l'intégralité des paires d'ions puisqu'on a autant d'ions B que de paires dans le volume disponible (V_d). Cette relation peut aussi s'écrire :

$$\begin{aligned}\frac{C_{AB}}{C_A} &= 4\pi r^2 \Delta r \cdot \mathcal{N} \cdot C_B \\ &\Leftrightarrow \quad K = \frac{C_{AB}}{C_A \cdot C_B} = 4\pi r^2 \Delta r \cdot \mathcal{N}\end{aligned}$$

Cette étude a été réalisée sans considérer explicetement le potentiel d'interaction entre les molécule constituant la paire d'ions. À présent, plaçons-nous dans le cas où il existe une interaction notée $\mathcal{W}$ que l'on peut assimiler au potentiel de force moyenne. Dans ces conditions, la constante d'équilibre peut s'exprimer comme :

$$K = \frac{C_{AB}}{C_A \cdot C_B} = \frac{1}{C_B}\frac{\rho_1}{\rho_0}$$

où ρ_1 représente la fraction de A participant à une paire et ρ_0 la fraction de A non liée. En mécanique statistique cette relation s'écrit de la manière suivante.

$$K = \frac{C_{AB}}{C_A \cdot C_B} = \frac{1}{C_B}\frac{V_d \mathcal{N} \int \exp\left(-\frac{pV}{kT}\right) dV \int\int \exp\left(\frac{-\mathcal{H}(\mathbf{r}^N, \mathbf{p}^N)}{kT}\right) d\mathbf{r}^{Nnc} d\mathbf{p}^{Nnc}}{V\mathcal{N} \int \exp\left(-\frac{pV}{kT}\right) dV \int\int \exp\left(\frac{-\mathcal{H}(\mathbf{r}^N, \mathbf{p}^N)}{kT}\right) d\mathbf{r}^N d\mathbf{p}^N}$$

ρ_1 est une distribution dans l'espace non contraint ($d\mathbf{r}^3, d\mathbf{r}^4, .., d\mathbf{r}^N$ et ρ_0) une distribution dans l'espace total ($d\mathbf{r}^1, d\mathbf{r}^2, .., d\mathbf{r}^N$). Étant donnée la relation (5.6), on peut écrire :

$$K = \frac{C_{AB}}{C_A \cdot C_B} = \frac{V_d}{VC_B} \exp\left(-\left[\frac{\mathcal{W}(\mathbf{r})}{kT}\right]\right)$$

Comme $4\pi r^2 \Delta r C_B \mathcal{N} V = V_d$, on a en intégrant :

$$K = \frac{C_{AB}}{C_A \cdot C_B} = \int_0^{d_l} 4\pi r^2 dr \cdot \mathcal{N} \cdot \exp\left(-\left[\frac{\mathcal{W}(r)}{kT}\right]\right)$$

avec d_l la borne supérieure d'intégration ou borne supérieure du volume disponible. Une fois la constante d'équilibre déterminée, il est possible d'accéder aux autres grandeurs thermodynamiques par les relations classiques de la thermodynamique :

$$\Delta_r G^o = -kT \ln K$$

$$\Delta_r S^o = -\frac{d\Delta_r G}{dT} = k\left[T\frac{1}{K}\frac{dK}{dT} + \ln K\right]$$

$$\Delta_r H^o = kT^2 \frac{1}{K}\frac{dK}{dT}$$

Comme le calcul de $\frac{1}{K}\frac{dK}{dT}$ donne :

$$\frac{1}{K}\cdot\frac{dK}{dT} = \frac{\int_0^{d_l} \frac{r^2\cdot\mathcal{W}}{kT^2}\exp\left(-\left[\frac{\mathcal{W}(r)}{kT}\right]\right)dr}{\int_0^{d_l} r^2\cdot\exp\left(-\left[\frac{\mathcal{W}(r)}{kT}\right]\right)dr}$$

nous obtenons alors :

$$\Delta G^o = -kT\ln\int_0^{d_l} 4\pi\mathcal{N}r^2\cdot\exp\left(-\left[\frac{\mathcal{W}(r)}{kT}\right]\right)dr \tag{5.21}$$

$$T\Delta S^o = kT\ln\int_0^{d_l} 4\pi\mathcal{N}r^2\cdot\exp\left(-\left[\frac{\mathcal{W}(r)}{kT}\right]\right)dr + \frac{\int_0^{d_l} r^2\cdot\mathcal{W}\exp\left(-\left[\frac{\mathcal{W}(r)}{kT}\right]\right)dr}{\int_0^{d_l} r^2\cdot\exp\left(-\left[\frac{\mathcal{W}(r)}{kT}\right]\right)dr}$$

$$\Delta H^o = \frac{\int_0^{d_l} r^2\cdot\mathcal{W}\exp\left(-\left[\frac{\mathcal{W}(r)}{kT}\right]\right)dr}{\int_0^{d_l} r^2\cdot\exp\left(-\left[\frac{\mathcal{W}(r)}{kT}\right]\right)dr}$$

5.3 Potentiel de force moyenne de la paire d'ions [$Ca^{2+}\cdots SO_4^{2-}$]

Cette étude se fera en deux parties. Dans la première partie, nous étudierons cette association en phase gazeuse afin de comparer les différentes méthodes entre elles. La comparaison avec une étude en phase aqueuse nous permettra de mettre en évidence les effet de solvant sur l'association.

5.3.1 Étude en phase gazeuse

Le calcul du profil de potentiel de force moyenne de la paire d'ion [$Ca^{2+}\cdots SO_4^{2-}$] a nécessitée l'échantillonnage de la distance séparant Ca^{2+} et SO_4^{2-} de 1 Å à 12 Å. L'utilisation

d'un incrément de 0.05 Å a impliqué la mise en place de 220 simulations de type dynamique moléculaire. Les caractéristiques des simulations sont les suivantes : un temps d'équilibration de 50 ps et un temps d'acquisition de 100 ps. Les algorithmes et le champ de forces utilisés pour les simulations ont été précisés dans le chapitre (2). Comme on l'a expliqué précédemment, nous allons comparer les quatre méthodes développées dans la section (5.2). Celles ci sont définies avec leur nomenclature dans le tableau (5.1).

La figure (5.4.a) montre que toutes les méthodes donnent des positions quasi identiques pour le puits de potentiel. Néanmoins, nous observons de grandes différences concernant la forme des profils obtenus. Les méthodes FCFS et FCS montrent des puits localisés respectivement à 2.3 Å et 2.9 Å. En outre, ces profils ont une forme identique. Les méthodes de perturbation FEP et FEP^m donnent des résultats similaires. Elles indiquent un puits dont la position est identique à celle obtenue par la méthode FCS avec des formes différentes. Cette différence entre les méthodes de perturbation et de force contrainte est due à la grandeur calculée, puisque dans le cas des méthodes de perturbation, on évalue une différence d'énergie $\Delta\mathcal{H}(\lambda)$ tandis qu'avec la méthode de force contrainte, on évalue une force qui n'est pas la dérivée de la différence d'énergie $\Delta\mathcal{H}(\lambda)$ calculée par les méthodes de perturbation. La similitude des résultats obtenus pour les méthodes FEP et FEP^m s'explique par l'absence de solvant. Dans ces conditions, le déplacement des atomes composant les deux molécules de la paire d'ions n'implique pas de rapprochement en direction des molécules de solvant. Nous verrons par la suite que les choses sont différentes lorsqu'on calcule le potentiel de force moyenne de cette même paire en milieu solvaté.

Ce résultat peut être corroboré par la détermination du potentiel de force moyenne utilisant la méthode PMFK [17]. Le profil de potentiel de force moyenne peut alors être déterminé en pratiquant une simulation de potentiel de force moyenne dans l'ensemble NpT. En effet, on simule une boîte de 500 paires d'ions afin d'avoir une statistique suffisante pour obtenir une distribution radiale convenable. Nous vérifions que le puits calculé avec la méthode PMFK est proche de celui calculé avec les autres méthodes. La figure (5.4.b) représente les résultats obtenus par les méthodes FEP^m, FCFS et PMFK. Elle montre une position du puits de potentiel identique pour les trois méthodes et des formes de profils différentes. On en déduit que les méthodes FEP et FC sont toutes aussi performantes que la méthode PMFK, et peuvent très

Formules opérationnelles

Potentiel de force moyenne

FEP
$$\mathcal{W}(d_{ji}) = -kT \sum_{\lambda} \ln \left\langle \exp\left(- \left[\frac{\mathcal{H}_{\lambda'}(\mathbf{r}^N, \mathbf{p}^N) - \mathcal{H}_{\lambda}(\mathbf{r}^N, \mathbf{p}^N)}{kT} \right] \right) \right\rangle_{\lambda}$$

FEP^m
$$\mathcal{W}(d_{ji}) = -kT \sum_{\lambda} \ln \left\langle \exp\left(- \left[\frac{\mathcal{H}_{\lambda'}(\mathbf{r}^N, \mathbf{p}^N) - \mathcal{H}_{\lambda}(\mathbf{r}^N, \mathbf{p}^N)}{kT} \right] \right) \right\rangle_{\lambda}$$

FCFS
$$\mathcal{W}(d_{ji}) = - \int_0^1 \frac{\partial r_{ij}}{\partial \lambda} \left\langle r_{ji}(t) \cdot \frac{\left(\mathbf{r}_{ji}^{nc}(t + \Delta t) \right)^2 - d_{ji}^2}{2 \cdot (\Delta t)^2 \cdot \mathbf{r}_{ji}(t) \cdot \mu \cdot \mathbf{r}_{ji}^{nc}(t + \Delta t)} \right\rangle_{\lambda} d\lambda$$

FCS
$$\mathcal{W}(d_{ji}) = - \int_0^1 \frac{\partial r_{ij}}{\partial \lambda} \left\langle \frac{d_{ji} - r_{ji}^{nc}(t + \Delta t)}{\Delta t^2 \cdot \mu} \right\rangle_{\lambda} d\lambda$$

PMFK
$$\mathcal{W}^n \left(\mathbf{r}^1, ..., \mathbf{r}^n \right) = -kT \ln g(r)$$

Grandeurs thermodynamiques d'associations

$$\Delta G^o = -kT \ln \int_0^{d_l} 4\pi \mathcal{N} r^2 \cdot \exp\left(- \left[\frac{\mathcal{W}(r)}{kT} \right] \right) dr$$

$$T\Delta S^o = kT \ln \int_0^{d_l} dr 4\pi \mathcal{N} r^2 \cdot \exp\left(- \left[\frac{\mathcal{W}(r)}{kT} \right] \right) dr + \frac{\int_0^{d_l} dr r^2 \cdot \mathcal{W} \exp\left(- \left[\frac{\mathcal{W}(r)}{kT} \right] \right) dr}{\int_0^{d_l} dr r^2 \cdot \exp\left(- \left[\frac{\mathcal{W}(r)}{kT} \right] \right) dr}$$

TAB. 5.1 – FEP : méthode de perturbation FEP classique dans laquelle on perturbe la distance entre les deux ions constituant la paire étudiée. FEP^m méthode de perturbation FEP où l'on recalcule les positions et les vitesses des atomes à l'état pertubé. Cette méthode a été décrite dans le chapitre (4). FCS : méthode de force contrainte sans linéarisation du binôme pour le calcul du multiplicateur de Lagrange dans la procédure SHAKE [20]. Cette méthode est explicitée en détail dans le chapitre (9) et dans la section (5.2). FCFS : méthode de force contrainte avec linéarisation du binôme pour le calcul du multiplicateur de Lagrange dans la procédure SHAKE. Cette procédure, appelée FAST SHAKE [56] est détaillée dans le chapitre (9). PMFK : relation obtenue à partir de l'expression de la distribution radiale formulée par Kirkwood [54]. Cette relation est discutée dans la section (5.2). $\mathcal{W}$ indique le potentiel de force moyenne. $\mathcal{H}$ est l'hamiltonien. λ est le paramètre de couplage. r_{ji}^{nc} est la distance non contrainte entre les espèces j et i, calculée aux cours de la simulation. d_{ji} est la distance contrainte fixée entre les espèces j et i de la paire d'ions étudiée. Δt est le pas d'intégration. μ est la masse réduite du système formé par les deux espèces constituant la paire. d_l est la borne supérieure d'intégration pour le calcul des grandeurs thermodynamiques.

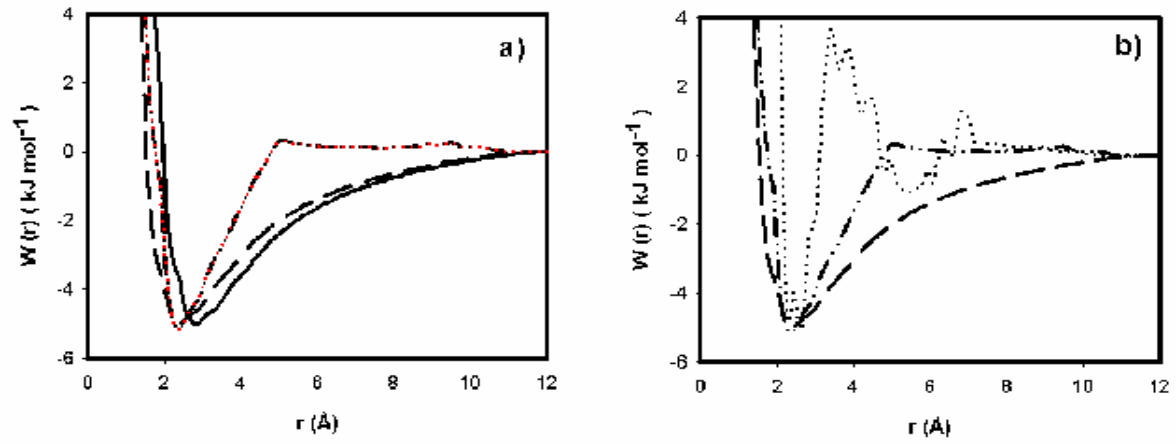

FIG. 5.4 – Profils de potentiel de force moyenne pour la paire d'ions [Ca^{2+} $\cdots$ SO_4^{2-}] en phase gazeuse. a) FCS (– – –), FCFS (—), FEP (···) et FEP^m (···). b) FCS (– – –), FEP^m,(– · · –) ainsi que la méthode PMFK(···). La signification de ces acronymes est donnée dans le tableau (5.1).

bien se substituer à cette dernière. Les résultats du calcul des grandeurs thermodynamiques pour chaque méthode sont regroupés dans le tableau (5.2). Les relations du tableau (5.1) peuvent être appliquées si les ions constituant la paire sont sphériques. Le calcul de l'asphéricité $\mathcal{A}$ pour l'ion SO_4^{2-} défini dans la section (2.6) donne une valeur proche de 0. On peut donc l'assimiler à une sphère.

	r_{min} (Å)	$logK$	$\Delta_r G^o$ (kJ mol^{-1})	$\Delta_r H^o$ (kJ mol^{-1})	$T\Delta_r S^o$ (kJ mol^{-1})
FCFS	2.8	2.2	-2.0	-9.0	-6.8
FCS	2.3	2.2	-2.0	-8.8	-6.9
FEP^m	2.4	2.3	-2.1	-9.1	-7.0
PMFK	2.6	2.9	-2.7	-7.9	-5.2

TAB. 5.2 – Grandeurs thermodynamiques d'association entre les ions Ca^{2+} et SO_4^{2-} en phase gazeuse. Les quatre méthodes étudiées sont représentées.

Ces résultats ont été déterminés par l'application des formules données dans le tableau (5.1). Ceux-ci sont calculés par intégration du potentiel de force moyenne entre les bornes $r = 0$

et $r = d_l$, où r est la distance séparant les deux ions. d_l représente la borne supérieure d'intégration. Nous avons testé plusieurs distances d_l telles que $d_l > r_{min}$, où r_{min} est la position du puits de potentiel. On montre que pour $d_l > (r_{min} + 1)$ Å les résultats sont indépendants de la valeur de la borne d'intégration. Néanmoins, il est nécessaire que la distance d_l soit supérieure à r_{min} où l'interaction entre les deux ions est maximale. Ce tableau montre que les méthodes sont équivalentes pour le calcul de $\Delta_r G^o$ et $\mathrm{T}\Delta_r S^o$ avec une déviation pour la méthode PMFK.

5.3.2 Étude en phase aqueuse

Protocole de simulation

Dans ce cas, les simulations ont été réalisées en utilisant un incrément de distance de 0.044 Å. Les distances (entre les ions formant la paire) échantillonnées sont comprises entre 1.0 Å et 12.0 Å , soit 220 simulations. Les algorithmes pour générer les configurations sont les mêmes que ceux exposés dans le chapitre (2). Ces simulations ont été réalisées dans l'ensemble NpT, ce qui nous a obligé à implémenter une version NpT du potentiel de force moyenne dans le code de dynamique moléculaire de DL_POLY_MD [27], puisque seul le calcul dans l'ensemble NVE était disponible. Le temps d'équilibration pour chaque simulation est de 60 ps tandis que le temps d'acquisition est de 50 ps, soit 24 ns de temps simulé pour la totalité de la perturbation. Le potentiel de force moyenne est évalué à partir des 5 000 configurations stockées durant la phase d'acquisition. Les charges atomiques de l'ion sulfonate ont été évaluées par la procédure CHELPG [28] avec le logiciel GAMESS [28]. Cette paire a été simulée avec 512 molécules d'eau (modèle TIP3P) dans une boîte cubique de longueur $L = 24$ Å.

Détermination des grandeurs thermodynamiques d'association

La méthode utilisée pour déterminer le potentiel de force moyenne est la méthode FEP^m. En effet, l'utilisation de la méthode classique FEP est problématique puisque lorsqu'on déplace les centres de masse des ions constituant la paire d'ions dans l'état de référence pour obtenir l'état perturbé, la probabilité d'un contact avec les molécules de solvant n'est pas négligeable. Cela entraîne des dérives énergétiques sur certaines configurations nuisant au calcul de $\Delta_r G^o$: cette méthode est alors difficilement utilisable. Néanmoins, il est possible de réduire ces dérives en diminuant l'incrément de distance utilisé dans la perturbation, c'est-à-dire en augmentant le nombre de simulations. On peut aussi utiliser la méthode de la force contrainte

afin de remédier à ce problème, puisque dans ce cas, le calcul de potentiel de force moyenne ne nécessite aucun déplacement d'atome. On vérifie alors que les méthodes FEP^m et FCS donnent des résultats similaires.

La figure (5.5.a) représente le profil de potentiel de force moyenne entre l'ion Ca^{2+} et l'ion SO_4^{2-} en phase gazeuse et en phase aqueuse. Elle montre deux puits de potentiel dans la phase aqueuse, p_1 à 2.9 Å et p_2 à 3.5 Å, tels que $E(p_1) < E(p_2)$ (E est l'énergie du puits considérée). En outre, ces puits sont beaucoup plus bas en énergie que le puits de potentiel obtenu en phase gazeuse. On en déduit que les interactions entre les ions Ca^{2+} et SO_4^{2-} sont plus favorables en milieu solvaté. Il existe donc des interactions [$Ca^{2+} \cdots H_2O$] et [$SO_4^{2-} \cdots H_2O$] permettant de stabiliser davantage la paire d'ions. Par ailleurs, la distance (relative au puits) entre les deux ions est plus grande en phase aqueuse qu'en phase gazeuse. Cette augmentation de distance s'explique par la présence de molécules d'eau entre les deux ions. La figure (5.6) permet de visualiser la position de la molécule d'eau entre les ions Ca^{2+} et SO_4^{2-}. En effet, la figure (5.5.b) montre qu'entre les deux puits de potentiel p_1 et p_2, le nombre de ponts de liaisons hydrogène est maximal n_{max}. Cela suggère, que ces deux puits forment une zone où les interactions sont maximales. Étant donnée la position de n_{max}, aucun puits n'est privilégié.

Nous avons aussi comparé la méthode de force contrainte et la méthode de perturbation. La figure (5.7) montre que l'ensemble des méthodes donne des résultats similaires en ce qui concerne la zone d'interaction maximale. Les profils obtenus par les méthodes de force contrainte ont des formes différentes et ne donnent pas de second puits de potentiel. En outre, les puits obtenus par les méthodes de force contrainte sont plus bas en énergie que ceux obtenus par la méthode FEP^m, mais les positionnements des puits (le premier pour FEP^m) sont identiques. Cette comparaison entre les méthodes montre que l'on ne peut tirer de conclusion d'ordre structural sur la présence du second puits de potentiel obtenu par FEP^m. Néanmoins, étant donnée la valeur de l'énergie du second puits, il faut en tenir compte dans les calculs des grandeurs thermodynamiques. Ces derniers sont rassemblés dans le tableau (5.3).

Les résultats du tableau (5.3) montrent que les méthodes FCS et FEP^m donnent des résultats en bon accord avec les valeurs expérimentales, contrairement à la méthode FCFS. Ainsi,

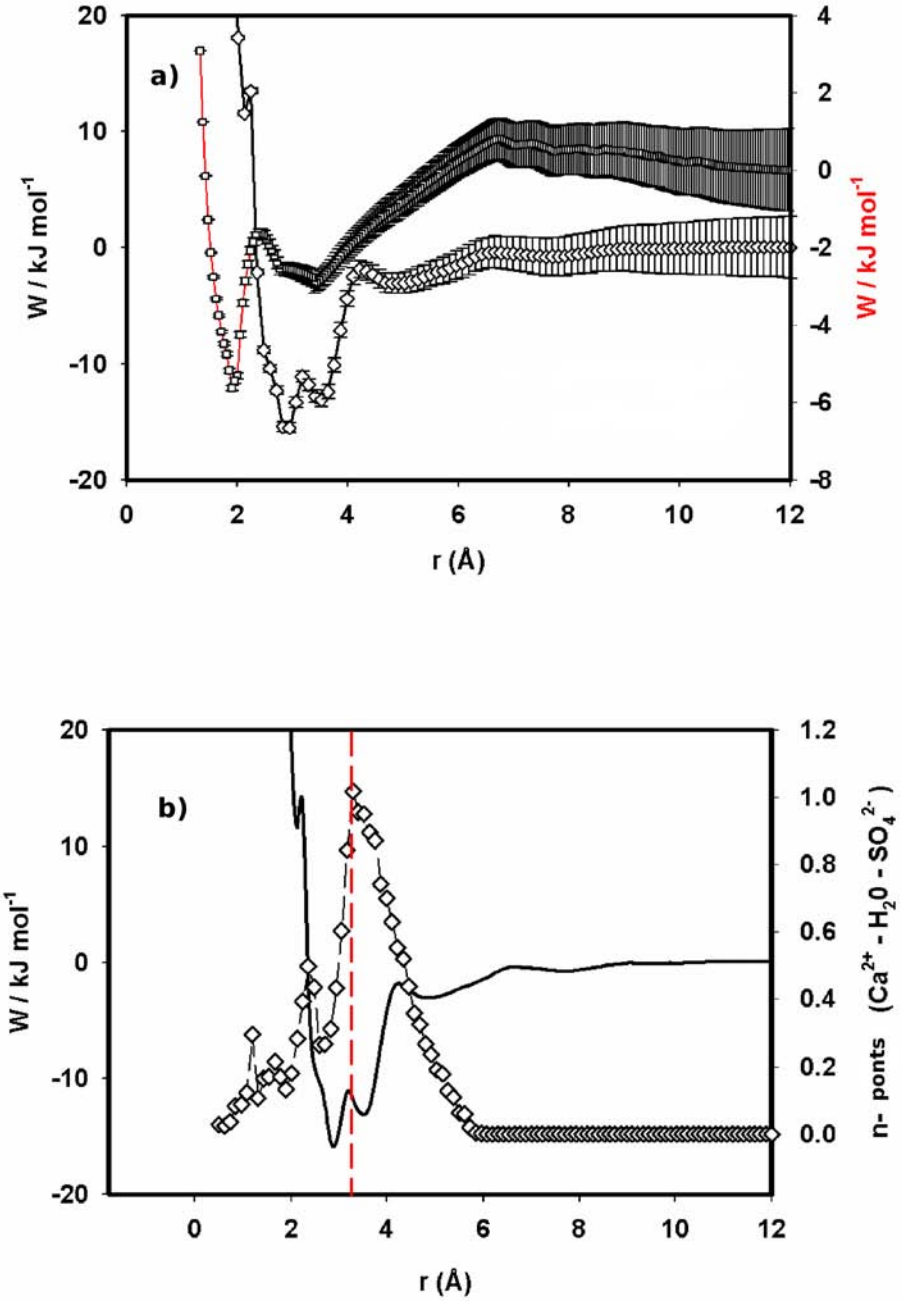

FIG. 5.5 – a) Profil de potentiel de force moyenne pour la paire d'ions $[Ca^{2+} \cdots SO_4^{2-}]$ en phase gazeuse (—) représenté sur l'axe de droite. Profil de potentiel de force moyenne pour la paire d'ions $[Ca^{2+} \cdots SO_4^{2-}]$ en phase aqueuse (—) représenté sur l'axe de gauche. b) (—) profil de potentiel de force moyenne en milieu aqueux représenté sur l'axe de gauche. (◊) nombre moyen de ponts de liaisons hydrogène pour le triplet $[Ca^{2+}, H_2O, SO_4{}^{2-}]$ représenté sur l'axe de droite. La courbe (—) marque la valeur n_{max} défini comme le nombre maximal de ponts de liaisons hydrogène entre les ions Ca^{2+} et SO_4^{2-} et le puits de potentiel.

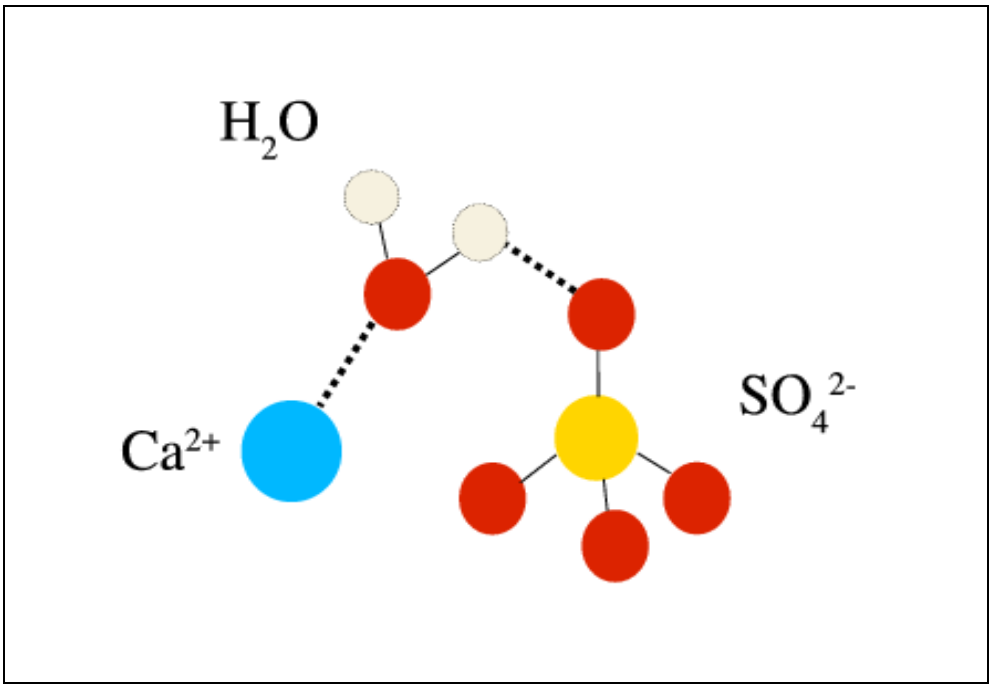

FIG. 5.6 – Visualisation du pont de liaison hydrogène entre les ions Ca^{2+} et SO_4^{2-}.

	r_{min} (Å)	$logK$	$\Delta_r G^o$ (kJ mol^{-1})	$\Delta_r H^o$ (kJ mol^{-1})	$T\Delta_r S^o$ (kJ mol^{-1})
FCFS	2.8	4.6	-26.2	0.003	-26.2
FCS	2.9	1.9	-11.0	0.3	11.3
FEP^m	2.9	1.5	-8.5	4.0	12.5
Exp. [55]		2.4	-13.7	3.4	17.1

TAB. 5.3 – Grandeurs thermodynamiques d'association entre les ions Ca^{2+} et SO_4^{2-} en phase aqueuse. Les trois méthodes étudiées sont représentées. Les résultat de la méthode FEP n'a pas été choisi à cause des dérives énergétiques importantes. (Exp.) indique la valeur expérimentale.

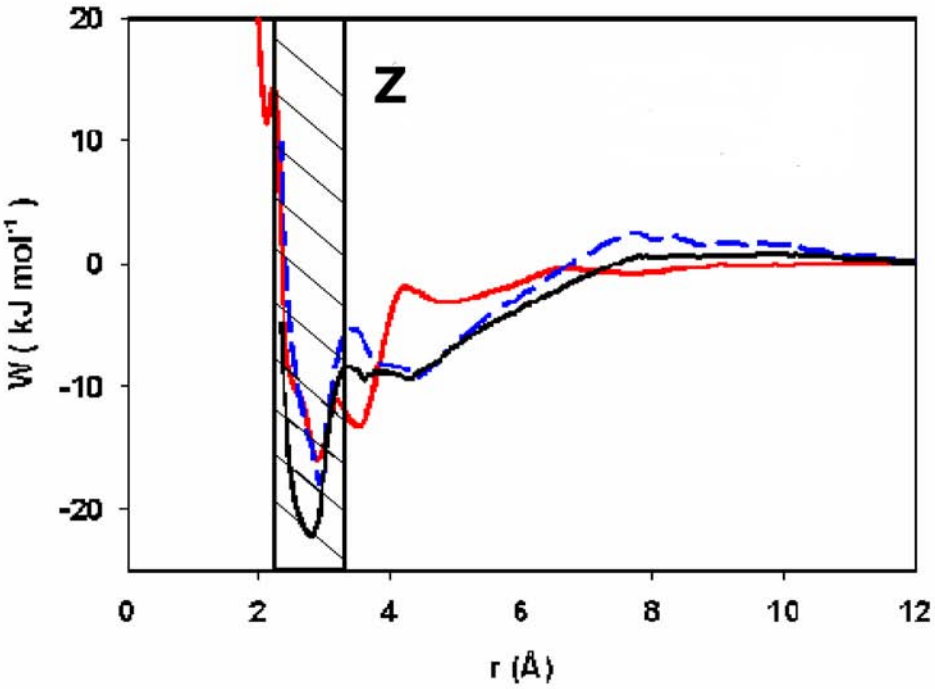

FIG. 5.7 – Profils de potentiel de force moyenne pour la paire d'ions $[Ca^{2+} \cdots SO_4^{2-}]$ en milieu solvaté. Nous avons représenté les trois méthodes étudiées. (—) représente la méthode FCFS. (—) représente la méthode FEP^m. (—) représente la méthode FCS. La zone Z indique la région permettant de différencier les méthodes pour le calcul des grandeurs absolues.

il semblerait que l'ensemble des méthodes FCS, FCFS et FEP donne des résultats équivalents pour le positionnement de la zone où se situe le minimum énergétique. Par contre, l'existence d'une discrimination pour le calcul des grandeurs thermodynamiques semble évidente. En effet, les méthodes FEP^m et FCS sont celles qui donnent les meilleurs résultats. Pour la méthode FEP^m, il a fallu choisir une borne supérieure d'intégration qui tient compte des deux puits. Si l'on ne considère que le premier puits de potentiel (zone Z sur la figure (5.7), les résultats obtenus sont moins bons puisqu'on obtient $K = 12.0$ mol^{-1} L, $\Delta G = -6.15$ kJ mol^{-1} et $\Delta S = 15.0$ kJ mol^{-1}, alors que pour la méthode FCS cette zone d'intégration donne de bons résultats. On explique ce phénomène par la forme et la profondeur des puits de potentiel. En effet, pour la méthode FCS, le puits obtenu est plus profond que celui calculé avec la méthode FEP^m. Par conséquent, pour obtenir des résultats comparables à FCS, il faut compenser cette différence d'énergie : celle-ci est compensée si on considère le second puits de potentiel. Nous avons mené une étude concernant la borne supérieure d'intégration. On montre qu'au delà d'une distance limite, le résultat est indépendant de la valeur de cette borne. Cette valeur limite se situe à environ 1 Å du puits de potentiel.

Nous allons, à présent, appliquer les méthodes FCS et FEP^m pour le calcul de grandeurs thermodynamiques d'association impliquant le calixarène et différents substrats dans le cas de complexes d'insertion et à sphère externe. Nous appliquerons la méthodologie pour des substrats n'impliquant aucune complexation avec le calixarène.

5.4 Potentiel de force moyenne pour différents types de complexes du calixarène

5.4.1 Mise en œuvre

Dans cette partie, nous avons étudié la complexation de trois substrats : l'ion tétraméthylammonium, l'ion lanthane et la molécule de méthanol conduisant à trois associations différentes : complexe d'insertion [6], complexe à sphère externe [5] et absence de complexation [57]. Les simulations ont été réalisées dans l'ensemble NpT. Les configurations ont été générées avec une version modifiée de DL_POLY [27] permettant de calculer le potentiel de force moyenne dans l'ensemble NpT et d'utiliser la méthode FCS. Les algorithmes utilisés sont ceux décrits dans le chapitre (2). Les systèmes sont composés du macrocycle, du substrat, d'ions Na^+ et de 900 molécules d'eau dans une boîte cubique de 30 Å de coté. L'équilibration et l'acquisition sont de 100 ps. L'incrément de distance est 0.044 Å, ce qui correspond à 260 simulations si la distance échantillonnée varie de 0.5 Å à 12 Å, soit au total 52 ns de simulation. On vérifie, en outre, que les ions étudiés sont de forme sphérique, afin de calculer les grandeurs thermodynamiques.

5.4.2 Calcul du PMF

La figure (5.8) représente les profils de potentiel de force moyenne pour les trois substrats étudiés. Cette figure montre un complexe à sphère externe dans le cas de l'ion lanthane (puits localisé à l'extérieur de la cavité)et un complexe d'insertion dans le cas de l'ion tétraméthylammonium (puits localisé à l'intérieur de la cavité). En outre, on remarque que la molécule de méthanol n'est pas complexé car le profil de potentiel de force moyenne n'indique aucun puits. Ces résultats sont en accord avec les hypothèses avancées par Bonal et *al* [1]. Les grandeurs d'association sont rassemblées dans le tableau (5.4). Ce tableau montre la qualité des méthodes utilisées puisqu'on obtient un bon accord avec les valeurs expérimentales. Nous vérifions, une fois de plus, l'indépendance des résultats obtenus par rapport à la valeur de la borne supérieure d'intégration d_l. Pour cela, nous avons représenté sur la figure (5.9) $\Delta_r G^o$ et $T\Delta_r S^o$ en fonction de d_l pour le complexe de l'ion tétraméthylammonium. Cette courbe montre, qu'à partir de $\Delta r = 3.0$ Å, les valeurs de ΔG et $T\Delta S$ convergent. Nous aurions obtenu de meilleurs résultats avec $\Delta r = 1.15$ Å, mais nous avons choisi de prendre une valeur qui donne des résultats convenables pour les deux complexes, et ainsi s'affranchir d'une quel-

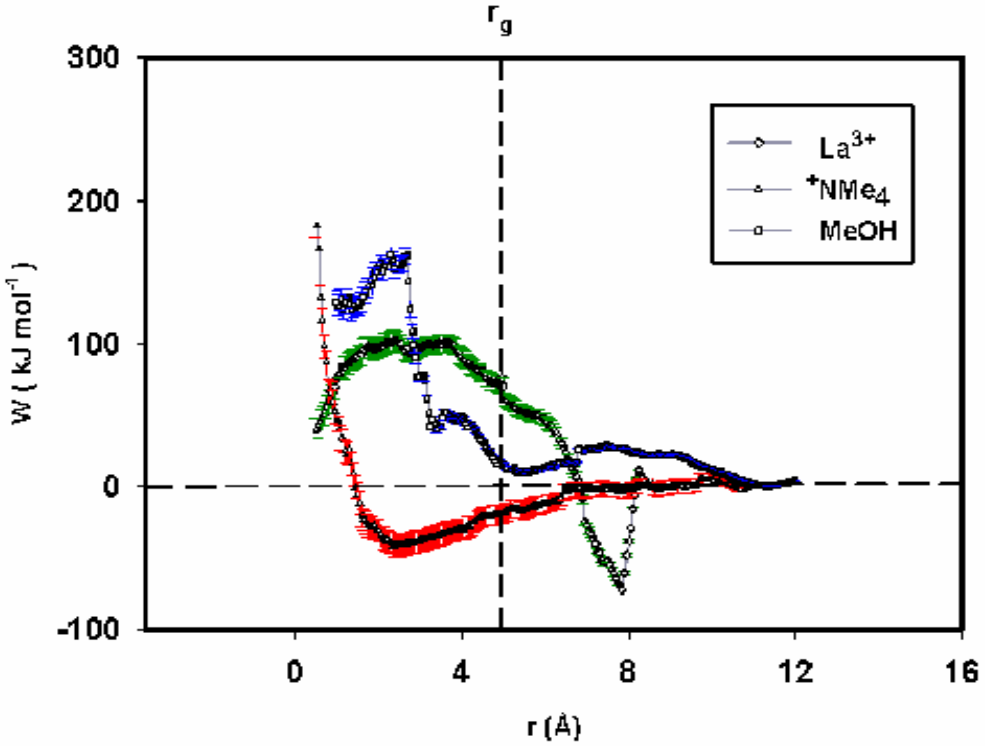

FIG. 5.8 – Profils de potentiel de force moyenne pour les paires d'ions : [Cal^{4-} $\cdots$ La^{3+}] (—), [Cal^{4-} $\cdots$ Me_4N^+] (—) et [Cal^{4-} $\cdots$ Me-OH] (—). r_g représente la valeur du rayon de giration du macrocycle, que nous avons défini dans la section (2.6).

conque dépendance vis-à-vis de Δr. L'obtention des puits de potentiel à ces positions vient valider les simulations déjà effectuées sur ces mêmes systèmes [5, 33, 6, 57] et les conclusions qui ont été émises. Cela confirme que les configurations obtenues dans ces simulations sont réellement des configurations d'équilibre correspondant à un minimum de potentiel pour le complexe.

	r_{min} (Å)	K (mol^{-1} L)	$\Delta_r G^o$ (kJ mol^{-1})	$\Delta_r H^o$ (kJ mol^{-1})	$T\Delta_r S^o$ (kJ mol^{-1})
			FEP^m		
[Cal^{4-} $\cdots$ La^{3+}]	7.85	45985	-26.6	8.1	34.7
[Cal^{4-} $\cdots$ Me_4N^+]	2.35	33296	-25.8	-26.1	-0.3
[Cal^{4-} $\cdots$ Me-OH]	–	0.0	∞	∞	∞
			FCS		
[Cal^{4-} $\cdots$ La^{3+}]	7.92	47221	-26.7	7.6	34.3
[Cal^{4-} $\cdots$ Me_4N^+]	2.35	38129	-26.1	-26.9	-0.8
[Cal^{4-} $\cdots$ Me-OH]	–	0.0	∞	∞	∞
			Exp [1]		
[Cal^{4-} $\cdots$ La^{3+}]		16765	-24.1	9.2	33.3
[Cal^{4-} $\cdots$ Me_4N^+]		25101	-25.1	-26.0	-0.9
[Cal^{4-} $\cdots$ Me-OH]		0.0	∞	∞	∞

TAB. 5.4 – Grandeurs thermodynamiques d'association calculées pour les trois complexes. Ces résultats sont obtenus en utilisant les méthodes FEP^m et FCS. Nous avons utilisé pour ce calcul une valeur de $d_l = (r_{min} + 1)$ Å.

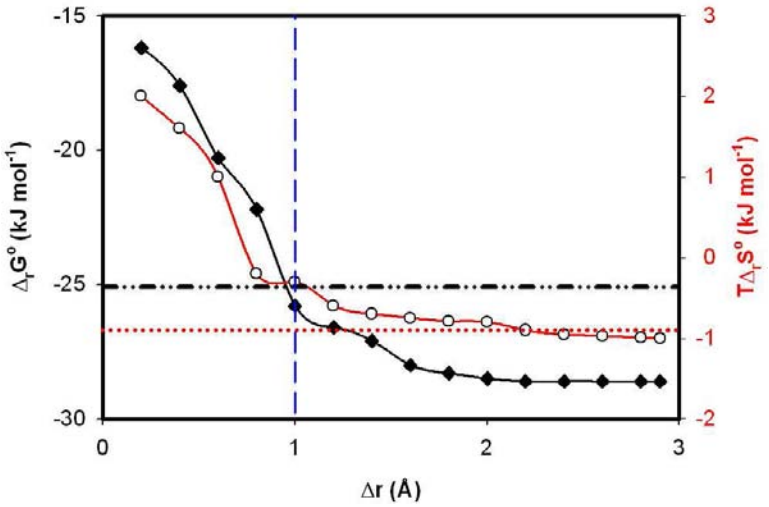

FIG. 5.9 – (—,♦) évolution de $\Delta_r G^o$ représentée sur l'axe de gauche en fonction de Δr tel que $d_l = r_{min} + \Delta\ r$. (—,◯) évolution de $T\Delta_r S^o$ représentée sur l'axe de droite. en fonction de Δr. r_{min} est la position du puits de potentiel obtenue pour la paire [$Cal^{4-} \cdots Me_4N^+$]. La méthode utilisée pour calculer le potentiel de force moyenne est la méthode FEP^m. (···) représente la valeur expérimentale de $T\Delta_r S^o$. (-···-) représente la valeur expérimentale de $\Delta_r G^o$. (- - -) indique la valeur de Δr utilisée pour le calcul des grandeurs d'association.

Étude structurale et énergétique de la complexation de l'ion La^{3+}

Interactions électrostatiques : on montre que le complexe externe formé est un complexe gouverné par les interactions électrostatiques. En effet, au puits de potentiel, l'énergie électrostatique est de -485.8 ± 12.0 kJ mol^{-1} et les interactions de dispersion-répulsion sont de 11.7± 2.1 kJ mol^{-1}. Lorsque l'on trace ces énergies en fonction de la distance contrainte (figure (5.10)), on ne constate aucune correspondance avec la forme du puits de potentiel. En effet, les interactions coulombiennes diminuent en fonction de la distance tandis que l'énergie dispersive diminue et l'énergie répulsive augmente jusqu'à 3 Å puis celles-ci se stabilisent. On peut penser qu'il existe un autre phénomène couplé à ces interactions électrostatiques suceptible de stabiliser davantage le complexe et d'expliquer la forme du profil de potentiel de force moyenne.

Ponts de liaisons hydrogène : dans ce type de complexe, les liaisons hydrogène jouent un rôle fondamental [58] : nous avons donc privilégié cette piste. Nous avons calculé le nombre de

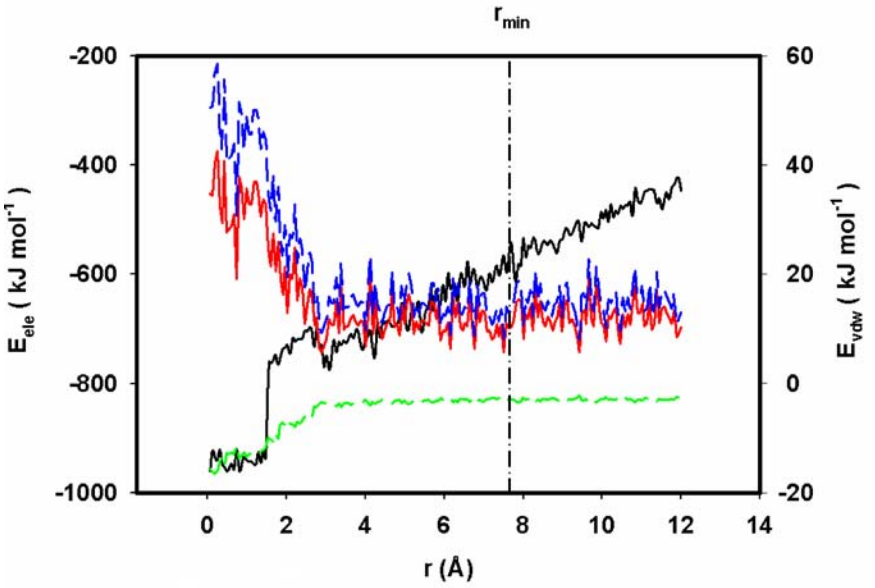

FIG. 5.10 – Interactions électrostatiques (—), dispersives (—), répulsives (—) et de dispersion-répulsion (—) entre le calixarène et l'ion La^{3+} en fonction de la distance séparant leurs centres de masse.

tubes de liaisons hydrogène pour le triplet $[SO_3^- \cdots nH_2O \cdots La^{3+}]$ en fonction de la distance contrainte (figure (5.11)). Ces tubes correspondent en faite un réseaux de liaisons hydrogène. Le résultat est édifiant, puisque la position du nombre maximal de ponts de liaisons hydrogène correspond exactement à la position du puits de potentiel. Pour résumer, le complexe formé par le calixarène et l'ion La^{3+} est un complexe à sphère externe dont la force motrice réside dans l'établissement des interactions électrostatiques entre le calixarène et l'ion La^{3+} et la présence de ponts de liaisons hydrogène pour le triplet $[SO_3^- \cdots H_2O \cdots La^{3+}]$. Par ailleurs, les liaisons hydrogène étant majoritairement d'origine électrostatique, on peut affirmer sans équivoque que le complexe ainsi formé est un complexe gouverné par les interactions électrostatiques.

Phénomènes de désolvatation : il a été montré que des phénomènes de désolvatation étaient à l'origine du contrôle entropique pour ce type de complexe [5]. On montre que la complexation de l'ion La^{3+} entraîne la désolvatation partielle de la cavité. Par contre, la première couche d'hydratation de l'ion La^{3+} reste intacte comme le montre la figure (5.12).

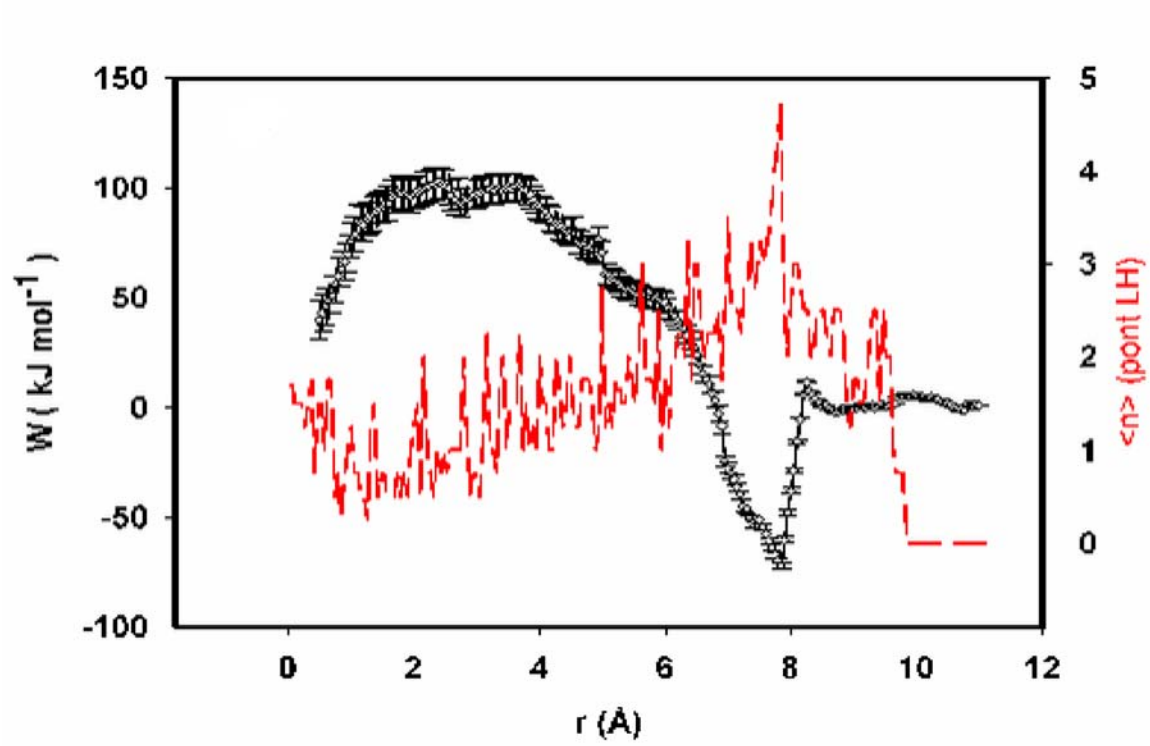

FIG. 5.11 – (—) Profil de potentiel de force moyenne pour la paire [$cal^{4-} \cdots La^{3+}$]. (—) Nombre de ponts de liaisons hydrogène entre les groupements sulfonate, l'ion lanthane et les molécules d'eau (représenté sur l'axe de droite).

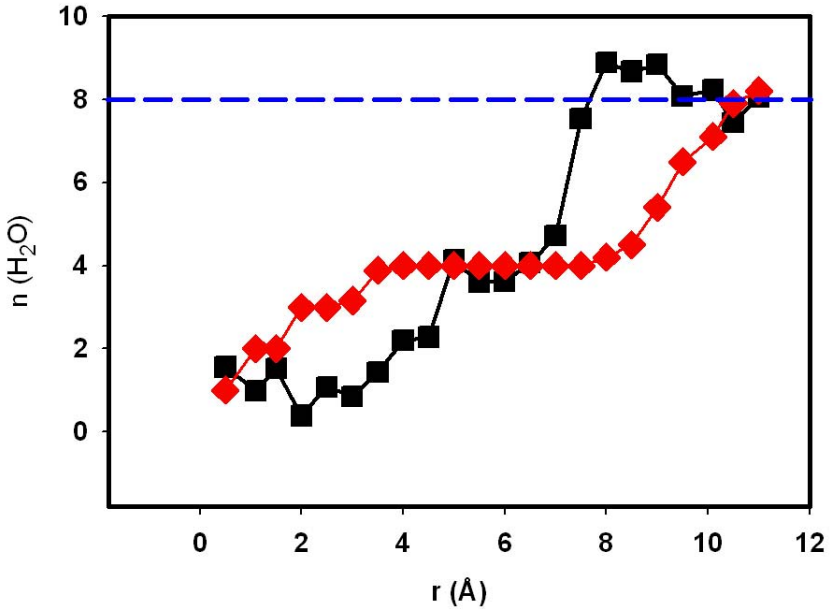

FIG. 5.12 – Nombre de molécules d'eau dans la première couche de solvatation de l'ion La^{3+} (■) et dans la cavité (♦). La courbe (- - -) représente le nombre de molécules d'eau dans la première couche d'hydratation de l'ion lanthane pour l'ion seul en solution (8) et le nombre de molécules d'eau dans la cavité pour la calixarène seul en solution (8).

Étude structurale et énergétique de la complexation de l'ion Me_4N^+

Interactions de dispersion-répulsion L'étude de la complexation de l'ion tétraméthylammonium montre un comportement différent. En effet, il s'agit d'un complexe d'insertion où la nature hydrophobe de l'ion Me_4N^+ empêche la formation de ponts de liaisons hydrogène. Étant donnée la nature du complexe formé, les interactions de dispersion-répulsion ont un rôle prépondérant dans le formation de ce complexe [5, 6]. Ainsi, lorsqu'on trace l'énergie coulombienne et l'énergie de dispersion-répulsion entre le macrocycle et l'ion tétraméthylammonium en fonction de la distance séparant ces deux ions, le profil de l'énergie de dispersion-répulsion a une forme superposable au profil de potentiel de force moyenne (figure (5.13)). Cela implique que les interactions de type dispersion-répulsion peuvent être considérées comme les interactions qui contrôlent la formation du complexe d'insertion.

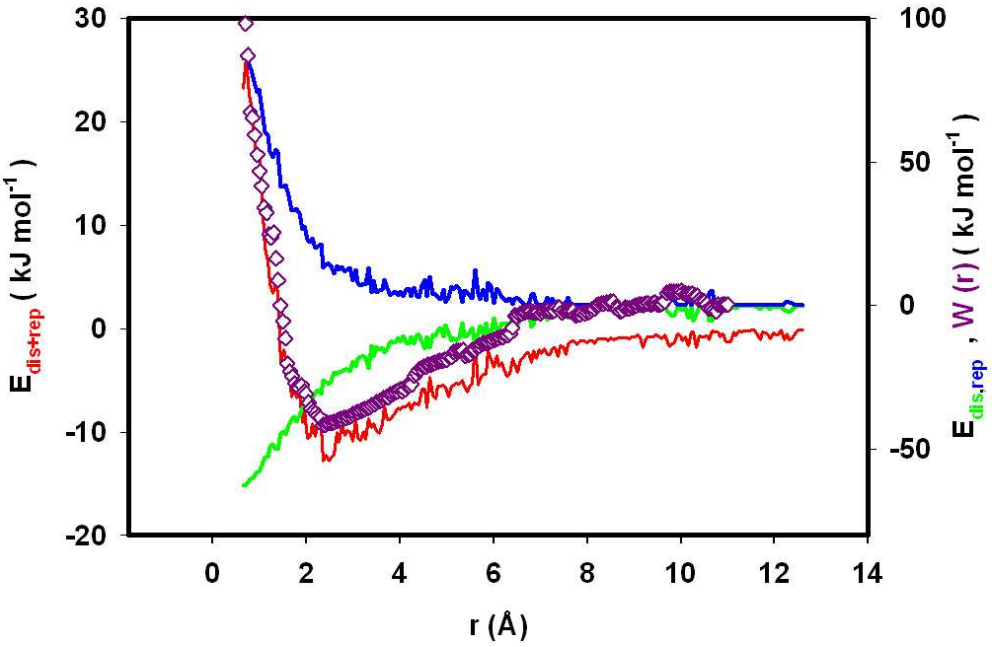

FIG. 5.13 – Interactions entre le calixarène et l'ion tétraméthylammonium en fonction de la distance séparant les centres de masse du macrocycle et du substrat. Les interactions dispersives (—) sont représentées sur l'axe de droite. Les interactions répulsives (—) sont représentées sur l'axe de droite. Les interactions de dispersion-répulsion (—) sont représentées sur l'axe de gauche. (◇) Potentiel de force moyenne entre le macrocycle et l'ion tétraméthylammonium représenté sur le premier axe de droite.

Phénomènes de désolvatation : nous constatons que la cavité est totalement désolvatée en présence de l'ion tétraméthylammonium. Dans le cas de l'ion La^{3+}, cette désolvatation n'est que partielle. On peut corréler cette désolvatation aux variations d'entropie. La variation d'entropie est moins favorable dans le cas de la complexation de l'ion Me_4N^+ par rapport à l'ion lanthane (*cf* tableau (5.4)). Il faut prendre en considération la perte de degré de liberté de l'ion Me_4N^+, qui est entropiquement défavorable. En effet, l'entropie défavorable liée à l'insertion du cation est compensée par l'entropie favorable liée à la désolvatation de la cavité. Comme l'ion La^{3+} forme un complexe de sphère externe, la perte de degré de liberté est beaucoup plus faible. Par conséquent, cette perte de degré de liberté ne vient pas compenser la désolvatation partielle de la cavité. Cette absence de compensation permet en partie d'expliquer l'entropie favorable dans le cas de l'ion La^{3+}.

Pourquoi le méthanol n'est-il pas complexé ?

Le potentiel de force moyenne entre le calixarène et le méthanol a été calculé pour comprendre l'absence de complexation observée expérimentalement [4], et confirmer les résultats des simulations précédentes (section (3.4)). La figure (5.8) montre que le méthanol n'est pas complexé. Nous avons expliqué [57] ce comportement en montrant que le coût énergétique qu'engendrerait la désolvatation du méthanol dans le cas d'une complexation ne serait pas compensée par les interactions entre le macrocycle et le méthanol. La figure (5.14.a) montre que, dans le cas de la formation d'un complexe d'insertion (c'est-à-dire une distance entre les centres de masse inférieure à 2.5 Å), la première et la deuxième sphère d'hydratation sont partiellement désolvatées. Or, le coût énergétique engendré par cette désolvatation est trop important pour être compensé par les interactions de dispersion-répulsion entre les cycles benzéniques et le groupement méthyle du méthanol. La figure (5.14.b) montre l'évolution des interactions entre le méthanol et le calixarène en fonction de la distance de contrainte. Les interactions entre le méthanol et les molécules d'eau de la première couche de solvatation de Me-OH sont également représentées en fonction de la distance de contrainte. Cette dernière courbe montre que l'énergie nécessaire à la désolvatation de la première couche ne peut être compensée par les interactions entre le macrocycle et la molécule de méthanol.

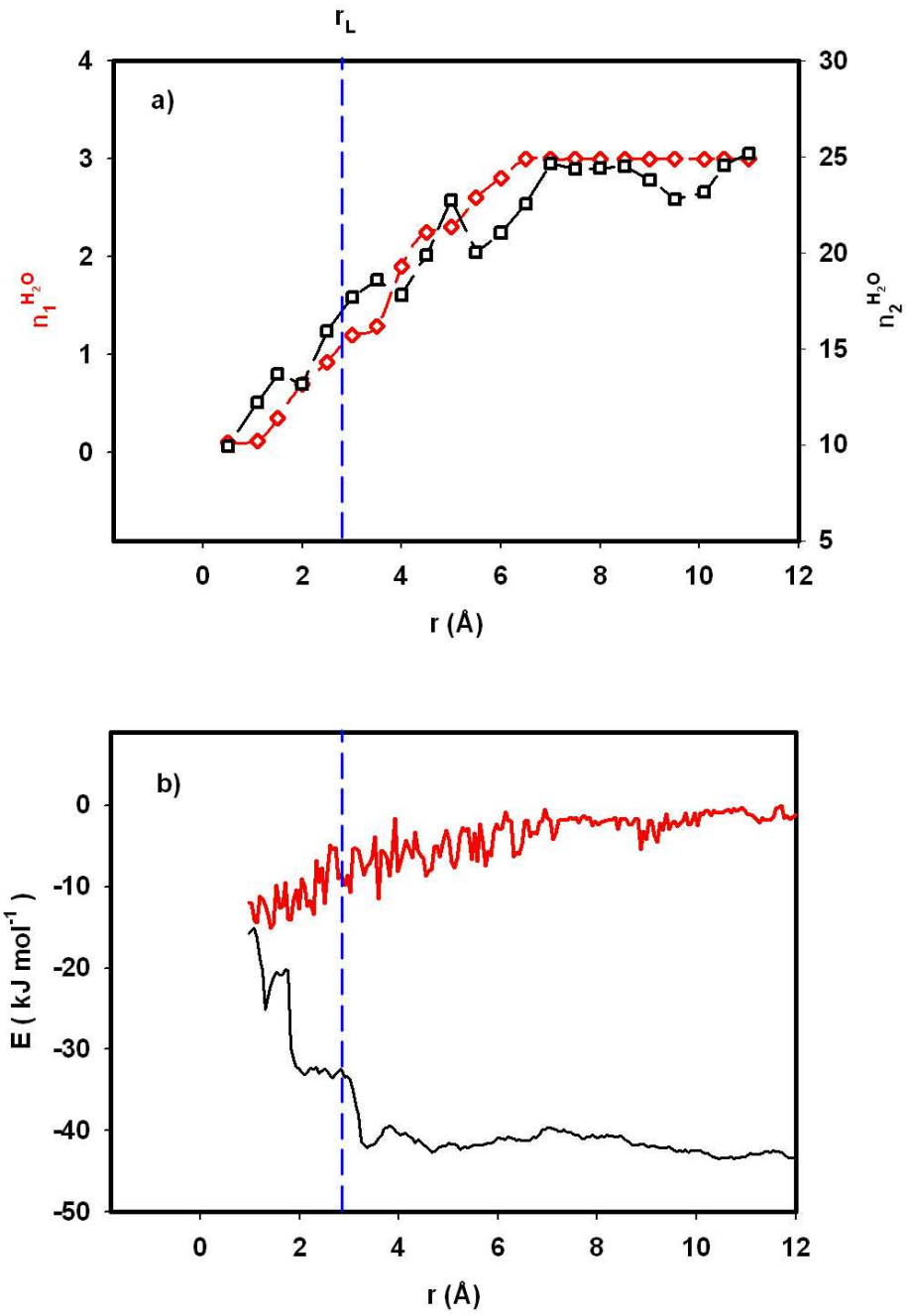

FIG. 5.14 – a) Évolution du nombre de molécules d'eau dans la première et la deuxième couche de solvatation du méthanol en fonction de la distance séparant les centres de masses du substrat et du macrocycle. La courbe (- - -) représente la distance limite d'insertion pour une molécule de cette taille. (◊) indique le nombre de molécules d'eau dans la première couche de solvatation, représenté sur l'axe de gauche. (□) indique le nombre de molécules d'eau dans la deuxième couche de solvatation, représenté sur l'axe de droite. b) Évolution des interactions entre le méthanol et le macrocycle (—). Évolution des interactions entre le méthanol et les molécules d'eau constituant la première couche de solvatation (—).

5.5 Résumé

Méthodologie

Ce travail nous a permis d'explorer les méthodes capables de déterminer les grandeurs thermodynamiques absolues. C'est Prue [59] en 1969 qui a mis au point une méthode permettant de déterminer la constante d'association d'une paire d'ions. Nous avons étendu cette méthode à la détermination des trois fonctions thermodynamiques $\Delta_r G^o$, $\Delta_r H^o$ et $\Delta_r S^o$. En effet, cette méthode est essentiellement basée sur la détermination du potentiel de force moyenne (introduite pour la première fois par Kirkwood [54] en 1935). La précision du calcul dépend de la précision avec laquelle on détermine le profil de potentiel de force moyenne.

La détermination du potentiel de force moyenne peut se faire de plusieurs façons. La méthode de perturbation et la méthode de la force contrainte sont celles que nous avons étudiées. La première est une application de la méthode FEP. Dans la seconde, on évalue la force contrainte entre les deux espèces. La contrainte de distance étant maintenue par l'algorithme SHAKE, il suffit d'évaluer le viriel de contrainte par la méthode des multiplicateurs de Lagrange.

Ces deux méthodes présentent deux lacunes si l'on considère les approximations usuelles. En effet, si on déplace les atomes constituant les deux espèces de la paire étudiée, il est possible de trop s'approcher d'une molécule de solvant, ce qui entraînera irrémédiablement une dérive énergétique. La seconde lacune réside dans la détermination des multiplicateurs de Lagrange. En effet, on linéarise en général le binôme obtenu par cette méthode. Or, on montre qu'il existe une différence non négligeable si l'on résoud rigoureusement l'équation du second degré.

Pour contourner ces problèmes, nous avons, dans le premier cas, utilisé la méthode mise au point dans le chapitre (4) en recalculant l'ensemble des positions et des vitesses à l'état perturbé (FEP^m). Ce calcul a l'avantage d'éviter les problèmes de dérives énergétiques. Dans le deuxième cas, nous avons résolu rigoureusement l'équation obtenue par la méthodes des multiplicateurs de Lagrange.

Application

Pour la complexation des ions La^{3+} et Me_4N^+ par le calixarène, nous avons obtenu des résultats satisfaisants puisque les grandeurs calculées sont en accord avec les grandeurs expérimentales. Le bon accord entre les valeurs expérimentales et calculées a permis de valider la position du puits de potentiel. Cette étude valide ainsi les simulations moléculaires classiques (sans contrainte de distance) et les conclusions émises auparavant [5, 33, 6, 57], puisque la position du puits de potentiel correspond à la distance moyenne entre le macrocycle et le substrat calculée dans des simulations classiques.

Le tableau suivant résume les caractéristiques essentielles des deux complexes étudiés. Ce tableau montre que seule l'évaluation de $\Delta_r S^o$ et $\Delta_r H^o$ permet de discriminer les deux processus. Ainsi, nous avons montré que pour un complexe de sphère externe, ce sont les interactions électrostatiques qui prédominent. En plus des interactions macrocycle-substrat, il faut tenir compte des ponts de liaisons hydrogène entre les ions La^{3+} et SO_3^- et les molécules d'eau qui participent activement à la stabilisation du complexe. Pour un complexe d'insertion, les interactions de dispersion-répulsion sont prédominantes. En outre, l'étude de la solvatation de la cavité montre une désolvatation totale de la cavité dans le cas de l'ion Me_4N^+ et partielle dans le cas de l'ion La^{3+}.

	$[cal^{4-} \cdots La^{3+}]$	$[cal^{4-} \cdots Me_4N^+]$
Contrôle	entropique	enthalpique
Type	sphère externe	d'insertion
Interactions	électrostatique	dispersion-répulsion
Désolvatation^{cav}	partielle	totale
Liaisons Hydrogène	oui	non

TAB. 5.5 – Résumé des caractéristiques des complexes formés. (^{cav}) indique la cavité du calixarène.

Enfin, la détermination du potentiel de force moyenne entre le méthanol et le calixarène a permis de démontrer l'absence de complexation. On montre que le coût énergétique de la désolvatation de la première couche d'hydratation ne peut être compensé par les interactions entre le calixarène et le méthanol.

Chapitre 6

Méthodologie pour la prise en compte des interactions cations$-\pi$

6.1 Introduction et problématique

Les interactions non-covalentes jouent un rôle important en reconnaissance moléculaire et en chimie supramoléculaire. Contrairement aux interactions plus conventionnelles telles que les liaisons hydrogène ou les interactions de van der Waals, les interactions cation$-\pi$ ont été sous estimées et donc peu étudiées. Ce type d'interaction met en jeu des électrons π délocalisés d'un cycle benzénique et d'un cation. Dougherty *et al* [60] ont montré l'importance de ces interactions dans les phénomènes de reconnaissance moléculaire. Néanmoins, la majorité des études effectuées pour la compréhension de ces interactions a été menée en phase gazeuse. Cepandant, les processus que l'on étudie ont lieu en milieu aqueux, où les molécules d'eau sont en compétition avec le cation pour établir ces interactions cation-π. Récemment, Morel *et al* [61] ont mesuré les grandeurs d'association entre le calix[4]arène-*p*-sulfonaté et l'ion césium en solution aqueuse à 25^oC. La force motrice de cette association est l'interaction cation-π entre les cycles benzéniques et l'ion Cs^+. Ce résultat a été montré par Dougherty [60] en phase gazeuse. Il est à noter que ce type d'interaction n'a jamais été mis en évidence dans l'eau. Ainsi, il nous est apparu intéressant d'étudier ce système au niveau moléculaire par le biais de simulations numériques et d'essayer de calculer les grandeurs d'association par un calcul de potentiel de force moyenne.

Jusqu'à présent, nous avons étudié des complexes mettant en jeu des interactions conventionnelles de type dispersion-répulsion et électrostatiques. Nous avons vu dans le premier chapitre que les méthodes de dynamique moléculaire ne traitent pas explicitement les interactions électroniques. Or, ces types d'interactions interviennent dans de nombreux processus

physico-chimique. Ainsi, si l'on veut tenir compte de ces interactions, il est nécessaire d'utiliser des méthodes complémentaires qui permettent la prise en compte des électrons. Parmi celles-ci, nous trouvons les méthodes de dynamique moléculaire quantique et les méthodes mixtes QM/MM. Ces méthodes sont limitées à des systèmes de petites tailles pour ne pas calculer un nombre important d'intégrales multiélectroniques.

Dans ces conditions, il nous est apparu nécessaire de développer une méthodologie associant des simulations quantiques et moléculaires pour essayer d'accéder aux grandeurs thermodynamiques. Notre objectif n'est pas d'étudier d'un point de vue structural le complexe, mais de développer une méthode de calcul de potentiel force moyenne utilisant la dynamique moléculaire classique et tenant compte des interactions cation-π. Cette entreprise a pu être réalisée grâce au concours de P. Archirel, A .Boutin et B. Lêvy du Laboratoire de Chimie Physique, UMR CNRS 8000, de l'Université Paris-Sud.

On présentera dans ce chapitre la méthodologie utilisée pour tenir compte des interactions entre le cation monovalent Cs^+ et les électrons π des cycles benzéniques. Ensuite, on développera l'aspect pratique de la méthode ainsi que le calcul du potentiel de force moyenne. On terminera par la présentation des résultats.

6.2 Méthodologie

Ces interactions cation-π ont été très largement sous-estimées et ce n'est que depuis une quinzaine d'années, avec l'avènement de la chimie supramoléculaire, que celles-ci ont commencé à être étudiées avec un grand interêt. L'un des pionniers dans la compréhension de ces interactions est Dougherty [60]. En effet, en s'appuyant sur des résultats de calculs quantiques et expérimentaux, en phase gazeuse, il a suggéré un modèle physique pour ces interactions cation-π. Ces études n'ont pas été réalisées en milieux aqueux (milieu naturel pour les processus biologiques) mais en phase gazeuse, afin d'éviter toute interaction entre les molécules d'eau et le cycle benzénique. En effet, une molécule d'eau peut être en compétition avec le cation pour établir ce type d'interactions. Néanmoins, il utilise les travaux de Kebarle [62] pour comparer les valeurs des énergies obtenues à celles calculées pour des systèmes de référence (eau$\cdots$cation) et (eau$\cdots$molécule). L'étude en phase gazeuse est intéressante dans la

mesure où on peut déterminer avec exactitude par une décomposition de Kitaura-Morokuma le poids relatif de chaque type d'interaction existant entre les deux espèces étudiées (énergie électrostatique, énergie de répulsion, énergie de polarisation et énergie de transfert de charge). Ainsi, il a montré que l'interaction cation-π est à 60 % d'origine électrostatique (en accord avec les travaux de Kollman [63] concernant les complexes arène-Na^+). Les interactions non-électrostatiques (40 %) correspondent essentiellement à un terme de polarisation et de transfert de charge. En outre, il exclut toute interaction ion-quadrupôle en prétextant l'établissement d'une distance trop élevée en comparaison de celle observée pour des interactions cation-π (5 Å contre 2.5 Å). Afin de vérifier ces conclusions, nous avons effectué une décomposition de Kitaura-Morokuma sur le système [$C_6H_6 \cdots Cs^+$] (unité structurale partielle du complexe étudié) sur une base HF 6-31 G** avec correction BSSE. Nous obtenons la décomposition énergétique suivante :[1] E^{Ele}= -39.4 kJ mol^{-1} ,E^{Rep}= 11.8 kJ mol^{-1} ,E^{Pol}= -15.4 kJ mol^{-1}, E^{TC}= -12.0 kJ mol^{-1}, E^{CE}= 4.7 kJ mol^{-1} soit E^{Tot}= -50.3 kJ mol^{-1}. Les termes de polarisations et de transferts de charge représentent 32.9 % de l'énergie totale contre 47.3 % pour l'énergie électrostatique. En d'autres termes, les contributions non-électrostatiques sont loin d'être négligeables. Ce calcul nous permet de corroborer les résultats obtenus par Dougherty et *al* [60] pour d'autres cations.

L'étude en phase aqueuse nécessite un calcul de décomposition supplémentaire pour les couples [$C_6H_6 \cdots H_2O$], afin d'évaluer une éventuelle compétition entre l'eau et le cation comme le suggère Dougherty. Les résultats de ce calcul sont rassemblés dans le tableau (6.1). Ce tableau montre la présence d'une compétition entre l'eau et l'ion césium, puisque les énergies de transfert de charge et de polarisation sont du même ordre de grandeur. Ainsi, dans un milieu solvaté, il est difficile de mettre en évidence les interactions cation-π. Ceci explique pourquoi il existe très peu de données sur les systèmes cation-π en mileu aqueux d'un point de vue expérimental et théorique. Si l'on veut tenir compte de ces interactions dans des simulations moléculaires, il est nécessaire d'utiliser un champ de forces susceptible de rendre compte des phénomènes de polarisations et de transferts de charge. À notre connaissance, seuls les champs de forces de deuxième génération sont capables de prendre en compte les phénomènes de polarisation (coûteux en temps de calculs), mais pas les transferts de charge. Par contre,

[1] Ele=électrostatique, Rep=répulsion, Pol=polarisation,TC=transfert de charge, CE=énergie de corrélation

	$[C_6H_6 \cdots Cs^+]$ (kJ mol^{-1})	$[C_6H_6 \cdots H_2O]$ (kJ mol^{-1})
E^{Ele}	-39.4	-42.1
E^{Rep}	-11.8	-9.1
E^{Pol}	-15.4	-14.2
E^{TC}	-12.0	-10.1
E^{CE}	4.7	5.1
E^{Tot}	-50.3	-70.4

TAB. 6.1 – Énergie issue de la décomposition de Kitaura-Morokuma dans une base HF 6.31 G**. Les significations des abréviations sont : Ele=électrostatique, Rep=Répulsion, Pol=Polarisation,TC=Transfert de charge, CE=Énergie de corrélation.

des simulations de type dynamique quantique telles que la méthode CPMD (*Car Parinello Molecular Dynamic*) sont capables de prendre en compte ces interactions. Ces méthodes sont cependant très couteuses en temps de calcul et sont limitées à un nombre restreint d'électrons. Elles ne sont donc pas envisageables dans notre cas, puisque notre but est d'accéder à des grandeurs d'association qui requièrent un échantillonnage important des configurations.

Afin de tenir compte des interactions cation-π dans nos simulations de type dynamique moléculaire, il nous est apparu judicieux de concevoir un potentiel intermoléculaire analytique prenant en compte à la fois les interactions entre l'ion césium et le calixarène ($I_{[Cal^{4-}\cdots Cs^+]}$), les interactions entre l'ion césium et un cluster d'eau ($I_{[H_2O\cdots Cs^+]}$) , les interactions entre le calixarène et le cluster d'eau ($I_{[Cal^{4-}\cdots H_2O]}$) et enfin les interactions entre molécules d'eau ($I_{[H_2O\cdots H_2O]}$) dans le cluster. En effet, les simulations de dynamique moléculaire avec l'utilisation d'un champ de forces classique comme celui d'AMBER ne permettent pas de simuler ces interactions cation-π, par contre, les interactions non covalentes conventionnelles sont modélisées convenablement. L'utilisation d'une simulation de ce type permet d'obtenir une structure de solvatation en accord avec la littérature pour le césium. Ainsi, pour une distance fixée entre

l'ion césium et le centre de masse du calixarène, on peut obtenir une configuration équilibrée en dynamique moléculaire, par le biais d'une simulation de potentiel de force moyenne. Si on applique à une partie de cette configuration (cluster d'eau + césium + calixarène) un calcul d'énergie quantique, l'énergie obtenue comprendra nécessairement les effets cation-π. Ces effets ne seront présents que pour des distances suffisamment proches entre l'ion césium et le calixarène. Un calcul d'optimisation énergétique est à proscrire puisque la distance doit absolument être contrainte si on veut obtenir un profil d'énergie libre. Ainsi, si on veut déterminer le potentiel de force moyenne entre le calixarène et l'ion césium, il faut corriger les calculs moléculaires par un calcul quantique. La méthodologie permettant cette correction est la suivante :

1 Détermination du potentiel de force moyenne par le biais de simulations moléculaires classiques
2 choix de l'environnement à traiter en mécanique quantique
3 calcul de l'énergie quantique E^{QM} et de l'énergie moléculaire E^{MM}
4 détermination de la variation d'énergie $\Delta E = E^{QM} - E^{MM}$
5 ajustement de ΔE par une fonction appropriée ΔE^{Fit}
6 correction du potentiel de force moyenne moléculaire
7 détermination des grandeurs thermodynamiques d'association.

Cette méthode nous permet d'obtenir un potentiel global tenant compte de l'ensemble des interactions intermoléculaire. Ce potentiel ne peut pas se substituer au potentiel classique utilisé en dynamique moléculaire qui est caractérisé par des interactions de paires. Finalement, cette méthodologie est spécifique au calcul de potentiel de force moyenne. La différence entre la partie moléculaire et la partie quantique rend compte des interactions qui ne sont pas prises en compte par le champ de forces utilisé. Cette partie «résiduelle» peut traduire des effets quantiques et non pas exclusivement des interactions cation-π.

6.2.1 Description de la méthode

Système

La procédure permettant de déterminer le potentiel de force moyenne a été exposée dans le chapitre précédent. Chaque simulation est caractérisée par une distance contrainte (d_c) entre

le centre de masse du calixarène et l'ion Cs^+. Nous considérons pour chaque simulation la dernière configuration, que l'on estime être à l'équilibre. Pour chacune d'elle, nous devons alors déterminer la région de l'espace des phases qui sera traitée par un calcul de mécanique quantique. Pour cela, nous utilisons comme critère de sélection le rayon de la première sphère de solvatation de l'ion césium et du centre de masse des groupements sulfonate. La figure (6.1.b) représente les distributions radiales du centre de masse des groupements sulfonate et de l'ion césium seul en solution et en présence du macrocycle. Cette figure montre que les rayons de la première sphère de solvatation de l'ion Cs^+ et du centre de masse de $[SO_3^-]$ sont identiques. Ainsi, pour les courtes distances, il suffit de considérer les molécules d'eau à l'intérieur de la sphère dont le centre est le centre de masse du système [calixarène + césium] et de rayon égal au rayon de la première sphère de solvatation des groupements sulfonate déterminé à partir des distributions radiales représentées sur la figure (6.1.b). La figure (6.1.a) permet de visualiser un cluster d'eau aux courtes distances. Aux longues distances, on ne pourra plus utiliser ce critère sphérique puisqu'on risque d'avoir une discontinuité (figure (6.2)). Pour remédier à ce problème, on utilise plutôt un critère elliptique.

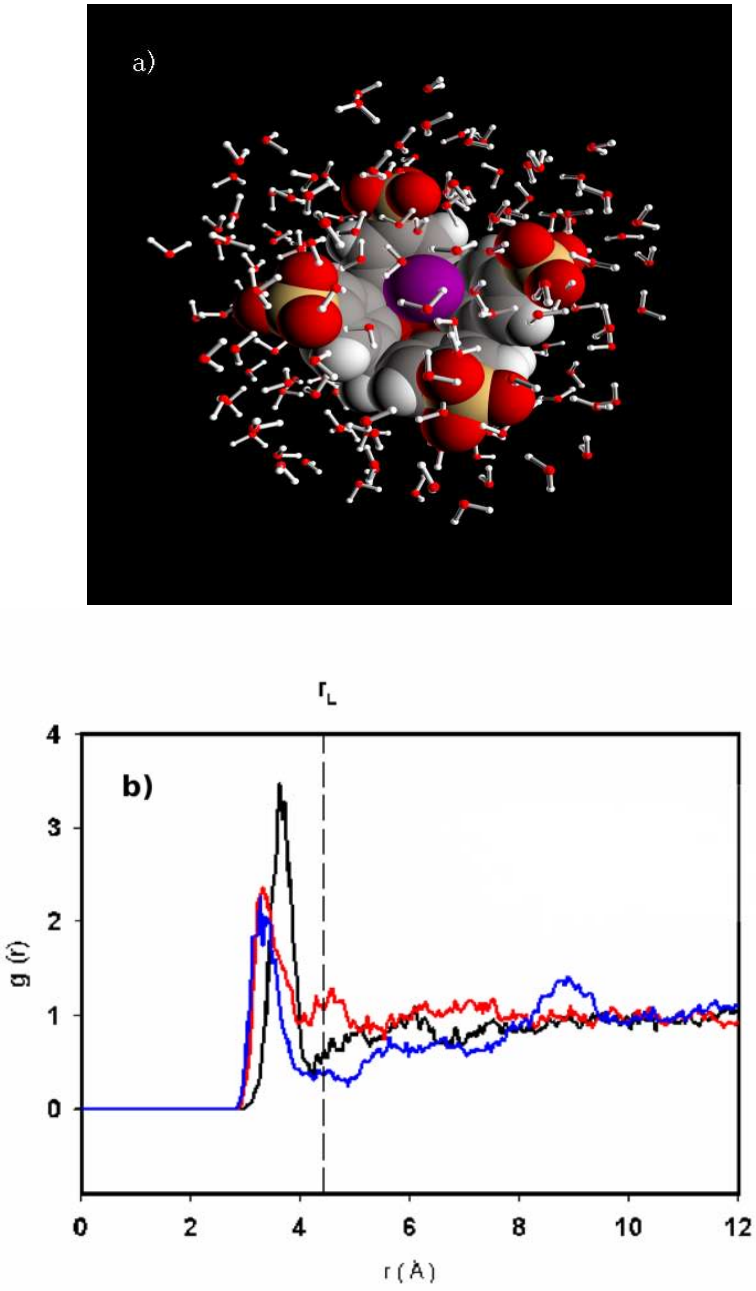

FIG. 6.1 – a) Visualisation du complexe dans le cluster d'eau. b) (—) Distribution radiale entre le centre de masse des groupements sulfonate et les atomes d'oxygène de l'eau. (—) Distribution radiale entre l'ion césium et les atomes d'oxygène de l'eau en présence du macrocycle et seul en solution (—). Ces distributions radiales correspondent au cas où la distance entre le calixarène et l'ion Cs^+ est de 2.0 Å. r_L indique la borne supérieure que l'on utilise dans notre critère de sélection.

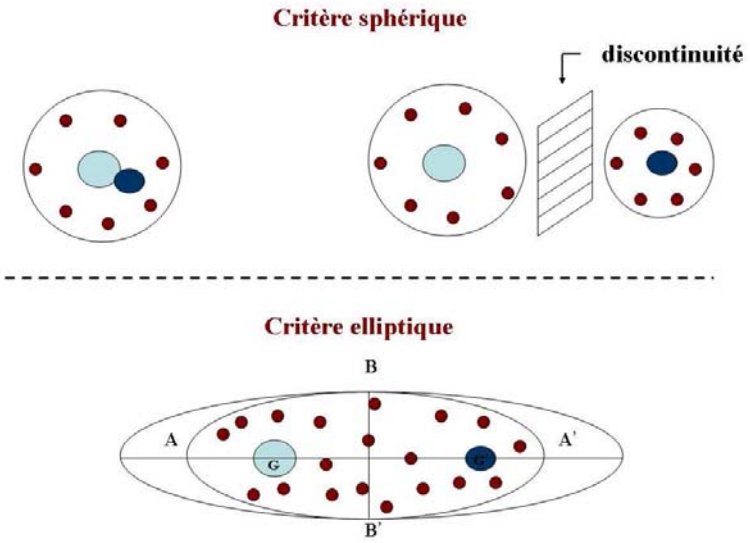

FIG. 6.2 – Visualisation des deux critères choisis pour déterminer l'ensemble qui sera traité en mécanique quantique. (•) et (•) représentent les centres de masse du calixarène et de l'ion Cs^+. (•) représente les molécules d'eau. $AA^{'} = r^{c} + r^{Cal^{4-}} + r^{Cs^{+}}$ représente le grand axe de l'ellipse ; $BB^{'} = r^{Cal^{4-}} + r^{Cs^{+}}$ représente le petit axe ; r est la dstance focale ; G et $G^{'}$ sont les deux foyers de l'ellipse correspondant aux centres de masse du calixarène et de l'ion Cs^+.

Une fois cet ensemble déterminé, on peut alors choisir les paramètres nécessaires aux calculs quantiques. Étant donné le nombre de molécules à traiter, un calcul DFT sera privilégié par rapport à un calcul RHF. Nous revenons sur la théorie de la DFT et ses avantages dans le chapitre (10).

Calcul du terme correctif

L'énergie quantique calculée par DFT comprend les termes d'interaction suivants :

$$\begin{aligned} E^{QM} &= E^{QM}_{[Cal^{4-}\cdots Cs^{+}]} + E^{QM}_{[Cs^{+}\cdots H_2O]} + E^{QM}_{[Cal^{4-}\cdots H_2O]} + E^{QM}_{[H_2O\cdots H_2O]} \\ &+ {}^{0}E^{QM}_{H_2O} + {}^{0}E^{QM}_{Cs^{+}} + {}^{0}E^{QM}_{Cal^{4-}} + {}^{'}E^{QM}_{H_2O} + {}^{'}E^{QM}_{Cs^{+}} + {}^{'}E^{QM}_{Cal^{4-}} \end{aligned} \tag{6.1}$$

Les six dernièrs termes correspondent à des énergies quantiques intramoléculaires. Les énergies ${}^{0}E^{QM}$ représentent les énergies des molécules seules en phase gazeuse sans interaction intermoléculaire, et ${}^{'}E^{QM}$ les termes relatifs à tous les autres effets quantiques (c'est-à-dire

les phènomènes de polarisation, de transfert de charge, etc) liés à l'environnement chimique des molécules. L'énergie intramoléculaire du césium en dynamique moléculaire étant nulle, on a :

$$E^{MM} = E^{MM}_{[Cal^{4-}\cdots Cs^{+}]} + E^{MM}_{[Cs^{+}\cdots H_2O]} + E^{MM}_{[Cal^{4-}\cdots H_2O]} + E^{MM}_{[H_2O\cdots H_2O]} + E^{MM}_{[H_2O]} + E^{MM}_{[Cal^{4-}]} \quad (6.2)$$

Dans le champ de forces utilisé $E^{MM}_{[H_2O]} = 0$. Afin d'obtenir un potentiel intermoléculaire, nous allons soustraire les parties intramoléculaires dans les expressions des relations (6.1) et (6.2). Un calcul préalable de tous les termes ${}^0E^{QM}$ est alors nécessaire. On obtient alors une énergie E_2^{QM} et E_2^{MM}telle que :

$$\begin{aligned} E_2^{QM} &= E^{QM} - \left({}^0E^{QM}_{H_2O} + {}^0E^{QM}_{Cs^{+}} + {}^0E^{QM}_{Cal^{4-}}\right) && (6.3)\\ E_2^{MM} &= E^{MM}_{[Cal^{4-}\cdots Cs^{+}]} + E^{MM}_{[Cs^{+}\cdots H_2O]} + E^{MM}_{[Cal^{4-}\cdots H_2O]} + E^{MM}_{[H_2O\cdots H_2O]} && (6.4) \end{aligned}$$

Ainsi, la différence énergétique que l'on peut ajuster ne contiendra que des termes intermoléculaires et des termes ${}'E^{QM}$ que l'on considére comme des termes intermoléculaires qui représentent les termes de polarisation, transfert de charge, etc.

$$\begin{aligned} E^{Fit} &= E_2^{QM} - E_2^{MM} = E^{QM}_{[Cal^{4-}\cdots Cs^{+}]} + E^{QM}_{[Cs^{+}\cdots H_2O]} + E^{QM}_{[Cal^{4-}\cdots H_2O]} + E^{QM}_{[H_2O\cdots H_2O]}\\ &- \left(E^{MM}{}_{[Cal^{4-}\cdots Cs^{+}]} + E^{MM}{}_{[Cs^{+}\cdots H_2O]} + E^{MM}{}_{[Cal^{4-}\cdots H_2O]} + E^{MM}{}_{[H_2O\cdots H_2O]}\right)\\ &+ {}'E^{QM}_{H_2O} + {}'E^{QM}_{Cs^{+}} + {}'E^{QM}_{cal^{4-}} \quad (6.5) \end{aligned}$$

Le terme d'ajustement apparaît comme un terme de correction à l'énergie intermoléculaire $E^{MM}_{corr} = E^{Fit} + E^{MM}$. Comme cet ajustement est obtenu par une différence énergétique de plusieurs configurations qui ne diffèrent que par la distance contrainte d_c entre le centre de masse du calixarène et celui de l'ion césium, on peut écrire :

$$E^{MM}_{corr}(d_c) = E^{Fit} + E^{MM} \quad (6.6)$$

Par souci d'homogénéité avec le reste du document, on préfèrera écrire :

$$\mathcal{U}_{corr}^{Inter}(d_c) = E^{Fit} + E^{MM} \tag{6.7}$$

Comme on l'a vu précédement, nous utilisons un critère elliptique pour éviter tout problème de discontinuité. Cependant, le nombre de molécules d'eau choisi n'est pas constant d'une fenêtre[2] à l'autre : il faut donc normer l'énergie. Mais le choix d'une normatlisation convenable est assez délicat car l'énergie n'est pas proportionnelle au nombre de molécules d'eau mais dépend essentiellement de la distance entre le calixarène et le césium et de la structure de solvatation.

En effet, si l'énergie pour une configuration à N_1 molécules d'eau vaut E_1, alors l'énergie pour une configuration à $N_2 = N_1 + \Delta N$ molécules d'eau n'est pas égale à $E_1 \frac{N_2}{N_1}$. La valeur de E_2 dépend essentiellement de la position des ΔN molécules d'eau.
Comme la détermination du potentiel de force moyenne met en jeu des différences énergétiques, il suffit de faire en sorte que la différence du nombre de molécules d'eau entre deux états soit toujours la même. Pour cela, il faut donc que la différence de structure entre deux configurations successives soit extrêmement faible, ce qui est le cas ici puisqu'on travaille avec des variations de distance de l'ordre de 0.045 Å.
Dans ces conditions, la variation d'énergie ne dépend que de d_c et non plus du nombre de molécules d'eau constituant le système. Cette seule dépendance est possible si et seulement si on considère que les molécules d'eau supplémentaires se trouvent en périphérie des sphères de solvatation, afin de perturber le moins possible les énergies $[Cal^{4-} \cdots H_2O]$ et $[Cs^+ \cdots H_2O]$.
Soit d_c^1 une distance entre le calixarène et l'ion césium, l'énergie correspondante est alors égale à :

$$\begin{aligned} E(d_c^1) &= \left[E^{QM}_{[Cal^{4-}\cdots Cs^+]} + E^{QM}_{[Cs^+\cdots H_2O]} + E^{QM}_{[Cal^{4-}\cdots H_2O]} +' E^{QM}_{Cs^+} +' E^{QM}_{Cal^{4-}} +^0 E^{QM}_{Cs^+} +^0 E^{QM}_{Cal^{4-}}\right] \\ &+ \left[' E^{QM}_{H_2O} + E^{QM}_{[H_2O\cdots H_2O]} +^0 E^{QM}_{H_2O}\right] \\ &= E^I(d_c^1, N_1) + E^{II}(N_1) \end{aligned} \tag{6.8}$$

[2]Terme employé pour parler d'un état λ dans le formalisme de perturbation.

L'énergie E^I dépend de d_c^1 et de N_1. L'énergie E^{II} ne dépend que de N_1. Soit d_c^2 la distance entre le calixarène et l'ion césium telle que $d_c^2 = d_c^1 + \Delta d_c$ (avec $\Delta d_c = 0.05$ Å), l'énergie de cette configuration est alors égale à :

$$E(d_c^2) = E^I(d_c^2, N_2) + E^{II}(N_2) \sim \ E^I(d_c^2, N_1) + E^{II}(N_1) + E^{II}(\Delta N) \tag{6.9}$$

où ΔN est la variation du nombre de molécules d'eau entre les deux configurations. Les interactions prises en compte dans le terme $E^{II}(\Delta N)$ concernent seulement les interactions entre les molécules d'eau en périphérie du cluster. On peut donc considérer que les interactions entre le calixarène et ces molécules d'eau (de même pour l'ion Cs^+) sont quasi nulles. Ainsi, ce terme peut être évalué à partir d'un calcul quantique sur un simple cluster d'eau : $E^{II}(\Delta N) \ = \ \Delta N.(E^{Tot}(H_2O)/N^{Tot})$, où N^{Tot} est le nombre de molécules d'eau considéré. E^{Tot} est l'énergie totale du cluster d'eau composé de N^{Tot} molécules. En écrivant l'équation (6.9), on fait l'hypothèse que l'énergie E^I ne dépend que de la distance contrainte (du moins pour l'énergie entre deux configurations successives). Ainsi, la différence d'énergie est égale à :

$$\begin{aligned}
\Delta E(d_c^1, d_c^2) &= E(d_c^2) - E(d_c^1) - \Delta N.\left(\frac{E^{Tot}(H_2O)}{N^{Tot}}\right) \\
&= \left[E^I(d_c^2, N_1) + E^{II}(N_1) + E^{II}(\Delta N)\right] - \left[E^I(d_c^1, N_1) + E^{II}(N_1)\right] \\
&- \ \Delta N.\left(\frac{E^{Tot}(H_2O)}{N^{Tot}}\right) \\
&= E^I(d_c^2, N_1) - E^I(d_c^1, N_1) \\
&= E^I(d_c^2) - E^I(d_c^1)
\end{aligned} \tag{6.10}$$

Cette démarche n'est valide que si les molécules d'eau supplémentaires (ΔN) n'interfèrent pas ou très peu sur les interactions $[Cs^+ \cdots H_2O]$ et $[Cal^{4-} \cdots H_2O]$, impliquant ainsi une différence struturale et énergétique très faible entre les états d_c et d_c + Δd_c . Ainsi, nous n'ajusterons pas la différence d'énergie $\Delta E^{QM/MM} = E^{QM} - E^{MM}$, mais plutôt la différence d'énergie donnée par l'équation (6.11) :

$$\Delta\Delta E^{QM/MM}(d_c^i, d_c^j) = \Delta E^{QM/MM}(d_c^i) - \Delta E^{QM/MM}(d_c^j) \tag{6.11}$$

Écriture du Potentiel de force moyenne

Comme nous l'avons vu dans le chapitre (5), le potentiel de force moyenne peut être obtenu par la méthode de perturbation. En effet, on peut écrire que :

$$\mathcal{W}(\mathbf{r}^N) = \int \langle F_c \rangle \, d\mathbf{r}^N = \int \left\langle \frac{\partial \mathcal{U}(\mathbf{r}^N)}{\partial \mathbf{r}^N} \right\rangle d\mathbf{r}^N = \int \left\langle \frac{\partial \mathcal{U}(\mathbf{r}^N)}{\partial \lambda} \right\rangle_\lambda d\lambda = \Delta G^{TI} = \Delta G^{FEP} \quad (6.12)$$

et donc :

$$\mathcal{W}(\mathbf{r}^N) = -kT \sum_\lambda \ln \left\langle \exp\left(-\frac{(\mathcal{U}_\lambda(\mathbf{r}^N) - \mathcal{U}_{\lambda'}(\mathbf{r}^N))}{kT} \right) \right\rangle_\lambda \quad (6.13)$$

Cette relation montre clairement que l'état perturbé doit nécessairement être évalué à partir de l'ensemble des configurations de l'état de référence (λ). Écrivons cette relation en utilisant le formalisme précédent :

$$\mathcal{W}^{MM}(\mathbf{r}^N) = -kT \sum_\lambda \ln \left\langle \exp\left(-\frac{\left(\mathcal{U}^{MM}_\lambda(\mathbf{r}^N) - \mathcal{U}^{MM}_{\lambda'}(\mathbf{r}^N)\right)}{kT} \right) \right\rangle_\lambda \quad (6.14)$$

Il s'agit ici du potentiel de force moyenne non corrigé, évalué à partir de simulations de dynamique moléculaire. Le potentiel corrigé peut alors s'écrire :

$$\begin{aligned} \mathcal{W}^{Tot}(\mathbf{r}^N) &= -kT \sum_\lambda \ln \left\langle \exp\left(-\frac{\left(\mathcal{U}^{Tot}_\lambda(\mathbf{r}^N) - \mathcal{U}^{Tot}_{\lambda'}(\mathbf{r}^N)\right)}{kT} \right) \right\rangle_\lambda \\ = \quad -kT \sum_\lambda \ln \Bigg\langle \exp\left(-\frac{\left(\mathcal{U}^{MM}_\lambda(\mathbf{r}^N) - \mathcal{U}^{MM}_{\lambda'}(\mathbf{r}^N)\right)}{kT} \right) &\exp\left(-\frac{\left(\mathcal{U}^{Corr}_\lambda(\mathbf{r}^N) - \mathcal{U}^{Corr}_{\lambda'}(\mathbf{r}^N)\right)}{kT} \right) \Bigg\rangle_\lambda \end{aligned} \quad (6.15)$$

$\mathcal{U}^{Corr}_\lambda$ correspond à la différence entre l'énergie quantique et moléculaire. Comme la moyenne ne concerne pas le terme $\mathcal{U}^{Corr}_\lambda$, on peut le sortir de cette moyenne et écrire :

$$\begin{aligned}\mathcal{W}^{Tot}(\mathbf{r}^N) &= -kT\sum_{\lambda}\ln\left\langle\exp\left(-\frac{\left(\mathcal{U}_{\lambda}^{MM}(\mathbf{r}^N)-\mathcal{U}_{\lambda'}^{MM}(\mathbf{r}^N)\right)}{kT}\right)\right\rangle_{\lambda}\\ &- kT\sum_{\lambda}\ln\exp\left(-\frac{\left(\mathcal{U}_{\lambda}^{Corr}(\mathbf{r}^N)-\mathcal{U}_{\lambda'}^{Corr}(\mathbf{r}^N)\right)}{kT}\right)\end{aligned}$$

On peut alors écrire que :

$$\mathcal{W}^{Tot}(\mathbf{r}^N)=\mathcal{W}^{MM}(\mathbf{r}^N)+\mathcal{W}^{Corr}(\mathbf{r}^N)\ avec\ \mathcal{W}^{Corr}(\mathbf{r}^N)=\sum_{\lambda}\left[\mathcal{U}_{\lambda'}^{Corr}(\mathbf{r}^N)-\mathcal{U}_{\lambda}^{Corr}(\mathbf{r}^N)\right]$$

Dans toutes les relations ci-dessus, les termes portant l'exposant ^{MM} indiquent les termes issus de calculs sur des configurations générées par des simulations de dynamique moléculaire, ceux portant l'exposant ^{Tot} indiquent les termes ^{MM} corrigés par mécanique quantique et ceux avec l'exposant ^{Corr} indiquent les termes de corrections. La relation (6.15) montre clairement que le terme de correction est évalué en considérant l'espace des phases $\Gamma(\lambda)$ de l'état λ. Par conséquent, on ne peut pas utiliser les énergies issues du calcul quantique direct pour calculer la différence d'énergie puisque l'énergie obtenue est calculée en considérant l'espace des phases $\Gamma(\lambda^{'})$ de l'état $\lambda^{'}$. Dans ce contexte, $\mathcal{W}^{Corr}$ s'écrit :

$$\begin{aligned}\mathcal{W}^{Corr}(\mathbf{r}^N) &= \sum_{\lambda}\Delta\Delta E^{QM/MM}(\mathbf{r}_{\lambda}^N,\mathbf{r}_{\lambda'}^N)\\ &= \sum_{\lambda}\left[\Delta E^{QM/MM}(\mathbf{r}_{\lambda}^N)-\Delta E^{QM/MM}(\mathbf{r}_{\lambda'}^N)\right]\end{aligned}$$

Ici, $\mathbf{r}_{\lambda}^N$ est l'espace des phases correspondant à l'état λ et $\mathbf{r}_{\lambda'}^N$ est l'espace des phases correspondant à l'état $\lambda^{'}$. Cette relation ne peut pas être utilisée car l'espace de phase pour les deux états n'est pas le même. La seule relation que l'on peut écrire est :

$$\begin{aligned}\mathcal{W}^{Corr}(\mathbf{r}^N) &= \sum_{\lambda}\Delta\Delta E^{QM/MM}(\mathbf{r}_{\lambda}^N)\\ &= \sum_{\lambda}\left[\Delta E^{QM/MM}(\mathbf{r}_{\lambda}^N)-\Delta E^{QM/MM}(\mathbf{r}_{\lambda'}^N)\right]\end{aligned}$$

Pour contourner ce problème, il suffit d'utiliser l'ajustement de la différence d'énergie qui ne dépend plus de λ ou de l'espace des phases $\Gamma(\lambda)$ mais seulement de la distance contrainte :

$$\boxed{\mathcal{W}^{Tot}(\mathbf{r}^N) = \mathcal{W}^{MM}(\mathbf{r}^N) + \mathcal{W}^{Fit}(\mathbf{r}_c) \quad avec \quad \mathcal{W}^{Fit}(\mathbf{r}_c) = \mathcal{T}^F \left(\sum_{\lambda} \left[\mathcal{U}_{\lambda'}^{Corr}(\mathbf{r}_{\lambda'}^N) - \mathcal{U}_{\lambda}^{Corr}(\mathbf{r}_{\lambda}^N) \right] \right)} \tag{6.16}$$

où $\mathcal{T}^F$ n'est autre que la fonction d'ajustement.

6.3 Mise en œuvre

6.3.1 Protocole

Cette méthodologie est générale et peut donc être applicable à tout système. Nous avons choisi le système calixarène-césium pour les raisons citées précédemment. Pour calculer le potentiel de force moyenne par une dynamique moléculaire classique, nous avons réalisé 250 simulations, échantillonnant ainsi des distances variant de 0.5 Å à 12 Å avec un incrément de distance de 0.046 Å. Un calcul classique par la méthode FEP nous oblige à nous placer à λ et à déplacer seulement les atomes du calixarène et le césium, ne pouvant éviter ainsi des chocs avec les molécules d'eau. Afin de minimiser ces accidents, on utilise la méthode FEP^m développée dans le chapitre (4). Il est possible de réduire le nombre de simulations si on utilise directement le calcul de la force contrainte. Néanmoins, cette méthode nécessite le calcul du multiplicateur de Lagrange durant l'intégration des équations de mouvement. Donc, si on veut incorporer les corrections dans la dynamique moléculaire, il faut déterminer les paramètres d'un potentiel de paires entre chaque atome constituant les deux ions étudiés, ce que la méthodologie mise au point ici ne permet pas. En effet, on détermine seulement un potentiel intermoléculaire entre deux ions et non pas des potentiels interatomiques. Ainsi, la méthode de perturbation est la méthode la mieux adaptée à notre démarche.

Chaque simulation à été réalisée avec le protocole suivant : 50 ps d'équilibration et 100 ps d'acquisation, chaque simulation i débutant avec la configuration de fin de la simulation $i-1$ d'où le faible temps d'équilibration. Au total, nous avons simulé 37.5 ns, ce qui a nécessité 4 mois de calculs sur un biprocesseur Intel Xeon 3.2 GHz. Le champ de forces et les algorithmes

de dynamique sont détaillés dans la section (2.6).

Les calculs quantiques ont été réalisés par DFT en utilisant la fonctionelle B3LYP. Ces calculs ont été possibles grâce au logiciel GAUSSIAN [64] disponible sur le centre de calculs de l'IDRIS (Institut de Développement et des Ressources en Informatique Scientifique). L'utilisation de 8 processeurs a été nécessaire, ce qui a permis de réaliser les calculs d'énergie en 6 mois au lieu de 4 ans sur un monoprocesseur standard.

6.3.2 Résultats

Ajustement et calcul potentiel de force moyenne

La figure (6.3) représente la courbe à ajuster et la courbe résultant de cet ajustement. La fonction retenue pour ce calcul est une fonction sinusoïdale amortie par une exponentielle :

$$\Delta E^{Fit}(\mathbf{r}^N) = A0.\exp\left(-\frac{r}{A1}\right)\sin\left(\frac{\pi.(r-A2)}{A3}\right) - \frac{A4}{x^{12}}$$

Ici, les termes $A0, A1, A2, A3$ *et* $A4$ sont les paramètres d'ajustement de la fonctions $\mathcal{W}^{Fit}$.

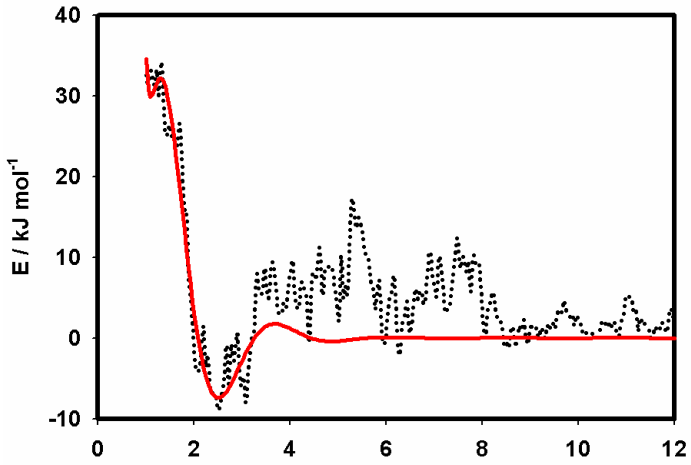

FIG. 6.3 – Ajustement de la différence d'énergie entre l'énergie moléculaire et l'énergie quantique. La différence a été ajustée par l'intermédiaire d'une fonction sinusoïdale amortie. (—) représente le résultat de l'ajustement et (· · ·) représente la courbe à ajuster.

Le premier terme de cette équation correspond à une partie attractive et le second à une partie répulsive. L'ajustement utilisé traduit bien le puits de potentiel mais présente des lacunes sur la partie attractive puisqu'on observe un décalage par rapport à la courbe à ajuster. Or, le puits de potentiel et la partie répulsive sont les parties qu'il est nécessaire de reproduire correctement puisque l'on intégre dans cette zone pour obtenir les grandeurs thermodynamiques. Une fois l'ajustement réalisé, l'utilisation de la relation (6.16) nous permet ainsi d'obtenir directement le potentiel de force moyenne corrigé.

La figure (6.4.a) représente le potentiel de force moyenne obtenu uniquement par dynamique moléculaire et celui obtenu en tenant compte des interactions quantiques. Cette figure montre la nécéssité de calculs quantiques parallèlelement aux calculs de dynamique moléculaire. La profil ajusté présente un puits de potentiel à 2.6 Å (cette distance représente la distance entre le centre de masse du calixarène et celui de l'ion césium). Le tableau (6.2) montre les valeurs calculées pour $\Delta_r H^o$, $\Delta_r G^o$ et $\Delta_r S^o$, en bon accord avec les valeurs expérimentales. Cette concordance permet donc de valider la méthodologie. Ces grandeurs ont été déterminées par l'utilisation des formules démontrées dans le chapitre (5), en prenant comme borne d'intégration supérieure une distance de r_{min}.

La figure (6.4.b) indique les valeurs de $\Delta_r H^o$, $\Delta_r G^o$ et $\Delta_r S^o$ en fonction de la valeur de la borne supérieure d'intégration. Cette figure montre qu'au delà de 5 Å, les grandeurs thermodynamiques tendent vers une valeur limite. À $d_l = r_{min} + 1$ Å, on améliore la valeur de la variation d'entropie et on diminue la précision du résultat sur $\Delta_r G^o$. On en déduit la nécessité d'effectuer une étude en fonction de r afin de déterminer la borne supérieure, néanmoins les valeurs $\Delta_r H^o$, $\Delta_r G^o$ et $\Delta_r S$ doivent évoluer très peu en fonction de d_l.

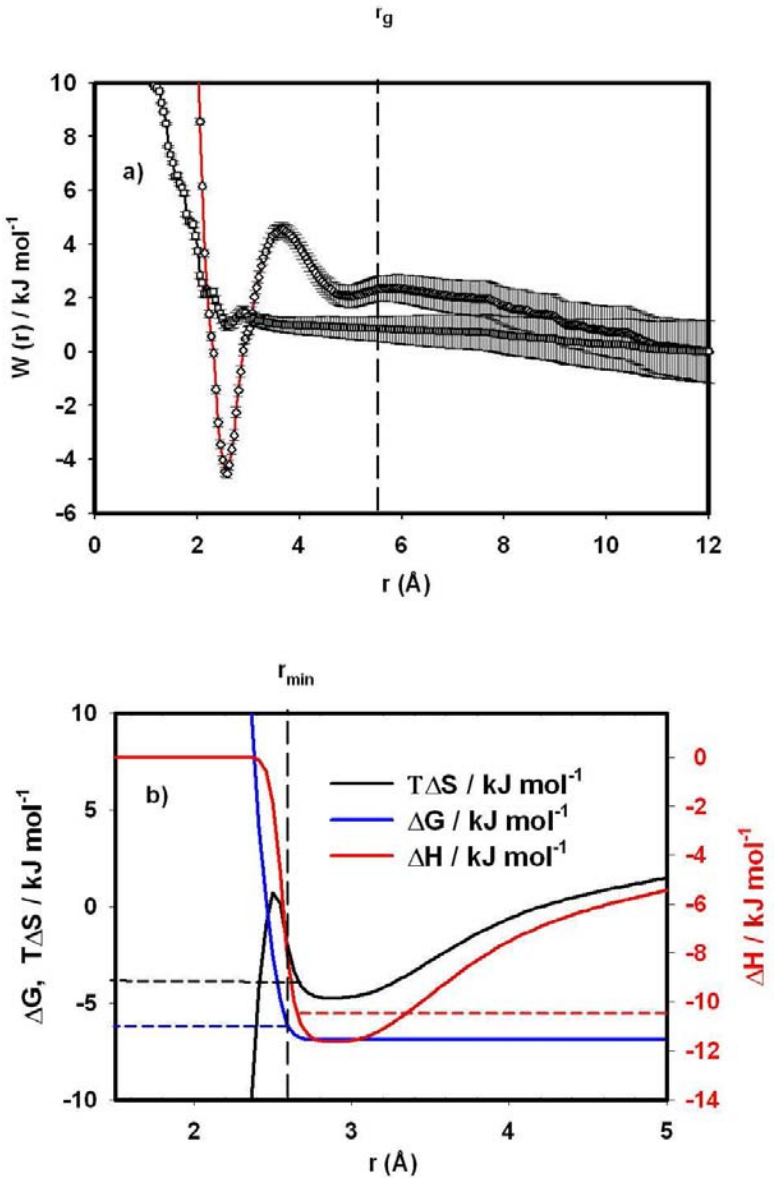

FIG. 6.4 – a) Potentiel de force moyenne corrigé (—,◊) et non corrigé (—,□). b) Intégration des potentiels de force moyenne permettant de déterminer les grandeurs d'association en fonction de la valeur de la borne d'intégration. Les traits en pointillés représentent les valeurs obtenues par microcalorimètrie.

	r_{min} (Å)	K (mol^{-1} L)	$\Delta_r G^o$ (kJ mol^{-1})	$\Delta_r H^o$ (kJ mol^{-1})	$T\Delta_r S^o$ (kJ mol^{-1})
MM/QM	2.8	12.2	-6.2	-10.1	-3.9
MM	–	0.0	∞	∞	∞
Exp. [61]	–	14.6	-6.6	-10.9	-4.3

TAB. 6.2 – Grandeurs thermodynamiques d'association obtenues par la méthode FEP^m. Nous avons choisit, comme borne supérieure d'intégration la position du puits de potentiel.

Origine du puits

À une distance de 2.6 Å, l'ion Cs^+ est à environ 5 Å de chacun des quatres cycles benzéniques. Or, Dougherty [60] a montré par des calculs en phase gazeuse que les interactions cation-π entre deux espèces se manifestent à une distance de 2.5 Å. À cette distance, ce sont les forces de polarisation et de transfert de charge qui prédominent, c'est-à-dire les forces en $\frac{1}{r^n}$ tel que $n > 2$. En outre, il montre que dans le cas d'interactions de type quadrupolaire, cette distance est ramenée à 5 Å. Une distance de 5 Å entre les cycles benzéniques et l'ion Cs^+ peut donc conduire à trois hypothèses :

1- La présence d'interactions quadrupolaires et donc l'absence d'interactions cation-π (selon Dougherty [60]).
2- Les interactions quadrupolaires ont un poids aussi important que les processus de polarisation et de transfert de charge, ce qui est en contradiction avec le modèle physique proposé par Dougherty pour les interactions cation-π.
3- Il s'agit, effectivement d'interactions cation-π. La présence de molécules d'eau et des quatres cycles benzéniques ainsi que la forme conique du calixarène sont à l'origine de cette augmentation de distance.

Afin de vérifier la troisième hypothèse, on calcule les distributions radiales de l'ion Cs^+ seul et en présence du complexe. De leur intégration, on déduit une désolvatation partielle de l'ion Cs^+ (figure (6.5.b)). En outre, la cavité est complètement désolvatée (figure (6.5.a)) ce qui permet aux interactions cation-π de se développer sans interférer avec les molécules d'eau. La distribution radiale du centre de masse du calixarène semble indiquer la présence de molécules

d'eau dans la cavité. Cette présence n'est pas réaliste étant donnée la taille de l'ion Cs^+. En effet, comme le centre de masse du macrocycle se trouve au dessus de l'ion Cs^+, le calcul de la distribution radiale comprend aussi les molécules d'eau appartenant à la sphère de solvatation de l'ion césium. Les figures (6.6.b) et (6.6.c) permettent de visualiser la désolvatation de la cavité.

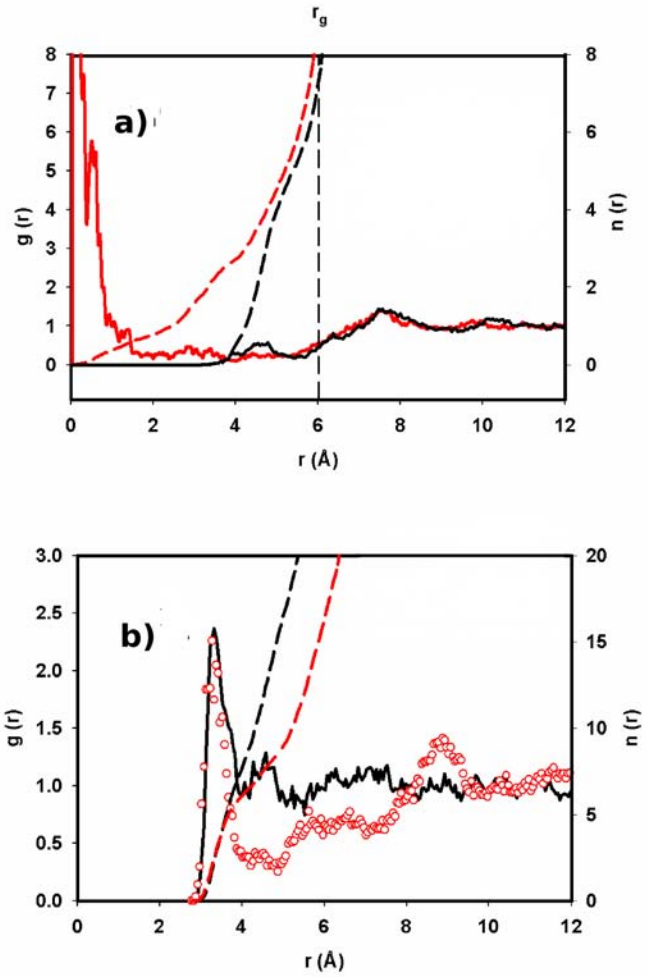

FIG. 6.5 – a) (—) Distribution radiale entre le centre de masse du macrocycle seul en solution et les atomes d'oxygène des molécules d'eau. (—) Distribution radiale entre le centre de masse du macrocycle en présence du complexe et les atomes d'oxygène de l'eau. r_g indique le rayon de giration du centre de masse du calixarène b) (—) Distribution radiale entre l'ion césium seul en solution les atomes d'oxygène des molécules d'eau. (—) Distribution radiale entre l'ion césium en présence du macrocycle les atomes d'oxygène des molécules d'eau. Sur les deux figures les courbes en pointillés représentent les courbes intégrales des distributions radiales (on utilise le même code de couleur que pour les distributions radiales).

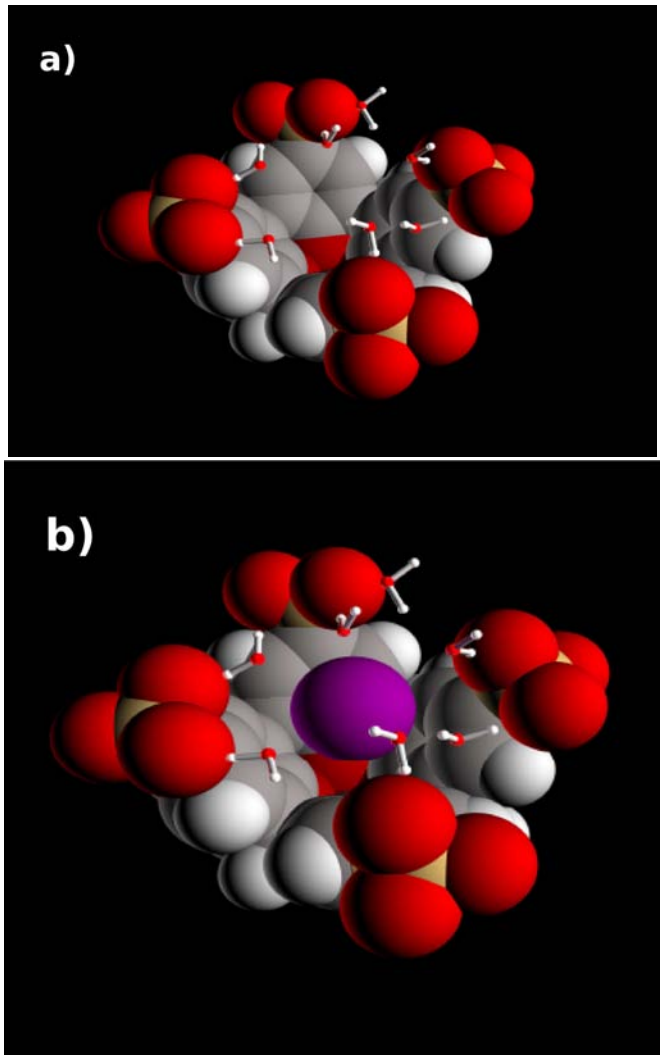

FIG. 6.6 – a) Visualisation de la désolvatation de la cavité. b)Visualisation de la solvatation de l'ion césium.

Néanmoins, il peut exister une compétition dans les interactions entre les oxygènes des groupements sulfonate et l'ion césium d'une part, mais aussi entre l'ion césium et les molécules d'eau présentes dans la première couche d'hydratation des sulfonates d'autre part (*cf* figure (6.6.c)). Ainsi, ces interactions de type électrostatique ne permettent pas à l'ion césium de s'insérer plus bas dans la cavité hydrophobe. Il existe donc une sorte de compensation d'interaction susceptible d'expliquer ce positionnement. En outre, si l'ion césium établissait des ponts de liaisons hydrogène avec les groupements sulfonate, cela renforcerait les interactions

de type électrostatique. Il n'en est rien : un calcul du nombre moyen de ponts de liaisons hydrogène [$Cs^+ \cdots H_2O \cdots SO_3^-$] indique une valeur sensiblement nulle. Par contre, il existe des interactions spécifiques entre l'ions Cs^+ et les molécules d'eau. On peut montrer l'importance des interactions dues aux molécules d'eau en optimisant la géométrie de la configuration correspondant au puits de potentiel par un calcul quantique en enlevant les molécules d'eau. Ce calcul montre en effet une diminution de la distance puisqu'elle est égale à 2.8 Å. Cela montre donc l'importance des molécules d'eau et des groupements sulfonate dans le positionnement de l'ion césium dans la cavité.

6.4 Résumé

Dans le chapitre (5), nous avons appliqué les méthodologies de calcul du potentiel de force moyenne à des complexes d'insertion et de sphère externe. Nous avons reproduit correctement les grandeurs d'association, témoignant d'une certaine efficacité des algorithmes. Nous avons déterminé le potentiel de force moyenne et les grandeurs thermodynamiques d'association pour des complexes dont les interactions prédominantes sont de nature électrostatique ou de dispersion-répulsion. Les complexes formés avec le calixarène mettant en jeu des interactions quantiques ne peuvent être simulés par des simulations de type moléculaire. Pour prendre en compte de telles interactions dans le calcul de potentiel de force moyenne, nous avons mis au point une méthodologie couplant les méthodes de dynamique moléculaire et de mécanique quantique.

Cette méthode consiste dans un premier temps à échantillonner l'ensemble des distances par des simulations de type dynamique moléculaire, puis de choisir judicieusement la partie de chaque configuration qui sera traitée par un calcul quantique. En ajustant la différence d'énergie E^{QM}-E^{MM}, on peut alors déterminer le potentiel de force moyenne. Les résultats obtenus par cette méthode sont en très bon accord avec les valeurs expérimentales mesurées au laboratoire, ce qui permet de valider notre méthode. Il a été impossible de tester cette méthode sur d'autres systèmes puisque le temps de calcul s'avère très élevé.

En outre, une étude structurale sur le complexe formé au puits de potentiel nous a permis de montrer une désolvatation complète de la cavité. L'ion césium, quant à lui, conserve sa couche de solvatation. Par ailleurs, nous avons montré que la position du puits de potentiel est liée aux interactions cation-π mais aussi aux interactions entre l'ion césium et les groupements sulfonate du macrocycle et aux interactions entre l'ion Cs^+ et les molécules d'eau de sa première sphère d'hydratation.

Chapitre 7

Conclusions et perspectives

7.1 Conclusions

La simulation numérique de type moléculaire est devenue un outil performant pour la description microscopique d'un système physico-chimique et la caractérisation énergétique des processus tels que les réactions d'association. La mise en œuvre d'une simulation requiert l'utilisation de plusieurs approximations et de protocoles. Lorsqu'un système comporte de nombreux groupements chimiques de natures différentes, le champ de forces utilisé doit être validé sur des critères énergétiques. Il est alors essentiel que les simulations moléculaires respectent, dans le cadre d'études de réaction d'association, la sélectivité du récepteur vis-à-vis de substrats différents. Cette étape est primordiale afin que la description moléculaire du système soit cohérente. Cette démarche nécessite une confrontation permanente des résultats des simulations aux données expérimentales ($\Delta_r G^0, \Delta_r H^0$).

Alors que la simulation de type dynamique moléculaire sur des récepteurs et substrats en phase aqueuse peut être réalisée sans difficulté majeure, le calcul de grandeurs thermodynamiques d'association est plus délicat et nécessite un développement algorithmique approprié. La mécanique statistique est alors un outil puissant pour arriver aux formules opérationnelles des propriétés thermodynamiques. Cette thèse nous a permis de mettre au point un certain nombre d'algorithmes pour le calcul de variations de grandeurs d'associations ($\Delta\Delta_r G^0, \Delta\Delta_r H^0, \Delta\Delta_r S^0$) et la détermination des propriétés thermodynamiques absolues d'association directement comparables à l'expérience. Nous avons également mis au point une méthodologie sur le calcul du potentiel de force moyenne pour prendre en compte des interactions cation-π. Une grande partie de ma thèse fut consacrée au développement de codes de calcul pour le calcul des grandeurs thermodynamiques d'association. La puissance de calcul

demandée pour obtenir la convergence de ces propriétés m'a ensuite conduit à développer des codes parallèles [65].

Une fois les caractérisations énergétiques effectuées, nous avons donné une vision structurale des complexes de sphère externe et des complexes d'insertion en étudiant précisément les phénomènes d'hydratation de l'ensemble des espèces impliqués dans la complexation. Cette relation entre l'aspect microscopique et la caractérisation énergétique des complexes a permis de donner des interprétations sur l'évolution des grandeurs thermodynamiques d'association. À chaque fois que cela a été possible, nous avons associé une description microscopique aux grandeurs énergétiques. Cependant, je pense que cette caractérisation moléculaire est beaucoup plus simple à mettre en œuvre au niveau algorithmique et qu'elle n'est pertinente que si tout le travail de développement algorithmique concernant le calcul de grandeurs énergétiques est validé. Je tiens réellement à mentionner cette remarque puisqu'elle a représenté le fil directeur de cette thèse.

Les codes que nous avons développés pour ce travail sont basés sur des considérations de thermodynamique statistique et m'ont conduit à revisiter certains algorithmes tels que celui de Berendsen et SHAKE pour l'ensemble NpT. Pour toutes ces raisons, j'ai voulu développer quelques idées et démonstrations dans la deuxième partie de cette thèse. Ces chapitres ne sont pas indispensables pour la compréhension des chapitres de la première partie mais témoignent du temps que j'ai consacré à la compréhension de ces algorithmes. Je rappelle maintenant les principaux résultats issus de ma thèse.

Dans le troisième chapitre, nous nous sommes intéressés à la complexation des alcools et diols linéaires par un calixarène. Nous avons montré que le type de complexe formé est un complexe d'insertion où la molécule d'alcool est partiellement insérée dans la cavité du calixarène. Nous montrons que l'alcool est beaucoup moins inséré dans la cavité qu'avec un cation de type tétraalkylammonium. Nous avons attribué cette particularité au groupement hydroxyde et à son hydratation. Nous avons montré une corrélation directe entre l'enthalpie d'association expérimentale et le nombre de groupements insérés dans la cavité du macrocycle. Nous avons également montré le rôle important joué par les molécules d'eau lors de l'élaboration de ponts

de liaisons hydrogène entre le groupement hydroxyde de l'alcool et les groupements sulfonate du calixarène. Nous avons expliqué les variations d'enthalpie et d'entropie d'association entre l'alcool (butanol) et le diol correspondant (1-4 butanediol) par des différences d'hydratation des deux groupements hydroxyde.

Le quatrième chapitre présente les différentes méthodes pour le calcul de variations de grandeurs thermodynamiques d'association. Le calcul de la variation d'entropie est établi pour la première fois avec la méthode FEP dans l'ensemble NpT. Nous avons développé une nouvelle méthodologie propre à la dynamique moléculaire permettant de recalculer les positions et vitesses des espèces perturbées. Nous avons appliqué cette méthode sur le calcul de grandeurs d'associations relatives sur des complexes de sphère externe du calixarène avec des ions inorganiques de masse et de charge différentes. Nous avons évalué les contributions de dispersion-répulsion et électrostatiques dans le calcul lorsque la charge atomique était perturbée. Nous avons reproduit de manière satisfaisante les grandeurs expérimentales, en particulier la variation d'entropie, propriété qui n'est pas facile à reproduire compte tenu qu'elle implique l'hamiltonien total dans sa formule opérationnelle.

Dans le cinquième chapitre, nous avons utilisé le concept de potentiel de force moyenne pour le calcul de grandeurs d'association absolues. Nous avons comparé les méthodes de perturbation et de force contrainte sur une paire d'ions en phases gazeuse et aqueuse, et étendu cette comparaison sur l'association du calixarène avec un ion inorganique de type lanthanide et un cation tétraméthylammonium. Nous avons démontré que le potentiel de force moyenne permet de distinguer deux types de complexation (sphère externe et d'inclusion) en localisant le puits d'enthalpie libre à l'extérieur et à l'intérieur de la cavité du macrocycle. Nous avons également confirmé l'absence de complexation avec le méthanol et montré que la désolvatation de la première couche d'hydratation du méthanol en cas de complexation ne serait pas compensée au niveau énergétique par les interactions dispersion-répulsion entre le méthanol et le calixarène. Le calcul des grandeurs thermodynamiques d'association à partir des potentiels de forces moyennes s'est avéré performant avec une bonne reproduction des valeurs expérimentales. Le point important réside également dans le fait d'avoir reproduit la valeur absolue et le signe de $\Delta_r H^o$ et de $\Delta_r S^o$ pour les deux types de complexes ($\Delta_r H^o < 0,\ \Delta_r S^o \sim 0$

pour un complexe d'insertion et $\Delta_r H^o > 0,\ \Delta_r S^o > 0$ pour un complexe de sphère externe).

Enfin, le dernier chapitre applicatif concerne la mise en œuvre d'une nouvelle méthodologie pour la mise en évidence d'interactions cation-π. Alors que le potentiel de force moyenne calculé par un champ de forces moléculaire ne donne aucun puits de potentiel entre l'ion césium et le calixarène, les données expérimentales prévoient une complexation entre ces deux espèces avec une enthalpie d'association négative suggérant une insertion de ce cation dans la cavité du macrocycle. Nous avons alors établi un protocole permettant d'ajouter au potentiel de force moyenne une contribution établie à partir de calculs DFT sur les configurations représentatives. Nous avons montré que la prise en compte de cette contribution «quantique» a permis d'établir un puits de potentiel à l'intérieur de la cavité. Les grandeurs thermodynamiques d'association résultant du potentiel de force moyenne modifiée sont reproduites convenablement avec une valeur négative de l'enthalpie d'association.

Les chapitres suivants constituent la seconde partie de ma thèse et sont consacrés à des réflexions sur certaines notions et formalisme liés à la dynamique moléculaire. Une présentation de la méthode DFT est également proposée dans cette partie.

7.2 Perspectives

Concernant la méthologie il est possible détendre les méthodes de perturbation au calcul de tension interfaciale et d'enthalpie de surface dans le cas de système hétŕogène tels que les systmes liquide-vapeur ou bien encore liquide-liquide. Par ailleurs, la méthologie FEP^m mise au point peut permettre le calcul précis du potentiel chimique de l'eau dans le cas de systèmes confinés.

Deuxième partie

Théorie

Sommaire - II

8 **Mécanique statistique** 181

9 **Simulation numérique de type dynamique moléculaire** 195

10 **Rappels de la théorie de la DFT** 239

Bibliographie

Les ouvrages utilisés pour rédiger cette partie sont :

1- [17, 18, 35, 66]

2- Stöcker, H. ; *Toutes les mathématiques et les bases de l'informatique* **Dunod**, (1999).

3- Viot, P. ; *Simulation numérique en physique statistique* **Cours commun aux DEA, Jussieu**, (2004).

4- Walter, A. ; *Mathématiques pour la physique et les physiciens !* **H&K Éditions**, (2002).

5- Walter, R. ; *Principes d'analyse mathématique* **EDISCIENCE international**, (1995).

6- Landau, L.D. ; Lifshitz, E. ; *Mécanique* **Édition Mir**, (1970).

7- Landau, L.D. ; Lifshitz, E. ; *Statistical physics* **Pergamon New York**, (1994).

8- Aslangul, C. ; *Physique statistique de la matière molle (II)* **DEA de Physique des Liquides**, (2002).

9- Skeel, R.D. ; Tupper, P. *Mathematical Issues in Molecular Dynamics*, (2005).

10- Smith, W. ; Forester, T.R. *The DL_POLY_2 USER MANUAL, version 2.13* CCLRC, Daresbury Laboratory, (2001).

11- Desbiens, N. *Développement de potentiels d'interactions cation-eau basés sur un nouveau modèle de potentiel à charges fluctuantes* **Rapport de stage de DEA**, (2003).

12- Rivail, J.L. ; *Élément de chimie quantique à l'usage des chimistes* **CNRS EDITIONS**, (1999).

13- Hutter, J. ; *Introduction to Ab Initio Molecular Dynamics* **Lecture Notes**, (2002).

Chapitre 8

Mécanique statistique

8.1 Description d'un système 181
8.1.1 État microscopique 181
8.1.2 État macroscopique 182
8.2 Facteur de Boltzmann et fonction de partition 182
8.2.1 Facteur de Boltzmann 183
8.2.2 Probabilité de Boltzmann 186
8.3 Notion d'ensemble statistique 187
8.4 Fonction de partition dans l'ensemble NVT 188
8.5 Fonction de partition dans l'ensemble NpT 190
8.6 Fonctions d'état 193

8.2. Description d'un système

D'une manière générale, c'est la mécanique quantique qui permet d'étudier les propriétés microscopiques d'un système. Il existe, néanmoins, des cas où l'on peut utiliser la mécanique newtonnienne classique.

8.2.1 État microscopique

En mécanique classique, l'état d'une particule ponctuelle est caractérisé par sa position et son impulsion, c'est-à-dire sa vitesse à un instant donné. Si on considère un système de N particules, son état instantané est déterminé par la donnée de $6N$ paramètres ($3N$ pour les positions et $3N$ pour les vitesses). Le mouvement du système est donc décrit par l'évolution temporelle de son état, c'est-à-dire par ses $6N$ fonctions du temps. Ainsi, à chaque grandeur

physique que l'on peut mesurer sur le système, est associée un fonction de variables $\mathbf{r}^N$ et $\mathbf{p}^N$ avec une éventuelle dépendance temporelle qui donne la valeur de cette grandeur. C'est, ainsi, qu'à l'énergie totale est associée la fonction de Hamilton que l'on peut écrire de la manière suivante, s'il n'existe pas de dépendance temporelle :

$$\mathcal{H}(\mathbf{r}^N, \mathbf{p}^N) = \sum_{i=1}^{N} \frac{\mathbf{p}_i^2}{2m_i} + \mathcal{U}(\mathbf{r}^N) \tag{8.1}$$

où $\mathcal{U}(\mathbf{r}^N)$ n'est autre que l'énergie potentielle. Le premier terme représente l'énergie cinétique du système. On dit que l'espace à $6N$ dimensions est *l'espace des phases* dans lequel chaque état possible du système est représenté par un point de coordonnées $(\mathbf{r}^N, \mathbf{p}^N)$. Le mouvement du système se traduit alors par un déplacement du point representatif dans l'espace des phases où il décrit une trajectoire caractérisée par la donnée de ses fonctions.

8.2.2 État macroscopique

Si le nombre de paramètres nécessaires à la description du système est grand, il est alors impossible de connaître exactement son état microscopique. On fait alors appel aux probabilités pour définir son état macroscopique dont la donnée sera suffisante pour déterminer les propriétés macroscopiques du système. Dans le cadre de la mécanique classique, les états microscopiques possibles forment un continuum représenté par l'espace des phases : chaque état sera donc repéré par l'ensemble des $6N$ variables continues. L'état macroscopique est alors caractérisé par une densité de probabilité dans l'espace des phases, c'est-à-dire par une fonction qui dépend des $6N$ variables définissant le système. Cette fonction s'écrit $W(\mathbf{r}^N, \mathbf{p}^N)$ et donne la probabilité (équation (8.2)) pour qu'à un instant donné l'état microscopique du système a son point représentatif dans le volume infinitésimal $d\mathbf{r}^N d\mathbf{p}^N$ au point $(\mathbf{r}^N, \mathbf{p}^N)$ de l'espace des phases :

$$dP(\mathbf{r}^N, \mathbf{p}^N) = W(\mathbf{r}^N, \mathbf{p}^N) d\mathbf{r}^N d\mathbf{p}^N \tag{8.2}$$

8.3 Facteur de Boltzmann et fonction de partition

Le facteur de Boltzmann est l'une des grandeurs les plus importantes en chimie physique. En effet, elle indique que dans le cas d'un système pouvant occuper des états d'énergie diffé-

rents E_1, E_2, E_3, ..., la probabilité p_j que le système occupe l'état d'énergie E_j dépend de l'exponentielle de l'énergie (équation (8.3)).

$$p_j \propto \exp\left(-\frac{E_j}{k_B T}\right) \tag{8.3}$$

k_B est la constante de Boltzmann. La somme des probabilités doit être égale à 1. La constante de normalisation des probabilités ($\sum_j p_j = 1$) est $\frac{1}{Q}$ telle que

$$Q = \sum_j \exp\left(-\frac{E_j}{k_B T}\right) \tag{8.4}$$

Le passage entre l'état quantique et l'état moléculaire est possible si l'on considère l'énergie du système comme continue. Dans ce cas, on peut remplacer l'énergie électronique par l'hamiltonien du système (section 2.2), si et seulement si, ces deux variables possèdent les mêmes propriétés. Il est, alors, possible de remplacer la sommation discrète par une intégrale sur le volume (équation (8.5)).

$$Q = \int\int \exp\left(-\frac{\mathcal{H}(\mathbf{r}^N, \mathbf{p}^N)}{k_B T}\right) d\mathbf{r}^N d\mathbf{p}^N \tag{8.5}$$

$\mathbf{r}^N$ et $\mathbf{p}^N$ sont respectivement les vecteurs position et les vecteurs impulsion des N particules constituant le système. La grandeur Q est appelée fonction de partition. Cette fonction va permettre d'établir le lien «micro-macro». Elle définit l'état du système et permet de déterminer les propriétés thermodynamiques.

8.3.1 Facteur de Boltzmann

Soit un système en équilibre macroscopique (les conditions extérieures au système sont maintenues constantes : p, V, ...). Ce système peut très bien être un liquide, un solide ou un gaz. D'un point de vue mécanique, ce système peut être décrit à partir du nombre de particules qui le constituent, du volume et des interactions entre les particules. Ce système est alors caractérisé par son hamiltonien et ses fonctions d'ondes associées. Ces grandeurs dépendent des coordonnées de toutes les particules du système. Ainsi, pour ce système à N corps, l'équation de Shrödinger est

$$\widehat{\mathcal{H}}\Psi_j = E_j\Psi_j,\ j = 1, 2, 3, ...$$

Ici, j est le nombre d'états énergetiques pouvant être pris par le système (état propre). $E_j(N, V)$ est l'énergie correspondant à l'état j qui dépend à la fois de N et de V. L'objectif est de déterminer la probabilité que le système soit dans un état j d'énergie $E_j(N, V)$. Pour ce faire, nous allons considérer un grand nombre de systèmes identiques en contact thermique avec un réservoir de chaleur à la température T. Tous les systèmes auront le même nombre de particules N, le même volume V et la même température T. Mais, ils seront probablement dans des états quantiques différents (états microscopiques) bien qu'étant dans des états macroscopiques identiques. Une telle collection de système est alors appelée ensemble statistique.

Le nombre de systèmes dans l'état j d'énergie $E_j(N, V)$ sera noté a_j et le nombre total de système $\mathcal{A}$. Déterminons, à présent, le nombre de systèmes dans chaque état. Soient deux états 1 et 2 caractérisés par les énergies $E_1(N, V)$ et $E_2(N, V)$. Les populations relatives des deux états 1 et 2 doivent dépendre de E_1 et E_2. On a donc

$$\frac{a_2}{a_1} = f\left(E_1, E_2\right) \tag{8.6}$$

Ici, a_2 et a_1 représentent le nombre de systèmes dans les états 1 et 2. La forme de la fonction f reste à déterminer. Une énergie est une grandeur relative à un état de référence, c'est-à-dire un état d'énergie nulle. En outre, si l'origine des énergies est la même dans les deux états, on peut alors considérer que f est fonction de la différence d'énergie entre ces deux états. Ce qui nous permet de réécrire la relation (8.6).

$$\frac{a_2}{a_1} = f\left(E_1, E_2\right) = f\left(E_1 - E_2\right) \tag{8.7}$$

L'équation (8.7) doit être satisfaite pour deux états d'énergie quelconque. On peut donc écrire la relation (8.8). En tenant compte de la relation (8.9), on peut remplacer $\frac{a_2}{a_1}$ et $\frac{a_3}{a_2}$ par les relation (8.7) et obtenir l'équation (8.10).

$$\frac{a_3}{a_2} = f\left(E_2 - E_3\right) \quad et \quad \frac{a_3}{a_1} = f\left(E_1 - E_3\right) \tag{8.8}$$

$$\frac{a_3}{a_1} = \frac{a_2}{a_1} \cdot \frac{a_3}{a_2} \tag{8.9}$$

$$f(E_1 - E_3) = f(E_1 - E_2).f(E_2 - E_3) \tag{8.10}$$

À première vue, une fonction qui vérifie cette relation n'est pas aisée à déterminer, mais sachant que :

$$\exp(x + y) = \exp(x).\exp(y)$$

nous voyons donc que $f(E)$ doit être de la forme :

$$f(E) = \exp(\beta E)$$

Où β est une constante arbitraire. Vérifions, à présent, si ce type de fonction est bien compatible avec l'équation (8.10). Pour ce faire, introduisons la fonction $f(E)$ dans l'équation (8.10)

$$\begin{aligned} \exp(\beta(E_1 - E_3)) &= \exp(\beta(E_1 - E_2))\exp(\beta(E_2 - E_3)) \\ &= \exp(\beta(E_1 - E_3)) \end{aligned}$$

De l'équation (8.7), on en déduit que :

$$\frac{a_2}{a_1} = \exp(\beta(E_1 - E_2)) \tag{8.11}$$

Les états 1 et 2 n'ont rien de particulier, on peut donc généraliser l'équation (8.11) au calcul du rapport de populations de deux états d'énergie quelconque et ainsi écrire que :

$$\frac{a_n}{a_m} = \exp(\beta(E_m - E_n)) \quad avec \quad n > m$$

La forme de cette équation implique donc que les populations a_m et a_n soient de la forme :

$$a_j = C\exp(-\beta(E_j)) \tag{8.12}$$

où j représente les états m et n et C une constante.

8.3.2 Probabilité de Boltzmann

L'équation (8.12) contient deux grandeurs C et β dont nous devons déterminer les expressions. C semble facile à déterminer. En effet, la somme des a_j s'écrit :

$$\sum_j a_j = C \sum_j \exp(-\beta E_j)$$

La somme des a_j doit donc être égale à $\mathcal{A}$. On en déduit l'expression de C :

$$C = \frac{\sum_j a_j}{\sum_j \exp(-\beta E_j)}$$

Si on introduit ce résultat dans l'équation (8.12), nous obtenons :

$$\frac{a_j}{\mathcal{A}} = \frac{\exp\left(-\beta E_j\right)}{\sum_j \exp(-\beta E_j)} \tag{8.13}$$

Le rapport $\dfrac{a_j}{\mathcal{A}}$ n'est autre que la proportion des systèmes dans l'état j d'énergie E_j. Lorsque $\mathcal{A}$ est grand, $\dfrac{a_j}{\mathcal{A}}$ est alors une *probabilité*. L'équation (8.13) devient donc :

$$p_j = \frac{\exp\left(-\beta E_j\right)}{\sum_j \exp(-\beta E_j)} \tag{8.14}$$

où p_j est la probabilité qu'un système choisi de façon aléatoire dans l'ensemble soit dans l'état j d'énergie $E_j(N,V)$. L'équation (8.14) a une importance considérable puisque, si on tient compte du fait que l'énergie dépend de N et V, alors on peut poser :

$$Q(N,V,\beta) = \sum_j \exp(-\beta E_j(N,V)) \tag{8.15}$$

où $Q(N,V,\beta)$ est appelée fonction de partition. On montre que $\beta = \dfrac{1}{k_B T}$. On obtient l'équation (8.16) où $p_j(N,V,T)$ est appelée probabilité de Boltzmann.

$$p_j(N,V,T) = \frac{\exp\left(-\frac{E_j(N,V)}{k_B T}\right)}{Q(N,V,T)} \tag{8.16}$$

Ces relations ont été développées en considérant des sommations discrètes. À présent, nous allons considérer des sommations continues et remplacer l'énergie totale électronique par l'hamiltonien du système [17], qui ne dépend que des impulsions et des coordonnées de chaque particule présente dans le système. On peut alors écrire la fonction de partition et la probabilité de Boltzmann (équations (8.17) et (8.18)) :

$$Q(N,V,T) = \int\int \exp\left(-\frac{\mathcal{H}(\mathbf{r}^N,\mathbf{p}^N)}{k_B T}\right) d\mathbf{r}^N d\mathbf{p}^N \tag{8.17}$$

$$p_j(N,V,T) = \frac{\exp\left(-\frac{\mathcal{H}(\mathbf{r}^N,\mathbf{p}^N)}{k_B T}\right)}{\int\int \exp\left(-\frac{\mathcal{H}(\mathbf{r}^N,\mathbf{p}^N)}{k_B T}\right) d\mathbf{r}^N d\mathbf{p}^N} \tag{8.18}$$

8.4 Notion d'ensemble statistique

La notion d'ensemble apparaît naturellement en thermodynamique statistique. En effet, il s'agit d'une assemblée (fictive) constituée d'un très grand nombre de systèmes, tous identiques physiquement et tous situés dans le même contexte physique. Quatre contextes physiques jouent un rôle important.

Par définition, *l'ensemble microcanonique* est associé à la situation où le système physique est isolé, c'est-à-dire qu'il n'y a ni échange de particules ni échange d'énergie avec le système extérieur. Dans ces conditions, le système est caractérisé par les variables suivantes : le volume V du système, l'énergie totale E du système et le nombre N de particules. Ce cas très particulier sert en fait de base de raisonnement pour des situations moins triviales. Cette ensemble n'est pas l'ensemble naturel d'observations expérimentales. Ainsi, pour se rapprocher de l'expérience, on préfèrera travailler dans d'autres ensembles.

L'ensemble canonique (NVT) est caractérisé par les variables suivantes : le volume V du système, la température T et le nombre N de particules. Pour ce système, la fonction de partition est donnée par la relation (8.19), où la sommation discrète ou continue parcourt l'ensemble des configurations α du système.

$$Q(V,\beta,N) = \sum_{\alpha} \exp(-\beta E(\alpha)) \tag{8.19}$$

L'ensemble grand canonique (μVT) est caractérisé par le volume V constant, la température et le potentiel chimique constants. Dans ces conditions, la fonction de partition vaut :

$$\Xi(V,\beta,\mu) = \sum_{N=0}^{\infty} \sum_{\alpha_N} \exp(-\beta(E_N(\alpha_N) - \mu_N) \tag{8.20}$$

Dans *l'ensemble isobare-isotherme (NpT)*, le système est caractérisé par l'ensemble des variables suivantes : la pression p, la température T et le nombre total N de particules. la fonction de partition s'écrit

$$Q(p,\beta,N) = \sum_{V} \sum_{\alpha} \exp\left(-\beta\left(E_N(\alpha_N) + pV\right)\right) \tag{8.21}$$

8.5 Fonction de partition dans l'ensemble NVT

Le calcul de grandeurs thermodynamiques à partir des simulations est essentiellement basé sur les concepts de la thermodynamique statistique. En effet, c'est la connaissance des fonctions de partition qui permet le calcul de ces grandeurs.

Une fonction de partition est une quantité permettant de décrire la distribution des particules par rapport aux différents états quantiques possibles. Cela est rendu possible grâce à la distribution de Boltzmann. La fonction définie dans l'équation (8.19) peut être décomposée en une fonction de partition idéale correspondant au gaz parfait Q^{id}_{NVT} et une fonction de partition d'excès traduisant l'écart à l'idéalité Q^{ex}_{NVT}.

$$Q_{NVT} = Q^{id}_{NVT} Q^{ex}_{NVT}$$

Ces deux grandeurs s'expriment par les relations (8.22) et (8.23).

$$Q^{id}_{NVT} = \frac{1}{N!}\,\frac{1}{h^{3N}} \tag{8.22}$$

$$Q^{ex}_{NVT} = \int\int \exp\left(\frac{-\mathcal{H}(\mathbf{p}^N,\mathbf{r}^N)}{kT}\right) d\mathbf{r}^N d\mathbf{p}^N \tag{8.23}$$

Dans ces deux relations, N est le nombre total de molécules, h est la constante de Planck, $\mathcal{H}$ est l'hamiltonien du système qui s'exprime par l'équation (8.24), $\mathbf{p}$ est le vecteur impulsion, $\mathbf{r}$ le vecteur position, k la constante de Boltzmann, et T la température du système.

$$\mathcal{H}(\mathbf{p}^N, \mathbf{r}^N) = \sum_{i=1}^{N} \frac{|\mathbf{p}_i|^2}{2m} + \mathcal{U}(\mathbf{r}^N) \tag{8.24}$$

Nous voyons donc que l'hamiltonien du système peut se diviser en deux composantes :

- une partie cinétique $\sum_{i=1}^{N} \frac{|\mathbf{p}_i|^2}{2m}$
- une partie relative à l'énergie potentielle du système $\mathcal{U}(\mathbf{r}^N)$.

Le terme d'éxcès peut donc être à nouveau divisé en deux termes :

$$Q_{NVT}^{ex} = \int \exp\left(-\beta \sum_{i=1}^{N} \frac{|\mathbf{p}_i|^2}{2m}\right) d\mathbf{p}^N \int \exp\left(-\beta \mathcal{U}(\mathbf{r}^N)\right) d\mathbf{r}^N \tag{8.25}$$

Ici, β est la température inverse $\frac{1}{kT}$. Le terme de gauche est une intégrale indéfinie que l'on évalue par la relation (8.26)

$$\begin{aligned} \int_0^{+\infty} \exp\left(-ax^2\right) dx &= \frac{1}{2}\left(\frac{\pi}{a}\right)^{\frac{1}{2}} \\ \int_{-\infty}^{+\infty} \exp\left(-ax^2\right) dx &= \left(\frac{\pi}{a}\right)^{\frac{1}{2}} \end{aligned} \tag{8.26}$$

Pour les trois degrés de libertés et pour les N molécules, on peut écrire que :

$$Q_{NVT}^{ex} = [2mkT\pi]^{\frac{3N}{2}} \int \exp\left(-\beta \mathcal{U}(\mathbf{r}^N)\right) d\mathbf{r}^N \tag{8.27}$$

En insérant la relation (8.27) dans la relation (8.23), nous obtenons l'équation (8.28) qui se simplifie pour donner la relation (8.29)

$$Q_{NVT} = \frac{1}{N!} \frac{1}{h^{3N}} [2mkT\pi]^{\frac{3N}{2}} \int \exp\left(-\beta \mathcal{U}(\mathbf{r}^N)\right) d\mathbf{r}^N \tag{8.28}$$

$$Q_{NVT} = \frac{1}{N!}\left(\frac{[2mkT\pi]}{h^2}\right)^{\frac{3N}{2}} \int \exp\left(-\beta\mathcal{U}(\mathbf{r}^N)\right) d\mathbf{r}^N \tag{8.29}$$

Nous obtenons alors la relation finale

$$Q_{NVT} = \frac{1}{N!}\,\frac{1}{\Lambda^3} \int \exp\left(\frac{-\mathcal{U}(r^N)}{kT}\right) d\mathbf{r}^N \tag{8.30}$$

avec

$$\Lambda = \left(\frac{h^2}{2mkT\pi}\right)^{\frac{1}{2}} \tag{8.31}$$

où Λ est la longueur d'onde de De Broglie.

8.6 Fonction de partition dans l'ensemble NpT

8.6.1 Entropie microcanonique

D'après le postulat fondamental, les divers états microscopiques accessibles (état microscopique compatible avec les valeurs macroscopiques de paramètres extérieurs) sont équiprobables. Par conséquent, si le nombre de ces états est Ω, la valeur correspondante $S^\star$ ($\star$ signifie que l'on travaille dans l'ensemble microcanonique) de l'entropie statistique est :

$$S^\star = k \ln \Omega \tag{8.32}$$

8.6.2 Distribution statistique $p - T$

Équilibre thermique et mécanique entre deux systèmes

Considérons deux systèmes A et B séparés par un piston mobile perméable à la chaleur, l'ensemble forme un système isolé dont l'énergie E_0 et le volume V_0 sont fixés à δE_0 *et* δV_0 près . Les énergies E_A et E_B ainsi que les volumes V_A et V_B sont libres de fluctuer tant que les relations de l'équation (8.33) sont vérifiées. Dans ce cas, on montre facilement que l'équilibre thermique et l'équilibre mécanique sont réalisés si la température et la pression sont les mêmes dans les deux compartiments ($T_A^\star = T_B^\star,\ P_A^\star = P_B^\star$).

$$E_0 = E_A + E_B$$
$$V_0 = V_A + V_B \tag{8.33}$$

Système en contact avec un réservoir d'énergie et de volume

Supposons, à présent, que le système B soit beaucoup plus gros que A tant en nombre de molécules qu'en volume. Les variations de son énergie et de son volume par suite des échanges avec A n'affectent alors pratiquement pas sa température et sa pression. B constitue alors une sorte de réservoir $\mathcal{R}$ d'énergie et de volume qui impose sa température et sa pression. Soit T et p les paramètres imposés au système A. La probabilité d'observer pour A les caractéristiques indiquées, quel que soit par ailleurs l'état du réservoir, est proportionnelle au nombre d'états accessibles dans lequel A a le volume V (à dV près) et se trouve dans l'état ℓ. Ce nombre est donc simplement le nombre d'états accessibles au réservoir lorsque le volume est $(V_0 - V)$ à dV près et que l'énergie est $(E_0 - E_\ell)$. Par rapport à l'ensemble NVT, nous devons tenir compte ici de la variation de volume. Le volume devient donc une variable, comme indiqué dans la relation (8.34).

$$W_\ell(V)dV \propto \Omega_{\mathcal{R}}(E_0 - E_\ell(V), V_0 - V)dV \tag{8.34}$$

Ici, $W_\ell(V)dV$ représente la probabilité pour que le volume de A soit compris entre V et $V+dV$ et que son état microscopique soit ℓ d'énergie $E_\ell(V)$ lorsqu'il est en contact thermique et mécanique avec le réservoir à la température T et la pression p. L'expression précédente peut aussi s'écrire si on introduit l'entropie microcanonique par

$$W_\ell(V)dV \propto dV \exp \frac{1}{k} S^\star_{\mathcal{R}}(E_0 - E_\ell(V), V_0 - V)$$

Si $E_\ell(V)$ et V sont petits par rapport à E_0 et V_0 et en développant $S^\star_{\mathcal{R}}$, on obtient la relation suivante :

$$S^\star_{\mathcal{R}}(E_0 - E_\ell(V), V_0 - V) = S^\star_{\mathcal{R}}(E_0, V_0) - E_\ell(V) \left(\frac{\delta S^\star_{\mathcal{R}}}{\delta E}\right)_{E_0,V_0} - V \left(\frac{\delta S^\star_{\mathcal{R}}}{\delta V}\right)_{E_0,V_0} + ...$$

En utilisant les relations (8.35), nous obtenons l'équation (8.36).

$$\begin{aligned} \frac{1}{T} &= \frac{\delta S^\star_{\mathcal{R}}}{\delta E} \\ \frac{p}{T} &= \frac{\delta S^\star_{\mathcal{R}}}{\delta V} \end{aligned} \tag{8.35}$$

$$\begin{aligned} dH = TdS + VdP &= \frac{\delta H}{\delta S}dS + \frac{\delta H}{\delta P}dP \\ dS = \frac{dU}{T} + \frac{PdV}{T} &= \frac{\delta S}{\delta U}dU + \frac{\delta S}{\delta V}dV \end{aligned} \tag{8.36}$$

En substituant convenablement, on obtient l'éqution suivante :

$$\begin{aligned} S^\star_{\mathcal{R}}(E_0 - E_\ell(V), V_0 - V) = S^\star_{\mathcal{R}}(E_0, V_0) \\ -\frac{E_l(V)}{T} - p\frac{V}{T} \end{aligned}$$

Ainsi, la probabilité cherchée est proportionnelle à (en regroupant les facteurs indépendants de ℓ)

$$W_\ell(V)dV \propto dV\, C\, \exp\left(-\frac{1}{kT}(\frac{E_l(V)}{T} + p\frac{V}{T})\right)$$

où C est déterminé par normalisation de la probabilité

$$\sum_{(l)} W_\ell(V)dV = 1$$

C vaut donc :

$$C = \int \sum_l \exp\left(-\frac{1}{kT}(\frac{E_\ell(V)}{T} + p\frac{V}{T})\right) dV$$

C n'est autre que la fonction de partition dans l'ensemble NpT, et $W_\ell(V)dV$ la probabilité de Boltzmann dans ce même ensemble. Si on considère l'hamiltonien du système, on peut écrire :

$$Q_{NpT} = \int \exp\left(-\frac{pV}{kT}\right) dV \int\int \exp\left(-\frac{H_(\mathbf{r}^N, \mathbf{p}^N)}{k_B T}\right) d\mathbf{r}^N d\mathbf{p}^N$$

En reprenant le raisonnement de la section (8.5), nous obtenons la relation :

$$Q_{NpT} = \frac{1}{N!}\,\frac{1}{\Lambda^3} \int \exp\left(-\frac{pV}{kT}\right) dV \int \exp\left(-\frac{\mathcal{U}(\mathbf{r}^N)}{kT}\right) d\mathbf{r}^N \tag{8.37}$$

Dans cet ensemble statistique, le nombre de particules, la pression et la température sont maintenus constants. Dans ces conditions, seul le volume varie.

8.7 Fonctions d'état

Nous postulons que l'énergie moyenne de l'ensemble n'est autre que l'énergie mesurée du système, c'est-à-dire l'énergie interne U. Ainsi, on peut écrire que :

$$\begin{aligned} U_{NVT} =< E >_{NVT} &= \sum_i p_j(N,V,T) E_j(N,V) \\ &= \sum_i \frac{E_j(N,V) \exp\left(-\frac{E_j(N,V)}{kT}\right)}{Q_{NVT}} \end{aligned}$$

Il est possible d'écrire cette relation en fonction de Q_{NVT}. En effet, si nous dérivons le terme $\ln Q_{NVT}$ par rapport à T en gardant V et N constant, nous obtenons :

$$\begin{aligned} \left(\frac{\partial \ln Q_{NVT}}{\partial T}\right)_{N,V} &= \frac{1}{Q_{NVT}} \left(\frac{\delta Q_{NVT}}{\delta T}\right) \\ &= \frac{1}{Q_{NVT}} \sum_i E_j(N,V) \frac{1}{kT^2} \exp\left(-\frac{E_j(N,V)}{kT}\right) \\ &= \frac{1}{kT^2} \sum_i \frac{E_j(N,V) \exp\left(-\frac{E_j(N,V)}{kT}\right)}{Q_{NVT}} \end{aligned}$$

$$Q_{NVT} = \sum_i \exp\left(-\frac{E_j(N,V)}{kT}\right)$$

On a obtient alors la relation :

$$U_{NVT} = kT^2 \left(\frac{\partial \ln Q_{NVT}}{\partial T} \right)_{N,V} \tag{8.38}$$

Or

$$\begin{aligned} \frac{\partial}{\partial T}\left(\frac{A}{T}\right) &= -\frac{A}{T^2}\frac{1}{T}\frac{\partial A}{\partial T} \\ &= -\frac{A}{T^2} - \frac{S}{T} \\ &= -\frac{U}{T^2} \end{aligned}$$

On a alors :

$$\left(\frac{A}{T}\right) = -\int \frac{U}{T^2} dT$$

et donc :

$$A_{NVT} = -kT \ln Q_{NVT} \tag{8.39}$$

Dans l'ensemble NpT, les relations pour G_{NpT} et H_{NpT} sont :

$$\begin{aligned} G_{N,p,T} &= -kT \ln Q_{NpT} \\ S_{N,p,T} &= \left(-\frac{\partial G(N,p,T)}{\partial T} \right)_{N,p} \\ &= k \ln Q_{NpT} + kT \left(\frac{\partial \ln Q_{NpT}}{\partial T} \right)_{N,p} \\ H_{N,p,T} &= G_{NpT} + TS_{NpT} \\ &= kT^2 \left(\frac{\partial \ln Q_{NpT}}{\partial T} \right)_{N,p} \end{aligned}$$

Chapitre 9

Simulation numérique de type dynamique moléculaire

9.1 Déterminisme et imprédictibilité 195
9.1.1 Définition 195
9.1.2 Sensibilité aux conditions initiales 196
9.1.3 Formalisme de Liouville 197
9.2 Algorithme de Berendsen 201
9.3 Sommation d'Ewald 213
9.3.1 Calcul dans l'espace de Fourier 215
9.3.2 Calcul dans l'espace rèel 218
9.3.3 Terme d'auto-interaction et terme intramoléculaire 221
9.4 Algorithme de Contrainte 223
9.4.1 Définition 223
9.4.2 Multiplicateurs de Lagrange 224
9.4.3 Méthode SHAKE 226
9.5 Méthode de non équilibre 235

9.1. Déterminisme et imprédictibilité

9.1.1 Définition

On dit qu'un système mécanique est *déterministe* lorsque la connaissance des forces qui agissent sur lui et de son état à un instant quelconque permet de déterminer son état à tout autre instant. L'état donné est le plus souvent l'état initial, caractérisé par les valeurs initiales des degrés de liberté et des moments conjugués, mais cela n'est pas une condition

nécessaire. La connaissance de l'état final permettrait de restituer les états antérieurs (principe du renversement du temps). Néanmoins, l'irréversibilité fondamentale de l'évolution de tout phénomène physique réel en accord avec le deuxième principe de la thermodynamique impliquant une orientation du temps du passé vers le futur, empêche ce retour dans le passé. Ainsi, le déterminisme implique que tout soit connu : passé, présent et futur.

9.1.2 Sensibilité aux conditions initiales

Pour démontrer la sensibilité aux conditions initiales, prenons le cas simple d'un pendule évoluant sans frottement, dans un champ de pesanteur. L'hamiltonien, qui s'identifie dans le cas à l'énergie mécanique, s'écrit :

$$\mathcal{H} = \frac{p_\theta^2}{2ml^2} - mglcos\theta \quad \text{et} \quad p_\theta = \left[2ml^2(\mathcal{H} + mglcos\theta)\right]^{1/2}$$

Deux cas peuvent se produire :

1- $-mgl < \mathcal{H} < mgl$. Le pendule oscille entre deux valeurs opposées de l'angle, θ_1 et $-\theta_1$ tel que $\theta_1 \arccos(-\mathcal{H}/mgl)$, ce qui traduit l'annulation de p_θ.

2- $-mgl < \mathcal{H} \geq mgl$. Le moment conjugué p_θ est une fonction périodique de θ, qui ne s'annule jamais. Dans ce cas, le mouvement est révolutif.

On distingue donc deux mouvements dans l'espace des phases, l'un qui est oscillatoire et l'autre révolutif. On voit que les conditions initiales permettent de déterminer sans ambiguïté le type de mouvement du pendule dans l'espace des phases. Ce cas simple est un bon exemple de *déterminisme*. Or, la nature est bien plus complexe que ce pendule. En effet, pour la plupart des systèmes dynamiques en physique et/ou en chimie, le nombre de constantes du mouvement est inférieur au nombre de degrés de liberté. Cela se traduit, dans l'espace des phases, par une complexité des courbes puisque l'on observe une *imbrication* des zones correspondantes aux différents types de mouvement. On comprend alors qu'une légère imprécision sur les données des conditions initiales a une grande influence sur le résultat obtenu au point de rendre l'analyse *imprédictive*, bien que le système soit déterministe. Des conditions initiales très proches peuvent donner des résultats différents. De tels systèmes, déterministes mais imprédictifs car sensibles aux conditions initiales, sont dits *chaotiques*. C'est H. Poincaré, en

1892, qui a mis en évidence cette imprédictibilité par l'étude de l'instabilité des systèmes astrophysiques et la nécessité de nouvelles analyses. D'une manière générale, les sources du chaos sont attribuées aux termes non linéaires dans les équations qui régissent le mouvement. Ces non-linéarités sont par exemples, le frottement visqueux proportionnel au carré de la vitesse (forme quadratique dans le cas de forces exercées par des fluides en mouvement), ou bien encore les forces en $1/r^2$. Il est possible de démontrer ce lien entre les caractéristiques d'un système à l'instant t et ceux à l'instant $t = 0$. Pour cela, nous devons utiliser le formalisme de Liouville.

9.1.3 Formalisme de Liouville

La distribution des points de l'espace des phases est décrite par une distribution de probabilité $\mathcal{W}^N(\mathbf{r}^N, \mathbf{p}^N, t)$ où N est le nombre de particules du système considéré, $\mathbf{r}^N$ représente l'ensemble des coordonnées spatiales et $\mathbf{p}^N$ les quantités de mouvement. L'évolution temporelle de cette distribution est donnée par l'équation de Liouville. Celle-ci est déterminée en considérant que la probabilité «s'écoule» dans l'espace des phases comme un fluide de densité $\mathcal{W}$ obéissant à la loi de conservation locale.

$$\frac{\partial \mathcal{W}}{\partial t} + div(\mathcal{W}) \tag{9.1}$$

Démonstration de la loi de conservation locale

Cette démonstration débute en rappelant le caractère conservatif de la masse. En mécanique newtonnienne, la masse d'un système est conservative, c'est-à-dire que dans le bilan de masse d'un système, entre les instants voisins t et $t + \delta t$, le terme de production de masse est nul (équation (9.2)).

$$dM = \delta M^r + \delta M^p \quad \text{avec } \delta M^p = 0 \tag{9.2}$$

Autrement dit, la variation de masse d'un système ne peut être attribuée qu'à la masse reçue à travers la surface qui la délimite. Ce caractère conservatif de la masse ayant été introduit, nous pouvons définir le *vecteur courant volumique de masse* comme le vecteur donné par la l'équation (9.3) :

$$\mathbf{J}_m = \rho \mathbf{v} \tag{9.3}$$

où ρ est la masse volumique du système et $\mathbf{v}$ la vitesse de la particule P au point A où elle se trouve. Le débit de masse q_m est la masse de la quantité qui traverse toute surface pendant l'unité de temps. Soit dS un élément de surface de normale $\mathbf{n}$. Pendant la durée élémentaire dt, la quantité de matière qui traverse la surface élémentaire dS, est contenue dans un cylindre de base dS et de hauteur $\mathbf{v}dt\mathbf{n} = v_n dt$. Par conséquent, le débit de masse élémentaire δq_m est :

$$\delta q_m = \frac{\delta m}{dt} = \frac{\rho dV}{dt} = \frac{\rho v_n dt dS}{dt} = \rho v_n dS = \mathbf{J}_m \mathbf{n} dS$$

Il en résulte que pour une surface quelquonque S, on obtient :

$$q_m = \int_S \mathbf{J}_m \mathbf{n} dS = \int_S \rho v_n \mathbf{n} dS$$

Dans ces conditions, le débit de masse à travers une surface est le flux du courant volumique de masse à travers cette surface. Considérons un système de volume V délimité par une surface S fixe. La masse étant une grandeur conservative, sa variation pendant la durée dt est due à la masse qui pénètre dans le système à travers S pendant cette même durée :

$$dM = \delta M^r \quad \text{avec} \quad dM = d\int_V \rho dV \quad \text{et} \quad \delta M^r = dt \int_S \delta q_m = -dt \int_S \mathbf{J}_m \mathbf{n} dS$$

Le signe $-$ vient du fait que $\mathbf{J}_m$ est dirigé vers l'extérieur du système, or nous évaluons la masse entrante donc de signe opposé à $\mathbf{n}$ qui est la normale sortante. À présent, si on utilise la formule d'Ostrogradsky en permutant les opérateurs d/dt et intégration, nous obtenons :

$$\begin{aligned} dM &= \delta M^r \rightarrow d\int_V \rho dV = -dt \int_S \mathbf{J}_m \mathbf{n} dS = -dt d \int_V div(\mathbf{J}_m) \\ &\rightarrow \int_V (\frac{\partial \rho}{\partial t} + div(\mathbf{J}_m)) dV = 0 \end{aligned}$$

Cette relation se simplifie en l'équation (9.4) qui est l'équation de conservation locale.

$$\frac{\partial \rho}{\partial t} + div(\mathbf{J}_m) = 0 \leftrightarrow \frac{\partial \rho}{\partial t} + div(\rho \mathbf{v}) = 0$$

Propriété de la densité de probabilité.

Cette démonstration permet de justifier l'écriture de la relation (9.1). L'analogie avec $\mathcal{W}$ nous permet d'écrire si on développe l'opérateur divergence :

$$\frac{\partial \mathcal{W}}{\partial t} + \sum_{i=1}^{N} \left[\frac{\partial}{\partial \mathbf{r}_i}(\mathcal{W}\dot{\mathbf{r}}_i) + \frac{\partial}{\partial \mathbf{p}_i}(\mathcal{W}\dot{\mathbf{p}}_i) \right] = 0$$

soit

$$\frac{\partial \mathcal{W}}{\partial t} + \sum_{i=1}^{N} \left[\dot{\mathbf{r}}_i \frac{\partial \mathcal{W}}{\partial \mathbf{r}_i} + \dot{\mathbf{p}}_i \frac{\partial \mathcal{W}}{\partial \mathbf{p}_i} + \mathcal{W} \left(\frac{\partial \dot{\mathbf{r}}_i}{\partial \mathbf{r}_i} + \frac{\partial \dot{\mathbf{p}}_i}{\partial \mathbf{p}_i} \right) \right] = 0$$

Or, selon les équations canoniques du mouvement (relations (2.12)), le deuxième terme s'annule :

$$\frac{\partial \dot{\mathbf{r}}_i}{\partial \mathbf{r}_i} + \frac{\partial \dot{\mathbf{p}}_i}{\partial \mathbf{p}_i} = \frac{\partial}{\partial \mathbf{r}_i} \left(\frac{\partial \mathcal{H}}{\partial \mathbf{p}_i} \right) - \frac{\partial}{\partial \mathbf{p}_i} \left(\frac{\partial \mathcal{H}}{\partial \mathbf{r}_i} \right) = 0$$

Nous obtenons ainsi l'équation de *Liouville* :

$$\frac{\partial \mathcal{W}}{\partial t} + \sum_{i=1}^{N} \left[\dot{\mathbf{r}}_i \frac{\partial \mathcal{W}}{\partial \mathbf{r}_i} + \dot{\mathbf{p}}_i \frac{\partial \mathcal{W}}{\partial \mathbf{p}_i} \right] = 0 \tag{9.4}$$

Cette relation s'écrit en considérant les équations de mouvement :

$$\frac{\partial \mathcal{W}}{\partial t} + \sum_{i=1}^{N} \left[\frac{\partial \mathcal{H}}{\partial \mathbf{p}_i} \frac{\partial \mathcal{W}}{\partial \mathbf{r}_i} - \frac{\partial \mathcal{H}}{\partial \mathbf{r}_i} \frac{\partial \mathcal{W}}{\partial \mathbf{p}_i} \right] = 0 \tag{9.5}$$

Cette équation représente finalement l'évolution temporelle de la distribution $\mathcal{W}$. Définissions l'opérateur de Liouville :

$$i\mathcal{L} = \sum_{i=1}^{N} \left[\frac{\partial \mathcal{H}}{\partial \mathbf{p}_i} \frac{\partial}{\partial \mathbf{r}_i} - \frac{\partial \mathcal{H}}{\partial \mathbf{r}_i} \frac{\partial}{\partial \mathbf{p}_i} \right] \tag{9.6}$$

Cela nous permet de réécrire la relation (9.4) :

$$\frac{\partial \mathcal{W}}{\partial t} + i\mathcal{L}\mathcal{W} = 0 \tag{9.7}$$

Cette équation est résolue par séparation des variables, ce qui permet d'obtenir :

$$\mathcal{W}(\mathbf{r}^N, \mathbf{p}^N, t) = \exp(-i\mathcal{L}t)\mathcal{W}(\mathbf{r}^N, \mathbf{p}^N, 0) \tag{9.8}$$

Ainsi, le formalisme de Liouville[1] a permis de montrer clairement la dépendance aux conditions initiales dont on a précédement discutée.

Pour terminer avec le formalisme de Liouville, nous devons noter l'élégante manière de démontrer la conservation du volume (un algorithme respectant cette conservation est dit symplectique). Appliquons pour cela la relation relative à la conservation locale du volume :

$$\frac{\partial V}{\partial t} + \sum_{i=1}^{N} \left[\frac{\partial \mathcal{H}}{\partial \mathbf{p}}_i \frac{\partial V}{\partial \mathbf{r}_i} - \frac{\partial \mathcal{H}}{\partial \mathbf{r}}_i \frac{\partial V}{\partial \mathbf{p}_i} \right] = 0 \tag{9.9}$$

Comme V est indépendant de $\mathbf{r}_i$ et $\mathbf{p}_i$, nous obtenons alors :

$$\frac{\partial V}{\partial t} = 0 \Rightarrow V = Cte \tag{9.10}$$

Le volume V lié à un espace $\mathbf{r}^N, \mathbf{p}^N$ doit être constant au cours du temps, ainsi tout algorithme de propagation doit être symplectique. On peut aussi démontrer cette relation en utilsant les volumes infinétisimaux comme le fait Diu dans son ouvrage [66]. Ce dernier utilise alors un changement de variable limité à l'ordre un.

$$q_i^i = q_i + \dot{q}_i \delta t \quad \text{et} \quad p_i^i = p_i + \dot{p}_i \delta t \tag{9.11}$$

Le calcul de $d\tau'$ à partir de $d\tau$ se ramène à un changement de variable multiple tel que :

[1] Il est possible de retrouver les équations de l'algorithme de Verlet en utilisant l'identité de Trotter puisque les opérateurs $\mathcal{L}_r$ et $\mathcal{L}_p$ ne commutent pas (le développement complet de la démonstration est fait dans l'ouvrage *Understanding Molecular Simulation - From Algorithms to Applications* [35]).

$$d\tau = \mathcal{D}_J d\tau^{'}$$

où $\mathcal{D}_J$ est le déterminant fonctionnel ou *jacobien* tel que :

$$\mathcal{D}_J = \begin{vmatrix} \frac{\partial q_1^{'}}{\partial q_1} & \cdots & \cdots & \frac{\partial q_1^{'}}{\partial q_N} & \cdots & \cdots & \frac{\partial q_1^{'}}{\partial p_N} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \frac{\partial q_N^{'}}{\partial q_1} & \cdots & \cdots & \frac{\partial q_N^{'}}{\partial q_N} & \cdots & \cdots & \frac{\partial q_N^{'}}{\partial p_N} \\ \frac{\partial p_1}{\partial q_1} & \cdots & \cdots & \frac{\partial p_1}{\partial q_N} & \cdots & \cdots & \frac{\partial p_1}{\partial p_N} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \frac{\partial p_N^{'}}{\partial q_1} & \cdots & \cdots & \frac{\partial p_N^{'}}{\partial q_N} & \cdots & \cdots & \frac{\partial p_N^{'}}{\partial p_N} \end{vmatrix}$$

À partir de l'opérateur de Liouville et des équations canoniques, on démontre aisément [66] que $\mathcal{D}_J = 1$. Ce changement de variable est aussi applicable à l'intégrale :

$$\int\int\int d\tau = \int\int\int \mathcal{D}_J d\tau^{'} \quad \text{comme} \quad \mathcal{D}_J = 1 \quad \Rightarrow V = V^{'}$$

Pour résumer, un bon algorithme de résolution des équations de mouvement doit :

1- être stable (ne pas subir de dérive énergétique aux temps longs)
2- être un bon compromis entre temps et précision de calcul
3- conserver l'énergie mécanique et les moments conjugués
4- être symplectique
5- respecter le renversement du temps.

9.2. L'algorithme de Berendsen

Précédemment, nous avons vu que les algorithmes de résolution des équations du mouvement doivent conserver l'énergie mécanique et le volume. Comme pour tous les systèmes étudiés, le nombre de particules est constant (système isolé sans création ni perte de matière), on se trouve donc dans l'ensemble statistique NVE. Or, cet ensemble n'est pas l'ensemble naturel d'observation. Ainsi, pour se rapprocher de la réalité, on préfèrera travailler dans d'autres ensembles tels que NVT et NpT. En outre, les systèmes naturels sont non-isolés,

principalement en équilibre avec un thermostat comme l'air ambiant. Cela nécessite donc une formulation différente des algorithmes précédemment établis afin de maintenir la température et/ou le volume et/ou la pression constante. La méthode la plus simple pour maintenir la température à une température T_o constante consiste à mettre à l'échelle les vitesses du système, c'est-à-dire de les multiplier par le facteur :

$$\sqrt{\frac{T}{T_o}} \text{ tel que } \left\langle \sum_{i=1}^{N} \frac{\mathbf{p}_i^2}{m_i} \right\rangle = 2 \langle \mathcal{K} \rangle = 3Nk_BT$$

Cette méthode respecte la distribution des vitesses de Maxwell-Boltzmann. Mais, il est difficile de lui trouver une interprétation physique. Des méthodes plus rigoureuses existent et sont très bien décrites dans l'ouvrage «*Computer Simulations of Liquids*» de Allen et Tildesley [18]. En fait, on peut décomposer l'ensemble des méthodes en deux sous-ensembles. Dans le premier sous-ensemble, on ajoute des termes supplémentaires à l'hamiltonien comme le volume qui est alors considéré comme une variable dynamique [67] (NpT), ou bien encore un degré de liberté s supplémentaire pour représenter le bain thermique. Cette variable s a pour effet de renormaliser le temps [68] (NVT). Pour le deuxième sous-ensemble, on ajoute des termes supplémentaires aux équations de Newton tels que des termes de friction et/ou stochastiques [10, 69, 70]. Néanmoins, les algorithmes les plus utilisés dans l'ensemble NpT sont ceux de Berendsen [10] et de Hoover et Nosé [68, 69]. En ce qui nous concerne, nous utiliserons celui de Berendsen pour les ensembles NpT et NVT puisqu'il modifie légèrement les vitesses et oscille faiblement autour de la température désirée. Dans l'ensemble canonique, l'énergie fluctue car on a un échange de chaleur avec un thermostat, si bien que l'énergie devient une variable interne et non plus un paramètre extérieur. Dans ce cas, c'est la température qui est constante et qui devient un paramètre extérieur. Par conséquent, un bon algorithme NVT doit minimiser les fluctuations de température. La méthode des contraintes mise au point par Andersen [67] et modifiée pour être utilisable avec l'algorithme de *Verlet Leap Frog* par Brown [70] est une méthode dont les fluctuations de température sont plus importantes. Néanmoins, cette méthode, si elle est modifée pour être utilisable avec l'algorithme de Gear, *Predictor Corrector* (plus stable que l'algorithme de *Verlet Leap Frog*), et si le terme de friction est évalué de manière explicite avec une mise à l'échelle, permet de faibles modifications des vitesses et un maintien de la température beaucoup plus efficace. L'algorithme utilisé étant celui de

Berendsen, nous allons discuter davantage l'origine du facteur de mise à l'échelle. La démonstration ayant été réalisée dans l'article [10], nous reviendrons juste sur la partie «silencieuse» de celle-ci. Cet algorithme permet de coupler le système et le thermostat par l'intermédiaire d'un terme stochastique $R(t)$ et d'un terme dissipatif $-m_i\gamma_i v_i$. Le formalisme de Lagrange et de Newton ne permet pas de traiter les forces de frottement qui sont fonctions de la vitesse. Les forces dissipatives sont plutôt du domaine de la mécanique des milieux continus. Cependant, le formalisme lagrangien permet de traiter les systèmes dissipatifs simples par un artifice. En effet, considérons un système perdant de l'énergie par frottement, effet Joule ou tout autre processus, l'artifice consiste à le coupler de manière appropriée à un système mirroir fictif qui absorbe formellement cette énergie, de telle manière que l'énergie totale des deux systèmes reste invariante au cours du temps. Dans l'algorithme de Berendsen, le terme miroir n'est autre que le terme stochastique, ainsi l'équation de mouvement s'écrit :

$$m_i\dot{\mathbf{v}}_i = \mathbf{f}_i - m_i\gamma_i\mathbf{v}_i + R(t)$$

$R(t)$ est la variable stochastique liée au théorème de fluctuation-dissipation (qui se démontre en résolvant l'équation de Langevin) :

$$\langle R_i(t)R_j(t+\tau)\rangle = 2m_i\gamma_i k_B T_o\delta(\tau)\delta_{ij}$$

Ici δ_{ij} est l'indice de Kronecker, γ_i le coefficient de friction, τ représente l'échelle de temps caractéristique de l'évolution de la particule ou bien le temps à partir duquel la vitesse moyenne de la particule s'amortit, $\delta(\tau)$ est la fonction de Dirac, résultat du bruit blanc. Le point de départ de la démonstration est le *principe de moindre action* proposé par Lagrange et conduisant aux équations de Lagrange-Euler. Avant de détailler l'algorithme de Berendsen, un autre rappel de mécanique analytique est nécessaire. Précédemment, nous avons donné l'expression de lagrangien, que nous rappelons ici :

$$\mathcal{L}\left(\{x\},\{\dot{x}\},t\right) = \mathcal{E}_k\left(\{\dot{x}\},t\right) - U\left(\{x\},t\right)$$

Pour toute trajectoire $x(t)$, on définit l'*action* S par l'intégrale

$$S = \int_{t_1}^{t_2} \mathcal{L}\left(\{x\}, \{\dot{x}\}, t\right) dt \tag{9.12}$$

Le *principe de moindre action* dit que la trajectoire physique effectivement suivie $X(t)$ est telle que S est minimale (extrémale, en général). Finalement, ce principe de moindre action n'est autre qu'une formulation différente du célèbre *principe variationnel* énoncé par Hamilton en 1834, analogue au *principe d'économie naturelle* de Fermat. Ce principe de moindre action est, en fait, le résultat d'une nouvelle approche des problèmes de mécanique. Au lieu de déterminer la position et la vitesse d'une particule à l'instant t, connaissant son état initial, Lagrange se demande quelle est la trajectoire effectivement suivie par la particule si, partant de x_1 à l'instant t_1, elle arrive en x_2 à l'instant t_2. Soit $X(t)$ la vraie trajectoire et $x(t)$ une trajectoire voisine de $X(t)$ partant elle aussi de x_1 à t_1 et aboutissant à x_2 à t_2, on a alors :

$$\begin{aligned} x(t) \quad &= \quad \mathcal{X}(t) + \delta x(t), \quad \dot{x}(t) = \dot{\mathcal{X}}(t) + \delta\dot{x}(t), \quad \delta\dot{x}(t) = \frac{d}{dt}\delta x(t) \\ &\qquad \text{en considérant que } \quad \delta x(t_1) = \delta x(t_2) = 0 \end{aligned}$$

Au premier ordre en δx, la variation de S est :

$$\delta S = \int_{t_1}^{t_2} \left[\left(\frac{\partial \mathcal{L}}{\partial x} \right) \delta x(t) dt + \frac{\partial \mathcal{L}}{\partial \dot{x}} \delta\dot{x}(t) \right] dt$$

Si on intègre le deuxième terme par parties ($U = \partial\mathcal{L}/\partial\dot{x}$ et $\dot{V} = \delta\dot{x}(t)$) et en tenant compte du fait que $\delta x(t_1) = \delta x(t_2) = 0$, on obtient :

$$\delta S = \int_{t_1}^{t_2} \left(\frac{\partial \mathcal{L}}{\partial x} - \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{x}} \right) \right) \delta x(t) dt$$

Le principe de moindre action affirme que δS doit être nulle quelle que soit la variation infinitésimale $\delta x(t)$. Par conséquent, l'équation de mouvement, c'est-à-dire l'équation qui détermine la trajectoire effectivement suivie, est l'équation de *Lagrange-Euler* :

$$\frac{\partial \mathcal{L}}{\partial x} = \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{x}} \right)$$

Remarque : avec cette équation de mouvement généralisée nous retrouvons l'équation de mouvement habituelle donnée par Newton.

À présent, généralisons cette équation pour un système constitué de $3N$ degrés de liberté $(\mathbf{r}_i, \dot{\mathbf{r}}_i)$, $\quad i = 1, ..., N \quad \mathbf{r} = x\mathbf{e}_x + y\mathbf{e}_y + z\mathbf{e}_z$. Le lagrangien est une fonction $\mathcal{L}\left(\{\mathbf{r}_i\}, \{\dot{\mathbf{r}}_i\}, t\right)$ des variables $\{\mathbf{r}_i\}$ et $\{\dot{\mathbf{r}}_i\}$ et les équations de mouvement sont données par l'ensemble des équations de mouvement de *Lagrange-Euler* :

$$\frac{\partial \mathcal{L}}{\partial \mathbf{r}_i} = \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{\mathbf{r}}_i}\right) \quad i = 1, ..., 3N$$

Le Lagrangien s'écrit comme :

$$\mathcal{L}\left(\{\mathbf{r}_i\}, \{\dot{\mathbf{r}}_i\}\right) = \sum_i^N \frac{\dot{\mathbf{r}}_i \dot{\mathbf{r}}_i m_i}{2} - \sum_i^N \sum_{j>i} u(r_{ij}) = \mathcal{E}_k\left(\{\dot{\mathbf{r}}_i\}, t\right) - U\left(\{\mathbf{r}\}, t\right)$$

Par conséquent, l'équation de *Lagrange-Euler* peut s'écrire :

$$\frac{\partial \mathcal{L}}{\partial \mathbf{r}_i} = \frac{d\mathcal{E}_k\left(\{\dot{\mathbf{r}}_i\}, t\right)}{dt} \quad i = 1, ..., 3N \tag{9.13}$$

L'équation (9.13) est le point de départ de l'algorithme de Berendsen. Il évalue cette quantité par un taux d'accroissement :

$$\frac{d\mathcal{E}_k\left(\{\dot{\mathbf{r}}_i\}, t\right)}{dt} = \frac{d\mathcal{E}_k}{dt} = \lim_{\Delta t \to 0}\left[\left(\sum_{i=1}^{3N} \frac{1}{2} m_i v_i^2(t + \Delta t) - \sum_{i=1}^{3N} \frac{1}{2} m_i v_i^2(t)\right) / \Delta t\right] \tag{9.14}$$

Étant donné la longueur de la démonstration, je ne donnerai que les grandes lignes. L'utilisation de développement de Taylor à l'ordre 2 des vitesses permet de reformuler la relation (9.14), comme :

$$\frac{d\mathcal{E}_k}{dt} = \sum_{i=1}^{3N} \frac{1}{2} m_i \left[\lim_{\Delta t \to 0} \frac{\Delta v_i^2}{\Delta t}\right] + \sum_{i=1}^{3N} m_i \left[v_i(t)\dot{v}_i(t)\right] \tag{9.15}$$

tel que $\Delta v_i = v_i(t + \Delta t) - v_i(t) = \int_t^{t+\Delta t} \dot{v}_i(t')dt'$

soit en intégrant l'équation de Langevin :

$\Delta v_i = \frac{1}{m}\int_t^{t+\Delta t} \left[f_i(t') - m_i\gamma_i v_i(t') + R(t')\right] dt'$

Le terme au carré peut aussi se développer comme :

$$\begin{aligned}\Delta v_i^2 &= \int_t^{t+\Delta t}\int_t^{t+\Delta t} v_i(t^{'})v_i(t^{''})dt^{'}dt^{''} \qquad (9.16)\\ &= \frac{1}{m_i^2}\int_t^{t+\Delta t}\int_t^{t+\Delta t}\Big[f_i(t^{'})f_i(t^{''}) - f_i(t^{'})m_i\gamma_i v_i(t^{''}) + m_i^2\gamma_i^2 v_i(t^{'})v_i(t^{''}) - f_i(t^{''})m_i\gamma_i v_i(t^{'})\\ &+ f_i(t^{'})R(t^{''}) - m_i\gamma_i v_i(t^{'})R(t^{'}) + R(t^{'})f_i(t^{''}) - R(t^{'})v_i(t^{''})m_i\gamma_i + R(t^{'})R(t^{''})\Big]dt^{'}dt^{''}\end{aligned}$$

L'équation (9.16) met en évidence deux parties : une partie stochastique T_s et une partie non stochastique T_{ns}. Le terme non stochastique qui contient les termes de vitesses et de forces dont l'intégrale peut être discrétisée (puisque Δt est suffisamment petit) est donné par :

$$T_{ns} = \frac{1}{m_i^2}\left[f_i(\chi_1)f_i(\chi_2) - f_i(\chi_1)m_i\gamma_i v_i(\chi_2) + m_i^2\gamma_i^2 v_i(\chi_1)v_i(\chi_2) - f_i(\chi_2)m_i\gamma_i v_i(\chi_1)\right]\Delta t^2$$

avec

$t < \chi_1, \chi_2 < t + \Delta t$

Dans la mesure où ce terme intervient à l'intérieur d'un taux d'accroissement tel que $\delta t \to 0$, il s'annule pour ne laisser que le terme stochastique. Cela nous permet d'écrire :

$$\begin{aligned}\lim_{\Delta t\to 0}\frac{1}{\Delta t}\sum_{i=1}^{3N}\frac{1}{2}m_i\Delta v_i^2 &= \lim_{\Delta t\to 0}\frac{1}{\Delta t}\sum_{i=1}^{3N}\frac{1}{2m_i}\left[\int_t^{t+\Delta t}\int_t^{t+\Delta t}\Big(f_i(t^{'})R(t^{''}) - m_i\gamma_i v_i(t^{'})R(t^{'})\right.\\ &+ \left. R(t^{'})f_i(t^{''}) - R(t^{'})v_i(t^{''})m_i\gamma_i + R(t^{'})R(t^{''})\Big)dt^{''}dt^{'}\right] \qquad (9.17)\end{aligned}$$

Comme les forces et les vitesses ne sont pas corrélées à la force aléatoire, le produit est donc nul, si bien que l'équation (9.17) peut s'écrire :

$$\lim_{\Delta t\to 0}\frac{1}{\Delta t}\sum_{i=1}^{3N}\frac{1}{2}m_i\Delta v_i^2 = \lim_{\Delta t\to 0}\frac{1}{\Delta t}\sum_{i=1}^{3N}\frac{1}{2m_i}\left[\int_t^{t+\Delta t}\int_t^{t+\Delta t} R(t^{'})R(t^{''})dt^{''}dt^{'}\right]$$

Le théorème de fluctuation-dissipation nous permet d'écrire que :

$$\lim_{\Delta t\to 0}\frac{1}{\Delta t}\sum_{i=1}^{3N}\frac{1}{2}m_i\Delta v_i^2 = \lim_{\Delta t\to 0}\frac{1}{\Delta t}\sum_{i=1}^{3N}\frac{1}{2m_i}\left[\int_t^{t+\Delta t}\int_t^{t+\Delta t} 2m_i\gamma_i k_B T_o\delta(t^{''} - t^{'})dt^{''}dt^{'}\right]$$

Soit, après simplification,

$$\lim_{\Delta t\to 0}\frac{1}{\Delta t}\sum_{i=1}^{3N}\frac{1}{2}m_i\Delta v_i^2=\lim_{\Delta t\to 0}\frac{1}{\Delta t}\sum_{i=1}^{3N}\left[\int_t^{t+\Delta t}\int_t^{t+\Delta t}\gamma_i k_B T_o\delta(t^{''}-t^{'})dt^{''}dt^{'}\right]$$

Si on choisit un terme de friction égal pour toutes les particules constituant le système, $\gamma_i=\gamma$, on peut alors écrire :

$$\sum_{i=1}^{3N}k_B T_o\delta(t^{''}-t^{'})\gamma_i=k_B T_o\delta(t^{''}-t^{'})\sum_{i=1}^{3N}\gamma_i=3N\gamma k_B T_o\delta(t^{''}-t^{'})$$

On a alors :

$$\lim_{\Delta t\to 0}\frac{1}{\Delta t}\sum_{i=1}^{3N}\frac{1}{2}m_i\Delta v_i^2=\lim_{\Delta t\to 0}\frac{1}{\Delta t}\left[\int_t^{t+\Delta t}\int_t^{t+\Delta t}3N\gamma k_B T_o\delta(t^{''}-t^{'})dt^{''}dt^{'}\right]$$

Finalement, il ne reste plus que le terme suivant à évaluer :

$$\int_t^{t+\Delta t}\int_t^{t+\Delta t}\delta(t^{''}-t^{'})dt^{''}dt^{'}$$

Une des propriétés de la fonction de Dirac est que :

$$\int_a^b f(t)\delta(t-\tau)dt=f(\tau)$$

Dans ces conditions, on peutt écrire :

$$\lim_{\Delta t\to 0}\frac{1}{\Delta t}\sum_{i=1}^{3N}\frac{1}{2}m_i\Delta v_i^2=\lim_{\Delta t\to 0}\frac{1}{\Delta t}\left[\int_t^{t+\Delta t}3N\gamma k_B T_o dt^{''}\right]=\lim_{\Delta t\to 0}\frac{1}{\Delta t}\left[3N\gamma k_B T_o\Delta t\right]$$

et donc :

$$\lim_{\Delta t\to 0}\frac{1}{\Delta t}\sum_{i=1}^{3N}\frac{1}{2}m_i\Delta v_i^2=3N\gamma k_B T_o \tag{9.18}$$

À présent, multiplions l'équation de Langevin par v_i et sommons sur l'ensemble des atomes :

$$\sum_{i=1}^{3N} m_i v_i \dot{v}_i = \sum_{i=1}^{3N} v_i f_i - \sum_{i=1}^{3N} m_i v_i^2 \gamma + \sum_{i=1}^{3N} R_i(t) v_i$$

Le dernier terme étant nul, on a :

$$\sum_{i=1}^{3N} m_i v_i \dot{v}_i = \sum_{i=1}^{3N} v_i f_i - 2\gamma \mathcal{E}_k \tag{9.19}$$

On peut écrire en reprenant l'équation (9.14) :

$$\frac{d\mathcal{E}_k}{dt} = \sum_{i=1}^{3N} v_i f_i + 2\gamma \left(\frac{3N}{2} k_B T_o - \mathcal{E}_k \right)$$

Or l'énergie cinétique peut s'écrire de la manière suivante :

$$\mathcal{E}_k = \frac{3N}{2} k_B T \quad \Rightarrow \quad \frac{d\mathcal{E}_k}{dt} = \sum_{i=1}^{3N} v_i f_i + 2\gamma \frac{3N}{2} k_B (T_o - T) \tag{9.20}$$

Le premier terme n'est autre que la dérivée de l'énergie potentielle en fonction du temps. Or, la conservation de l'énergie mécanique du système nous permet d'écrire que :

$$\mathcal{H} = \mathcal{E}_k + U_p = Cte \Rightarrow \frac{d\mathcal{E}_k}{dt} + \frac{dU_p}{dt} = 0$$

Ainsi, le deuxième terme apparaît comme un paramètre de couplage décrivant le couplage du système avec un réservoir thermique. On peut l'interpréter comme un terme correctif qui viendrait corriger la température en cas d'éloignement par rapport à la température fixée T_o. On peut alors écrire que :

$$\frac{d\mathcal{E}_k}{dt} = -\frac{dU_p}{dt} + \left(\frac{d\mathcal{E}_k}{dt} \right)_{\mathcal{R}es} = -\frac{dU_p}{dt} + \frac{3N}{2} k_B \left(\frac{dT}{dt} \right)_{\mathcal{R}es}$$

Par identification, on a :

$$\left(\frac{dT}{dt} \right)_{\mathcal{R}es} = 2\gamma (T_o - T) = \frac{T_o - T}{\tau_T}$$

τ_T est la constante de temps caractéristique de ce couplage. Comme les simulations effectuées n'utilisent pas de terme stochastique, il faut déterminer une autre équation de mouvement. Pour cela, utilisons la relation (9.20). Pour une particule avec un degré de liberté :

$$\frac{d\mathcal{E}_k}{dt} = m_i v_i \dot{v}_i = v_i f_i + 2\gamma \frac{3N}{2} k_B (T_o - T)$$

Si on divise par v_i, on a :

$$m_i \dot{v}_i = f_i + \gamma \frac{3Nk_B}{v_i}(T_o - T)$$

Or, on a :

$$2\mathcal{E}_k = m_i v_i^2 = 3Nk_B T \Rightarrow \frac{3Nk_B}{v_i} = \frac{m_i v_i}{T}$$

Nous obtenons alors la nouvelle équation de mouvement :

$$m_i \dot{v}_i = f_i + m_i \gamma v_i \frac{(T_o - T)}{T}$$

Le deuxième terme apparaît comme un terme de renormalisatation des vitesses. Comme on l'a vu dans le chapitre (2), la résolution des équations de mouvement utilise des développements limités. On a :

$$\begin{aligned} r_i(t+\delta t) &= r_i(t) + v_i(t)\Delta t + \frac{f_i(t)}{2m_i}\Delta t^2 + \gamma v_i(t)\frac{(T_o - T)}{2T}\Delta t^2 \Rightarrow \\ r_i(t+\delta t) &= r_i(t) + v_i(t)\Delta t(1 + \gamma\frac{(T_o - T)}{2T}\Delta t) + \frac{f_i(t)}{2m_i}\Delta t^2 \end{aligned}$$

Ainsi, le facteur de renormalisation est :

$$\lambda = 1 + \frac{\Delta t}{4\tau_T}\left(\frac{T_o - T}{T}\right)$$

Ce calcul revient à pratiquer une dynamique classique ($m_i \dot{v}_i = f_i$) avec une renormalisation des vitesses à chaque pas par λ. Nous avons précédemment établi que, pour tout système

conservatif isolé, il y a une invariance de l'énergie mécanique, c'est-à-dire de l'hamiltonien, ce qui se traduit par :

$$\frac{dH}{dt} = \frac{d\mathcal{E}_k}{dt} + \frac{dU}{dt} = 0 \Rightarrow \frac{d\mathcal{E}_k}{dt} = -\frac{dU}{dt}$$

Or, nous avons démontré que :

$$\frac{d\mathcal{E}_k}{dt} = m_i v_i \dot{v}_i = v_i f_i + 2\gamma \frac{3N}{2} k_B (T_o - T) = -\frac{dU}{dt} + 2\gamma \frac{3N}{2} k_B (T_o - T) \tag{9.21}$$

Cette équation (9.21) montre que cet algorithme ne conserve pas l'énergie totale[2]. Par contre, les moments conjugés sont conservés. Le deuxième inconvénient de cet algorithme est qu'il ne génère pas rigoureusement de distribution canonique. Néanmoins, il s'agit d'un algorithme stable numériquement (aucune dérive d'énergie aux temps longs) ne nécessitant pas de corrections *ad hoc* [10]. L'algorithme de Nosé-Hoover a l'avantage de ne pas présenter ces deux problèmes mais les fluctuations autour de la température d'équilibre avant équlibration sont plus importantes dans ce cas. Un algorithme intéressant est celui d'Andersen [67] qui ne présente aucun des inconvénients relatés ci-dessus. Cependant, la renormalistation des vitesses s'effectue par l'intermédiaire de $\chi = (T_o/T)^{0.5}$ obtenu en supposant que la température désirée est liée aux vitesses non projetées, ce qui est une égalité non justifiée. Mais il est possible de passer outre ce problème si on calcule de manière explicite ce paramètre χ puisqu'il est lié au coefficient de friction.

Nous pouvons coupler ce thermostat à un pressiostat basé sur le même principe. En effet, la variation de la pression avec le temps donne une relation similaire à l'équation (9.21), c'est-à-dire :

$$\left(\frac{dP}{dt}\right)_{Res} = \frac{P - P_o}{\tau_p}$$

[2]On peut écrire l'énergie totale comme la somme de l'énergie des états microscopiques E_l et de celle du thermostat E_L. La fonction de partition dans NVT étant proprtionnelle à $\exp -\frac{E_l}{k_B T}$, cela implique que tous les états microscopiques sont représentés quelles que soient leurs énergies. Cela est dû au fait que le thermostat met un «complément de l'énergie» pour maintenir l'énergie totale constante [66], ce qui n'est pas le cas ici.

La démarche pour maintenir la pression constante est identique à la précédente. Elle réside dans la recherche d'un facteur de mise à l'échelle permettant de maintenir la pression constante. Pour cela, Berendsen utilise une partie de l'algorithme d'Andersen[3]. Andersen suppose que le fluide peut être simulé dans un boîte de volume variable. Ce fluide peut alors être compressé par un piston. Dans ces conditions, le travail exercé par le piston vaut $\alpha V = P_E V$ où P_E est la pression externe. En développant les équations de mouvement et le lagrangien du système, on montre en passant aux coordonnées réduites ($\rho_i = r_i V^{-1/3}$) la relation (9.22).

$$\frac{d\mathbf{r}_i}{dt} = \mathbf{v}_i + \frac{1}{3V}\frac{dV}{dt}\mathbf{r}_i \tag{9.22}$$

Comme la pression est donnée en fonction de l'énérgie cinétique et du viriel, on obtient en considérant la relation (9.22), l'expression suivante pour la pression.

$$P = \frac{2}{3V}\left(\mathcal{E}_k + \frac{1}{2}\sum_{i<j}\mathbf{r}_{ij}\cdot\mathbf{f}_{ij}\right) = \frac{2}{3V}\left(\mathcal{E}_k - \Xi\right)$$

Des variations de la pression entraînent des variations du terme du viriel et donc des fluctuations des distances intermoléculaires. Si on veut maintenir une pression constante, il faut renormaliser les coordonnées à chaque pas. Cela se fait par une action sur la dérivée temporelle de $\mathbf{r}_i$ par l'ajout d'un terme supplémentaire, qui a la dimension d'une vitesse :

$$\frac{d\mathbf{r}_i}{dt} = \mathbf{v}_i + \xi\mathbf{r}_i \tag{9.23}$$

La comparaison entre les équations (9.22) et (9.23) permet d'écrire :

$$\xi = \frac{1}{3V}\frac{dV}{dt} \tag{9.24}$$

La terme permettant de lier $\frac{dV}{dt}$ et $\frac{dP}{dt}$ est la compressibilité isotherme :

[3]Il faut savoir que le pressiostat de Berendsen a été mis en place par ce dernier pour remédier au choix arbitraire de la masse du piston M dans l'algorithme d'Andersen [67], masse qui permettait d'influer sur le temps de décroissance de fluctuation du volume.

$$\beta_T = -\frac{1}{V}\frac{dV}{dP} = -\frac{1}{V}\frac{dV}{dt}\frac{dt}{dP} \Rightarrow \frac{dP}{dt} = -\frac{1}{\beta_T V}\frac{dV}{dt} = -\frac{3\xi}{\beta_T} \Rightarrow \xi = -\beta_T\frac{(P_o - P)}{3\tau_p}$$

Finalement, on peut écrire que :

$$\frac{d\mathbf{r}_i}{dt} = \mathbf{v}_i - \beta_T\frac{(P_o - P)}{3\tau_p}\mathbf{r}_i$$

Un développement de Taylor à l'ordre deux pour $\mathbf{r}_i$ donne dans le cadre de la résolution des équations de mouvement :

$$\begin{aligned} r_i(t+\delta t) &= r_i(t) + v_i(t)\Delta t - \beta_T\frac{(P_o - P)}{3\tau_p}\mathbf{r}_i(t)\Delta t + \frac{f_i(t)}{2m_i}\Delta t^2 \Rightarrow \\ r_i(t+\delta t) &= r_i(t)(1 - \beta_T\Delta t\frac{(P_o - P)}{3\tau_p}) + v_i(t)\Delta t\frac{f_i(t)}{2m_i}\Delta t^2 \end{aligned}$$

Cela revient à renormaliser les positions et le volume de la boîte par le facteur μ défini par la relation (9.25).

$$\mu = 1 - \beta_T\Delta t\frac{(P_o - P)}{3\tau_p} \tag{9.25}$$

Ainsi, l'application de λ sur les vitesses des particules et de μ sur leurs positions et le volume de la boîte permet de maintenir la température et la pression constantes.

L'algorithme de Berendsen, bien que possédant quelques imperfections, reste un algorithme stable avec de très faibles fluctuations autour de la température d'équilibre. Pour terminer, soulignons que l'énergie cinétique instantanée obtenue avec cet algorithme est dépendante de la température. Or, la dérivée de celle-ci par rapport à T doit être absolument nulle afin de repecter l'indépendance de l'hamiltonien par rapport à T. En effet, on sait que la dérivée $\frac{\partial \mathcal{H}}{\partial T}$ est nulle, cela implique que $\frac{\partial \mathcal{U}}{\partial T} = -\frac{\partial \mathcal{E}_k}{\partial T}$, or $\frac{\partial \mathcal{U}}{\partial T} = 0$, ce qui implique $\frac{\partial \mathcal{E}_k}{\partial T} = 0$. Or, dans l'algorithme de Berendsen, cette égalité n'est pas justifiée alors qu'elle est dans l'algorithme de Nosé.

9.3. Sommation d'Ewald

Pour des potentiels d'interaction à longue distance comme le potentiel de Coulomb, les interactions entre toutes les particules du système doivent être prises en compte. Pour des systèmes sans conditions aux limites (réseaux à dimension finie), comme des gouttes d'eau par exemple, le calcul des interactions coulombiennes se réduit au potentiel classique de Coulomb en $\frac{1}{r}$. Dans le cas de système où l'on applique les conditions périodiques afin de simuler des systèmes infinis (voir la section (2.6.2)), il ne suffit pas de calculer les interactions entre les particules de la cellule centrale uniquement, mais aussi avec toutes les particules situées dans les images périodiques de l'unité centrale :

$$U = \frac{1}{2}\sum_{i,j}^{N}\sum_{i,j}^{N}\sum_{n}^{'}\frac{q_i q_j}{|\mathbf{r}_{ij} - \mathbf{n}L|} \tag{9.26}$$

Ici, $\mathbf{n}$ est le vecteur unitaire de la maille (motif élémentaire) de longueur L, $\sum_n^{'}$ indique que $i \neq j$ pour $n = 0$. Le potentiel évoluant en $\frac{1}{r}$, il est faiblement convergent puisqu'il évolue comme une série asymptotique. Un calcul direct par la relation (9.26) conduit à une imprécision sur le résultat du calcul. Étant donnée l'importance de ces interactions, ce type d'imprécisions n'est pas envisageable. Pour résoudre ce problème, il suffit de décomposer le potentiel en série de Fourier dont les termes sont plus rapidement convergents.

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty}\left[a_k \cos\left(\frac{k\pi x}{L}\right) + b_k \sin\left(\frac{k\pi x}{L}\right)\right]$$

La série converge en tous les points de continuité vers $f(x)$. Aux sauts de discontinuité, elle converge vers la valeur moyenne des limites à droite et à gauche. Ce développement en série n'est possible que pour des fonctions périodiques. Or le potentiel coulombien étant en $\frac{1}{r}$ il n'est pas périodique. Par conséquent, nous devons pratiquer une transformation de Fourier discrète (décomposition en série de Fourier pour des fonctions non périodiques) où le potentiel est divisé en deux parties : une partie évaluée dans l'espace réel et une autre dans l'espace de Fourier. Le potentiel étant une solution de l'équation de Poisson ($\nabla^2\phi(\mathbf{r}) = -4\pi\rho(\mathbf{r})$), l'étude peut être menée en considérant la densité de charge du système[4]. On peut alors décomposer

[4]la notion de densité de charge permet de décrire la charge ponctuelle avec les «outils du continu», cela implique l'utilisation d'un delta de Dirac : $\rho(x) = \sum_{i=1}^{n} q_i \delta(x - x_i)$.

la densité de charge en deux parties :

$$\rho(r) \quad = \quad \rho^{Re}(r) + \rho^{Im}(r)$$

Interprétation physique : Cette méthode consiste à écranter toutes les charges q_i du réseau par des distributions gaussiennes de charges ρ_i^{ecran} telles que :

$$\rho_i^{ecran}(r) = -q_i \left(\frac{\alpha}{\sqrt{\pi}}\right)^3 \exp\left[-\alpha^2 r^2\right]$$

α est un paramètre positif qui détermine la largeur de la distribution et r représente la position du centre de la distribution. Le choix s'est porté sur une distribution de type gaussienne car il s'agit d'une distribution qui offre une certaine flexibilité mathématique. Ainsi, les charges gaussiennes d'écrantage créent un potentiel calculé dans l'espace image, tandis que les charges écrantées sont évaluées dans l'espace réel (cette méthode peut être apparentée à la méthode des images en électrostatique). Cela revient à écrire que :

$$\rho(r) \quad = \quad \rho^{Re}(r) + \rho^{Im}(r) = [\rho(r) + \rho_i^{ecran}(r)] - \rho_i^{ecran}(r)$$

On pose $\rho^{Im}(r) = -\rho_i^{ecran}(r)$. Il s'agit de la densité évaluée dans l'espace réciproque, puisque $\rho_i^{ecran}(r)$ converge lentement dans l'espace réel. De plus, on pose $\rho^{Re}(r) = \rho(r) + \rho_i^{ecran}(r)$. Il s'agit de la densité déterminée dans l'espace réel (convergence rapide dans cet espace). Pour obtenir l'énergie potentielle électrostatique, nous devons suivre le protocole suivant :

1- Résoudre l'équation de Poisson ($\nabla^2\phi(\mathbf{r}) \; = \; -4\pi\rho(\mathbf{r})$) pour accéder au potentiel électrostatique :
2- intégrer sur tout l'espace, le produit $\rho(\mathbf{r})\phi(\mathbf{r})$, pour déterminer l'énergie potentielle : $U(r) = \int\int\int_{R^3} \rho(r')\phi(r')dr'^3$.

Comme la densité de charge peut être décomposée en deux termes, il en va de même pour le potentiel électrostatique et l'énergie potentielle. Nous allons commencer par déterminer le terme lié à l'espace de Fourier.

9.3.1 Calculs dans l'espace de Fourier

Étant donné la périodicité de la densité de charge (rendue possible grâce à la description continue de la charge par l'intermédiaire d'une distribution de Dirac) $\rho(\mathbf{r}) = \rho(\mathbf{r} + \mathbf{n}L)$, elle peut être décomposée en série de Fourier :

$$\rho(\mathbf{r}) = \sum_{\mathbf{k}} \rho(\mathbf{k}) \exp\left(\frac{2\pi i \mathbf{n}_k \mathbf{r}}{L}\right) \text{ soit, en posant } \mathbf{k} = \frac{2\pi \mathbf{n}_k}{L}, \rho(\mathbf{r}) = \sum_{\mathbf{k}} \rho(\mathbf{k}) \exp\left(i\mathbf{k}\mathbf{r}\right)$$

De la même manière, on a :

$$\phi(\mathbf{r}) = \sum_{\mathbf{k}} \phi(\mathbf{k}) \exp\left(i\mathbf{k}\mathbf{r}\right)$$

L'équation de Poisson, reliant le potentiel et la densité de charge, nous permet d'écrire :

$$\sum_{\mathbf{k}} k^2 \phi(\mathbf{k}) \exp\left(i\mathbf{k}\mathbf{r}\right) = \sum_{\mathbf{k}} 4\pi\rho(\mathbf{k}) \exp\left(i\mathbf{k}\mathbf{r}\right)$$

ce qui donne :

$$\phi(\mathbf{k}) = \frac{4\pi}{k^2}\rho(\mathbf{k}) \tag{9.27}$$

Il faut, à présent, déterminer $\rho(\mathbf{k})$. Ce dernier étant un coefficient de la série, il vient :

$$\rho(\mathbf{k}) = \frac{1}{V}\int_V \exp -(i\mathbf{k}\mathbf{r})\rho(\mathbf{r})d\mathbf{r}$$

Étant donnée la forme de la distribution de charges écrantées dans le réseau périodique :

$$\rho(\mathbf{r}) = \sum_{j_q}\sum_{\mathbf{n}} q_{j_q}\left(\frac{\alpha}{\sqrt{\pi}}\right)^3 \exp\left[-\alpha^2\left|\mathbf{r} - (\mathbf{r}_{j_q} - \mathbf{n}L)\right|^2\right]$$

on peut écrire :

$$\rho(\mathbf{k}) = \frac{1}{V}\int_V \left(\exp -(i\mathbf{k}\mathbf{r})\sum_{j_q}\sum_{\mathbf{n}} q_{j_q}\left(\frac{\alpha}{\sqrt{\pi}}\right)^3 \exp\left[-\alpha^2\left|\mathbf{r} - (\mathbf{r}_{j_q} - \mathbf{n}L)\right|^2\right]\right) d\mathbf{r} \tag{9.28}$$

La relation (9.28) peut se simplifier si on utilise la *formule sommatoire de Poisson*, selon laquelle on a :

$$\sum_{\mathbf{n}} \exp\left(-\left[\mathbf{r}+n\right]^2 t\right) = \left[\frac{\sqrt{\pi}}{L} t^{-1/2}\right]^3 \sum_{\mathbf{k}} \exp\left(i\mathbf{kr}\right) \exp\left(-\frac{k^2}{4t}\right)$$

Il faut alors écrire la relation (9.28) de la manière suivante :

$$\rho(\mathbf{k}) = \frac{1}{V} \int_V \left(\exp -(i\mathbf{kr}) \sum_{j_q} q_{j_q} \left(\frac{\alpha}{\sqrt{\pi}}\right)^3 \sum_{\mathbf{n}} \exp\left[-\alpha^2 \left|(\mathbf{r}-\mathbf{r}_{j_q})+\mathbf{n}L\right|^2\right] \right) d\mathbf{r}$$

Ainsi, l'utilisation de la formule sommatoire de Poisson permet d'obtenir l'équation suivante :

$$\rho(\mathbf{k}) = \frac{1}{V^2} \int_V \left(\exp -(i\mathbf{kr}) \sum_{j_q} q_{j_q} \sum_{\mathbf{k}} \exp\left(i\mathbf{k}(\mathbf{r}-\mathbf{r}_{j_q})\right) \exp\left(-\frac{k^2}{4\alpha^2}\right) \right) d\mathbf{r}$$

Soit, en simplifiant par $\exp -(i\mathbf{kr})$:

$$\rho(\mathbf{k}) = \frac{1}{V^2} \sum_{j_q} q_{j_q} \int_V \left[\sum_{\mathbf{k}} \exp\left(-i\mathbf{kr}_{j_q}\right) \exp\left(-\frac{k^2}{4\alpha^2}\right) \right] d\mathbf{r}$$

Dans la mesure où l'on calcule la densité de charge dans l'espace réciproque pour $\mathbf{k}$ donné, on peut s'absoudre de la sommation sur $\mathbf{k}$ d'une part, et d'autre part comme il n'y a plus de grandeur dépendante de r, on peut écrire :

$$\rho(\mathbf{k}) = \frac{1}{V^2} \sum_{j_q} q_{j_q} \exp\left(-i\mathbf{kr}_{j_q}\right) \exp\left(-\frac{k^2}{4\alpha^2}\right) \int_V d\mathbf{r}$$

On a donc en calculant l'intégrale $\int_V d\mathbf{r} = V$

$$\rho(\mathbf{k}) = \frac{1}{V} \sum_{j_q} q_{j_q} \exp\left(-i\mathbf{kr}_{j_q}\right) \exp\left(-\frac{k^2}{4\alpha^2}\right)$$

La relation (9.27) permet d'accéder au potentiel dans l'espace réciproque :

$$\phi(\mathbf{k}) = \frac{4\pi}{k^2 V} \exp\left(-\frac{k^2}{4\alpha^2}\right) \sum_{j_q} q_{j_q} \exp\left(-i\mathbf{k}\mathbf{r}_{j_q}\right)$$

Le potentiel dans l'espace réel s'écrit alors :

$$\phi(\mathbf{r}) = \frac{4\pi}{V} \sum_{\mathbf{k}\neq 0} \sum_{j_q} \frac{q_{j_q}}{k^2} \exp\left(-\frac{k^2}{4\alpha^2}\right) \exp\left(i\mathbf{k}(\mathbf{r} - \mathbf{r}_{j_q})\right)$$

Cette relation n'est valable que pour $\mathbf{k} \neq 0$. Pour $\mathbf{k} = 0$, on prend $\phi(\mathbf{k}) = 0$. L'énergie potentielle se détermine par :

$$U(r)^{(k)} = \frac{1}{2} \sum_{l=1}^{N} q_l \phi(\mathbf{r}_l) \tag{9.29}$$

$$= \frac{2\pi}{V} \sum_{\mathbf{k}\neq 0} \sum_{j=1}^{N} \sum_{l=1}^{N} \frac{q_j q_l}{k^2} \exp\left(-\frac{k^2}{4\alpha^2}\right) \exp\left(i\mathbf{k}(\mathbf{r_l} - \mathbf{r}_j)\right) \tag{9.30}$$

Typiquement, on utilise 100 à 200 vecteurs $\mathbf{k}$ (Woodcock et Singer [71]). La détermination de cette énergie nécessite l'évaluation de trois sommes, ce qui est prohibitif en terme de temps de calcul. Pour abaisser le temps de calcul, il suffit de réduire la triple somme en une double somme (Heyes [72]) :

$$U(r)^{(k)} = \sum_{\mathbf{k}\neq 0} \frac{2\pi}{k^2 V} \exp\left(-\frac{k^2}{4\alpha^2}\right) \left| \sum_{j=1}^{N} q_j \exp\left(i\mathbf{k}(\mathbf{r}_j)\right) \right|^2 \tag{9.31}$$

Lors d'une simulation de type dynamique moléculaire, nous intégrons les équations de mouvement qui nécessitent la connaissance des forces conservatives exercées sur chaque atome i :

$$\mathbf{f}_i = -\frac{\partial}{\partial \mathbf{r}_i} U_{tot}(r) \tag{9.32}$$

La force électrostatique vaut :

$$\mathbf{f}_i = \mathbf{f}_i^{(k)} + \mathbf{f}_i^{(r)} + \mathbf{f}_{sc} \tag{9.33}$$

où le premier terme est celui étudié actuellement et représente le terme dans l'espace de Fourier, le deuxième terme est celui de l'espace réel et le troisième un terme correctif. La force relative à l'espace réciproque vaut :

$$\begin{aligned}\mathbf{f}_i^{(k)} &= -\frac{\partial}{\partial \mathbf{r}_i} U(r)^{(k)} \\ &= -q_i \sum_{\mathbf{k}\neq 0} \exp\left(i\mathbf{k}(\mathbf{r}_i)\right) \frac{4\pi \mathbf{k} i}{k^2 V} \exp\left(-\frac{k^2}{4\alpha^2}\right) \left| \sum_{j=1}^{N} q_j \exp\left(i\mathbf{k}(\mathbf{r}_j)\right) \right|\end{aligned}$$

En unité électrostatique, on a :

$$\mathbf{f}_i^{(k)} = -q_i \sum_{\mathbf{k}\neq 0} \exp\left(i\mathbf{k}(\mathbf{r}_i)\right) \frac{\mathbf{k} i}{k^2 V \epsilon_o} \exp\left(-\frac{k^2}{4\alpha^2}\right) \left| \sum_{j=1}^{N} q_j \exp\left(i\mathbf{k}(\mathbf{r}_j)\right) \right| \tag{9.34}$$

$$U(r)^{(k)} = \sum_{\mathbf{k}\neq 0} \frac{1}{2k^2 V \epsilon_o} \exp\left(-\frac{k^2}{4\alpha^2}\right) \left| \sum_{j=1}^{N} q_j \exp\left(i\mathbf{k}(\mathbf{r}_j)\right) \right|^2 \tag{9.35}$$

9.3.2 Calculs dans l'espace réel

Étant donnée la relation $\rho^{Re}(\mathbf{r}) = [\rho(\mathbf{r}) + \rho_i^{ecran}(\mathbf{r})]$ et l'aspect ponctuel de la charge[5], on peut écrire pour une charge q_j :

$$\rho_j^{Re}(\mathbf{r}) = \rho(\mathbf{r}) + \rho_i^{ecran}(\mathbf{r}) = q_j \delta(\mathbf{r}) - q_j \left(\frac{\alpha}{\sqrt{\pi}}\right)^3 \exp\left[-\alpha^2 r^2\right]$$

L'équation de Poisson, en coordonnées sphériques, permet d'écrire :

$$-\frac{1}{r}\frac{\partial^2 r\phi_j^{Re}(\mathbf{r})}{\partial r^2} = 4\pi\rho_j^{Re}$$

L'obtention du potentiel passe par une double intégration de l'équation de Poisson :

[5]Cet aspect ponctuel n'autorise pas de description continue de la distribution propre au domaine macroscopique (présence de discontinuité). Pour remédier à ce problème, on décrit la charge ponctuelle par une distribution continue : *le delta de Dirac.*

$$\begin{aligned}
\frac{\partial^2 r\phi_j^{Re}(\mathbf{r})}{\partial r^2} &= -4\pi r\rho_j^{Re} \\
\Rightarrow \frac{\partial r\phi_j^{Re}(\mathbf{r})}{\partial r} &= -\int_r^\infty 4\pi r\rho_j^{Re}dr \\
&= -q_j\int_r^\infty 4\pi r\delta(\mathbf{r})dr + q_j\left(\frac{\alpha}{\sqrt{\pi}}\right)^3\int_r^\infty 4\pi r\exp\left[-\alpha^2r^2\right]dr \\
&= -4\pi q_j\int_r^\infty r\delta(\mathbf{r})dr - 2q_j\left(\frac{\alpha}{\sqrt{\pi}}\right)\exp\left[-\alpha^2r^2\right]
\end{aligned}$$

Pour déterminer le premier terme, il suffit de rappeler le théorème suivant :

– **Théorème :** *Le Laplacien de la fonction radiale* $f : \mathbf{r} \rightarrow \frac{1}{r} = \frac{1}{||\mathbf{r}||}$ est :

$$\Delta\left[\frac{1}{r}\right] = -4\pi\delta$$

On peut alors écrire :

$$\frac{\partial r\phi_j^{Re}(\mathbf{r})}{\partial r} = q_j\int_r^\infty r\Delta\left[\frac{1}{r}\right]dr - 2q_j\left(\frac{\alpha}{\sqrt{\pi}}\right)\exp\left[-\alpha^2r^2\right]$$

En poursuivant le calcul, on a :

$$r\phi_j^{Re}(\mathbf{r}) = q_j\int_0^r\int_r^\infty r\Delta\left[\frac{1}{r}\right]dr^2 - 2q_j\left(\frac{\alpha}{\sqrt{\pi}}\right)\int_0^r\exp\left[-\alpha^2r^2\right]dr$$

Comme, $r\Delta\left[\frac{1}{r}\right] = \frac{\partial^2 1}{\partial r^2}$ [6] donc $\int_0^r\int_r^\infty \frac{\partial^2 1}{\partial r^2}dr^2 = 1$ on a :

$$r\phi_j^{Re}(\mathbf{r}) = q_j - 2q_j\left(\frac{\alpha}{\sqrt{\pi}}\right)\int_0^r\exp\left[-\alpha^2r^2\right]dr$$

Or,

[6] Si f est une fonction radiale deux fois dérivable sur une partie de $\mathbb{R}^3$, son laplacien se calcule par la formule : $\nabla f = \frac{1}{r}\frac{\partial^2}{\partial r^2}(rf)$.

$$erf(x) = \frac{2}{\sqrt{\pi}} \int_0^x \exp -t^2 dt$$

en posant $t = \alpha r^{'}$ on a $dt = \alpha dr^{'}$, il vient donc :

$$\int_0^r \exp\left[-\alpha^2 r^{'2}\right] dr^{'} = \int_0^{\alpha r} (\alpha)^{-1} \exp\left[-t^2\right] dt = (\alpha)^{-1} erf(\alpha r) \frac{\sqrt{\pi}}{2}$$

Le potentiel vaut donc :

$$\phi_j^{Re}(\mathbf{r}) = q_j \left(\frac{1 - erf(\alpha r)}{r} \right) = q_j \frac{erfc(\alpha r)}{r}$$

De la même manière que pour le potentiel dans l'espace de Fourier, nous avons :

$$U^{Re} = \sum_{i=1}^{N-1} \sum_{l=i+1}^{N} q_i q_l \frac{erfc(\alpha \left|\mathbf{r}_i - \mathbf{r}_l\right|)}{\left|\mathbf{r}_i - \mathbf{r}_l\right|}$$

Ce potentiel ne prend en compte que les interactions entre atomes n'appartenant pas à la même molécule. Pour les atomes de la même molécule, ce potentiel intervient pour les interactions entre atomes séparés par au moins quatre liaisons.

La force qui dérive de ce potentiel est donnée par :

$$\begin{aligned} \mathbf{f}_i{}^{Re} &= -\frac{\partial U^{Re}}{\partial \mathbf{r}_i} = -\frac{\partial}{\partial \mathbf{r}_i} \left[\sum_{i=1}^{N-1} \sum_{l=i+1}^{N} q_i q_l \frac{1 - erf(\alpha \left|\mathbf{r}_i - \mathbf{r}_l\right|)}{\left|\mathbf{r}_i - \mathbf{r}_l\right|} \right] \\ \mathbf{f}_i{}^{Re} &= \sum_{i=1}^{N-1} \sum_{l=i+1}^{N} q_i q_l \frac{\frac{\mathbf{r}_{il}}{r_{il}} \cdot erfc(\alpha \left|\mathbf{r}_i - \mathbf{r}_l\right|) + r_{il} \frac{2\alpha}{\sqrt{\pi}} \exp -(\alpha^2 r_{il}^2)}{r_{il}^2} \end{aligned}$$

On a $\frac{\mathbf{r}_{il}}{r_{il}} = \frac{\partial r_{il}}{\partial \mathbf{r}_i}$ puisque $\hat{\mathbf{r}}_{il} = \frac{\mathbf{r}_{il}}{r_{il}}$. Le deuxième terme s'obtient en effectuant les changements de variables : $t = \alpha r_{il} \rightarrow dt = \alpha dr_{il}$ *et* $dr_{il} = dr_i$ car l est constant. Le terme $\frac{2\alpha}{\sqrt{\pi}}$ provient de la définition de la fonction d'erreur. Comme $\int af = F,\ dF = af$.

On a, aprés simplification :

$$\mathbf{f}_i{}^{Re} = \sum_{i=1}^{N-1} \sum_{l=i+1}^{N} q_i q_l \frac{\mathbf{r}_{il}}{r_{il}^3} \left[erfc(\alpha \left|\mathbf{r}_{il}\right|) + \frac{2\alpha \left|\mathbf{r}_{il}\right|}{\sqrt{\pi}} \exp -(\alpha^2 \left|\mathbf{r}_{il}\right|^2) \right]$$

9.3.3 Terme d'autointeraction et terme intramoléculaire

Le potentiel que l'on veut évaluer est une somme d'interactions de paires *intermoléculaires*. Or, les gaussiennes utilisées écrantent chacune des charges du réseau. Ainsi, le potentiel calculé dans l'espace image comprend des termes *intramoléculaires* puisqu'on ne fait pas la distinction entre les paires d'atomes intramoléculaires et intermoléculaires. Quand on passe d'une triple somme à une double somme pour le calcul du potentiel ϕ^{Re}, on ne distingue plus les parties intramoléculaire et intermoléculaire dans les interactions. Par conséquent, il faut retrancher ces contributions intramoléculaires au potentiel total. Cette contribution pour une charge j contient deux termes :

1- La partie du potentiel créée par la charge écrantée j, ressentie par elle même. Cette contribution n'est autre que la limite du potentiel calculé dans l'espace réel quand $r \to 0$.
2- L'effet des écrans dus aux charges $i \neq j$ appartenant à la même molécule.

Auto-interaction :

$$\phi_j^{Re}(O) = \lim_{r\to 0}\left[q_j \frac{erfc(\alpha r)}{r}\right] = \lim_{r\to 0}\left[q_j \frac{2}{\sqrt{\pi}} \frac{\int_{\alpha r}^{\infty} \exp -t^2 dt}{r}\right]$$

En posant $t = \alpha r$, on a :

$$\phi_j^{Re}(O) = q_j \frac{2\alpha}{\sqrt{\pi}} \lim_{r\to 0}\left[\frac{\int_{\alpha r}^{\infty} \exp -t^2 dt}{t}\right]$$

$$\phi_j^{Re}(O) = q_j \frac{2\alpha}{\sqrt{\pi}}$$

Partie intramoléculaire :

Soit N_m le nombre d'atomes par molécules, la densité de charge totale relative aux gaussiennes d'écrantage des atomes $j \neq i$ dans l'espace réciproque ressentie par i dans une molécule donnée vaut :

$$\rho^{ecran} = \sum_{i=1,\neq j}^{N_m} q_i \left(\frac{\alpha}{\sqrt{\pi}}\right)^3 \exp\left[-\alpha^2 r^2\right]$$

Comme on l'a fait dans la section précédente, la double intégration de l'équation de Poisson donne :

$$\sum_{i=1,\neq j}^{N_m} q_i \frac{erf(\alpha \left|\mathbf{r}_i - \mathbf{r}_j\right|)}{\left|\mathbf{r}_i - \mathbf{r}_j\right|}$$

Le potentiel total, contenant le terme *self-term* ou *auto-correction* et le terme intramoléculaire, vaut :

$$\phi_j^{Self} = q_j \frac{2\alpha}{\sqrt{\pi}} + \sum_{i=1,\neq j}^{N_m} q_i \frac{erf(\alpha \left|\mathbf{r}_i - \mathbf{r}_j\right|)}{\left|\mathbf{r}_i - \mathbf{r}_j\right|}$$

De la même manière que précédement, nous obtenons l'énergie potentielle :

$$U_j^{Self} = \sum_{j=1}^{N} \left[q_j^2 \frac{\alpha}{\sqrt{\pi}} + \sum_{i=1,\neq j}^{N_m} q_i q_j \frac{erf(\alpha \left|\mathbf{r}_i - \mathbf{r}_j\right|)}{2\left|\mathbf{r}_i - \mathbf{r}_j\right|} \right]$$

Comme pour le potentiel de l'espace réel, la force vaut :

$$\mathbf{f}_l{}^{Self} = -\sum_{j=1}^{N} \sum_{i=1,\neq j}^{N_m} q_i q_j \frac{\mathbf{r}_{ij}}{r_{ij}^3} \left[erf(\alpha \left|\mathbf{r}_{ij}\right|) - \frac{\alpha \left|\mathbf{r}_{ij}\right|}{2\sqrt{\pi}} \exp -(\alpha^2 \left|\mathbf{r}_{ij}\right|^2) \right]$$

Ainsi, l'énergie électrostatique et la force électrostatique (sur un atome j) ont pour expression :

$$
\begin{aligned}
U^C &= \sum_{\mathbf{k}\neq 0} \frac{2\pi}{k^2 V} \exp\left(-\frac{k^2}{4\alpha^2}\right) \left| \sum_{j=1}^{N} q_j \exp\left(i\mathbf{k}(\mathbf{r}_j)\right) \right|^2 + \sum_{i=1}^{N-1} \sum_{l=i+1}^{N} q_i q_l \frac{erfc(\alpha \left|\mathbf{r}_i - \mathbf{r}_l\right|)}{\left|\mathbf{r}_i - \mathbf{r}_l\right|} \\
&- \sum_{j=1}^{N} \left[q_j^2 \frac{\alpha}{\sqrt{\pi}} + \sum_{i=1,\neq j}^{N_m} q_i q_j \frac{erf(\alpha \left|\mathbf{r}_i - \mathbf{r}_j\right|)}{2\left|\mathbf{r}_i - \mathbf{r}_j\right|} \right] \qquad (9.36)
\end{aligned}
$$

$$
\begin{aligned}
\mathbf{f}^C &= -q_i \sum_{\mathbf{k}\neq 0} \exp\left(i\mathbf{k}(\mathbf{r}_i)\right) \frac{4\pi \mathbf{k} i}{k^2 V} \exp\left(-\frac{k^2}{4\alpha^2}\right) \left| \sum_{j=1}^{N} q_j \exp\left(i\mathbf{k}(\mathbf{r}_j)\right) \right| \\
&+ \sum_{i=1}^{N-1} \sum_{l=i+1}^{N} q_i q_l \frac{\mathbf{r}_{il}}{r_{il}^3} \left[erfc(\alpha \left|\mathbf{r}_{il}\right|) + \frac{2\alpha \left|\mathbf{r}_{il}\right|}{\sqrt{\pi}} \exp -(\alpha^2 \left|\mathbf{r}_{il}\right|^2) \right] \qquad (9.37) \\
&- \sum_{j=1}^{N} \sum_{i=1,\neq j}^{N_m} q_i q_j \frac{\mathbf{r}_{ij}}{r_{ij}^3} \left[erf(\alpha \left|\mathbf{r}_{ij}\right|) - \frac{\alpha \left|\mathbf{r}_{ij}\right|}{2\sqrt{\pi}} \exp -(\alpha^2 \left|\mathbf{r}_{ij}\right|^2) \right]
\end{aligned}
$$

9.4. Algorithme de contrainte

9.4.1. Définition

Cette méthode permet d'introduire des contraintes de distances dans l'algorithme de dynamique moléculaire. Cela permait de réduire le nombre de degrés de liberté et donc de réduire le temps de calcul. La théorie des contraintes tient essentiellement du formalisme lagrangien.

Soit un système de N particules repérées par l'indice α variant de 1 à N. Si l'écriture de toutes les équations de la dynamique permet en principe, en y ajoutant les conditions initiales, de déterminer complétement le mouvement, alors cette résolution devient plus délicate lorsqu'il existe des contraintes. On définit une contrainte comme étant un ensemble de relations liant certains paramètres (positions, vitesses,...). Celles-ci ont l'avantage d'abaisser la dimensionnalité du système. Soit $\mathbf{r}$ l'entité sur laquelle porte la contrainte. Supposons que cette contrainte concerne l'ensemble des particules α, elle peut s'écrire par exemple, dans le cas simple où elle fixe la valeur de $\mathbf{r}$:

$$\mathbf{r}^\alpha = K \Leftrightarrow R_j(\mathbf{r}^\alpha) = 0 \text{ où } j \text{ est le nombre de contraintes}$$

Remarque : S'il existe une dépendance explicite au temps, on peut alors écrire $R_j(\mathbf{r}^\alpha, t) = 0$.

Comme la dynamique moléculaire est basée sur l'intégration des équations de mouvement de Newton, il faut donc trouver le moyen de résoudre ces équations de mouvement en tenant compte de ces contraintes. Cela revient à déterminer l'expression des forces contraintes. La relation de Newton s'écrit alors :

$$m\ddot{\mathbf{r}} = \mathbf{F} + \mathbf{G} \text{ où } \mathbf{G} \text{ est le vecteur force contrainte totale}$$

La détermination de la force contrainte se fait par la méthode des multiplicateurs de Lagrange.

9.4.2. Multiplicateurs de Lagrange

Cette méthode permet de déterminer les points stationnaires (point annulant la différentielle totale) d'une fonction indépendante ou liée $y = f(x_i)$. Considérons le cas d'une fonction liée c'est-à-dire une fonction où les n variables x_i sont liées par la contrainte :

$$R(x_1, ..., x_n) = 0$$

Les points stationnaires doivent remplir deux conditions :

$$dy = \sum_i^n \frac{\partial f}{\partial x_i} dx_i = 0$$
$$dR = \sum_i^n \frac{\partial R}{\partial x_i} dx_i = 0$$

De l'expression de dR, on extrait dx_n :

$$dx_n = \left[dR - \sum_{i=1}^{n-1} \frac{\partial R}{\partial x_i} dx_i \right] \frac{\partial x_n}{\partial R}$$

Cela nous permet de réécrire l'expression de dy :

$$dy = \sum_{i=1}^{n-1} \frac{\partial f}{\partial x_i} dx_i + \frac{\partial f}{\partial x_n} \left[dR - \sum_{i=1}^{n-1} \frac{\partial R}{\partial x_i} dx_i \right] \frac{\partial x_n}{\partial R}$$

$$\Rightarrow dy = \sum_{i}^{n-1} \frac{\partial f}{\partial x_i} dx_i - \frac{\partial f}{\partial x_n} \left[dR - \sum_{i=1}^{n-1} \frac{\partial R}{\partial x_i} dx_i \right] \frac{\partial R}{\delta x_n}$$

$$\Rightarrow dy = \sum_{i}^{n-1} dx_i \left(\frac{\partial f}{\partial x_i} + \frac{\partial f}{\partial x_n} \frac{\partial R}{\partial x_n} \frac{\partial R}{\partial x_i} dx_i \right) - \frac{\partial f}{\partial x_n} \frac{\partial R}{\partial x_n} dR$$

$$\Rightarrow dy + \frac{\partial f}{\partial x_n} \frac{\partial R}{\partial x_n} dR = \sum_{i}^{n-1} dx_i \left(\frac{\partial f}{\partial x_i} + \frac{\partial f}{\partial x_n} \frac{\partial R}{\partial x_n} \frac{\partial R}{\partial x_i} dx_i \right)$$

En posant :

$$\lambda = \frac{\partial f}{\partial x_n} \frac{\partial R}{\partial x_n}, \text{ on écrit que :}$$

$$dy + \lambda dR = \sum_{i}^{n-1} dx_i \left(\frac{\partial f}{\partial x_i} + \lambda \frac{\partial R}{\partial x_i} dx_i \right) = 0$$

Selon cette relation, seuls les $n-1$ accroissements dx_i sont indépendants. Dans ces conditions, les points stationnaires x_i^o et λ peuvent être déterminés en résolvant le système suivant :

$$\begin{cases} \dfrac{\partial f}{\partial x_i} + \lambda \dfrac{\partial R}{\partial x_i} = 0, \ i = 1....n-1 \\ \\ R(x_i) = 0 \end{cases}$$

À présent, généralisons cette relation à α contraintes tel que :

$$\begin{cases} R_\alpha(x_1, ..., x_n) = 0 \\ \\ \alpha = 1, ..., m \end{cases}$$

Dans ce cas, on introduit alors m multiplicateurs de Lagrange. On a alors à résoudre le système d'équations suivant, noté $\mathcal{S}$:

$$\begin{cases} \sum_{j=1}^{n}\sum_{k=1}^{m}\dfrac{\partial f}{\partial x_j}+\lambda_k\dfrac{\partial R_k}{\partial x_j}=0 \\ \\ \sum_{k=1}^{m}R_k(x_1,.....,x_n)=0 \end{cases}$$

9.4.3. Méthode SHAKE

D'apés ce qui a été fait, on peut considérer l'ensemble des contraintes (σ_k) sous la forme suivante :

$$\begin{cases} \sigma_k(\mathbf{r})=0 \\ \\ k=1,....,\ell \end{cases}$$

où l'indice k représente le nombre de contraintes holonomiques (*contraintes mettant en jeu des relations en coordonnées cartésiennes*). Si on considére une dépendance temporelle, on peut écrire, étant donné le système $\mathcal{S}$:

$$\sum_{i=1}^{N}\sum_{k=1}^{\ell}\frac{\partial f}{\partial x_i}+\lambda_k(t)\frac{\partial \sigma_k}{\partial x_i}=0$$

$$\sigma_k(\mathbf{r})=0$$

Considérons un système de N particules en interaction, on a :

$$\begin{cases} \mathbf{r}_i \rightarrow \text{Positions de l'atome i} \\ \\ \sigma_k(\mathbf{r}(t))=(\mathbf{r}_j(t)-\mathbf{r}_i(t))^2-d_{ij}^2 \end{cases}$$

$(\mathbf{r}_j(t)-\mathbf{r}_i(t))^2$ est la distance calculée et d_{ij}^2 la distance fixée. La dynamique du système peut être obtenue par la résolution des $3N$ équations lagrangiennes :

$$\begin{cases} \dfrac{d\nabla_{v_\alpha}L}{dt}=\nabla_{\mathbf{r}_\alpha} \\ \\ L=T-V \end{cases}$$

Si on divise le système global en un sous-système contraint $\mathbf{G}_i$ et un sous-système non contraint $\mathbf{F}_i$, la relation de Newton pour chaque atome i s'écrit :

$$\begin{aligned} m_i\ddot{\mathbf{r}}_i &= \mathbf{F}_i + \mathbf{G}_i \\ &= -\nabla_i V - \sum_{k=1}^{l} \nabla_i \sigma_k \lambda_k \end{aligned}$$

L'objectif est d'intégrer les équations de mouvement en y associant les contraintes afin de déterminer les mouvements des particules. Pour cela, on doit résoudre $3N + \ell$ équations différentielles à 3N + ℓ inconnues de type $\mathbf{r}(t), \lambda(t)$ telles qu'à l'instant origine les doublets $\{\mathbf{r}_i(0), \dot{\mathbf{r}}_i(0)\}$ sont connues sans ambiguïté. Comme le multiplicateur de Lagrange dépend du temps, on peut l'exprimer par un développement limité d'ordre n (au même titre que l'ensemble des positions dans l'algorithme de Verlet). On a alors :

$$\lambda_k(t) = \sum_{k=0}^{\infty} \lambda_k^n(t_o).\frac{(t-t_o)}{n!}$$

où n représente le degré de la dérivée et t_o le temps initial. Il apparaît évident que l'ensemble des positions va dépendre du multiplicateur de Lagrange. Nous pouvons alors écrire $\mathbf{r}_i(t, \lambda(t))$ en fonction de $\lambda_k^n(t_o)$ en effectuant un développement limité d'ordre 2 de $\mathbf{r}_i(t, \lambda(t))$. En effet, les relations de Verlet sont :

$$\begin{aligned} \mathbf{r}_i(\Delta t + t_o) &= \mathbf{r}_i(t_o) + \dot{\mathbf{r}}_i(t_o)\Delta t + \ddot{\mathbf{r}}_i(t_o)\frac{\Delta t^2}{2} \\ &= \mathbf{r}_i(t_o) + \dot{\mathbf{r}}_i(t_o)\Delta t + \frac{\mathbf{F}_i(t_o)}{m_i}\frac{\Delta t^2}{2} \end{aligned}$$

En posant $t = t_o + \Delta$, on peut écrire que :

$$\mathbf{r}_i(t) = \mathbf{r}_i(t_o) + \dot{\mathbf{r}}_i(t_o)(t-t_o) + \frac{\mathbf{F}_i(t_o)}{m_i}\frac{(t-t_o)^2}{2}$$

On va généraliser cette expression en insérant les contraintes sous forme de potentiel et non pas sous forme de gradient de potentiel.

$$\begin{aligned}\mathbf{r}_i(t,\{\lambda(t)\}) &= \mathbf{r}_i(t,\{\lambda^n(t_o)\}) = \mathbf{r}_i(t_o) + \dot{\mathbf{r}}_i(t_o)(t-t_o) \\ &+ \frac{1}{m_i}\sum_{n=2}^{\infty}\frac{(t-t_o)^n}{n!} \bullet \left[\mathbf{F}_i(t_o)^{n-2} - \sum_{k=1}^{\ell}\sum_{p=0}^{n-2}\lambda_k^p(t_o)\left[\nabla_i\sigma_k^{n-2-p}\right]_{t=t_o}\right]\end{aligned} \tag{9.38}$$

La détermination de cette relation nécessite certaines explications :

- La sommation commence à $n = 2$ puisque les dérivées d'ordre 0 et 1 correspondent aux deux premiers termes
- le dernier terme correspond à $\mathbf{r}_i(t_o)^n$
- comme $\mathbf{r}^2 = \dfrac{\mathbf{F}_i(t_o)}{m_i} \Rightarrow \mathbf{r}^n = \dfrac{\mathbf{F}_i(t_o)^{n-2}}{m_i}$.

Intéressons nous au troisième terme (double sommation). Ce terme correspond en fait à $\mathbf{G}_i(t_o)^{n-2}$. Or, on a vu que :

$$\mathbf{G}_i = -\sum_{k=1}^{\ell}\left(\nabla_i\sigma_k\right).\lambda_k$$

Le terme auquel on s'intéresse correspond à la dérivée d'ordre $n-2$ de $\mathbf{G}_i$. Pour déterminer son expression, il faut savoir que la dérivée d'ordre n d'un produit AB vaut :

$$(AB)^n = \sum_{k=0}^{n} A^k.B^{n-k}$$

Cela nous permet d'écrire que :

$$(\mathbf{G}_i)^{n-2} = -\sum_{k=1}^{\ell}\sum_{p=0}^{n-2}\left[\nabla_i\sigma_k\right]^{n-2-p}.{\lambda_k}^p$$

La dérivée d'ordre $n-2$ peut s'exprimer par :

$$\begin{aligned}\mathbf{F}_i^s(t_o) &= \left[\frac{d^s}{dt^s}\mathbf{F}_i(t)\right]_{t=t_o} \\ &= \sum_{j=1}^{N}\frac{\partial{\mathbf{F}_i}^{s-1}}{\partial\mathbf{r}_j}.\dot{\mathbf{r}}_j(t_o)\end{aligned}$$

On a alors :

$$\begin{aligned}
\mathbf{F}_i^s(t_o) &= \frac{d^s}{dt^s}\mathbf{F}_i(t) = \frac{\partial \mathbf{F}_i(t)^s}{\partial t^s} \\
&= \frac{\partial}{\partial t}\frac{\partial \mathbf{F}_i(t)^{s-1}}{\partial t^{s-1}}.\frac{\partial \mathbf{r}_j}{\partial \mathbf{r}_j} \\
&= \frac{\partial \mathbf{F}_i(t)^{s-1}}{\partial t}\frac{\partial \mathbf{r}_j}{\partial \mathbf{r}_j} \\
&= \frac{\partial \mathbf{F}_i(t)^{s-1}}{\partial \mathbf{r}_j}.\frac{\partial \mathbf{r}_j}{\partial t}
\end{aligned}$$

De la même manière, on a en remplaçant les dérivées partielles par l'opérateur ∇ :

$$\left[(\nabla_i \sigma_k)^s\right]_{t=t_o} = \left[\frac{d^s}{d^s t}\left(\nabla_i \sigma_k\right)\right]_{t=t_o} = \sum_{j=1}^{N} \nabla_j \left[\left(\nabla_i \sigma_k\right)^{s-1}\right]_{t=t_o} \dot{\mathbf{r}}_j(t_o)$$

A présent que l'on a exprimé la forme intégrée des équations de Newton, il faut déterminer les multiplicateurs de Lagrange et leurs dérivées (D.L.) pour accéder au mouvement complet. Pour cela, nous utilisons la définition de la contrainte : $\sigma_k(\mathbf{r}(t)) = (\mathbf{r}_j(t) - \mathbf{r}_i(t))^2 - d_{ij}^2$. Les dérivées de cette fonction sont donc nulles. On peut alors écrire que :

$$\sigma_k(\mathbf{r}(t_o))^{s+2} = \left[\frac{d^{s+2}\sigma_k(\mathbf{r}(t_o))}{dt^{s+2}}\right]_{t=t_o} = 0$$

L'interêt de dériver à l'ordre $s+2$ est de pouvoir dériver un terme d'ordre 2 à l'ordre s. Or, ce terme d'ordre 2 n'est autre que la dérivée seconde du vecteur position, qui represente la somme des forces. On évalue cette dérivée d'ordre $n+2$ de la même manière que $\mathbf{F}_i$ et $\mathbf{G}_i$. On a alors :

$$\begin{aligned}
d\sigma_k &= \sum_{j=1}^{N} \frac{\partial \sigma_k}{\partial \mathbf{r}_j}.d\mathbf{r}_j \\
&= \sum_{j=1}^{N} \left[\nabla_j \sigma_k\right].d\mathbf{r}_j \\
\Rightarrow \frac{d\sigma_k}{dt} &= \sum_{j=1}^{N} \left[\nabla_j \sigma_k\right]\frac{d\mathbf{r}_j}{dt}
\end{aligned}$$

Pour dériver cette expression à l'ordre $n+1$ et obtenir une dérivée à l'ordre $n+2$, il suffit de se souvenir que la dérivée d'un produit à l'ordre $n+1$ peut être donnée par l'expression : $(AB)^{n+1} = \sum_{k=0}^{n+1} A^k.B^{n+1-k}$. Ce qui nous permet d'écrire que :

$$\frac{d^{s+2}\sigma_k}{dt^{s+2}} = \left(\frac{d\sigma_k}{dt}\right)^{s+1} = \sum_{j=1}^{N}\sum_{\alpha=0}^{s+1} \mathbf{r}_j^{1+(s+1-\alpha)}(t_o)\cdot\left[(\nabla_j\sigma_k)^\alpha\right]_{t=t_o}$$

$$\Rightarrow \frac{d^{s+2}\sigma_k(\mathbf{r}(t_o))}{dt^{s+2}} = \sigma_k^{s+2}(\mathbf{r}(t_o)) = \sum_{j=1}^{N}\sum_{\alpha=0}^{s+1} \mathbf{r}_j^{s+2-\alpha}(t_o)\cdot\left[(\nabla_j\sigma_k)^\alpha\right]_{t=t_o} = 0$$

Or, $\mathbf{r}_j^{s+2-\alpha}(t_o)$ n'est autre que la dérivé seconde que l'on a dérivée $(s-\alpha)$ fois. Comme on connaît l'expression de la dérivée seconde (relation de Newton) :

$$\mathbf{r}_j(t)^2 = \frac{1}{m_j}\left[\mathbf{F}_j(t_o) - \sum_{k=1}^{\ell}\lambda_k(t_o)\left(\nabla_j\sigma_k\right)_{t=t_o}\right]$$

On peut alors, en dérivant cette expression $(s-\alpha)$ fois, réécrire $\frac{d^{s+2}\sigma_k(\mathbf{r}(\vec{t_o}))}{dt^{s+2}}$:

$$\begin{aligned}\frac{d^{s+2}\sigma_k(\mathbf{r}(\vec{t_o}))}{dt^{s+2}} &= 0.0 \\ &= \sum_{j=1}^{N}\sum_{\alpha=0}^{s+1}\frac{1}{m_j}\left[\mathbf{F}_j^{s-\alpha}(t_o) - \sum_{k=1}^{\ell}\sum_{\beta=0}^{s-\alpha}\lambda_k^\beta(t_o)\left[(\nabla_j\sigma_k)^{s-\alpha-\beta}\right]_{t=t_o}\right]\left[(\nabla_j\sigma_k)^\alpha\right]_{t=t_o}\end{aligned}$$

Pour la suite de notre étude, nous devons isoler les termes d'ordres supérieurs. Ces termes correspondent à $\alpha = 0$ et $\beta = s$.

$$\begin{aligned}\Rightarrow 0 &= \sum_{j=1}^{N}\frac{1}{m_j}\left[\mathbf{F}_j^s(t_o) - \sum_{k=1}^{\ell}\lambda_k^s(t_o)\left[(\nabla_j\sigma_k)\right]_{t=t_o}\right]\cdot\left[(\nabla_j\sigma_k)\right]_{t=t_o} \\ &+ \mathcal{F}\left[\left\{\lambda_p^{s-1}(t_o),...,\lambda_p^0(t_o)\right\},\left\{\mathbf{r}(t_o),\dot{\mathbf{r}}(t_o)\right\}\right] \qquad (9.39)\end{aligned}$$

Ici, $\mathcal{F}$ est fonction des dérivées de λ à un degré inférieur à s, des coordonnées et des moments. Si on se place dans le cadre du développement limité initial de degré 2 en $\mathbf{r}$, alors

la fonction $\mathcal{F}$ n'a plus aucune existence, si bien qu'on se retrouve seulement avec le premier terme. Dans ces conditions, la détermination des λ_k peut se faire par intégration numérique itérative de l'équation différentielle. De même, pour un développement complet, on peut de manière itérative déterminer l'ensemble des dérivées. On commence par $\{s = 0,\lambda\}$. Connaissant $\{s = 0,\lambda\}$ on détermine $\{s = 1,\lambda_1\}$, et ainsi de suite.

Si on détermine les multiplicateurs de Lagrange à partir de l'équation (9.39) de manière itérative, on n'est pas sûr de maintenir les contraintes à chaque pas d'intégration, si bien qu'on risque d'observer une dérive de cette distance. Comme la méthode est itérative, si une erreur systématique est commise sur la détermination des coordonnées (méthode de Verlet, par exemple) celle-ci sera propagée et amplifée à chaque pas d'intégration. Dans ces conditions, la contrainte ne pourra pas être maintenue. Le problème vient du fait que l'espace des contraintes est traité de la même manière que celui qui n'est pas contraint.

Pour remédier à ce problème, une procédure a été mise au point [20] permettant d'incorporer automatiquement les contraintes dans l'algorithme sans pour autant introduire d'erreurs supplémentaires sur la trajectoire. Supposons que l'algorithme utilisé soit celui de Verlet, l'erreur introduite sur les coordonnées (en utilisant la relation (9.39), c'est-à-dire la détermination de $n-2$ dérivées) est alors de l'ordre de $\Delta t^{(m+1)}$(dans le cas d'un développement de Taylor d'ordre 2, cela correspond au premier terme). Si cet algorithme correspond à la relation (9.39), alors on doit évaluer la dérivée d'ordre $n-2$ sur les forces. Comme $F \propto r^{(2)}$, cela revient à déterminer la dérivée d'ordre n sur l'ensemble des coordonnées $\mathbf{r}$. Ainsi, lors des intégrations successives, nous avons :

$$\sigma_k\left(\mathbf{r}^A(t)\right) = \mathcal{O}\left[(\Delta t)^{(m+1)}\right]$$

erreur introduite sur la contrainte

où $r^A(t)$ est la valeur obtenue à partir de l'algorithme donnant $r_i(t)$.
Au lieu d'utiliser cette méthode classique, on peut utiliser la relation (9.38). Cette relation, par la séparation du terme de plus haut degré, permet de séparer les $n-3$ premiers termes à t_o du terme $n-2$:

$$\left(\lambda_k^0(t_o),...,\lambda_k^{n-2}(t_o)\right) = \left(\lambda_k^0(t_o),...,\lambda_k^{n-3}(t_o)\right) + \left(\lambda_k^{n-2}(t_o)\right)$$

On peut alors remplacer le terme isolé $\lambda_k^{n-2}(t_o)$ par un jeu de paramètres γ_k tel que la relation $\sigma_k(\mathbf{r}(t,\gamma_k)) = 0$ soit satisfaite. Cela revient à diviser l'espace des phases en deux, un espace *contraint* et un espace *non contraint*. En appliquant cette subdivision à la détermination de l'ensemble des coordonnées, on a :

$$\begin{aligned} &\mathbf{r}(t_o+\Delta t;\{\lambda_k^{(o)}(t_o),...,\lambda_k^{(n-3)}(t_o,\gamma_k)\} \\ &= \\ &\mathbf{r}'(t_o+\Delta t;\{\lambda_k^{(o)}(t_o),...,\lambda_k^{(n-3)}(t_o)\} + \delta\mathbf{r}(t_o+\Delta t,\gamma_k) \end{aligned} \tag{9.40}$$

Ici, $\mathbf{r}'$ est indépendant de γ_k et correspond à l'espace non contraint, contrairement à $\delta\mathbf{r}$ qui correspond à l'espace des contraintes. En fait, le terme $\delta\mathbf{r}$ peut correspondre à un terme correctif. Le premier terme correspond à la détermination des coordonnées des atomes contraints dans l'espace non contraint, c'est-à-dire sans considérer le potentiel non contraint. Le deuxième terme doit, dans son expression, tenir compte du potentiel contraint. C'est donc $\delta\mathbf{r}$ qui permet de maintenir la contrainte et de corriger les erreurs engendrées par l'intégration de $\mathbf{r}'$. Ainsi, en tenant compte de la relation (9.40) on peut réécrire la relation $(\mathbf{r}_j(t) - \mathbf{r}_i(t))^2 - d_{ij}^2 = 0.0$.

$$\begin{aligned} & (\mathbf{r}_j(t) - \mathbf{r}_i(t))^2 - d_{ij}^2 = 0 \\ \Leftrightarrow\quad & (\mathbf{r}_j'(t) - \mathbf{r}_i{}'(t) + \delta\mathbf{r}_j(t) - \delta\mathbf{r}_i(t))^2 - d_{ij}^2 = 0 \\ \Leftrightarrow\quad & (\mathbf{r}_j'(t) - \mathbf{r}_i{}'(t))^2 + (\delta\mathbf{r}_j(t) - \delta\mathbf{r}_i(t))^2 \\ & +2.(\delta\mathbf{r}_j(t) - \delta\mathbf{r}_i(t)).(\mathbf{r}_j'(t) - \mathbf{r}_i'(t)) - d_{ij}^2 = 0 \\ \Leftrightarrow\quad & 2.(\delta\mathbf{r}_j(t) - \delta\mathbf{r}_i(t)).(\mathbf{r}_j'(t) - \mathbf{r}_i'(t)) \\ & +(\delta\mathbf{r}_j(t) - \delta\mathbf{r}_i(t))^2 = d_{ij}^2 - (\mathbf{r}_j'(t) - \mathbf{r}_i{}'(t))^2 \end{aligned} \tag{9.41}$$

Les $\delta\mathbf{r}_x(t)$ sont déterminées en sortant de la relation (9.41) le terme de plus haut degré (correspondant à λ_k^{n-2}). Cela nous permet d'écrire que :

$$\delta\mathbf{r}_i(t_o+\Delta t,\gamma_k) = -\frac{(\Delta t)^n}{m_i.n!}\sum_{k=1}^{\ell}\gamma_k.\left(\nabla_i.\sigma_k\right)_{t_o} \tag{9.42}$$

On se retrouve donc avec un système de ℓ équations à ℓ inconnues, susceptible d'être résolu sous forme matricielle de manière itérative. Il faut tout d'abord déterminer $\mathbf{r}_i(t_o + \Delta t)^{nc}$ pour tous les atomes y compris les atomes contraints (l'exposant ^{nc} indique des positions sur lesquelles on n'a pas appliqué la contrainte). Cela peut se faire par un D.L. d'ordre n.

$$\mathbf{r}_i^{nc}(t+\Delta t) = \mathbf{r}_i(t) + \dot{\mathbf{r}}_i(t)\cdot\Delta t + \frac{1}{m_i}\sum_{n=2}^{\infty}\frac{\Delta t^n}{n!}\cdot\left[\mathbf{F}_i(t)^{n-2}\right]$$

Ici, $\mathbf{F}_i(t)$ représente les forces non contraintes. Ensuite, on peut commencer la procédure qui détermine les coordonnées des atomes contraints, permettant le maintien de la contrainte. Pour cela, on utilise le fait que :

$$\mathbf{r}_i(t+\Delta t) = \mathbf{r}_i^{nc}(t+\Delta t) - \frac{(\Delta t)^n}{m_i.n!}\sum_{k=1}^{\ell}\gamma_k.\left(\nabla_i.\sigma_k\right)(t)$$

Ces expressions nous permettent de réécrire la relation (9.41) :

$$\begin{aligned}
-2\cdot\ & \left(\frac{(\Delta t)^n}{m_j.n!}\sum_{k=1}^{\ell}\gamma_k.\left(\nabla_j.\sigma_k\right)(t) - \frac{(\Delta t)^n}{m_i.n!}\sum_{k=1}^{\ell}\gamma_k.\left(\nabla_i.\sigma_k\right)(t)\right)\\
\cdot\ & \left(\mathbf{r}_j^{nc}(t+\Delta t) - \mathbf{r}_i^{nc}(t+\Delta t)\right)\\
+\ & \left[\frac{(\Delta t)^n}{m_j.n!}\sum_{k=1}^{\ell}\gamma_k.\left(\nabla_j.\sigma_k\right)(t) - \frac{(\Delta t)^n}{m_i.n!}\sum_{k=1}^{\ell}\gamma_k.\left(\nabla_i.\sigma_k\right)(t)\right]^2\\
=\ & d_{ij}^2 - \left(\mathbf{r}_j^{nc}(t+\Delta t) - \mathbf{r}_i^{nc}(t+\Delta t)\right)^2
\end{aligned}$$

$$\begin{aligned}
\Leftrightarrow\ & -2\cdot\left(\frac{(\Delta t)^n}{m_j.n!}\sum_{k=1}^{\ell}\gamma_k.\left(\nabla_j.\sigma_k\right)(t) - \frac{(\Delta t)^n}{m_i.n!}\sum_{k=1}^{\ell}\gamma_k.\left(\nabla_i.\sigma_k\right)(t)\right)\\
\cdot\ & \left(\mathbf{r}_j^{nc}(t+\Delta t) - \mathbf{r}_i^{nc}(t+\Delta t)\right)\\
+\ & \left[\frac{(\Delta t)^n}{m_j.n!}\sum_{k=1}^{\ell}\gamma_k.\left(\nabla_j.\sigma_k\right)(t) - \frac{(\Delta t)^n}{m_i.n!}\sum_{k=1}^{\ell}\gamma_k.\left(\nabla_i.\sigma_k\right)(t)\right]^2\\
=\ & d_{ij}^2 - \left(\mathbf{r}_j^{nc}(t+\Delta t) - \mathbf{r}_i^{nc}(t+\Delta t)\right)^2
\end{aligned}$$

On obtient ℓ équations à ℓ inconnues (multiplicateurs), qui peuvent se mettre sous forme matricielle. Les équations ainsi obtenues ne sont pas linéaires mais peuvent être résolues entièrement et efficacement par itérations : cela est justifié par le fait que pour $\Delta t \to 0$, les

termes γ_k sont petits.

Remarque : La détermination des multiplicateurs de Lagrange et des coordonnées contraintes se font au même moment lors des intégrations des équations de mouvement.

Application : algorithme de Verlet ou méthode des paramètres indéterminés Cet algorithme est un cas particulier du développement précédent, puisqu'on limite le D.L. à l'ordre 2 :

$$\begin{aligned} & -(\Delta t)^2 \sum_{k=1}^{\ell} \gamma_k . \left[\frac{\nabla_j}{m_j} - \frac{\nabla_i}{m_i}\right] . \sigma_k(t) \cdot \left(\mathbf{r}_j^{nc}(t+\Delta t) - \mathbf{r}_i^{nc}(t+\Delta t)\right) \\ + & \frac{1}{4} \cdot \left[\frac{(\Delta t)^2}{m_j} \sum_{k=1}^{\ell} \gamma_k . \left(\nabla_j . \sigma_k\right)(t) - \frac{(\Delta t)^2}{m_i} \sum_{k=1}^{\ell} \gamma_k . \left(\nabla_i . \sigma_k\right)(t)\right]^2 \\ = & \left(d_{ij}^2 - \left(\mathbf{r}_j^{nc}(t+\Delta t) - \mathbf{r}_i^{nc}(t+\Delta t)\right)^2\right) \end{aligned}$$

$$\begin{aligned} \Leftrightarrow & -(\Delta t)^2 \sum_{k=1}^{\ell} \gamma_k . \left[\frac{\nabla_j}{m_j} - \frac{\nabla_i}{m_i}\right] . \sigma_k(t) \cdot \left(\mathbf{r}_j^{nc}(t+\Delta t) - \mathbf{r}_i^{nc}(t+\Delta t)\right) \\ + & \frac{(\Delta t)^4}{4} \cdot \sum_{k=1}^{\ell} \sum_{k'=1}^{\ell} \gamma_k \gamma_{k'} \left[\frac{\nabla_j}{m_j} - \frac{\nabla_i}{m_i}\right] . \sigma_k(t) \cdot \left[\frac{\nabla_j}{m_j} - \frac{\nabla_i}{m_i}\right] . \sigma_{k'}(t) \\ = & \left(d_{ij}^2 - \left(\mathbf{r}_j^{nc}(t+\Delta t) - \mathbf{r}_i^{nc}(t+\Delta t)\right)^2\right) \end{aligned}$$

On se retrouve donc avec un système de ℓ équations du second degré à ℓ inconnues. Ce système peut être résolu par itérations successives lors de la dynamique moléculaire. Néanmoins, dans la plupart des simulations, on néglige les termes quadratiques, si bien que l'on se retrouve avec un système matriciel d'équations du premier degré.

Ciccotti et Ryckaert [20] utilisent le $\delta\mathbf{r}$ suivant :

$$\delta\mathbf{r}_i(t_o + \Delta t, \gamma_k) = -\frac{(\Delta t)^n}{m_i} \sum_{k=1}^{\ell} \gamma_k . \left(\nabla_i . \sigma_k\right)_{t_o}$$

alors que les calculs effectués montrent que :

$$\delta \mathbf{r}_i(t_o + \Delta t, \gamma_k) = -\frac{(\Delta t)^n}{m_i n!} \sum_{k=1}^{\ell} \gamma_k . \left(\nabla_i . \sigma_k\right)_{t_o}$$

soit un facteur $\frac{1}{n!}$ près. Néanmoins, il est possible de retrouver la relation de Ciccotti et Ryckaert, si on considère l'algorithme de Verlet [7] dans son intégralité, c'est-à-dire :

$$\mathbf{r}_i(t + \Delta t) = 2 \cdot \mathbf{r}_i(t) - \mathbf{r}_i(t - \Delta t) + \dot{\mathbf{r}}_i(t) \cdot \Delta t + \frac{\Delta t^2}{m_i} \cdot \left[\mathbf{F}_i(t) - \sum_{k=1}^{\ell} \gamma_k . \left(\nabla_i . \sigma_k\right)(t) \right]$$

Si l'on sépare les parties contrainte et non contrainte, cette relation s'écrit :

$$\mathbf{r}_i(t + \Delta t) = \mathbf{r}_i^{nc}(t + \Delta t) - \frac{\Delta t^2}{m_i} \cdot \left[\sum_{k=1}^{\ell} \gamma_k . \left(\nabla_i . \sigma_k\right)(t) \right]$$

Dans ces conditions, la matrice équation devient :

$$\begin{aligned}
\Leftrightarrow \quad & -2 \cdot (\Delta t)^2 \sum_{k=1}^{\ell} \gamma_k . \left[\frac{\nabla_j}{m_j} - \frac{\nabla_i}{m_i} \right] . \sigma_k(t) \cdot \left(\mathbf{r}_j^{nc}(t + \Delta t) - \mathbf{r}_i^{nc}(t + \Delta t) \right) \\
+ \quad & (\Delta t)^4 \cdot \sum_{k=1}^{\ell} \sum_{k'=1}^{\ell} \gamma_k \gamma_{k'} \left[\frac{\nabla_j}{m_j} - \frac{\nabla_i}{m_i} \right] . \sigma_k(t) \cdot \left[\frac{\nabla_j}{m_j} - \frac{\nabla_i}{m_i} \right] . \sigma_{k'}(t) \\
= \quad & \left(d_{ij}^2 - \left(\mathbf{r}_j^{nc}(t + \Delta t) - \mathbf{r}_i^{nc}(t + \Delta t) \right)^2 \right)
\end{aligned} \tag{9.43}$$

9.5. Méthode de non-équilibre

Dans le cas d'un couplage entre λ et t, la différence d'énergie entre deux états 0 et 1 est :

$$\begin{aligned}
\mathcal{H}_1(\tau(t_f)) \quad &= \quad \mathcal{H}_0(\tau(0)) + \int_0^{t_f} dt \left[\frac{\partial \lambda}{\partial t} \right] \frac{\partial \mathcal{H}(\tau(t))}{\partial \lambda} \\
&\Leftrightarrow \quad \mathcal{H}_1(\tau(t_f)) = \mathcal{H}_0(\tau(0)) + \mathcal{W}^{\,t_f}
\end{aligned}$$

Ici, $\mathcal{W}^{\,t_f}$ représente le travail nécessaire pour parcourir le chemin thermodynamique $0 \rightarrow 1$. Si t_f tend vers l'infini, on a alors l'équivalence suivante :

$$\mathcal{W}^{\,t_f \to \infty} = \Delta G = G_1 - G_0$$

Cette équation s'explique par le fait que la moyenne d'ensemble se fait sur un ensemble d'états ayant les mêmes états initiaux et finaux. Donc, si t_f tend vers l'infini, alors chaque état sera à l'équilibre et donc $\mathcal{W}^{\,t_f}$ sera identique pour chaque état. Ainsi, on aura :

$$< \mathcal{W}^{\,t_f \to \infty} >_{N_s} = \mathcal{W}^{\,t_f \to \infty}$$

On pourra alors écrire :

$$\Delta G = \mathcal{W}^{\,t_f \to \infty}$$

Si on subdivise le chemin thermodynamique en sous états, on aura alors :

$$\Delta G_{(0\,\to\,1)} = \mathcal{W}^{\,t_f \to \infty} = \sum_{\lambda(t)=0}^{\lambda(t_f)=1} \mathcal{W}(\lambda(t))$$

Le travail vaut donc :

$$\mathcal{W}(\lambda(t)) = \mathcal{H}(\lambda(t)) - \mathcal{H}(\lambda(t - \delta t))$$

On peut alors écrire que :

$$\mathcal{W}^{\,t_f \to \infty} = \sum_{t=0}^{t=t_f} \mathcal{H}_{t+1}(z(t)) - \mathcal{H}_t(z(t))$$

Cette méthode ne permet pas de déterminer la variation d'entropie puisque ΔS est la dérivée de ΔG par rapport à T, comme W ne dépend pas de T donc $\Delta S = 0$. Si on veut accéder à la variation d'entropie par une méthode similaire, il faut utiliser la relation de Jarzinsky [40] qui permet l'accès à des propriétés d'équilibre par des calculs sur des états hors équilibre. En effet, ces méthodes utilisent le même principe de couplage de λ par rapport au temps.

La principale différence réside dans le temps de simulation, puisque, dans ce cas, ces derniers sont très courts. On montre, que la variation d'enthalpie libre [35] s'écrit :

$$\Delta G_{(0\to 1)} = -kT \ln \left\langle \exp\left(-\frac{\mathcal{W}}{kT}\right)\right\rangle_0$$

Ici, la moyenne se fait sur un ensemble de simulations de type stochastique. Cette relation étant dérivable par rapport à T, nous devrions pouvoir déterminer une relation permettant de calculer la variation d'entropie :

$$\begin{aligned} T\Delta S_{(0\to 1)} &= kT \ln \left\langle \exp\left(-\frac{\mathcal{W}}{kT}\right)\right\rangle_0 \\ &- \left\langle \left[\mathcal{U}(\mathbf{r}^N)_0 + pV\right]\right\rangle_0 + \left(\frac{\left\langle \left[\mathcal{U}(\mathbf{r}^N)_0 + \mathcal{W} + pV\right] \exp\left(-\left[\frac{\mathcal{W}}{kT}\right]\right)\right\rangle_0}{\left\langle \exp\left(-\left[\frac{\mathcal{W}(\mathbf{r}^N)}{kT}\right]\right)\right\rangle_0}\right) \end{aligned}$$

Nous avons testé cette méthode sur deux systèmes simples :

- Calcul du potentiel chimique et de l'entropie molaire de l'eau
- calcul de la différence de variation d'entropie et d'enthalpie libre d'hydratation entre l'ion Ca^{2+} et l'ion Mg^{2+}.

Les résultats sont rassemblés dans le tableau suivant :

L'obtention de ces résultats a nécessité la mise en place de 3000 simulations avec $\tau = 3$ (nous avons aussi testé 1000 simulations avec $\tau = 10$)[7]. Nous avons généré les configurations par une dynamique de type stochastique avec un thermostat de Andersen [67] tel que $\delta t_{and} = 0.01$ ps dans l'ensemble NpT. Nous avons utilisé l'algorithme de *Verlet Leap Frog* avec un pas d'intégration de 0.01 ps.

Ces résultats montrent que cette méthode, bien qu'efficace pour évaluer la variation d'enthalpie libre, n'est pas très performante pour calculer la variation d'entropie. On peut expliquer cela par le fait que la détermination de l'entropie nécessite un échantillonnage conséquent de

[7] τ représente le nombre de configurations permettant d'écrire les moyennes d'ensemble.

	$\Delta\Delta G^o$, $\mu^\dagger$ (kJ mol^{-1})	$\Delta\Delta S^o$,$\overline{S}^\dagger$ (kJ mol^{-1})	$T\Delta\Delta H^o$, $\overline{H}^\dagger$ (kJ mol^{-1})
$H_20^\dagger$ (25^oC)	-43.2 $_5$	-10.1$_3$	-53.3$_6$
exp [73]	-45.5	7.6	-37.9
$Ca^{2+} \rightarrow Mg^{2+}$ (25^oC)	313.5 $_8$	-25.0 $_6$	338.5$_8$
Exp. [55]	347	24	323

TAB. 9.1 – Grandeurs thermodynamiques d'hydratation obtenues par l'égalité de Jarzinsky [40]. † indique les grandeurs évaluées pour l'eau. Exp. indique les valeurs expérimentales.

l'espace des phases [53], alors que cette méthode de non-équilibre est basée sur des temps de simulation assez courts.

Chapitre 10

Rappels sur la théorie de la DFT

Le concept fondamental de la DFT réside dans le fait que l'énergie d'un système électronique peut être exprimée en fonction de sa densité électronique qui ne dépend que de trois coordonnées spatiales (voire six si on considère deux populations de spins). Par contre, si on s'en tient à la description d'un système avec une fonction d'onde à plusieurs électrons N, celle-ci dépend alors de $3N$ variables pour un système sans spin. La théorie de la fonctionnelle de la densité electronique a été développée en 1964 et en 1965 par Hohenberg, Kohn et Sham [81, 82]. Elle consiste en la réduction du problème à plusieurs corps en un problème à un seul corps dans un champ effectif prenant en compte toutes les interactions. L'idée principale est que les propriétés exactes de l'état fondamental d'un système fermé, formé de noyaux positionnés dans des sites fixes et d'électrons les entourant, sont des fonctionnelles de la seule densité électronique. Ainsi, la DFT repose sur deux théorèmes fondamentaux :

1- L'énergie de l'état fondamental est une fonctionnelle unique de la densité électronique.
2- Le minimum de l'énergie totale du système correspond à la densité exacte de l'état fondamental.

L'approche de Kohn et Sham permet d'accéder à l'énergie totale du système (énergie cinétique incluse) bien que celle-ci laisse indéteminé le terme d'échange-corrélation. Cette approche consiste dans un premier temps à traiter de manière exacte le terme d'énergie cinétique en introduisant des orbitales. Dans un deuxième temps, le système est redéfini par rapport à un système d'électrons sans interaction et de même densité, de manière à faire apparaître les termes d'interaction comme des corrections aux autres termes. Pour traiter des systèmes polarisés en spin, nous devons séparer l'expression de la densité en deux parties.

$$\rho(\mathbf{r}) = \rho_{\uparrow}(\mathbf{r}) + \rho_{\downarrow}(\mathbf{r}) = \sum_{\mathbf{i=1}}^{\mathbf{N_e^{\uparrow}}} |\psi^{\uparrow}(\mathbf{r})|^{\mathbf{2}} + \sum_{\mathbf{i=1}}^{\mathbf{N_e^{\downarrow}}} |\psi^{\downarrow}(\mathbf{r})|^{\mathbf{2}} \tag{11.1}$$

où $\rho_{\uparrow}(\mathbf{r})$ et $\rho_{\downarrow}(\mathbf{r})$ désignent respectivement les densités électroniques relatives aux spin *up* et *down*. Dans le cas où les positions des N noyaux sont fixées, l'énergie totale du système peut alors s'exprimer de la manière suivante :

$$\begin{aligned} E_{tot}^{KS} &= \sum_i \left\langle \psi_i \left| \frac{-\nabla^2}{2} \right| \psi_i \right\rangle + \frac{1}{2} \int \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}' - \mathbf{r}|} d\mathbf{r} d\mathbf{r}' - \int \rho \mathbf{r} \sum_{I=1}^{N} \frac{Z_I}{|\mathbf{r} - \mathbf{R}_I|} d\mathbf{r} - \sum_{I<J}^{N} \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_I|} \\ &+ E_{xc}[\rho] \\ &\Leftrightarrow E_{tot}^{KS} = T_e^0 + E_H + E_{ext} + E_{xc}[\rho] \end{aligned} \tag{11.2}$$

Ici, T_e^0 est l'énergie cinétique du système d'électrons sans interaction, E_H désigne le terme de Hartree, E_{ext} inclut l'interaction coulombienne des électrons avec le noyau et celles des noyaux entre eux. Le terme $E_{xc}[\rho]$ comprend la déviation de l'énergie cinétique et les corrections au terme de Hartree, toutes deux dues aux corrélations entre électrons. Le terme Z désigne la charge du noyau. Déterminer l'état fondamental du système revient alors à résoudre, de manière auto-cohérente, un ensemble d'équations aux valeurs propres appelées équations de Kohn-Sham :

$$\left[\frac{-\nabla^2}{2} + \frac{\delta E_H}{\delta \rho(\mathbf{r})} + \frac{\delta E_{ext}}{\delta \rho(\mathbf{r})} + \frac{\delta E_{xc}}{\delta \rho(\mathbf{r})} \right] \left| \psi_i \right\rangle = \epsilon_i |\, \psi_i \rangle \,, \; i = 1, ..., N \tag{11.3}$$

La somme des termes $\dfrac{\delta E_H}{\delta \rho(\mathbf{r})} + \dfrac{\delta E_{ext})}{\delta \rho(\mathbf{r})} + \dfrac{\delta E_{xc}}{\delta \rho(\mathbf{r})}$ constitue un potentiel effectif V_{eff} qu'on peut qualifier de local car il ne dépend que de r. À ce niveau du calcul, tous les termes peuvent être déterminés excepté celui d'échange-corrélation. C'est par la manière de déterminer ce terme qu l'on va pouvoir différencier les multiples procédures DFT. Ces équations de Kohn et Sham ont une forme analogue à celle de Hartree-Fock. Elles se différencient par le fait qu'elles font intervenir la densité électronique totale du système. Les méthodes de résolution utilisent des bases d'orbitales atomiques gaussiennes analogues à celles utilisées dans un calcul Hartree-Fock et des expressions spécifiques pour le terme d'échange-corrélation correspondant à différents niveaux d'approximation. L'intérêt de ce type de calcul par rapport aux méthodes

ab initio réside dans le fait que, pour une base identique, le nombre de calculs est réduit et les effets de corrélation électronique sont pris en compte, contrairement aux méthodes Hartree qui nécessitent des calculs supplémentaires. Définissons, à présent, les termes d'échange et de corrélation.

- **L'effet d'échange** (corrélation de Fermi) résulte de l'antisymétrie de la fonction d'onde ψ. Il correspond au fait que deux électrons de même spin ont une probabilité nulle de se trouver au même entroit, ce qui se manifeste au niveau de la densité de paire par $\rho(r, r') = 0$. Cet effet est bien entendu relié au principe d'exclusion de Pauli. Il n'a pas lieu pour les électrons de spin opposé. De plus, il est pris en compte naturellement dans l'approximation Hartree-Fock par l'antisymétrie du déteminant de Slater. Il est à noter que cet effet est indépendant de la charge de l'électron.
- **La corrélation** (de Coulomb) est due à la charge de l'électron. Elle est reliée à la répulsion des électrons en $\frac{1}{r - r'}$. Contrairement à l'effet d'échange, elle est indépendante du spin. L'approximation Hatree-Fock néglige cet effet. Il faut effectuer un traitement post-Hartree-Fock pour y accéder.
- Le troisième effet est la correction de **self-interaction** qui conduit à une détermination correcte du nombre de paires d'électrons. Cette correction provient du fait que la fonction d'onde électronique est formulée en terme de particules indépendantes.

Afin d'inclure ces effets d'échange-corrélation dans l'énergie totale du potentiel effectif, on utilise le principe du *trou d'échange corrélation* qui traduit la diminution de densité électronique dans tout l'espace lié à la présence d'un électron en un point particulier (ce terme est toujours négatif). Il existe actuellement différents types d'approximations permettant de déterminer ce terme d'échange-corrélation.

L'approximation de densité locale *(LDA, local density approximation ; LSDA, local spin-density approximation)* est l'approximation sur laquelle repose pratiquement toutes les approches. Elle est basée sur le fait que dans le cas d'un gaz d'électrons homogène, l'énergie d'échange-corrélation peut être déterminée exactement. Cette approximation repose sur le fait que les termes d'échange-corrélation ne dépendent que de la valeur locale de $\rho(r)$. Par

conséquent, on néglige les effets des variations de densité. Cette approximation s'exprime par :

$$E_{xc}^{LDA} = \int \rho(\mathbf{r}) \epsilon_{xc}\left[\rho_{\downarrow,\uparrow}\right] d\mathbf{r} \tag{11.4}$$

où $\epsilon_{xc}\left[\rho_{\downarrow,\uparrow}\right]$ est l'énergie d'échange-corrélation par particule d'un gaz d'électrons uniforme qui a été paramétrée pour différentes valeurs de densité électronique.

L'approximation du gradient généralisé (*GEA*, *Gradient Expansion Approximation*) L'idée est qu'on peut tenir compte des variations locales de densité par l'intermédiaire du gradient de densité $\nabla\rho(\mathbf{r})$. Ainsi, l'approximation LDA apparaît comme le premier terme d'un développement en série de Taylor en fonction de ce gradient (*GEA*, *Gradient Expansion Approximation*). Cette méthode, bien que considérée comme une amélioration de la LDA, a donné des résultats surprenant puisqu'elle donne des valeurs positives aux trous d'échange. Ainsi le principe de trou d'échange-corrélation perd toute sa signification physique. Ce problème a été résolu en annulant tous les termes qui ne permettent pas au trou d'échange d'être négatif. Les fonctionnelles qui en résultent, sont appelées approximations de gradient généralisé *GGA*. En pratique, elles traitent séparément la partie d'échange et de corrélation.

L'énergie d'échange peut être décrite de la manière suivante :

$$E_{xc}^{GGA} = E_{xc}^{LDA} - \sum_{\sigma} \int d\mathbf{r} \rho_{\sigma}(\mathbf{r})^{4/3} \epsilon_{xc} F_x(x_\sigma) d\mathbf{r} \quad \text{avec} \quad x_\sigma = \frac{|\nabla\rho_\sigma|}{\rho_\sigma^{4/3}} \tag{11.5}$$

Le terme x_σ représente pour le spin σ le gradient de densité réduit. La puissance 4/3 lui confère un caractère adimensionnel. $\mathrm{F}_x(x_\sigma)$ est la fonctionnelle d'échange-corrélation. Une fonctionnelle développée par Becke [83], notée **B88**, a été conçue pour reproduire correctement le comportement asymptotique du terme d'échange. Elle a servi de base à de très nombreuses autres fonctionelles. Le point fort de cette méthode est qu'elle ne repose que sur un seul paramètre. Utilisée seule, la correction B88 estime aussi bien l'énergie que l'approximation LDA. Le point faible de cette méthode est d'ignorer le trou corrélation étant donné l'aspect prédominant du terme d'échange et la complexité du terme de corrélation. Perdew et Wang [84] (approximation notée **PW91**) ont considéré le problème autrement et ont décomposé l'énergie d'échange-corrélation en deux termes distincts. L'approximation PW91 se distingue de B88, principalement par le fait qu'elle traite l'échange-corrélation dans un même

cadre.

Les corrections des termes de corrélation sont beaucoup plus complexes à formuler que celles d'échange. De plus, leurs influences sur les propriétés électroniques et structurales sont moins significatives que celles concernant l'échange. Néanmoins, leur prises en compte est nécessaire pour l'obtention de résultats satisfaisants. Une des corrections les plus utilisées est celle proposée par Langreth et Mehl, et modifiée par Perdew afin de prédire les énergies de corrélation avec précision dans le cas d'atomes et de molécules.

Kohn et Sham ont établi une expression exacte de l'énergie d'échange-corrélation basée sur l'approximation de Hartree-Fock pour l'échange, le terme de corrélation restant inchangé par rapport à la LDA. Cette approche présente donc une séparation artificielle du terme d'échange et de corrélation, puisqu'on traite le trou d'échange comme non local (Hartree Fock) et le trou de corrélation comme local. Ainsi, le trou total perd de son autocohérence. En outre, ils ont montré que l'utilisation de cette fonctionnelle hybride (HF+LDA) fournit un comportement asymptotique correct du potentiel effectif (en 1/r). Or, un des problèmes de la LDA et de la GGA et qu'elles ne tiennent pas compte des effets de corrélation à longue portée. Il a été ainsi suggéré de reformuler la séparation d'échange/corrélation en incluant les effets à longue portée. Becke a donc utilisé l'échange exact (du moins une partie) pour formuler une énergie d'échange-corrélation. Il a ainsi proposé une expression à trois paramètres, **B3** :

$$E_{xc} = E_{xc}^{LDA} + a_0(E_x^{exacte} - E_x^{LDA}) + a_x \Delta E_x^{B88} + a_c \Delta E_c^{PW91} \tag{11.6}$$

Les coefficients a_0, a_x et a_c sont déterminés de manière semi-empirique par un ajustement sur des données expérimentales. E_x^{exacte} représente l'énergie totale exacte déterminée par un calcul Hartree-Fock. Grâce a cette approximation, les précisions sur les énergies sont bien meilleures que lorsqu'on utilise l'approximation GGA. C'est une variante de cette approximation proposée par Lee, Yang et Parr (**LYP**) connue sous le nom de **B3LYP** [85] que nous allons utiliser pour nos calculs d'énergie. Cette méthode donne des résultats extrêmement précis pour le calcul d'énergie. Cette méthode, bien que très séduisante, ne corrige que partiellement le problème de *self-interaction*, et ne met pas fin aux discussions concernant les problèmes d'échange-corrélation.

Bibliographie

[1] Bonal, C. ; Israeli, Y. ; Morel-Desrosiers, N. ; Morel, J.P. ; *J. Chem. Soc., Perkin. Trans. 2* **2001**, 1075.

[2] Douteau-Guevel, N. ; Coleman, A. W. ; Morel, J. P. ; Morel- Desrosiers, N. ; *J. Chem. Soc., Perkin. Trans. 2* **1999**, 629.

[3] Douteau-Guevel, N. ; Perret, F. ; Coleman, A. W. ; Morel, J. P. ; Morel-Desrosiers, N. ; *J. Chem. Soc., Perkin. Trans. 2* **2002**, 524.

[4] Perret, F. ; Morel-Desrosiers, N. ; Morel, J.P. ; *Supramol. Chem.* **2003**, *15*, 199.

[5] Mendes, A. ; Bonal, C. ; Israeli, Y. ; Morel-Desrosiers, N. ; Morel, J.P. ; Malfreyt, P. ; *J. Phys. Chem. B* **2002**, *106*, 4516.

[6] Ghoufi, A. ; Morel-Desrosiers, N. ; Morel, J.P. ; Malfreyt, P. ; *J. Phys. Chem. B* **2004**, *108*, 11744.

[7] Verlet, L. ; *Phys. Rev.* **1967**, *156*, 98.

[8] Gear, F. ; *Numerical initial value problems in ordinary differential equations.* **Pentice−Hall, Englewood Cliffs, NJ**, *3.2, 3.2.1, 8.7, E.1*, (1971).

[9] Hockney, R.W. ; *Methods. Comput. Phys.* **1970**, *9*, 136.

[10] Berendsen, H. J. C. ; Postma, J. P. M. ; Van Gunsteren, W.F ; DiNola, A. ; Haak, J. R. ; *J. Chem. Phys.* **1984**, *84*, 3684.

[11] Swope, W.C. ; Andersen, H.C. ; Berens, P.H. ; Wilson, K.R. ; *J. Chem. Phys.* **1982**, *88*, 6548.

[12] Gear, F. *Numerical integration of ordinary differential equations of various order.*, **Report ANL 7126, Argonne National Laboratory.** *3.2, 8.7, E..1*, (1966).

[13] Jorgensen, W. ; Maxwell, D. ; Tirado-Rives, J. ; *J. Am. Chem. Soc.* **1996**, *118*, 11225.

[14] Cornell, W. D. ; Cieplak, P. ; Bayly, C. I. ; Gould, I. R. ; Jr Merz, K. M. ; Ferguson, D.M. ; Fox, D.M. ; Spellemer, T. ; Caldwell, J.W. ; .Kolleman, P. ; *J. Am. Chem. Soc.* **1995**, *5179*, 117.

[15] Kong, C. ; *J. Chem. Phys.* **1973**, *59*, 2464.

[16] Waldman, M. ; Hagler, A. ; *J. Comput. Chem.* **1993**, *14*, 1077.

[17] Mc. Quarrie, D. A. ; *Statistical Mechanics* **University science book**, (2000).

[18] Allen, M. P. ; Tildesley, D. J. ; *Computer Simulation of Liquids* **Oxford : Clarendon Press**, (1989).

[19] Smith, E. R. ; *Proc. R. Soc. Lond. A* **1981**, *375*, 475.

[20] Ryckaert, J.P. ; Ciccotti, G. ; Berendsen, H.J.C. ; *J. Comput. Phys.* **1977**, *23*, 327.

[21] Jörgensen, W. L. ; Chandrasekhar, J. ; Madura, J. D. ; *J. Chem. Phys.* **1983**, *79*, 926.

[22] Tobias, D. J. ; Klein, M. L. ; *J. Phys. Chem.* **1996**, *100*, 6637.

[23] Yoshida, I. ; Yamamoto, N. ; Sagara, F., Ishii, D. ; Ueno, D. ; Shinkai, S. ; *Bull. Chem. Soc. Jpn.* **1992**, *65*, 1012.

[24] Scharff, J. P. ; Mahjoubi, M. ; Perrin, R. ; *New. J. Chem.* **1991**, *15*, 883.

[25] Arena, G. ; Cali, R. ; Lombardo, G.G., Rizzarelli, E. ; Sciotto, D. ; Ungaro, R. ; Casnati, A. ; *Supramol. Chem.* **1992**, *1*, 19.

[26] Street, W.B ; Tildesley, D.J. ; Saville, G. ; *Mol. Phys.* **1978**, *35*, 639.

[27] DL_POLY is a parallel molecular dynamics simulation package developed at the Daresbury Laboratory Project for Computer Simulation under the auspices of the EPSRC for the Collaborative Computational Project for Computer Simulation of Condensed Phases (CCP5) and the Advanced Reserach Computing Group (ARCG) at the Daresbury Laboratory.

[28] Schmit, M. W. ; Baldridge, K. B. ; Boatz, J. A. ; Elbert, S. T. ; Gordon, M. S. ; Jensen, J. J ; Koseki, S. ; Matsunaga, N. ; Nguyen, K. A. ; Su, S. ; Windus, T. L. ; Dupuis, M. ; Montgomery, J. A. ; *J. Comput. Chem.* **1993**, *65*, 14.

[29] Smith, W. ; Rapaport, D.C. ; *Mol. Sim.* **1992**, *9*, 25.

[30] Rudnick, J. ; Gaspari, G. ; *Science* **1987**, *237*, 384.

[31] Rapapport, D.C. ; *The art of molecular of dynamics simulation* **Cambridge University Press** (1995)

[32] Ota, N. ; Brünger, A. T. ; *Theor. Chem. Acc.* **1997**, *98*, 171.

[33] Ghoufi, A. ; Morel-Desrosiers, N. ; Morel, J.P. ; Malfreyt, P. ; *J. Phys. Chem.B* **2004**, *108*, 5095.

[34] Hansen, J.P. ; McDonald, I.R. ; in *Theory of Simple Liquids* Academic Press, New York, (1986).

[35] Frenkel, D. ; Smit, B. ; *Second edition of undersanding molecular simulation from algorithms and applications* **Academic Press**, *77*, (2002).

[36] Shirts, M.R ; Pietera, J.W. ; Swope, W.C. ; Pande, V.S. ; *J. Chem. Phys.* **2003**, *11*, 5740.

[37] Torrie, G.M. ; Valleau, J.P. ; *J. Comput. Phys.* **1977**, *23*, 187.

[38] de Koning, M. ; Antonelli, A. ; Yip, S. ; *Phys. Rev. Lett.* **1999**, *83*, 3973.

[39] Zwanzig, R.W ; *J. Chem. Phys.* **1954**, *22*, 1420.

[40] Jarcynsky, C. ; *Phys. Rev. Lett.* **1997**, *78*, 2690.

[41] Bennet, C.H. ; *J. Comput. Phys.* **1976**, *22*, 245.

[42] Watanabe, M ; Reinhart, W.P. ; *Phys. Rev. Lett.* **1990**, *65*, 3301.

[43] Torrie, G.M. ; Valleau, J.P. ; *Chem. Phys. Lett.* **1974**, *28*, 578.

[44] Mezei, M. ; Beveridge, D.L. ; *Ann. NY. Acad. Sci. USA* **1986**, *1*, 482.

[45] Aqvist, J. ; *J. Phys. Chem. B* **1990** *94*, 8021

[46] Buckner, J.K. ; Jorgensen, W.L. ; *J. Am. Chem. Soc.* **1989**, *111*, 2507.

[47] Rao. B.J. ; Singh, U.C. ; *J. Am. Chem. Soc.* **1989**, *111*, 3125 .

[48] Rizzo, R.C. ; Joregnsen, W.L. ; *J. Am. Chem. Soc.* **1999**, *121*, 4827.

[49] Swope, W.C. ; Andersen, H.C ; *J. Phys. Chem.* **1984**, *76*, 637.

[50] Mezei, M. ; Swaminathan, S. ; Beveridge, D.L. ; *J. Am. Chem. Soc.* **1978**, *100*, 3255.

[51] Pearlman, D.A. ; *J. Comput. Chem.* **1994**, *15*, 105.

[52] Lu, N. ; Kofke, D.A. ; Woolf, T. B. ; *J. Phys. Chem.B* **2003**, *107*, 5598.

[53] Peters, P. ; Oostenbrink, C. ; Van Dorp, A. ; van Gunsteren, W. ; *J. Chem. Phys.* **2004**, *120*, 2652.

[54] Kirkwood, J. G. ; *J. Chem. Phys.* **1935**, *3* 313.

[55] Burgess, M.A, in *Metal Ions in Solution* Ellis Horwood :Chichester, England, (1978).

[56] Kräutler, V. ; Van Gunsteren, W.F. ; Hünenberger, P.H. ; *J. Comput.Chem.* **2001**, *22, No.5*, 501.

[57] Ghoufi, A. ; Morel-Desrosiers, N. ; Morel, J.P. ; Malfreyt, P. ; *J. Phys. Chem. B* **2005**, *109*, 23579.

[58] Lehn, J.M. ; *La chimie supramoléculaire, concepts et perspectives.* **De Boeck Université**, (1997).

[59] Prue, J.E. ; *J. Chem. Ed.* **1969**, *46, No.1*, 12.

[60] MA, J.C. ; Dougherty, D.A. ; *Chem. rev.* **1997**, *97*, 1303.

[61] Morel, J.P. ; Morel-Desrosiers, N. ; *Org. Biom. Chem.* **2005**, *4*, 462.

[62] Sunner, J. ; Nishizawa, K. ; Kebarle, P.J. ; *J. Phys. Chem.* **1981**, *85*, 1814.

[63] Caldwel, J.W. ; Kollman, P.A. ; *J. Am. Chem. Soc.* **1995**, *117*, 4177.

[64] Frish, M.J. ; Head-Gordon, M. ; Schlegel, H.B. ; Raghavachari, K. ; Binkley, J.S. ; Gonzales, C. ; Defrees, D.J. ; Fox, D.J. ; Whiteside, R.A. ; Seeger, R. ; Melius, C.F. ; Baker, J. ; Martin, R.L. ; Kahn, L.R. ; Stewart, J.J.P. ; Fluder, E.M. ; Topiol, S. ; Pople, J.A. ; *GAUSSIAN 92, Gaussian.Inc, Pittsburgh PA* **1992**

[65] Gropp, W. ; Lusk, E. ; Skjellum, A ; *Using MPI, second edition* **The MIT Press 1999**

[66] Diu, B. ; Guthmann, C. ; Lederer, D. ; Roulet, B. ; *Physique statistique* **Hermann : Editeurs des sciences et des arts**, (1989).

[67] Andersen, H.C. ; *J. Chem. Phys.* **1980**, *72*, 2384.

[68] Nose, S. ; *J. Chem. Phys.* **1985**, *No.1 ; 511*, 81.

[69] Hoover, W.G. ; *Phys. Rev.* **1984**, *A31*, 1695.

[70] Brown, D. ; Clarke, J.H.R. ; *Mol. Phys.* **1983**, *No.5 ; 51*, 1243.

[71] Woodcock, L.V. ; Singer, K. ; *Trans. Faraday. Soc.* **1971**, *67*, 12.

[72] Heyes, D.M. ; *J. Chem. Phys.* **1981**, *74*, 1924.

[73] Ben-Naim A. ; Marcus, Y. ; *J. Chem. Phys.* **1984**, *81*, 2016.

[74] Kirkwood, J. G. ; *Theory of Liquids.* **B.J. Alder, ed., Gordon and Breach, New York**, (1968).

[75] Straatsma, T.P. ; Zacharias, M. ; McCammon, J.A. ; *J. Chem. Phys.* **1991**, *95, No.2*, 1175.

[76] Guerrault, X. ; Rousseau, B. ; Farago, J. ; *J. Chem. Phys.* **2004**, *13*, 6538.

[77] Patey, G.N. ; Valleau, J.P. ; *J. Chem. Phys.* **1975**,*63*, 2334.

[78] Torrie, G.M ; Valleau, J.P. ; *J. Comput. Phys.* **2001**, *23*, 187-199.

[79] McCammon, J. A. ; Harvey, S.C. ; *Dynamics of Proteins and Nucleic Acids.* **Cambridge : University press**, (1987).

[80] Straatsma, T.P. ; Zacharias, M. ; McCammon, J.A. ; *Chem. Phys. Lett.* **2001**, *22, No.5*, 501.

[81] Honenberg, P. ; Kohn, W. ; *Phys. Rev. B* **1964**, *136*, 864.

[82] Kohn, W. ; Sham, L.J ; *Phys. Rev. A* **1965**, *140*, 1133.

[83] Becke, A.D. ; *Phys. Rev. A* **1988**, *38*, 3098.

[84] Perdew, J.P. ; Wang, Y. ; *Phys. Rev. 3* **1991**, *45*, 13244.

[85] Becke, A.D. ; *J. Chem. Phys.* **1993**, *98*, 5648.

Index

A

alcools . . . 50
algorithme SHAKE . . . 223
approximation de Born-Oppenheimer . . . 24
asphéricité . . . 45

B

Berendsen
    algorithme . . . 201
Boltzmann . . . 182

C

calixarène . . . 45
chaos . . . 197
conditions périodiques aux limites . . . 40
conservation de l'énergie . . . 32
conservation des moments conjugués . . . 32
conservation locale . . . 197
contraintes . . . 224
correction à longue distance . . . 41
cutoff . . . 41

D

décomposition de Kitaura-Morokuma . . . 155
déterminisme . . . 195
DFT . . . 239
distribution de Maxwell-Boltzmann . . . 99
distribution radiale . . . 114

E

effet d'échange . . . 241
effet de corrélation . . . 241
ensemble μVT . . . 187
ensemble NpT . . . 188
ensemble NVT . . . 187
équations de mouvement . . . 31
espace des phases . . . 26, 182
état de non-équilibre . . . 235

F

fonction de partition . . . 183
fonctionnelle B3LYP . . . 243
force contrainte . . . 119

G

grandeurs absolues . . . 113, 125

H

hamiltonien . . . 30
hamiltonien quantique . . . 23
hypothèse de L. De Broglie . . . 23

I

interactions électrostatiques . . . 39
interactions cation-π . . . 151

J

jacobien de transformation canonique 79
Jarzinsky
    égalité 236

L

lagrangien 29
Liouville 197
    équation 199
    opérateur 199
longueur d'onde de De broglie 190

M

méthode FEP 80
    ΔG 80
    ΔS 84
méthode FEP^m 92
méthode TI 87
    ΔG 87
    ΔS 88
méthode TP 90
multiplicateurs de Lagrange 224

P

paramètre de couplage 74
Poisson
    équation de Poisson 214
    formule sommatoire 216
postulat de Schrödinger 23
potentiel de force moyenne 113, 114
potentiel intramoléculaire 37
potentiel Lennard-Jones 38
potentiel thermodynamique 193
Prue 125

R

rayon de giration 46
relation de Heisenberg 23

S

self interaction
    DFT 241
    sommation d'Ewald 222
sommation d'Ewald 213

T

théorème de fluctuation-dissipation 206

U

Umbrella sampling 115

V

Verlet
    algorithme 35

Z

Zwanzig 73