



**HAL**  
open science

# Vers une modélisation plus réaliste des systèmes biologiques

Fabien Archambault

► **To cite this version:**

Fabien Archambault. Vers une modélisation plus réaliste des systèmes biologiques. Autre. Université Henri Poincaré - Nancy 1, 2011. Français. NNT : 2011NAN10037 . tel-01746179

**HAL Id: tel-01746179**

**<https://hal.univ-lorraine.fr/tel-01746179v1>**

Submitted on 29 Mar 2018

**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.



## AVERTISSEMENT

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

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

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

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

## LIENS

Code de la Propriété Intellectuelle. articles L 122. 4

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

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

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

## AVERTISSEMENT

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

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

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

➤ Contact SCD Nancy 1 : [theses.sciences@scd.uhp-nancy.fr](mailto:theses.sciences@scd.uhp-nancy.fr)

## LIENS

Code de la Propriété Intellectuelle. articles L 122. 4

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

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

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

## ÉCOLE DOCTORALE SESAMES

SYNTHÈSES, EXPÉRIENCES, SIMULATIONS, APPLICATIONS : DE LA  
MOLÉCULE AUX ÉDIFICES SUPRAMOLÉCULAIRES

# THÈSE

pour obtenir le titre de

**Docteur en Sciences**

de l'Université de Nancy

**Mention : Chimie Informatique et Théorique**

Présentée et soutenue par Fabien ARCHAMBAULT

# Vers une modélisation plus réaliste des systèmes biologiques

Thèse dirigée par Christophe CHIPOT

soutenue le 5 juillet 2011

### Composition du jury :

*Rapporteurs :* Bertrand GUILLOT - Docteur, Université Pierre et Marie Curie  
Daniel BORGIS - Docteur, École Normale Supérieure de Paris

*Examineurs :* F. Javier LUQUE - Facultat de Farmàcia (Barcelona)  
Xavier ASSFELD - Professeur, Université Henri Poincaré  
François DEHEZ - Docteur, Université Henri Poincaré  
Christophe CHIPOT - Docteur, Université Henri Poincaré



## Remerciements

Je souhaite remercier tout particulièrement Daniel Borgis et Bertrand Guillot qui ont accepté d'être les rapporteurs de ce manuscrit de thèse. Je remercie également Xavier Assfeld et Javier Luque de m'avoir fait l'honneur d'être dans mon jury.

J'ai effectué cette thèse sous la direction de Christophe Chipot et François Dehez. Je les remercie chaleureusement pour toute l'aide et les connaissances transmises tout au long de cette thèse. Je les remercie aussi d'avoir autorisé les divers déplacements effectués tout au long de la thèse qui font découvrir d'autres groupes et d'autres environnements de travail.

Je souhaite remercier toutes les personnes qui étaient présentes pendant ces années au sein de l'équipe des dynamiques des assemblages membranaires mais aussi les autres groupes avec qui la discussion permettait toujours d'approfondir les connaissances ou de pouvoir avancer dans le projet.

Je voudrais finalement remercier Mélanie Jacquot et tous ceux qui ont cru en moi pour leur soutien précieux et leurs relectures de ce mémoire.



# Table des matières

<b>1</b>	<b>Champs de force et polarisation</b>	<b>1</b>
1.1	Champs de force additifs de paire . . . . .	4
1.2	La polarisation explicite est-elle nécessaire? . . . . .	9
1.3	Problématique et objectifs . . . . .	14
<b>2</b>	<b>Énergies intermoléculaires</b>	<b>17</b>
2.1	Méthodes quantiques . . . . .	18
2.1.1	Théorie des perturbations . . . . .	18
2.1.2	<i>Symmetry-Adapted Perturbation Theory</i> . . . . .	20
2.2	Méthodes empiriques . . . . .	22
2.2.1	Polarisabilités distribuées . . . . .	23
2.2.2	Charges Fluctuantes . . . . .	30
2.2.3	Oscillateurs de Drude . . . . .	33
<b>3</b>	<b>Développement d'un potentiel intermoléculaire polarisable simple et précis</b>	<b>37</b>
3.1	Référence quantique utilisée . . . . .	39
3.1.1	Question sur le $\delta HF$ . . . . .	39
3.1.2	Analogie entre SAPT et le potentiel intermoléculaire . . . . .	40
3.1.3	SAPT et les molécules chargées négativement . . . . .	41
3.2	Électrostatique . . . . .	41
3.2.1	Effets de pénétration . . . . .	44
3.2.2	Construction du modèle électrostatique . . . . .	46
3.2.3	Interactions électrostatiques . . . . .	50
3.3	Induction . . . . .	60
3.3.1	Construction du modèle de polarisabilité . . . . .	60
3.3.2	Énergie d'interaction d'induction . . . . .	64
3.3.3	Fonctions d'atténuation . . . . .	64
3.3.4	Énergie d'interaction d'induction entre molécules . . . . .	66
3.4	van der Waals . . . . .	75
3.4.1	Différentes expressions du potentiel de van der Waals . . . . .	75
3.4.2	Détermination des paramètres de van der Waals . . . . .	76



---

3.5	Énergie totale . . . . .	78
3.6	Vers les simulations de dynamique moléculaire . . . . .	89
3.7	Traitement des moments induits en dynamique moléculaire . . . . .	89
3.7.1	Le Lagrangien étendu . . . . .	89
3.7.2	Le traitement auto-cohérent . . . . .	90
3.8	Code de dynamique moléculaire : TINKER . . . . .	91
3.9	Cas d'une simulation de l'eau . . . . .	92
3.9.1	Nouveaux paramètres de van der Waals . . . . .	96
<b>4</b>	<b>Conclusion et perspectives</b>	<b>99</b>
<b>A</b>	<b>Fichiers d'entrée et de sortie SAPT</b>	<b>101</b>
A.1	Fichiers d'entrée . . . . .	101
A.2	Fichier de sortie . . . . .	104
<b>B</b>	<b>Détermination des paramètres de Tang and Toennies en mécanique quantique</b>	<b>109</b>
<b>C</b>	<b>Intégration des moments induits en dynamique moléculaire</b>	<b>111</b>

# Champs de force et polarisation

## Sommaire

<b>1.1 Champs de force additifs de paire</b> . . . . .	<b>4</b>
<b>1.2 La polarisation explicite est-elle nécessaire ?</b> . . . . .	<b>9</b>
<b>1.3 Problématique et objectifs</b> . . . . .	<b>14</b>

L'énergie potentielle  $V$  d'un système de  $N$  particules en interaction peut s'écrire comme une somme de potentiels  $V_n$ , à  $n$ -corps [1], où  $n = 1, 2, \dots$  :

$$\begin{aligned}
 V(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N) = & \sum_i^N V_1(\vec{r}_i) + \sum_{i<j}^N V_2(\vec{r}_i, \vec{r}_j) + \sum_{i<j<k}^N V_3(\vec{r}_i, \vec{r}_j, \vec{r}_k) \\
 & + \dots + \sum_{i<j<\dots<z}^N V_n(\vec{r}_i, \vec{r}_j, \dots, \vec{r}_z)
 \end{aligned} \tag{1.1}$$

Les vecteurs  $\vec{r}_i$  sont les vecteurs positions des  $N$  particules.  $V_1$  représente le potentiel intramoléculaire et  $V_2$  le potentiel d'interaction additif de paire. Les termes restants sont communément appelés potentiels à  $n$ -corps et résultent de l'interaction simultanée de  $n$  particules (avec  $n > 2$ ). Dans le cas de systèmes moléculaires, les interactions à  $n$ -corps sont en général d'un ordre de grandeur plus faible que les interactions de nature additive (à deux-corps). Toutefois, il est aujourd'hui admis qu'elles sont primordiales pour comprendre les propriétés de systèmes aussi simples que de petits agrégats d'eau [2-5]. En pratique, ces effets d'ordre supérieurs peuvent être partiellement inclus dans un potentiel effectif de paire [6] ( $V_2^{\text{effectif}}$ ) :

$$V(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N) \simeq \sum_i^N V_1(\vec{r}_i) + \sum_{i<j}^N V_2^{\text{effectif}}(\vec{r}_i, \vec{r}_j) \tag{1.2}$$

La plupart des potentiels empiriques utilisés pour l'étude de systèmes moléculaires en phase condensée reposent sur cette approximation. Le potentiel effectif de paire contient alors les effets moyens dus à la polarisation [6].

La polarisation électronique traduit la distorsion du nuage électronique d'une molécule sous l'effet d'un champ électrique créé par les molécules environnantes. En phase condensée, la polarisation participe d'un mécanisme auto-cohérent où chaque molécule est simultanément polarisée et polarisante. Pour illustrer schématiquement le processus de polarisation, considérons une isosurface de densité électronique pour une molécule d'eau isolée (figure 1.1). En présence d'une autre espèce chimique (ici un ion), générant un champ électrique, cette isosurface est déformée. Cette déformation traduit la polarisation électronique de la molécule d'eau. À champ constant, plus la molécule est polarisable, plus la surface est déformable.

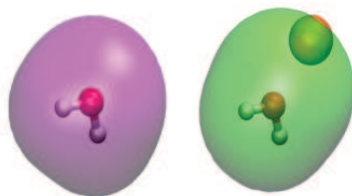


FIGURE 1.1 – Isosurface de densité électronique de la molécule d'eau perturbée par un ion (à droite) et non perturbée (à gauche).

Pour introduire de manière plus formelle les effets de polarisation, il convient d'exprimer le potentiel électrostatique par une série multipolaire (charges, dipôles, quadrupôles...) [7, 8]. Cette description, adéquat à longue distance, ne prend pas en compte les effets de pénétration électronique qui interviennent à proximité des noyaux (« sphère de divergence » [9]). À courte distance, l'expression multipolaire du potentiel électrostatique doit être complétée par un terme de pénétration. Les approches permettant de prendre en compte ces effets seront détaillées dans le chapitre suivant. Lorsqu'une molécule est polarisée par un champ électrique externe, l'ensemble des multipôles varie, générant ainsi des moments multipolaires induits. Pour un champ faible, l'approximation de la réponse linéaire, permet de relier le moment induit ( $\Delta\boldsymbol{\mu}$ ) au champ électrique local ( $\mathbf{E}$ ) :  $\Delta\boldsymbol{\mu} = \alpha\mathbf{E}$  où  $\alpha$  est un tenseur de polarisabilité moléculaire. Cette quantité traduit directement la capacité du nuage électronique d'une molécule à se déformer sous l'effet d'un champ électrique externe. Dans le cas où  $\boldsymbol{\mu}$  est un moment dipolaire,  $\Delta\boldsymbol{\mu}$  est un dipôle induit et  $\alpha$  est alors le tenseur de polarisabilité dipolaire. Si le nuage électronique peut se déformer de façon isotrope,  $\alpha$  se réduit à un scalaire. C'est typiquement le cas pour des ions comme par exemple le potassium où  $\bar{\alpha} = 5.736$  u.a. [10]. Dans le cas général, les électrons d'une

molécule n'ont pas la même liberté de mouvement dans les trois directions de l'espace et la polarisabilité ne peut plus être traitée comme un scalaire. Prenons l'exemple du benzène dont la polarisabilité dipolaire moyenne est de l'ordre de 69 u.a. [11]. Cette valeur ne permet pas de décrire simultanément la distorsion du nuage électronique de la molécule hors du plan (système  $\pi$ ) et dans le plan (système  $\sigma$ ) de la molécule. Il convient alors d'utiliser un tenseur d'ordre 3 reproduisant l'anisotropie de la polarisabilité du benzène :

$$\alpha = \begin{vmatrix} 48 & & \backslash \\ & 48 & \\ \backslash & & 79 \end{vmatrix} \text{ u.a.} \quad (1.3)$$

où la plus grande des valeurs, 79 u.a., traduit une polarisabilité plus importante le long de l'axe perpendiculaire au plan de la molécule.

Dans un système moléculaire en interaction, l'ensemble des moments induits peut interagir entre eux mais également avec les moments permanents de chaque molécule du système. Ces interactions sont à l'origine d'une énergie stabilisante : l'énergie de polarisation ou énergie d'induction. L'importance de cette contribution dépend directement de la polarisabilité des espèces mais également du pouvoir polarisant du milieu dans lequel elles se trouvent. D'une manière générale, l'énergie d'interaction intermoléculaire peut être directement obtenue à partir de méthodes quantiques. Les méthodes de décomposition comme, par exemple, Kitaura et Morokuma [12], la théorie des perturbations de Rayleigh-Schrödinger [13] ou *Symmetry-Adapted Perturbation Theory* [14] permettent d'accéder directement à l'énergie de polarisation. Dans la première méthode, la matrice de Fock est diagonalisée de façon à obtenir les contributions énergétiques de l'interaction. Les éléments diagonaux définissent les énergies électrostatiques, d'échange, de transfert de charge et de dispersion+polarisation. Dans la théorie de Rayleigh-Schrödinger (RS) les contributions d'électrostatique, d'induction et de dispersion sont obtenues en traitant l'énergie d'interaction comme une perturbation. La théorie SAPT est une extension de la théorie RS permettant également de décrire correctement les interactions à courte distance (contribution d'échange). L'ensemble de ces approches est limité par le temps de calcul nécessaire pour obtenir les énergies d'interaction et sont en pratique inapplicables pour étudier de grands assemblages moléculaires. Elles constituent en revanche un outil de choix pour détailler la physique des interactions intermoléculaires dans le cas de dimères voir de trimères, et à ce titre peuvent servir de référence dans la mise au point

de potentiels intermoléculaires modèle.

La modélisation de systèmes complexes, comme par exemple les protéines en phase condensée, requiert l'utilisation de potentiels intermoléculaires simples capables de décrire correctement la physique des interactions intermoléculaires et, pouvant être évalués rapidement. Dans le cadre de l'approximation de Born-Oppenheimer [15], le mouvement des électrons peut être dissocié de celui des noyaux du fait de leurs vitesses relatives. De plus, la distance intermoléculaire d'un système est très supérieure à la longueur d'onde de De Broglie ce qui permet de négliger les effets quantiques. Ces deux approximations sont à la base de la construction de potentiels intermoléculaires empiriques appelés encore champs de force. Ces derniers constituent un ensemble de fonctions mathématiques et de paramètres permettant d'évaluer l'énergie d'un système moléculaire à partir de la position des noyaux des atomes. Les champs de force standards reposent sur des potentiels additifs de paire effectifs dans lesquels les effets de polarisation sont inclus implicitement.

## 1.1 Champs de force additifs de paire

Il existe dans la littérature de nombreux champs de force empiriques permettant de modéliser des systèmes de nature différente. Un certain nombre d'entre eux ont été développés spécifiquement pour les systèmes biologiques : CHARMM [16, 17], AMBER [18], GRO-MOS [19], OPLS-AA [20]... L'écriture et la paramétrisation de ces champs de force diffèrent, mais ils possèdent tous des termes de valence et des termes d'interaction d'atomes non-liés. Dans le cas particulier du champ de force CHARMM [16], l'énergie potentielle  $U$  d'un système moléculaire en interaction contient plusieurs composantes qui modélisent les forces intra et intermoléculaires :

$$\begin{aligned}
 U = & \sum_{\text{liaisons}} K_L(r - r_0)^2 + \sum_{\text{angles}} K_\theta(\theta - \theta_0)^2 + \sum_{\text{dièdres}} \sum_n \frac{V_n}{2}(1 + \cos(n\phi - \gamma)) \\
 & + \sum_{\text{impropres}} K_\psi(\psi - \psi_0)^2 + \frac{1}{k^{1-4}} \sum_{\substack{\text{paires} \\ \text{non-liées}}} \left\{ \varepsilon_{ij} \left[ \left( \frac{R_{ij}^*}{r_{ij}} \right)^{12} - 2 \left( \frac{R_{ij}^*}{r_{ij}} \right)^6 \right] + \frac{1}{4\pi\varepsilon_0\varepsilon_1} \frac{q_i q_j}{r_{ij}} \right\}.
 \end{aligned} \tag{1.4}$$

**Interactions de valence :**

L'énergie de liaison est décrite par un potentiel harmonique, dans lequel l'énergie varie selon le carré du déplacement par rapport à la longueur de la liaison de référence  $r_0$  (en Å) avec  $r$  la longueur de la liaison, et  $K_L$  la constante de force de la liaison en  $\text{kcal.mol}^{-1}.\text{Å}^{-2}$ .

L'énergie de déformation angulaire est également décrite par un potentiel harmonique où  $\theta_0$  est la valeur de l'angle de référence en degré,  $\theta$  est la valeur de l'angle en degré et  $K_\theta$  est la constante de force de l'angle en  $\text{kcal.mol}^{-1}$ .

L'énergie de torsion associée à l'angle dièdre  $\phi$  (en degré) est une fonction périodique où les  $V_n$  sont des termes d'énergie exprimés en  $\text{kcal.mol}^{-1}$ ,  $n$  est le terme de multiplicité (nombre de minima entre 0 et  $360^\circ$ ) et  $\gamma$  le facteur de phase (en degré).

L'énergie associée à la déformation des angles dièdres impropres  $\psi$  (en degré) est décrite par un potentiel harmonique où  $\psi_0$  est l'angle dièdre de référence. Grâce à une constante de force  $K_\psi$  (en  $\text{kcal.mol}^{-1}.\text{Å}^{-2}$ ) élevée, ce terme permet de fixer la configuration d'un groupe d'atomes composé d'un atome central lié chimiquement à trois autres.

**Interactions d'atomes non-liés :**

Ces interactions impliquent les atomes d'une même molécule séparés par plus de deux liaisons chimiques ou les atomes de différentes molécules. Elles sont de deux natures : les interactions électrostatiques et de van der Waals. Lorsque les atomes sont séparés par exactement trois liaisons chimiques (termes dits « 1-4 »), les contributions entre atomes non-liés sont pondérées par le facteur  $k^{1-4}$ .

Les interactions électrostatiques entre deux atomes  $i$  et  $j$  sont calculées avec la loi de Coulomb où  $q_i$  et  $q_j$  sont les charges partielles portées par les atomes  $i$  et  $j$  et  $r_{ij}$ , en Å, est la distance entre les deux atomes  $i$  et  $j$ .

Les interactions de van der Waals sont décrites par un potentiel de Lennard-Jones en 6-12 où  $\varepsilon_{ij}$  et  $R_{ij}^*$  sont les paramètres de Lennard-Jones pour l'interaction entre les atomes  $i$  et  $j$  et  $r_{ij}$ , en Å, est la distance entre les deux atomes  $i$  et  $j$ .  $\varepsilon_{ij}$  correspond à la profondeur

du puits à la distance  $R_{ij}^*$ . Les valeurs  $\varepsilon_{ij}$  et  $R_{ij}^*$  sont obtenues à partir de paramètres atomiques ( $R_{ii}^*$ ,  $\varepsilon_{ii}$ ) et des règles de combinaison de Lorentz-Berthelot (équation 1.5).

$$\begin{aligned}\varepsilon_{ij} &= \sqrt{\varepsilon_{ii}\varepsilon_{jj}} \\ R_{ij}^* &= \frac{R_{ii}^* + R_{jj}^*}{2}\end{aligned}\tag{1.5}$$

Le potentiel de Lennard-Jones est composé d'un terme répulsif ( $r_{ij}^{-12}$ ) et d'un terme attractif ( $r_{ij}^{-6}$ ). Le terme en  $r_{ij}^{-6}$  représente les forces de dispersion ou forces de London, alors que le terme en  $r_{ij}^{-12}$  représente les forces d'échange-répulsion qui interdit aux électrons d'occuper la même région de l'espace.

### Stratégie de mise au point d'un champ de force :

La mise au point des paramètres des champs de force constitue une étape importante et délicate. Celle-ci repose sur la notion de types d'atomes définis non seulement par leur numéro atomique mais également par leur état d'hybridation et leur environnement chimique local. Quel que soit le système moléculaire dans lequel il apparaît, un atome d'un type donné sera toujours décrit par le même jeu de paramètres. Ce choix est fondé sur la notion de transférabilité qui fait l'hypothèse que les propriétés d'un type d'atome sont les mêmes d'une molécule à l'autre chimiquement apparenté. Il est évident, de part la simplicité des potentiels intermoléculaires utilisés dans les champs de force, que cette transférabilité n'est pas universelle et se limite souvent à une famille de molécules (par exemple les protéines) dans des conditions particulières (phase aqueuse à température et pression ambiantes). Par ailleurs, il est important de noter que l'optimisation des paramètres est une démarche essentiellement heuristique. À ce titre, il n'existe pas de stratégie unique clairement définie et je présente ici la philosophie générale sous-jacente aux différentes procédures existantes [21].

Les potentiels de valence pour les systèmes moléculaires sont généralement optimisés à partir de géométries (expérimentales ou obtenues par la mécanique quantique), de spectres vibrationnels, ou encore d'énergies de conformation, associés à des molécules de référence. Les valeurs d'équilibre de liaison, d'angle, de dièdre sont souvent optimisées pour reproduire des données géométriques en phase gazeuse obtenues par la mécanique quantique, la diffraction électronique ou encore par des expériences micro-ondes. Toutefois, les champs

de force dédiés à la modélisation de systèmes biologiques doivent prendre en compte les effets de l'environnement sur la géométrie. Les paramètres doivent donc être optimisés pour reproduire les géométries en phase condensée plutôt qu'en phase gazeuse [22]. Par exemple, il est possible d'utiliser les données cristallographiques présentes dans la CSD (*Cambridge Crystal Database*) [23]. Les paramètres géométriques sont alors optimisés à partir d'un ensemble de structures représentatives. Suivant le même objectif, les calculs de mécanique quantique peuvent être effectués en présence de molécules d'eau ou dans un solvant implicite [24].

La détermination des paramètres pour les potentiels entre atomes non-liés est généralement plus complexe car le nombre de données cibles est relativement restreint par rapport au nombre de paramètres à optimiser. De ce fait, il n'existe pas de solution unique limitant ainsi la mise en place d'algorithmes d'optimisation automatiques. La contribution de van der Waals aux champs de force empiriques est souvent considérée de moindre importance par rapport à la contribution électrostatique. Ce point de vue est la plus part du temps erroné. Des études ont montré que la contribution de van der Waals pouvait participer à plus de 50 % à l'énergie d'interaction moyenne dans le N-méthylacétamide (NMA) liquide [25] ou bien encore dans des cristaux de base d'acide nucléique [26]. Il apparaît donc clairement que l'optimisation des paramètres de van der Waals est essentielle pour la qualité des champs de force pour l'étude de biomolécules en phase condensée. Des avancées importantes dans l'optimisation des paramètres de van der Waals ont été réalisées avec le développement du champ de force OPLS [27]. Dans cette approche, les paramètres de van der Waals sont affinés pour reproduire des grandeurs telles que les énergies de vaporisation et les volumes moléculaires déterminés à partir de simulations Monte-Carlo. Des stratégies identiques ont également été appliquées pour développer les champs de force tels que CHARMM [16, 17] et AMBER [18]. Des méthodes alternatives basées seulement sur la mécanique quantique ont été proposées pour dériver les paramètres de van der Waals. En pratique, les approches quantiques seules ne peuvent rendre compte précisément des interactions de dispersion. De ce fait, les modèles obtenus décrivent assez mal les propriétés de systèmes en phase condensée.

La mise au point de modèles pour traiter les interactions électrostatiques est au centre du développement des champs de force. Pour les systèmes biologiques, la description est en général limitée à l'utilisation de charges ponctuelles atomiques. De nombreuses approches



ont été proposées dans la littérature pour dériver ces ensembles de charges [28] et toutes reposent, totalement ou en partie, sur la reproduction de données quantiques. Les champs de force standards tels que CHARMM, OPLS et AMBER utilisent principalement deux stratégies. Dans la première, l’optimisation des charges atomiques partielles est basée sur la reproduction du potentiel électrostatique autour d’une molécule calculé par la mécanique quantique [28–32] (méthodes ESP). De nombreux programmes tels que OPEP [33] permettent de réaliser cette opération. Le potentiel électrostatique de référence est déterminé en général en phase gazeuse ce qui limite a priori l’utilisation des charges atomiques ainsi obtenues. En pratique, les effets de polarisation dus à la phase condensée peuvent être partiellement inclus en utilisant par exemple, pour le calcul du potentiel, un niveau de théorie quantique (HF/6-31G\*) qui surestime artificiellement les moments dipolaires [31]. Une deuxième stratégie consiste à dériver des charges partielles capables de reproduire des énergies d’interaction et des distances d’équilibre pour des dimères impliquant une petite molécule avec elle-même ou une molécule d’eau [34]. Un des avantages de cette approche est de prendre en compte directement la polarisation électronique locale. Les effets de polarisation à  $n$ -corps sont en revanche inclus a posteriori par un ajustement sur les propriétés de la phase condensée.

Le développement des paramètres intra- et intermoléculaires n’est pas systématiquement découplé et certains champs de force ont comme dernière étape de leur procédure d’optimisation l’ajustement du potentiel de torsion [35] qui souvent absorbe les imperfections de paramétrisation des autres termes.

Cette stratégie générale d’ajustement ne permet pas de donner aux termes individuels du champ de force un sens physique rigoureux. En particulier, de nombreux effets électrostatiques pas ou peu décrits par le potentiel de Coulomb sont moyennés dans le potentiel de van der Waals, voire dans le potentiel de torsion. Ceci implique qu’un champ de force additif de paire n’est utilisable que dans son ensemble pour décrire des situations chimiques similaires à celles employées dans le processus de paramétrisation. Pour autant, ce cadre limitant n’a pas empêché les champs de force additifs d’être utilisés avec succès, en particulier dans le cas des protéines [36–40]. Par ailleurs, l’absence de traitement explicite de polarisation n’interdit pas de modéliser des systèmes en phase condensée. L’exemple le plus illustratif est sans aucun doute celui de l’eau. Le moment dipolaire de la molécule d’eau, en phase gazeuse, est de  $\sim 1.85$  D [41] alors qu’en phase aqueuse,

sous l'effet de la polarisation des molécules d'eau environnantes, il augmente et prend la valeur  $\sim 2.95$  D [42]. Les propriétés de l'eau liquide peuvent toutefois être modélisées par l'intermédiaire d'un potentiel additif de paires. De nombreux modèles, décrits de manière extensive dans la revue de Guillot [43], font appel à un jeu de charges partielles fixes représentatif du dipôle de l'eau en phase condensée. Les différentes propriétés chimiques de l'eau peuvent ainsi être reproduites par des potentiels additifs spécifiques bien qu'aucun de ces modèles n'ait la capacité de les décrire toutes précisément simultanément. Les champs de force additifs de paires ont été, sont et seront sans aucun doute utilisés avec succès pour modéliser une grande variété des processus chimiques. Ce constat conduit légitimement à se poser la question de la nécessité de développer des champs de force traitant explicitement les effets de polarisation à  $n$ -corps.

## 1.2 La polarisation explicite est-elle nécessaire ?

Depuis le début de l'utilisation des champs de force, la nécessité de traiter explicitement les effets de polarisation est source de nombreuses discussions [44]. Prenons l'exemple de petits solutés peu polaires dans un milieu faiblement polarisant, un potentiel additif de paires est tout à fait capable de reproduire bon nombre de propriétés physico-chimiques de ces systèmes, dont, par exemple, les énergies de solvation [45–48]. Pour modéliser de tels assemblages, il est peu pertinent d'alourdir inutilement la description des interactions intermoléculaires. Pour des assemblages plus complexes, l'approximation additive s'avère souvent suffisante pour rendre compte, si ce n'est quantitativement, au moins qualitativement de la physique de ces systèmes. Par exemple, pour le canal ionique de la gramicidine, il a clairement été montré que le traitement explicite de la polarisation ne fournit qu'une correction marginale au profil d'énergie libre associé à la translocation de cation sodium ou potassium au travers de cette protéine [49, 50].

Dans le cas de l'eau, le traitement explicite de la polarisabilité conduit à des conclusions mitigées. La plupart des modèles polarisables, bien qu'améliorant sensiblement la reproduction de certaines grandeurs ne permettent pas, à l'instar des modèles additifs, de reproduire l'ensemble du spectre des propriétés physicochimiques de l'eau. Il est toutefois difficile d'attribuer cette défaillance au seul problème de la polarisation. Le modèle électrostatique, le modèle polarisable bien souvent isotrope ou bien encore l'absence du

traitement explicite des autres effets à  $n$ -corps peuvent tout aussi bien être incriminés. Les auteurs d'un article récent ont réussi pour la première fois la prouesse de développer un potentiel capable de décrire aussi bien les propriétés du dimère que celles du liquide à partir de la seule mécanique quantique [51]. Ce travail ouvre sans aucun doute la voie vers la mise au point de nouveaux potentiels intermoléculaires « universels » à partir des principes de base de la mécanique quantique [52, 53].

Le traitement moyen de la polarisation dans les potentiels additifs peut apparaître naturel dans des milieux homogènes. En revanche, il est beaucoup plus difficile d'envisager qu'un même modèle additif puisse décrire simultanément l'interaction d'une molécule avec des environnements chimiques de polarité différente. Cette situation est fréquente et apparaît dès lors que les systèmes étudiés comportent des interfaces (mélange de liquides, membranes biologiques, protéines dans l'eau...). Le mélange eau-dichlorométhane montre [54] une diminution du moment dipolaire moyen de la molécule d'eau d'environ 30 % à proximité de l'interface entre les deux solvants et une augmentation d'environ 10 % pour le moment dipolaire moyen du dichlorométhane. Des résultats similaires ont été obtenus pour le mélange eau-tétrachlorométhane [55]. Ce type de systèmes peut être partiellement décrit par des modèles additifs [56–61]. Ces derniers permettent de reproduire les propriétés globales du mélange (densité, température critique, points d'ébullition...) mais en revanche n'assurent pas la description précise de l'interface. La description précise des interfaces est d'autant plus importante que les effets de polarisation sont grands. C'est le cas de l'étude de répartition d'ions à l'interface eau-air [62, 63], question centrale dans de nombreux processus chimiques (chimie de l'atmosphère, surface d'électrolytes, formation de micelles...). Prenons le cas de la translocation d'ions bromure ou iodure à l'interface eau-air. Un champ de force polarisable prédit que le minimum de l'énergie libre se situe au niveau de l'interface [64]. L'utilisation d'un potentiel additif montre en revanche une équiprobabilité de présence de l'anion dans la phase aqueuse. Il apparaît clairement que seule la description non additive permet d'expliquer pourquoi expérimentalement les anions s'accumulent préférentiellement au niveau de l'interface [65].

À l'intérieur des membranes biologiques, un important potentiel électrostatique positif [66, 67] est responsable de la différence de perméabilité entre des ions chargés négativement ou positivement [68–70]. Ce potentiel, aussi nommé potentiel dipolaire membranaire, joue un rôle important dans le contrôle de l'association des protéines

périphériques à la surface de la membrane ainsi que sur la structure et la fonction des protéines membranaires [71]. Des simulations de mono-couches lipidiques montrent sans ambiguïté que l'utilisation d'un potentiel polarisable, à la différence des modèles additifs existants, permet une reproduction du potentiel transmembranaire en très bon accord avec l'expérience [72, 73]. Des simulations de dynamique moléculaire de bicouches lipidiques polarisables publiées récemment, accréditent totalement les conclusions formulées à partir des monocouches [74, 75].

D'autres types d'interface, apparus avec l'émergence de nanosystèmes hybrides impliquant en particulier des interfaces solides et du matériel biologique (ADN, protéines, médicaments...) sont aujourd'hui l'objet d'étude de modélisation [76]. La description précise de processus d'association à de telles interfaces peut nécessiter un traitement rigoureux des changements de polarisation du substrat absorbé ainsi que de la surface, surtout lorsque celle-ci est fortement polarisable (nanotubes de carbone, nanoparticules...) [76–78].

Les champs de force additifs de paire sont en général considérés comme semi-quantitatifs [79, 80]. Certaines situations spécifiques peuvent néanmoins nécessiter d'être modélisées avec un grand degré de précision. Tel est le cas de certaines études pharmaceutiques où l'objectif est l'identification de molécules ayant une grande affinité pour une cible protéique donnée. Pour ce faire, des méthodes de *docking* [81] faisant appel à une fonction de score empirique sont en général utilisées. Toutefois pour des systèmes où les effets spécifiques liés à la polarisation sont importants ces méthodes peuvent s'avérer inefficaces. Une description plus précise des interactions protéines–ligands devient alors nécessaire pour obtenir une réponse adéquate. Prenons le cas de l'association de deux ligands chargés, la benzamidine et la diazamidine avec la trypsine [82]. Les différences d'énergie libre associées à la transformation alchimique benzamidine–diazamidine dans l'eau et dans la trypsine sont de l'ordre d'une dizaine de kcal/mol. La différence relative entre ces deux quantités n'est en revanche que de  $\sim 2$  kcal/mol. Jiao *et al.* ont montré que la modélisation de cette faible différence nécessitait l'utilisation d'un potentiel polarisable (ici AMOEBA [48, 83–85]) seul capable de rendre compte des effets spécifiques de l'eau et du site actif de la trypsine sur la distribution électronique des deux solutés chargés. Des conclusions similaires ont été obtenues par d'autres études d'association protéine–ligand [85–87].

Le développement de méthodes permettant de modéliser les structures de cristaux est d'un intérêt biologique fondamental, en particulier pour la prédiction d'isomorphes pour l'industrie pharmaceutiques [88]. Dans ce cadre, la description des effets de polarisation est d'une importance capitale car l'énergie d'induction contribue de 20 à 40 % à l'énergie électrostatique totale d'une maille cristalline [89]. Une description précise de l'énergie stabilisatrice comme la polarisation permet de prédire des paramètres structuraux de la maille sans données expérimentales [90] ou de reproduire des grandeurs mesurées expérimentalement à température ambiante [91–94] voir même à diverses pressions ou températures [95].

La modélisation précise des interactions intermoléculaires peut également être nécessaire pour décrire qualitativement certains processus chimiques comme la sélectivité des ions au travers des canaux ioniques [96, 97], et plus particulièrement, celle du cation potassium par rapport au cation sodium [98, 99]. Des simulations de canaux ioniques avec un champ de force polarisable [100–102] montrent l'importance des effets d'induction dans la description de la sélectivité. Pour le canal ionique KcsA  $K^+$ , les hélices  $\alpha$  possèdent un dipôle qui, dans le cadre d'un champ de force additif, stabilisent fortement les ions au sein du canal [103]. Au contraire, l'utilisation d'un champ de force polarisable prédit un effet beaucoup moins important de ce dipôle macromoléculaire [104] et une barrière totale de perméation de l'ion dans le canal d'environ 4 kcal/mol [105]. Cette description plus précise est en accord avec l'expérience [106] qui a montré que l'annulation des effets des hélices  $\alpha$  avait un rôle mineur dans la fonction de sélectivité du canal. En pratique, cette annihilation est réalisée par des mutations à l'intérieur du canal d'acides aminés neutres en acides aminés chargés positivement. Si elles sont bien choisies, ces mutations génèrent un dipôle global contre-balançant le dipôle macromoléculaire des hélices  $\alpha$ .

Un autre processus de transport qui a reçu beaucoup d'attention ces dernières années est celui qui se produit dans l'aquaporine [107–116]. Warshel et collaborateurs ont montré qu'un traitement explicite de la polarisation était nécessaire pour décrire quantitativement et qualitativement la translocation de l'ion  $H_3O^+$  [117]. En particulier, seul un modèle polarisable semble capable de reproduire à la fois l'interaction du proton et de l'eau et l'interaction du proton avec le motif asparagine-proline-alanine (NPA) [105, 112, 117–119] présent dans ces protéines.

D'une manière générale, la modélisation rigoureuse des interactions impliquant des

espèces fortement polarisables et fortement polarisantes est susceptible de requérir un traitement explicite de la polarisation. Le site actif des métalloprotéines impliquant souvent des acides aminés polarisables et de petits ions fortement polarisants illustre parfaitement cette situation. En particulier, Gresh et collaborateurs [86, 120–126] ont étudié des systèmes pour lesquels l’inclusion de transfert de charges entre le cation zinc ( $\text{Zn}^{2+}$ ) et les acides aminés sont importants pour décrire avec précision le site actif. Le traitement explicite de la polarisation (induction et transfert de charge) est nécessaire pour décrire les conformations obtenues expérimentalement de la protéine HIV-1 [126] possédant deux cations zinc au site d’interaction.

Jusqu’à récemment, les champs de force additifs de paires étaient utilisés dans le cadre de simulations de dynamique moléculaire de relativement courte durée (picosecondes, nanosecondes) permettant de modéliser un système à proximité d’une conformation locale, comme c’est le cas pour les simulations de protéines autour de la conformation obtenue par cristallographie. Aujourd’hui, avec l’avènement des supercalculateurs parallèles et des codes de dynamique moléculaire *scalables* appropriés, les temps caractéristiques accessibles se situent dans des échelles de temps plus grandes, de l’ordre de quelques microsecondes voir de la milliseconde [127] pour des systèmes composés de plusieurs dizaines de milliers d’atomes. Il est donc clair que l’espace conformationnel peut être décrit de façon plus extensive permettant ainsi d’accéder à d’importants changements de géométrie. Dans ce contexte, il est important de pouvoir modéliser précisément les différents états énergétiques du système ainsi que les barrières les séparant. Des études récentes effectuées par Freddolino *et al.* [128–130] ont mis en évidence que les champs de force additifs ne permettent pas toujours de décrire de telles transitions conformationnelles. Ils ont montré, en étudiant le repliement du domaine WW de la protéine humaine Pin1 que les champs de force CHARMM et AMBER n’étaient pas à même de reproduire fidèlement le chemin de repliement caractérisé expérimentalement. En particulier, la simulation reproduit un repliement préférentiel en hélices contrairement aux observations expérimentales mettant en évidence la présence de feuillets  $\beta$  [131]. Pour éviter de « mal » replier les protéines, il semble important, entre autres, d’utiliser un modèle de champ de force plus raffiné tel qu’un champ de force polarisable qui pourra aussi avoir des potentiels explicites pour les liaisons hydrogènes ou des sites non-atomiques [130]. Les phénomènes lents, de l’ordre de la  $\mu\text{s}$  tels que le repliement de protéines sont des cas d’études récents pour des simulations tout atomes [132–135] et de ce point de vue, il reste encore beaucoup à découvrir sur

l'importance du potentiel intermoléculaire.

Au travers de la liste non exhaustive présentée ici il est évident que la prise en compte explicite de la polarisation est nécessaire pour décrire quantitativement et/ou qualitativement de nombreux processus chimiques. Dans ce contexte, la mise au point de champs de force de nouvelle génération, plus généraux, plus précis et plus robustes pour des temps de simulation longs est plus que jamais d'actualité. Sur la scène internationale de nombreux efforts ont été effectués ces dernières années pour développer, en particulier des champs de force polarisables directement applicables pour la simulation de grands assemblages biologiques. Les nombreuses stratégies adoptées par les différents acteurs ont fait l'objet de nombreux articles de revue [38, 136–143].

### 1.3 Problématique et objectifs

La modélisation de grands assemblages moléculaires requiert l'utilisation de potentiels intermoléculaires simples et rapides à calculer. Le traitement explicite de la polarisation introduit dans le champ de force un degré de complexité supplémentaire. Pour que le surcoût d'évaluation des énergies d'interaction reste acceptable, il faut que le modèle de polarisation soit de conception relativement simple. De nombreuses approches font le choix d'utiliser un modèle rudimentaire (polarisabilité dipolaire isotrope ou équivalent) et dérivent les autres paramètres du champ de force sur la base de simulations de liquides. Ces stratégies garantissent une certaine efficacité en terme de temps de calcul mais n'assurent pas toujours de donner un sens physique individuel aux différents termes de l'énergie d'interaction.

L'objectif de mon travail de thèse est la mise au point d'un potentiel intermoléculaire polarisable alliant simplicité et précision. La stratégie mise en œuvre repose sur le développement terme à terme de l'énergie d'interaction à partir de calculs quantiques de référence impliquant de petites molécules isolées ou sous forme de dimères. Un des aspects importants de ce développement est la prise en compte explicite de l'anisotropie de la polarisabilité moléculaire.

Je présente dans le chapitre suivant la théorie des interactions intermoléculaires en

exposant dans un premier temps la théorie des perturbations à longue distance (Rayleigh-Schrödinger) puis en détaillant la théorie SAPT (*Symmetry-Adapted Perturbation Theory*) adéquate pour décomposer en termes physiques l'énergie d'interaction d'espèces proches les unes des autres. Dans ce même chapitre, après avoir introduit le formalisme des polarisabilités distribuées [9], je détaille quelques potentiels empiriques proposés dans la littérature pour traiter explicitement les effets de polarisation, en me limitant aux cas des systèmes d'intérêt biologique. Je discute en particulier des approches basées sur les moments induits [144], sur les oscillateurs de Drude [145] et enfin sur le formalisme des charges fluctuantes [146, 147]

Le modèle sera décomposé en trois contributions énergétiques différentes : l'électrostatique, l'induction et le van der Waals. Je présenterai, dans un premier temps, la référence quantique utilisée et, à partir des contributions quantiques, une comparaison au modèle empirique sera effectuée. Pour cela, j'exposerai les trois contributions séparément. Pour l'électrostatique, la dérivation du modèle passe par des charges atomiques ponctuelles obtenues dans le cas où la molécule est isolée puis cette molécule est mise en interaction avec une autre espèce ou elle-même pour comparer les énergies d'interaction quantiques et empiriques. Des fonctions *ad hoc* de pénétration électronique seront proposées pour modéliser le comportement à courte distance dans l'énergie d'interaction électrostatique. Pour l'induction, je parlerai, de la même façon, du cas où la molécule est isolée et pour laquelle le modèle est dérivé puis les énergies d'interaction d'induction de dimères modifiées par un potentiel d'atténuation. Enfin, une description sera faite de plusieurs formes de potentiel de van der Waals pour des molécules en interaction. Le comportement de chaque potentiel de van der Waals par rapport à la référence quantique sera argumenté. Lorsque les trois contributions énergétiques sont définies, l'énergie d'interaction totale calculée en mécanique quantique est comparée à celle obtenue par le modèle.

Le modèle, une fois défini, est complet pour effectuer des simulations de dynamique moléculaire. Je présenterai tout d'abord les méthodes permettant de calculer les moments induits en dynamique moléculaire. Puis, je présenterai le code de dynamique moléculaire, TINKER, utilisé pour intégrer le modèle de polarisabilité, ainsi que les modifications ajoutées afin de décrire celui-ci. Enfin, je montrerai les résultats des premières simulations de l'eau liquide, ainsi que les améliorations apportées pour corriger des approximations dues à l'approche utilisée pour dériver les paramètres.





# Énergies intermoléculaires

---

## Sommaire

---

<b>2.1 Méthodes quantiques</b> . . . . .	<b>18</b>
2.1.1 Théorie des perturbations . . . . .	18
2.1.2 <i>Symmetry-Adapted Perturbation Theory</i> . . . . .	20
<b>2.2 Méthodes empiriques</b> . . . . .	<b>22</b>
2.2.1 Polarisabilités distribuées . . . . .	23
2.2.2 Charges Fluctuantes . . . . .	30
2.2.3 Oscillateurs de Drude . . . . .	33

---

L'énergie d'interaction moléculaire peut être décomposée en trois contributions : l'énergie électrostatique, l'énergie d'induction et l'énergie de van der Waals. Parmi ces trois contributions, l'étude plus précise de l'énergie d'induction (ou de polarisation) sera discutée. Pour déterminer cette contribution, il existe différentes méthodes quantiques permettant de quantifier les énergies d'interaction (en particulier avec les théories perturbatives [14]). Les méthodes empiriques (polarisabilité distribuée, charges fluctuantes, oscillateurs de Drude...), quand à elles, permettent de modéliser l'énergie de polarisation. Dans la première partie de ce chapitre, les méthodes *ab initio* utilisant la théorie de la perturbation de Rayleigh-Schrödinger [13] seront présentées. Puis, la théorie, plus récente, adaptée aux interactions à courte distance, *Symmetry-Adapted Perturbation Theory* [14] (SAPT) sera décrite. À travers ces approches quantiques, la contribution à l'énergie d'induction sera définie parmi les autres contributions à l'énergie d'interaction totale. Dans la littérature des systèmes biologiques, trois grandes approches sont particulièrement utilisées pour modéliser l'énergie d'induction : les polarisabilités distribuées [9] (le plus souvent dipolaires isotropes), les charges fluctuantes [147] et les oscillateurs de Drude [145]. Dans la deuxième partie de ce chapitre, je développerai de façon détaillée le formalisme des polarisabilités distribuées [9] et la façon dont celles-ci permettent d'accéder à l'énergie d'induction. Le formalisme des polarisabilités distribuées

qui est la base du champ de force polarisable développé dans cette thèse sera présenté de façon détaillée. Ensuite, les approches des charges fluctuantes [147] et des oscillateurs de Drude [145] seront présentées brièvement à des fins de comparaison méthodologique. Ces trois approches permettent de modéliser explicitement les effets à  $n$ -corps et de prendre en compte l'influence de l'environnement.

## 2.1 Méthodes quantiques

### 2.1.1 Théorie des perturbations

Les énergies d'interaction de molécules sont généralement faibles par rapport aux autres énergies (liaison, torsion, dièdres...). L'énergie d'une liaison hydrogène est de l'ordre de la kilocalorie, alors que l'énergie de liaison des atomes ou de contrainte stérique, est souvent de l'ordre de dizaines de kilocalories. La théorie de la perturbation de Rayleigh-Schrödinger permet de trouver une solution approchée de l'énergie d'interaction en considérant celle-ci comme une perturbation faible à l'interaction totale. Dans le cas où les molécules sont assez éloignées, le recouvrement des nuages électroniques entre les molécules est négligeable ou nul. En prenant en compte cette approximation, London [148, 149] a été le premier à décrire cette théorie qui a ensuite été améliorée par d'autres scientifiques [150–154].

L'approximation de recouvrement nul proposée par London résulte de l'impossibilité d'échange entre les électrons pour des molécules éloignées. Si l'on suppose une fonction d'onde  $\Psi^A(1, 2, \dots, n_A)$  qui décrit la molécule  $A$  (une fonction des coordonnées des  $n_A$  électrons), et une fonction d'onde  $\Psi^B(1', 2', \dots, n'_B)$  décrivant les électrons de la molécule  $B$  (les « primes » permettent de séparer les électrons de la molécule  $A$  de ceux de la molécule  $B$ ). Il existe, par hypothèse, alors une région de l'espace pour laquelle la probabilité de trouver les électrons associés à la fonction d'onde  $\Psi^A$  (respectivement  $\Psi^B$ ) de la molécule  $A$  (respectivement  $B$ ) soit non-nulle. Les deux régions de l'espace sont, dans cette approximation distinctes. La fonction d'onde pour le système complet peut alors s'écrire comme un produit antisymétrisé  $\mathcal{A}\Psi^A\Psi^B$  où  $\mathcal{A}$  est un antisymétriseur pour les états  $A$  et  $B$ . La décomposition de ce produit fait apparaître des termes comme  $\Psi^A(1', 2, \dots, n_A) \times \Psi^B(1, 2', \dots, n'_B)$  dans lesquels l'électron 1 de la molécule  $A$  est échangé avec l'électron 2 de la molécule  $B$ . Le recouvrement entre le système sans échange et le

système où deux électrons sont échangés s'écrit :

$$\int \Psi^A(1, 2, \dots, n_A)^* \Psi^B(1', 2', \dots, n'_B)^* \times \Psi^A(1', 2, \dots, n_A) \times \Psi^B(1, 2', \dots, n'_B) \quad (2.1)$$

L'intégration sur les coordonnées de l'électron 1 est nulle car les deux fonctions d'onde ne se recouvrent pas. Ainsi chaque électron est défini dans son propre espace et aucun échange ne peut être défini. Il n'est donc pas nécessaire d'antisymétriser le produit pour obtenir le résultat final. Grâce à ces simplifications, les contributions à l'énergie d'interaction pour chaque molécule peuvent être déterminées. Ces équations sont développées dans le chapitre 4 du livre de Stone [9].

La théorie des perturbations est tout à fait valide dès lors que la distance entre les molécules est importante (par rapport à la taille des molécules). Par contre, l'approximation de recouvrement nul ne peut plus être supposée lorsque les molécules sont proches. En pratique, la contribution de recouvrement n'est pas toujours nulle, mais décroît de manière exponentielle avec la distance entre les molécules ce qui fait qu'elle est faible par rapport à l'énergie totale. Cependant, lorsque les molécules se rapprochent, le recouvrement n'est plus négligeable et doit être traité de manière appropriée (voir section suivante). De plus, dans la décomposition Rayleigh-Schrödinger, le potentiel électrostatique est calculé à partir d'un développement multipolaire [9]. À courte distance, l'approximation multipolaire n'est plus valide car elle ne permet pas de prendre en compte les effets de pénétration électronique voir n'a plus de sens dans le cas où l'interaction se situe dans la « sphère de divergence » [7–9]. L'énergie d'interaction ( $E_{\text{int}}$ ) dans la théorie des perturbations s'écrit, de façon simplifiée, comme la somme d'une énergie électrostatique ( $E_{\text{ele}}$ ), d'une énergie de dispersion ( $E_{\text{disp}}$ ) et d'une énergie d'induction ( $E_{\text{ind}}$ ) :

$$E_{\text{int}} = E_{\text{ele}} + E_{\text{disp}} + E_{\text{ind}} \quad (2.2)$$

Dans le cadre des simulations statistiques de systèmes biologiques, les molécules sont en général proches les unes des autres. Cette proximité fait que la contribution d'échange à l'énergie intermoléculaire ne peut plus être négligée. Pour que cette contribution apparaisse dans le développement perturbatif de l'énergie d'interaction, il est nécessaire d'autoriser l'échange des électrons des molécules en interaction. C'est précisément ce que permet de faire la théorie SAPT (*Symmetry-Adapted Perturbation Theory*).

### 2.1.2 *Symmetry-Adapted Perturbation Theory*

Par rapport à la théorie des perturbations à longue distance, un des principes à respecter pour les interactions à courte distance est le principe de Pauli. Lorsque deux molécules ont leur fonction d'onde qui se recouvre, alors le produit entre leurs fonctions d'ondes d'ordre zéro doit être antisymétrisé. Le principe de Pauli se traduit par :

$$|m_A n_B\rangle = \mathcal{A}|m_A\rangle|n_B\rangle \quad (2.3)$$

où  $|m_A n_B\rangle$  est antisymétrique pour toutes les permutations possibles entre les électrons et  $|m_A\rangle|n_B\rangle$  est antisymétrique seulement pour les échanges entre les électrons de  $A$  (ou de  $B$ ) au sein de la molécule.  $\mathcal{A}$  est un opérateur permettant d'antisymétriser les états  $m_A$  et  $n_B$ .

Pour respecter ce principe, dans la théorie SAPT (*Symmetry-Adapted Perturbation Theory*), l'hamiltonien total d'un dimère s'écrit  $\hat{H} = \hat{F} + \hat{V} + \hat{W}$ , où  $\hat{F} = \hat{F}_A + \hat{F}_B$  est la somme des opérateurs de Fock des monomères  $A$  et  $B$ ;  $\hat{V}$  est l'opérateur d'interaction intermoléculaire;  $\hat{W} = \hat{W}_A + \hat{W}_B$  est la somme des opérateurs de Møller-Plesset [155]. Ce dernier opérateur est défini comme  $\hat{W}_X = \hat{H}_X - \hat{F}_X$ , où  $\hat{H}_X$  est l'hamiltonien total du monomère  $X$ . L'énergie d'interaction,  $E_{\text{int}}$ , est alors écrite à partir d'un développement perturbatif :

$$E_{\text{int}} = \sum_{n=1}^{\infty} \sum_{j=0}^{\infty} (E_{\text{pol}}^{(nj)} + E_{\text{exch}}^{(nj)}) \quad (2.4)$$

Les indices  $n$  et  $j$  sont représentatifs de l'ordre utilisé pour développer les opérateurs  $\hat{V}$  et  $\hat{W}$ , respectivement. L'énergie de polarisation  $E_{\text{pol}}^{(nj)}$  est identique aux corrections obtenues avec la théorie des perturbations de Rayleigh-Schrödinger. Les corrections d'échange  $E_{\text{exch}}^{(nj)}$  proviennent de l'utilisation d'un antisymétriseur. Cette contribution permet de forcer les permutations nécessaires pour respecter le principe de Pauli de la fonction d'onde du dimère. Autrement dit, « la symétrie est adaptée » (*symmetry adaptation*).

Les corrections du terme de polarisation au premier ordre de l'opérateur  $\hat{V}$ ,  $E_{\text{pol}}^{(1j)}$ , décrivent l'énergie d'interaction électrostatique classique et sont notées  $E_{\text{elst}}^{(1j)}$ . Ces contributions se retrouvent aussi dans la théorie des perturbations à longue distance.

Les corrections de second ordre peuvent être décomposées en une partie d'induction et une partie de dispersion :

$$E_{\text{pol}}^{(2j)} = E_{\text{ind}}^{(2j)} + E_{\text{disp}}^{(2j)} \quad \text{et} \quad E_{\text{exch}}^{(2j)} = E_{\text{exch-ind}}^{(2j)} + E_{\text{exch-disp}}^{(2j)} \quad (2.5)$$

La composante d'induction est la somme des énergies d'interaction des moments multipolaires permanents d'un monomère avec les moments multipolaires induits de l'autre monomère. Le terme de dispersion provient de la corrélation induite par le mouvement des électrons d'un monomère sur les électrons de l'autre monomère. De manière similaire, la correction du terme de polarisation au troisième ordre peut être décomposée en :

$$E_{\text{pol}}^{(3j)} = E_{\text{ind}}^{(3j)} + E_{\text{ind-disp}}^{(3j)} + E_{\text{disp}}^{(3j)} \quad (2.6)$$

Le même traitement peut être appliqué pour la correction au terme d'échange. Une présentation plus détaillée des termes d'ordre trois est disponible dans les références [14, 156].

Rybak *et al.* [157] ont montré, par exemple, que la sommation des corrections d'échange et de polarisation d'ordre zéro en  $\hat{W}$  fournit une bonne approximation de l'énergie d'interaction supermoléculaire Hartree-Fock,  $E_{\text{int}}^{\text{HF}}$  :

$$E_{\text{int}}^{\text{HF}} = E_{\text{elst}}^{(10)} + E_{\text{exch}}^{(10)} + E_{\text{ind,resp}}^{(20)} + E_{\text{exch-ind,resp}}^{(20)} + \delta E_{\text{int,resp}}^{\text{HF}}, \quad (2.7)$$

où  $\delta E_{\text{int,resp}}^{\text{HF}}$ , défini par l'équation précédente, correspond aux termes d'ordres trois et quatre d'induction et d'échange induction. L'indice « resp » signifie réponse (*response*), qui traduit la prise en compte du couplage Hartree-Fock d'un système perturbé. En ajoutant la corrélation intramoléculaire équivalente au second ordre de théorie perturbative à plusieurs corps [158, 159], l'énergie d'interaction s'écrit :

$$E_{\text{int}}^{\text{SAPT2}} = E_{\text{int}}^{\text{HF}} + E_{\text{elst,resp}}^{(12)} + \varepsilon_{\text{exch}}^{(1)}(2) + {}^tE_{\text{ind}}^{(22)} + {}^tE_{\text{exch-ind}}^{(22)} + E_{\text{disp}}^{(20)} + E_{\text{exch-disp}}^{(20)}, \quad (2.8)$$

où l'énergie  $\varepsilon^{(1)}(2)$  est définie comme égale à la somme :  $\sum_{j=1}^2 E^{(1j)}$ .  ${}^tE_{\text{ind}}^{(22)}$  est la différence entre  $E_{\text{ind}}^{(22)}$  et  $E_{\text{ind,resp}}^{(20)}$ , qui correspond à la réponse du système en interaction sur l'énergie d'induction.  ${}^tE_{\text{exch-ind}}^{(22)}$  est une estimation de la partie d'échange,  ${}^tE_{\text{ind}}^{(22)}$  :

$${}^tE_{\text{exch-ind}}^{(22)} \approx E_{\text{exch-ind,resp}}^{(20)} \frac{{}^tE_{\text{ind}}^{(22)}}{E_{\text{ind,resp}}^{(20)}}. \quad (2.9)$$

À une distance interatomique  $R$  suffisamment grande pour négliger les termes d'échange, la théorie SAPT est similaire à la théorie des perturbations de Rayleigh-Schrödinger. De plus amples détails sur le développement de l'énergie perturbative sont disponibles dans la revue de Jeziorski *et al.* [14].

L'utilisation de cette théorie pour décomposer l'énergie d'interaction est de plus en plus généralisée pour obtenir des énergies de référence pour des modèles empiriques ou pour comprendre des phénomènes spécifiques comme la nature des interactions entre systèmes  $\pi$ . Bukowski *et al.* [160] ont, par exemple, utilisé les décompositions obtenues avec la théorie SAPT pour déterminer des paramètres pour le potentiel de liaison entre atomes. La théorie SAPT permet aussi de dériver des paramètres atomiques pour des potentiels empiriques [51–53, 91, 92, 161–164]. Dans ces approches, l'énergie d'interaction SAPT est comparée à des potentiels empiriques et, quand cela est possible, des paramètres atomiques sont dérivés à partir de deux espèces en interaction. C'est cette stratégie que j'ai choisie dans ma thèse pour dériver les paramètres d'un champ de force polarisable. Une autre étude avec la théorie SAPT est celle des interactions des systèmes  $\pi$  avec diverses espèces [165–171]. Le fait de pouvoir décomposer les énergies en contributions distinctes permet de comprendre des phénomènes d'association d'espèces chimiques. L'énergie de transfert de charge est un exemple de contribution accessible grâce à la décomposition SAPT [172] et cette contribution a fait l'objet d'un intérêt particulier dans des interactions telles que des ions avec des protéines [86, 173–176].

## 2.2 Méthodes empiriques

Les méthodes présentées dans la partie précédente permettent de décomposer l'énergie d'interaction de molécules de façon à obtenir, en particulier, l'énergie d'induction. Cela permet d'avoir une énergie de référence calculée *ab initio*. Par contre, de part le nombre de calculs à mettre en œuvre, l'utilisation de telles méthodes pour des systèmes autres que des fragments moléculaires n'est pas envisageable. Dans cette partie, je ne parlerai que de l'énergie d'induction pour la décomposition, car pour chaque méthode, le potentiel électrostatique et le potentiel de van der Waals sont définis séparément et reposent sur des principes utilisés dans les champs de force classiques. Ces dernières années, la littérature a été riche pour discuter des trois grandes familles de méthodes ajoutant explicitement de

la polarisation électronique [38, 137–140, 142, 143, 177]. En particulier, certains champs de force polarisables sont dédiés aux systèmes biologiques. Parmi les méthodes détaillées dans la littérature, je parlerai plus spécifiquement des polarisabilités distribuées puisque cette méthode est celle qui sera employée pour développer le champ de force polarisable présenté dans cette thèse. Puis, je parlerai des charges fluctuantes et de leurs applications actuelles. Enfin, la dernière méthode que j’aborderai sera celle des oscillateurs de Drude et de leurs développements récemment publiés.

### 2.2.1 Polarisabilités distribuées

Pour introduire la notion de polarisabilité distribuée, que je vais utiliser dans le modèle présenté dans cette thèse, il est nécessaire de définir les interactions électrostatiques. Il existe une grande variété de modèles d’interaction électrostatique. Leur complexité s’échelonne des simples distributions de charges aux développements multipolaires multicentriques complétés par des polarisabilités. Les méthodes permettant de déterminer ces modèles sont très nombreuses [28–32].

La façon la plus simple d’obtenir des distributions de charges centrées sur les atomes est d’utiliser l’analyse de population de Mulliken [29]. De tels modèles sont obtenus très rapidement car la plupart des programmes *ab initio* et semi-empiriques les calculent systématiquement. Ces distributions de charges sont pourtant rarement utilisées car elles reproduisent très mal le dipôle moléculaire et le potentiel électrostatique autour de la molécule [32, 178, 179]. Une meilleure méthode consiste à ajuster les charges afin de reproduire le potentiel électrostatique d’une fonction d’onde *ab initio* [180].

Dans le formalisme du développement multipolaire, l’interaction électrostatique est décrite par une série en puissance de  $1/R$  dont les coefficients correspondent aux différents moments : charges, dipôles, quadrupôles, etc. Il est bien connu qu’un tel développement monocentrique ne converge pas à courte distance [9]. Ce problème a été résolu en réalisant un développement multicentrique dans lequel les moments sont associés à des régions de la molécule à des atomes ou à des groupes d’atomes.

Une façon d’obtenir ces modèles consiste à utiliser l’analyse des multipôles distribués (DMA) proposée par Stone [7, 181]. Le point de départ de cette méthode est une matrice



densité développée sur une base de fonctions gaussiennes.

$$\begin{aligned}\rho(\mathbf{r}) &= \sum_{tu} \rho_{tu} \phi_t^A(\mathbf{r}) \phi_u^B(\mathbf{r}) \\ \phi_t^A(\mathbf{r}) &= f_t(\mathbf{r} - \mathbf{A}) \exp \left[ -\alpha(\mathbf{r} - \mathbf{A})^2 \right]\end{aligned}\tag{2.10}$$

où  $f_t$  représente les forces sur les particules. Étant donné les propriétés de ces deux fonctions, chaque produit de fonctions primitives placées sur deux centres  $\mathbf{A}$  et  $\mathbf{B}$  peut se ramener à une densité de charges centrée en un seul point  $\mathbf{P} = \frac{(\alpha\mathbf{A} + \beta\mathbf{B})}{\alpha + \beta}$ , sur lequel est effectué le développement multipolaire. Les multipôles obtenus sont ensuite translatés sur un site atomique ou arbitraire (par exemple le centre de liaison) le plus proche. La position des sites sur lesquelles est effectué le développement multipolaire dépend des valeurs des exposants  $\alpha$  et  $\beta$  des fonctions primitives et il s'ensuit que le résultat varie notablement avec la base.

Cette dépendance peut être réduite si l'analyse se fait non plus selon les propriétés des fonctions constituant la base, mais en partageant l'espace moléculaire selon des propriétés physiques. La théorie topologique de Bader [182] des atomes dans la molécule (AIM) consiste à partager l'espace en régions disjointes  $\Omega_A$ . Ces régions sont séparées par des surfaces,  $S_A(\mathbf{r})$ , où la densité électronique de charge  $\rho(\mathbf{r})$  satisfait à la condition de flux nul :  $\nabla\rho(\mathbf{r}) \cdot \mathbf{n}(\mathbf{r}) = 0$   $\mathbf{r} \in S_A(\mathbf{r})$ , où  $\mathbf{n}(\mathbf{r})$  est un vecteur normal à la surface  $S_A(\mathbf{r})$ .

Selon le formalisme d'analyse des multipôles distribués, l'énergie d'interaction électrostatique d'un ensemble de molécules est définie comme :

$$U_{\text{ele}} = \frac{1}{2} \sum_A \sum_{B \neq A} Q_t^A T_{tu}^{AB} Q_u^B\tag{2.11}$$

où  $Q_t^A$  et  $Q_u^B$  sont les composantes multipolaires (charges, dipôles, quadrupôles...) des sites localisés sur les molécules A et B, respectivement,  $T_{tu}^{AB}$  est la matrice d'interaction dont les éléments dépendent de la distance entre les sites, de leur orientation relative et de l'ordre des composantes des multipôles. Les formules permettant de calculer ces éléments de matrice ont été dérivées et sont disponibles dans la littérature (table F.1 de l'ouvrage de Stone [9]).

Comme nous l'avons mentionné précédemment, les potentiels intermoléculaires se limitent le plus souvent au second terme de l'équation 1.1, qui correspond à des interactions

de paire, et les effets à  $N$  corps y sont pris en compte de façon moyenne. Cependant, pour modéliser précisément les interactions d'espèces polarisables, l'inclusion explicite de la polarisabilité est indispensable pour décrire la distorsion de la distribution de charges sous l'effet du champ électrique créé par les molécules voisines. Cela est également nécessaire pour développer des modèles valables aussi bien en phase gazeuse qu'en phase condensée. Par exemple, le dipôle de l'eau passe de 1.85 D dans la phase gazeuse à environ 2.95 D dans la phase liquide et seul un modèle polarisable permet de reproduire une telle variation.

L'introduction de polarisabilités dans les potentiels d'interaction a principalement été réalisée pour des études de l'eau et de systèmes ioniques [183–189]. Dans la plupart des cas, la polarisabilité est introduite dans les potentiels à l'aide de polarisabilités dipolaires locales ou de charges variant en valeur et/ou en position.

Un meilleur modèle, d'un point de vue théorique, consiste à utiliser des polarisabilités distribuées ou des polarisabilités d'ordre plus élevé obtenues par les calculs de chimie quantique. Plusieurs méthodes de partition existent. Stone [190] a développé un formalisme général qui introduit des sites de polarisabilité décrivant le changement de charges, dipôles, quadrupôles... en réponse à un potentiel, un champ électrique, un gradient de champ... Ce modèle offre une vue détaillée des effets d'induction grâce à des polarisabilités locales et non locales permettant, en particulier, aux charges atomiques partielles de fluctuer d'un site à l'autre. Le Sueur et Stone [191] ont exposé des méthodes pratiques pour calculer des polarisabilités distribuées. Bader [192, 193] a appliqué sa théorie pour calculer des polarisabilités locales. Plus récemment, Ángyán et Chipot [28] ont implanté le calcul des polarisabilités distribuées selon la définition de Stone mais en utilisant la théorie topologique de Bader des atomes dans la molécule et ont montré que cette méthode conduit à des résultats transférables et stables vis-à-vis de la base de calcul.

Ces modèles de polarisabilités distribuées utilisés conjointement avec des multipôles distribués constituent des modèles électrostatiques très précis. Il est cependant nécessaire de tronquer de tels développements afin de réduire les temps de calcul dans le contexte de simulations. Dans ce travail, nous présentons des calculs utilisant des développements tronqués à l'ordre 1 pour modéliser les polarisabilités distribuées.

En ce qui concerne les polarisabilités distribuées, plusieurs méthodes de relocalisa-

tion [194, 195] ont récemment été proposées afin de réduire considérablement le nombre de composantes. En conservant les termes de flux de charges totalement distribués, la relocalisation des composantes charge–dipôles et dipôle–dipôle s’effectue selon l’expression :

$$\alpha_{\alpha\beta}^A = \frac{1}{2}(\alpha_{\alpha\beta}^{AB} + \alpha_{\alpha\beta}^{BA}) + \mathbf{r}_{\alpha}^{AB}(\alpha_{\mathbf{q}\beta}^{AB} - \alpha_{\mathbf{q}\beta}^{BA}) + \mathbf{r}_{\beta}^{AB}(\alpha_{\alpha\mathbf{q}}^{AB} - \alpha_{\alpha\mathbf{q}}^{BA}) \quad (2.12)$$

où  $\mathbf{r}^{AB} = \mathbf{R}^A - \mathbf{R}^B$  représente le vecteur entre les sites A et B et  $(\alpha, \beta) = 00, 10, 11c, 11s$ .

La formule générale pour recalculer les composantes de la polarisabilité moléculaire totale a été donnée par Stone [190]. Dans notre cas, en se limitant à une distribution de composantes de flux de charges  $\alpha_{\mathbf{q}\mathbf{q}}^{AB}$ , charge–dipôle  $\alpha_{\mathbf{q}\beta}^{AB}$ , dipôle–charge  $\alpha_{\alpha\mathbf{q}}^{AB}$  et dipôle–dipôle  $\alpha_{\alpha\beta}^{AB}$  centrées en  $\mathbf{r}_{\alpha}^A$  et  $\mathbf{r}_{\beta}^B$ , l’expression est la suivante :

$$\alpha_{\alpha\beta} = \sum_A \sum_B (\mathbf{r}_{\alpha}^A \alpha_{\mathbf{q}\mathbf{q}}^{AB} \mathbf{r}_{\beta}^B + \mathbf{r}_{\alpha}^A \alpha_{\mathbf{q}\beta}^{AB} + \alpha_{\alpha\mathbf{q}}^{AB} \mathbf{r}_{\beta}^B + \alpha_{\alpha\beta}^{AB}) \quad (2.13)$$

La conservation de la charge moléculaire totale est réalisée par le fait que les polarisabilités de flux de charges et charge–dipôle satisfont à la contrainte  $\sum_A \alpha_{\mathbf{q}\mathbf{t}}^{AB} = 0$  où  $t = 00, 10, 11c, 11s$ . Les notations prises sont en coordonnées sphériques avec la polarisabilité de flux de charge ( $\alpha_{00}^{AB}$ ) qui est l’ordre zéro du développement et les polarisabilités dipolaires ont une correspondance avec celles en coordonnées cartésiennes tel que :

$$\begin{aligned} \alpha_{10}^{AB} &= \alpha_{zz}^{AB} \\ \alpha_{11c}^{AB} &= \alpha_{xx}^{AB} \\ \alpha_{11s}^{AB} &= \alpha_{yy}^{AB} \end{aligned} \quad (2.14)$$

Pour les ordres supérieurs de la polarisabilité, cette correspondance n’est pas aussi directe. Il est possible d’aller voir l’annexe E du livre de Stone [9].

Les polarisabilités distribuées ainsi définies peuvent servir pour calculer l’énergie d’induction du système en mécanique moléculaire. Il est nécessaire de calculer tous les moments induits du système en interaction. Cela peut-être réalisé de manière itérative jusqu’à convergence de ces moments induits ou par inversion de matrice [196, 197].

Le calcul des moments induits d’une molécule A nécessite l’évaluation du champ extérieur créé par les molécules environnantes :

$$\Delta Q_{t'}^{a'} = -\alpha_{t't}^{a'a} V_{t'}^{a'} \quad (2.15)$$

où  $\Delta Q_{t'}^{a'}$  représente la composante induite d'ordre  $t'$  sur le site  $a'$ ,  $\alpha$  est la matrice de polarisabilités et  $V_{t'}^{a'}$  le champ total créé par les molécules voisines sur le multipôle d'ordre  $t'$  du site  $a'$ , soit :

$$V_{t'}^{a'} = \sum_{B \neq A} T_{t'u}^{a'b} Q_u^b \quad (2.16)$$

où  $Q_u^b$  est un vecteur contenant les moments des différents atomes des molécules B et  $T_{t'u}^{a'b}$  est le tenseur d'interaction.

Le caractère auto-cohérent de ce calcul provient du fait que, si les molécules B sont également polarisables, l'évaluation de leurs moments induits requiert la connaissance de ceux des molécules de type A, et donc, d'une manière tout à fait générale,

$$\Delta Q_{t'}^a = - \sum_{B \neq A} \alpha_{t't}^{aa'} T_{t'u}^{a'b} (Q_u^b + \Delta Q_u^b). \quad (2.17)$$

Une fois les moments induits déterminés, l'énergie d'induction se calcule comme une somme d'interactions de paires entre les moments induits d'une molécule A et les moments permanents d'une molécule B :

$$U_{indA} = \frac{1}{2} \sum_{B \neq A} \Delta Q_t^a T_{tu}^{ab} Q_u^b. \quad (2.18)$$

### La catastrophe de polarisation

Dans l'approche présentée ici, l'interaction se fait au travers du tenseur d'interaction ( $\mathbf{T}_{ij}$ ). En couplant les interactions entre les dipôles induits, à courte distance un phénomène de « catastrophe de polarisation » peut se produire. Cet effet fut décrit en premier lieu par Applequist *et al.* [198] et Thole [199]. Il est facilement démontrable, mathématiquement, qu'un tel comportement peut se produire dans le cas d'un système diatomique (A et B) ayant une polarisabilité  $\alpha_A$  et  $\alpha_B$ . La polarisabilité moléculaire, reliée au moment dipolaire moléculaire ( $\boldsymbol{\mu} = \boldsymbol{\mu}_A + \boldsymbol{\mu}_B$ ) par le champ électrique, possède deux composantes, l'une parallèle et l'autre perpendiculaire à l'axe de liaison entre A et B,

$$\begin{aligned}\alpha_{\parallel} &= \frac{\alpha_A + \alpha_B + \frac{4\alpha_A\alpha_B}{r^3}}{1 - \frac{4\alpha_A\alpha_B}{r^6}} \\ \alpha_{\perp} &= \frac{\alpha_A + \alpha_B - \frac{2\alpha_A\alpha_B}{r^3}}{1 - \frac{\alpha_A\alpha_B}{r^6}}\end{aligned}\quad (2.19)$$

La composante parallèle devient infinie lorsque la distance entre les deux atomes tend vers  $(4\alpha_A\alpha_B)^{1/6}$  (voir page 128 du livre de Stone [9]). Afin d'éviter cette singularité, dans le cas d'un système diatomique, des polarisabilités suffisamment petites sont considérées. Ceci permet au facteur  $4\alpha_A\alpha_B/r^6$  d'être inférieur à 1 pour des distances typiques interatomiques ( $> 1 \text{ \AA}$ ).

Lorsque le système est plus complexe, il existe d'autres approches pour éviter ce comportement à courte distance, en particulier les approches de Thole [199], de Jensen [200] et de Tang et Toennies [201–203]. Ces trois méthodes seront détaillées ici mais il existe également d'autres schémas numériques, comme ceux de Masia *et al.* [204] ou de Elking *et al.* [205], par exemple.

### Atténuation de Thole

Dans l'approche proposée par Thole [199], les interactions entre les moments induits à courte distance sont atténués. Cette atténuation est nécessaire car, à courte distance, les charges (multipôles) ponctuelles ne peuvent plus être considérées comme exactes (troncation du développement multipolaire). Dans la procédure de Thole, un terme d'atténuation est introduit en modifiant la distance par  $s_{ij} = 1.662(\alpha_i\alpha_j)^{1/6}$ . Ceci permet, dans le cas de deux atomes de polarisabilité  $1 \text{ \AA}^3$ , d'avoir un rayon de densité de charge de  $1.662 \text{ \AA}$ . Le tenseur d'interaction dipolaire, en coordonnées cartésiennes, revient à :

$$\mathbf{T}_{ij} = (4v^3 - 3v^4)\frac{1}{r^3}\mathbf{I} - v^4\frac{3}{r^5}\begin{pmatrix} x^2 & xy & xz \\ yz & y^2 & yz \\ zx & zy & z^2 \end{pmatrix}\quad (2.20)$$

où  $v = r/s_{ij}$ .  $\mathbf{T}_{ij}$  est inchangé dans le cas où  $r$  est plus grand que  $s_{ij}$ . Le paramètre de  $1.662 \text{ \AA}$  a été déterminé par Thole afin de reproduire au mieux les polarisabilités moléculaires pour un jeu de 16 molécules (voir table 1 dans la référence [199]). De nombreuses publications [85, 206] montrent que cette approche donne de bons résultats

en dynamique moléculaire.

### Atténuation de Jensen

La détermination de l'atténuation nécessaire pour éviter la catastrophe de polarisation à courte distance dans l'approche de Jensen *et al.* [200] se base sur les dépendances en fréquence des polarisabilités. Jensen *et al.* [221] ont calculé cette grandeur pour une série de 184 molécules et ont ainsi proposé un moyen d'atténuer à courte distance le tenseur d'interaction  $T$  en modifiant la distance entre les sites :

$$r^* = \sqrt{r^2 + \frac{\pi}{4a}} \quad (2.21)$$

où  $r^*$  est la nouvelle distance entre le site polarisable et le site polarisant. La grandeur  $a$  est déterminée à partir de calculs quantiques, d'après l'écart quadratique moyen entre la polarisabilité calculée et la polarisabilité modèle. Divers travaux utilisent cette approche pour atténuer l'énergie d'induction à courte distance pour des systèmes variés tels que les cristaux, les solutions, les systèmes  $\pi$  ou des colorants [162, 163, 222–227].

### Atténuation de Tang et Toennies

L'atténuation proposée par Tang et Toennies [201–203] était initialement formulée pour les paramètres de van der Waals. La formulation générale s'écrit :

$$f_n(r) = 1 - \exp(-ar) \sum_{k=0}^n \frac{(ar)^k}{k!} \quad (2.22)$$

où  $f_n(r)$  est la fonction permettant de modifier l'interaction, à une distance  $r$ , entre les sites polarisables. Le paramètre  $a$  est déterminé de manière quantique ou en l'adaptant au cas considéré. L'ordre  $n$  du développement dépend du terme de dispersion utilisé dans le modèle de van der Waals ( $C_6$ ,  $C_8$ ,  $C_{10}$ ...). Millot et Stone [228] ont reformulé cette écriture pour l'adapter à l'interaction de moments polarisables modifiant ainsi l'énergie d'induction. L'équipe de Stone [89–92, 161] a, en particulier, utilisé cette reformulation afin de calculer l'énergie d'induction dans un modèle polarisable. Il est particulièrement intéressant de noter que le paramètre  $a$  peut être obtenu à partir d'un simple calcul quantique [91, 92]. D'autres auteurs se sont servis de l'approche de Tang et Toennies pour contrebalancer l'énergie d'induction [164, 229–242].

### 2.2.2 Charges Fluctuantes

Suivant les groupes, cette approche est nommée différemment, mais repose sur le même principe. Dans la littérature, les noms de charges fluctuantes [243, 244]; d'« équilibration du potentiel chimique (ou de l'électronégativité) » [245–253]; ou d'« équilibration de charge » [147, 254–259] sont utilisés.

L'approche des charges fluctuantes modifie le potentiel du champ de force non-polarisable 1.4. Cette modification intervient sur le potentiel de Coulomb. Ainsi, l'interaction de Coulomb varie en fonction de l'environnement. Pour définir ce potentiel, prenons un développement de Taylor à l'ordre deux de l'énergie requise pour modifier la charge  $q$  localisée sur un atome :

$$U(q) = E^0 + \chi^0 q + \frac{1}{2} J q^2 \quad (2.23)$$

où  $\chi^0$  et  $J$  peuvent être identifiés [243] en définissant le potentiel d'ionisation [ $IP = U(1) - U(0)$ ] et l'affinité électronique [ $EA = U(-1) - U(0)$ ]. Cette identification se fait en considérant la variation du potentiel entre deux états électroniques et, en particulier, dans le cas où le développement est valide pour une variation de  $\pm 1e$ . La réécriture des coefficients se traduit donc comme :

$$\chi^0 = \frac{IP + EA}{2} \quad (2.24)$$

$$J = IP - EA. \quad (2.25)$$

L'équation 2.24 est la définition de l'électronégativité donnée par Mulliken [260]. Le terme linéaire du développement de Taylor,  $\chi^0$ , peut être considéré comme étant égal à l'électronégativité de l'atome. Quant au second terme du développement,  $\frac{1}{2}J$ , il décrit la dureté de l'atome [261]. En faisant l'hypothèse que cette formule peut être utilisée pour un ensemble de molécules, le potentiel permettant de décrire l'énergie des charges du système ( $\mathbf{q}$ ), dans le modèle d'égalisation des électronégativités vaut :

$$U(\mathbf{q}) = \sum_i \left( E_i^0 \chi_i^0 q_i + \frac{1}{2} J_{ii} q_i^2 \right) + \sum_i \sum_{j>i} J_{ij}(\mathbf{r}_{ij}) q_i q_j. \quad (2.26)$$

Les effets de second ordre, décrits avec le terme  $J_{ij}$ , dépendent de la distance et peuvent être assimilés à des effets de pénétration à courte distance entre atomes. Ceux-ci ont un comportement en  $1/r_{ij}$  à longue distance et un comportement modifié à plus courte

distance, dû à l'interaction déstabilisante électron–électron. Cet effet d'écran à courte distance provient du fait que les charges atomiques devraient être modélisées grâce à une densité de charge  $\rho(\mathbf{r})$ , plutôt qu'au travers de charges ponctuelles. Ainsi le coefficient d'ordre 2 se réécrit :

$$J_{ij}(\mathbf{r}) = \int \frac{\rho_i(\mathbf{r}_i)\rho_j(\mathbf{r}_j)}{|\mathbf{r}_i - \mathbf{r}_j - \mathbf{r}|} d\mathbf{r}_i d\mathbf{r}_j \quad (2.27)$$

Pour décrire la pénétration électronique, certains champs de force [223, 262–265] utilisent une distribution de charge non-sphérique à la place des charges ponctuelles. L'utilisation de cette description plus réaliste a toutefois un coût sur le temps de calcul.

Le potentiel,  $U(\mathbf{q})$ , est alors minimisé de manière à obtenir les charges portées par chaque atome. De plus, afin d'assurer la neutralité du système, sa charge totale ( $q_{\text{tot}}$ ) est contrainte :

$$\sum_i q_i = q_{\text{tot}} \quad (2.28)$$

La contrainte de minimisation peut être obtenue par la méthode de multiplicateur de Lagrange ( $\lambda$ ) :

$$U(\mathbf{q}) = U(\mathbf{q}) - \lambda \left( \sum_i q_i - q_{\text{tot}} \right) \quad (2.29)$$

Pour tous les sites atomiques  $i$  définis, l'énergie,  $U$ , obtenue peut être minimisée en fonction de toutes les charges ( $q_i$ ). De plus, en supposant que la molécule est neutre ( $q_{\text{tot}} = 0$ ) :

$$\left( \frac{\partial U}{\partial q_i} \right) - \lambda = 0 \quad \forall i \quad (2.30)$$

Étant donné que les  $i$  multiplicateurs de Lagrange doivent être identiques sur l'ensemble du système moléculaire, il s'ensuit que :

$$\left( \frac{\partial U}{\partial q_i} \right) = \left( \frac{\partial U}{\partial q_j} \right) \quad \forall i, j \quad (2.31)$$

En identifiant  $\partial U / \partial q$  à une électronégativité, la minimisation de l'énergie par rapport aux charges revient à égaliser cette énergie pour chaque atome :

$$\chi_i \equiv \left( \frac{\partial U}{\partial q_i} \right) = \chi_i^0 + J_{ii}q_i + \sum_{j \neq i} J_{ij}(\mathbf{r}_{ij})q_j \quad (2.32)$$



À présent, l'électronégativité ( $\chi_i$ ) d'un atome  $i$  dans une molécule est définie comme une grandeur différente de l'électronégativité de l'atome isolé ( $\chi_i^0$ ). Cette électronégativité est dépendante de sa charge et des charges environnantes, de sa dureté et des interactions avec les autres atomes au travers de  $J_{ij}(\mathbf{r}_{ij})$ . Pour résumer,  $\chi_i$  décrit le déplacement des charges autour d'une molécule afin d'en minimiser l'énergie. Parr *et al.* [261] ont aussi remarqué que  $\mu = -\partial U/\partial q$ , ce qui revient à égaliser l'électronégativité ou le potentiel chimique.

En généralisant l'équation 2.26 à une collection de molécules, l'énergie d'interaction peut être écrite comme :

$$U(q, \mathbf{r}) = \sum_{\alpha} \left( \sum_{i \in \alpha} \chi_i^0 q_i + \frac{1}{2} \sum_{i \in \alpha} \sum_{j \in \alpha} q_i q_j J_{ij}(\mathbf{r}_{ij}) - E_{\alpha}^{\text{gp}} \right) + \sum_{\alpha} \sum_{\beta > \alpha} \sum_{i \in \alpha} \sum_{j \in \beta} q_i q_j J_{ij}(\mathbf{r}_{ij}) \quad (2.33)$$

où la sommation sur  $\alpha$  et  $\beta$  sont les molécules,  $i$  et  $j$  sont les atomes. Le terme  $E_{\alpha}^{\text{gp}}$  est l'énergie d'une molécule  $\alpha$  en phase gazeuse.

L'équation 2.33 remplace alors le terme de Coulomb ( $q_i q_j / r_{ij}$ ) dans les champs de forces additifs (voir équation 1.4). Les charges,  $q_i$ , sont ainsi traitées comme des variables indépendantes et la fluctuation des charges en réponse à l'environnement reproduit la polarisation. Dans l'approche des charges fluctuantes, la charge totale de la molécule est conservée. De ce fait, la charge portée par une molécule ne peut être transférée à une autre. Pourtant, les effets de transfert de charge dans certains systèmes sont communément décrits [12, 86, 172–176, 220, 249, 266–271]. Il apparaît donc important de ne pas contraindre la charge totale à la molécule. Cependant, en enlevant cette contrainte, les charges fluctuantes mènent à une surestimation de l'énergie de polarisation au fur et à mesure de l'augmentation de la taille de la molécule [272]. Cette sur-polarisation s'explique par le faible coût énergétique nécessaire pour faire fluctuer les charges le long des liaisons. Afin d'éviter la sur-polarisation, une méthode de transfert de charge interatomique (*atom-atom charge transfer*, AACT) [273] a été proposée. Dans cette méthode, le développement de Taylor ne se fait plus sur les charges atomiques, mais sur le terme de transfert de charge entre paires.

### 2.2.3 Oscillateurs de Drude

Les oscillateurs de Drude [274] ont été décrits pour la première fois afin de modéliser simplement les propriétés dispersives des matériaux (pages 382–400). Cette méthode consiste à modéliser la polarisation par l’intermédiaire des dipôles ; grâce à des paires de charges ponctuelles reliées par des ressorts harmoniques (voir figure 2.1). Par exemple, dans le cas de la molécule d’eau, la charge fictive est reliée à l’atome d’oxygène par un ressort. La charge totale de la molécule est assurée par le fait que la charge de l’atome d’oxygène soit réduite par la valeur de la charge portée par la particule fictive. Le reste de la molécule est inchangé par rapport à un modèle de champ de force additif (TIP3P, SPC...). Ainsi, la polarisation est reproduite par le déplacement relatif de la charge fictive par rapport à l’environnement ou par rapport à l’atome de référence. Dans la littérature, cette méthode est aussi appelée « *shell model* » ou « *charge-on-spring* » [139, 145, 275, 276].

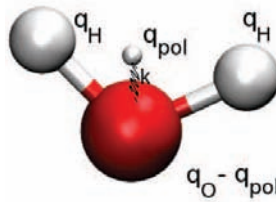


FIGURE 2.1 – Molécule d’eau dans le modèle de Drude. La charge fictive,  $q_{pol}$ , est reliée par un ressort de force  $k$  à l’atome d’oxygène. La charge partielle de ce dernier est réduite de  $q_{pol}$ .

Dans ce modèle, la constante de force du ressort et les charges ponctuelles sont fixes tout au long de la simulation. Les paramètres portés par les atomes et les contributions intramoléculaires sont déterminés afin de reproduire au mieux certaines valeurs expérimentales (moment dipolaire, enthalpie de vaporisation, densité...) [139, 275]. Pour plus de détails, Yu *et al.* [275] proposent dans la partie C un cheminement pour dériver les paramètres pour un modèle de type « *charge-on-spring* ».

Pour illustrer le modèle des oscillateurs de Drude, considérons un atome neutre (ex. : un gaz rare). Si la charge sur l’atome fictif est  $-q_i$ , alors la charge portée par l’atome neutre est  $+q_i$ . La distance,  $\mathbf{d}_i$ , entre ces deux charges ponctuelles génère un dipôle de magnitude :

$$\boldsymbol{\mu}_i = -q_i \mathbf{d}_i \quad (2.34)$$

Pour décrire l'énergie d'interaction, décomposons l'énergie des dipôles induits en trois termes : un terme d'interaction avec le champ environnant statique (stat), un terme d'interaction entre dipôles induits ( $\mu\mu$ ) et un terme de polarisation (pol). De cette décomposition, l'énergie d'induction vaut :

$$U_{\text{ind}} = U_{\text{stat}} + U_{\mu\mu} + U_{\text{pol}} \quad (2.35)$$

L'énergie de polarisation correspond, dans ce cas, à celle des oscillateurs harmoniques du système (dans le cas d'un système mono-atomique, seul le ressort entre la charge fictive et l'atome est à considérer) :

$$U_{\text{pol}} = \frac{1}{2} \sum_{i=1}^N k_i d_i^2 \quad (2.36)$$

où  $U_{\text{pol}}$  est définie pour une collection de  $N$  atomes polarisables ayant un ressort de raideur  $k_i$  et un déplacement  $d_i = |\mathbf{d}_i|$ . La polarisabilité atomique,  $\alpha_i$ , est définie en fonction de la constante de force  $k$  du ressort, grâce à la relation  $\alpha_i = q_i^2/k_i$  [137]. Par rapport au cas des polarisabilités distribuées (section 2.2.1), la polarisabilité n'est plus un tenseur, mais un scalaire.

Pour une collection de  $N$  atomes, l'énergie d'interaction entre les dipôles induits,  $U_{\mu\mu}$ , est la somme entre les interactions charge–charge des quatre sites :

$$U_{\mu\mu} = \frac{1}{2} \sum_{i=1}^N \sum_{j \neq i}^N q_i q_j \left[ \frac{1}{|\mathbf{r}_{ij}|} - \frac{1}{|\mathbf{r}_{ij} - \mathbf{d}_j|} - \frac{1}{|\mathbf{r}_{ij} + \mathbf{d}_i|} + \frac{1}{|\mathbf{r}_{ij} - \mathbf{d}_j + \mathbf{d}_i|} \right] \quad (2.37)$$

Dans ce modèle, il est intéressant de noter que seuls les tenseurs d'interaction charge–charge sont utilisés, donc les approches PME (*particle Mesh-Ewald*) [277–280] développées spécifiquement pour ces interactions sont toujours valides. Ces gains, non négligeables par rapport aux modèles utilisant des dipôles, (calcul des interactions dipôles–dipôles et charges–dipôles) sont annulés par l'ajout d'une particule fictive sur chaque site atomique et donc augmente le nombre d'interactions électrostatiques à calculer [143].

La dernière énergie,  $U_{\text{stat}}$ , représente les interactions des dipôles induits avec le champ permanent. Chaque site possédant une charge contribue à cette énergie, soit :

$$U_{\text{stat}} = - \sum_{i=1}^N q_i [\mathbf{r}_i \cdot \mathbf{E}_i^0 - (\mathbf{r}_i + \mathbf{d}_i) \cdot \mathbf{E}_i^{0'}] \quad (2.38)$$

Dans cette expression,  $\mathbf{E}_i^0$  (respectivement  $\mathbf{E}_i^{0'}$ ) est le champ permanent autour de l'atome  $i$  (respectivement la particule de Drude liée à l'atome  $i$ ) à la distance  $\mathbf{r}_i$  (respectivement  $\mathbf{r}_i + \mathbf{d}_i$ ).

Un des développements majeurs pour les oscillateurs de Drude est l'implantation dans le champ de force CHARMM du modèle d'eau polarisable [244, 281]. Dans le modèle initial de la molécule d'eau, appelé SWM4-DP [145] (*simple water model with four sites and Drude polarizability*), la particule de Drude, reliée à l'atome d'oxygène, est chargée positivement. Le fait de considérer la particule comme chargée positivement n'est pas représentatif de l'interprétation physique de cette particule qui est sensée représenter les degrés de liberté du système électronique. Par conséquent, un nouveau modèle nommé SWM4-NDP [282] (*negative Drude particle*) et ayant un modèle électronique inversé par rapport au précédent, a été développé. Dans celui-ci, contrairement aux modèles « shell » standards, le potentiel 6-12 de Lennard-Jones est associé à l'atome d'oxygène et non plus à la particule de Drude. Ce modèle possède l'avantage de conduire à de très bonnes propriétés (enthalpie de vaporisation, constante diélectrique, constante d'auto-diffusion) pour des simulations d'eau liquide à température et pression ambiantes. Cependant, certaines propriétés restent mal reproduites, comme la fonction de distribution radiale (premier et second pics indiquant un manque de structure) et une polarisabilité moléculaire trop faible ( $0.978 \text{ \AA}^3$ ) par rapport à la valeur expérimentale ( $1.44 \text{ \AA}^3$ ). Ceci peut être expliqué par le fait que les paramètres importants (charges et constantes de raideur du ressort) de ce modèle sont déterminés de manière à reproduire au mieux les grandeurs expérimentales, sans tenir compte du résultat quantique (voir protocole dans les articles [283, 284]). Malgré cela, depuis plusieurs années, les oscillateurs de Drude permettent d'obtenir de très bons résultats pour un faible coût calculatoire, pour des simulations d'ions, d'alcanes, de protéines... [72, 73, 139, 143, 214, 218, 285–290].



# Développement d'un potentiel intermoléculaire polarisable simple et précis

---

## Sommaire

---

<b>3.1</b>	<b>Référence quantique utilisée . . . . .</b>	<b>39</b>
3.1.1	Question sur le $\delta HF$ . . . . .	39
3.1.2	Analogie entre SAPT et le potentiel intermoléculaire . . . . .	40
3.1.3	SAPT et les molécules chargées négativement . . . . .	41
<b>3.2</b>	<b>Électrostatique . . . . .</b>	<b>41</b>
3.2.1	Effets de pénétration . . . . .	44
3.2.2	Construction du modèle électrostatique . . . . .	46
3.2.3	Interactions électrostatiques . . . . .	50
<b>3.3</b>	<b>Induction . . . . .</b>	<b>60</b>
3.3.1	Construction du modèle de polarisabilité . . . . .	60
3.3.2	Énergie d'interaction d'induction . . . . .	64
3.3.3	Fonctions d'atténuation . . . . .	64
3.3.4	Énergie d'interaction d'induction entre molécules . . . . .	66
<b>3.4</b>	<b>van der Waals . . . . .</b>	<b>75</b>
3.4.1	Différentes expressions du potentiel de van der Waals . . . . .	75
3.4.2	Détermination des paramètres de van der Waals . . . . .	76
<b>3.5</b>	<b>Énergie totale . . . . .</b>	<b>78</b>
<b>3.6</b>	<b>Vers les simulations de dynamique moléculaire . . . . .</b>	<b>89</b>
<b>3.7</b>	<b>Traitement des moments induits en dynamique moléculaire . . . . .</b>	<b>89</b>
3.7.1	Le Lagrangien étendu . . . . .	89
3.7.2	Le traitement auto-cohérent . . . . .	90
<b>3.8</b>	<b>Code de dynamique moléculaire : Tinker . . . . .</b>	<b>91</b>
<b>3.9</b>	<b>Cas d'une simulation de l'eau . . . . .</b>	<b>92</b>

Il existe plusieurs méthodes [9] pour décomposer l'énergie d'interaction de manière physique. Parmi celles-ci, des potentiels empiriques permettent de décrire de façon simplifiée l'énergie d'interaction d'assemblages moléculaires. Pour obtenir les paramètres des potentiels empiriques, l'utilisation de la mécanique quantique seule s'est avérée, jusqu'à présent, insuffisante pour des systèmes en phase condensée. Dans le cas particulier présenté ici, les contributions énergétiques physiques développées dans la théorie SAPT sont simplifiées en deux ou trois fonctions *ad hoc*. Du fait de cette simplification une partie de l'information exacte est perdue. Par contre, il est montré [164] qu'un potentiel empirique est à même de reproduire une somme d'énergies obtenues en mécanique quantique. Par exemple, les contributions d'échange et de dispersion sont assez bien reproduites par certains potentiels de van der Waals (voir section 3.4). Dans une certaine mesure, cette sommation peut-être considérée au détriment de la précision sur l'énergie d'interaction. Par exemple, en dynamique moléculaire, seules deux fonctions (Coulombienne et de van der Waals) sont utilisées pour décrire le potentiel intermoléculaire. Dans le cas où une plus grande précision serait demandée, l'ajout de potentiels à  $N$ -corps (potentiel d'induction, de transfert de charge...) ou l'utilisation de fonctions différentes permettent de décrire de façon plus fidèle les énergies quantiques. En contre partie, cette précision a un coût qui doit être pris en compte pour de grands assemblages moléculaires. Par exemple, le calcul d'un potentiel de Buckingham en exponentiel pour la partie répulsive sera toujours plus long à calculer qu'une simple puissance 12 pour un Lennard-Jones 6-12. Il est donc nécessaire de faire des compromis entre les potentiels considérés et la précision requise.

La base du modèle présenté ici vient de l'expérience que possède l'équipe de Nancy pour développer, pour des molécules isolées, des modèles de multipôles décrivant le potentiel électrostatique environnant mais aussi des polarisabilités distribuées qui reproduisent l'énergie d'induction. Pour générer ces modèles, le logiciel OPEP<sup>1</sup> est disponible.

Nous présenterons dans la suite de ce chapitre le niveau de théorie ainsi que la référence quantique qui vont être utilisés pour développer le modèle d'interaction. Pour cela, des dimères de diverses espèces (eau, ions, benzène...) seront considérés. Puis, nous développerons le modèle qui contient un potentiel électrostatique et un potentiel d'induction pour

1. <http://www.edam.uhp-nancy.fr/Opep>

lesquels les paramètres seront obtenus au travers de calcul sur des molécules isolées. Pour ces deux potentiels nous verrons des interactions spécifiques de dimères formées à partir de l'eau, d'ions ou de benzène. Le troisième potentiel considéré est un potentiel de van der Waals pour lequel nous choisirons une fonction *ad hoc* pour être comparée à la référence quantique.

### 3.1 Référence quantique utilisée

Le moyen choisi ici pour valider le modèle d'interaction de molécules est la théorie SAPT. Les énergies obtenues grâce à cette approche seront sommées de façon à permettre la comparaison avec le modèle de mécanique moléculaire dérivé.

Le niveau de théorie employé pour obtenir la géométrie de la molécule isolée est MP2(Full)/6-311++G(2d,2p) [291, 292], qui permet d'obtenir des géométries au minimum fiables pour un coût de calcul raisonnable [293]. Dans cette écriture, MP2(Full) signifie que le calcul sera effectué à l'ordre 2 des perturbations selon Møller-Plesset [155], 6-311++G(2d,2p) est la base de fonctions d'ondes pour lesquelles les orbitales 2d et 2p seront prises en compte pour les atomes lourds et ++ indique que des fonctions diffuses seront ajoutées. L'obtention des propriétés moléculaires et intermoléculaires a lieu au niveau MP2(Full)/Sadlej ou, de façon équivalente, en SAPT2/Sadlej. La base de Sadlej [294–297] possède une corrélation de haut niveau comparable à une fonction triple dzêta mais ne nécessite qu'un temps de calcul équivalent à celui d'une fonction double dzêta. De plus, il a été observé que cette base fournissait des propriétés électroniques (électrostatique et induction) en très bon accord avec l'expérience [298].

Pour les molécules en interaction, la suite de programmes SAPT [14] est utilisée. Le détail des fichiers d'entrée et de sortie de ce code est fourni dans l'annexe A.

#### 3.1.1 Question sur le $\delta HF$

Dans la décomposition SAPT, une grandeur énergétique nommée  $\delta HF$  est incluse. Cette grandeur correspond au reste des termes d'ordres supérieurs non développés dans le calcul SAPT, bien que sa nature exacte ne soit pas clairement identifiée. Elle embrasse des contributions d'origines induction, échange et répulsion. Dans la théorie SAPT [156],



elle est évaluée comme :

$$\delta E_{int,r}^{HF} = E_{int}^{HF} - E_{pol}^{(10)} - E_{exch}^{(10)} - E_{ind,resp}^{(20)} \quad (3.1)$$

où  $E_{int}^{HF}$  est l'énergie d'interaction supermoléculaire Hartree-Fock ;  $E_{pol}^{(10)}$  est l'interaction électrostatique non-corrélée des moments multipolaires des monomères ;  $E_{exch}^{(10)}$  est l'échange électronique non-corrélé et  $E_{ind,resp}^{(20)}$  est la réponse corrélée de l'induction. En observant cette contribution, il est difficile de dire quelle est sa nature. Dans la documentation disponible avec la suite de programmes SAPT, il est écrit que cette grandeur représente les ordres supérieurs d'induction. Ceci est aussi repris dans d'autres schémas de décomposition de l'énergie [91, 169]. Il a longtemps été question de la place exacte de cette contribution dans le modèle dérivé à Nancy. Il s'agissait notamment de savoir si le modèle choisi pouvait prendre en compte cette contribution, ou si celle-ci était de nature à être prise en compte dans le potentiel de van der Waals. Après plusieurs essais sur son intégration dans diverses interactions, il s'est avéré que le modèle d'induction permettait de prendre en compte ces effets implicitement [164].

### 3.1.2 Analogie entre SAPT et le potentiel intermoléculaire

Après avoir levé le doute sur le terme  $\delta HF$ , les contributions équivalentes entre le potentiel de mécanique moléculaire et les contributions SAPT sont (pour les notations voir annexe A) :

**Électrostatique** ( $\Delta\mathcal{U}_{ele}^{MM}$ ) : contribution électrostatique à l'énergie d'interaction. Ce potentiel peut être modifié par un facteur traduisant les effets de pénétration ( $a$ ). Le potentiel de Coulomb considéré est basé sur un jeu de charges ponctuelles. Le potentiel électrostatique est comparé à  $\Delta\mathcal{U}_{ele}^{SAPT}$ .

**Induction** ( $\Delta\mathcal{U}_{ind}^{MM}$ ) : contribution d'induction à l'énergie d'interaction. Ce potentiel peut être modifié par une fonction d'atténuation ( $f$ ). Le potentiel d'induction considéré s'écrit sous la forme d'un développement multipolaire de la polarisabilité. Le potentiel d'induction est comparé à  $\Delta\mathcal{U}_{ind}^{SAPT} + \Delta\mathcal{U}_{exc-ind}^{SAPT} + \delta HF$ .

**van der Waals** ( $\Delta\mathcal{U}_{vdW}^{MM}$ ) : contribution de van der Waals à l'énergie d'interaction. Ce potentiel sera comparé à  $\Delta\mathcal{U}_{disp}^{SAPT} + \Delta\mathcal{U}_{exc-disp}^{SAPT}$ .

### 3.1.3 SAPT et les molécules chargées négativement

Dans le cadre des calculs SAPT, la convergence pour des systèmes chargés négativement n'est pas garantie [299]. Ceci est expliqué, brièvement, par le fait que les fonctions chimiques chargées négativement sont très diffuses et donc la convergence des opérateurs SAPT est très lente. Il est donc possible de converger vers des minima locaux au lieu d'obtenir le minimum global lors de la résolution de ces opérateurs. Dans la suite de la thèse, nous aborderons l'étude de systèmes anioniques en interaction. Il est ainsi légitime de se poser la question de la convergence ou non du calcul.

Pour vérifier si la convergence est atteinte, une première méthode simple consiste à refaire le calcul avec une autre base et de vérifier que l'on obtient un résultat cohérent avec le résultat précédent. Ainsi, lors de chaque étude utilisant une molécule chargée négativement, un calcul est effectué avec une base différente [6-311++G(2d,2p)].

Cette condition permet de s'assurer que la convergence est atteinte, mais peut ne pas être suffisante puisque les deux calculs SAPT peuvent ne pas avoir convergé. Afin de la confirmer, une comparaison avec un calcul RVS (Restricted Valence Shell) [300] est effectuée. Ce calcul est limité au niveau de théorie Hartree-Fock, mais permet de comparer le résultat obtenu par la théorie SAPT (au niveau MP2) avec un calcul non corrélé qui converge dans tous les cas. De plus, si possible, des résultats à d'autres niveaux de théorie issus de la littérature seront utilisés pour étayer la convergence du calcul SAPT. En particulier, l'article de Kim *et al.* [165] est intéressant car les auteurs utilisent des anions en interaction avec des systèmes  $\pi$ .

À présent que nous connaissons les contributions utilisées pour la référence, regardons la dérivation du potentiel intermoléculaire.

## 3.2 Électrostatique

Dans l'approche présentée ici, nous imposons une simplicité et une compacité au modèle. Ici, le potentiel électrostatique se limite à l'utilisation de charges ponctuelles. Malgré le niveau d'approximation, ce choix s'avère efficace dans de nombreux cas pour reproduire le potentiel électrostatique [30, 32, 178, 179]. Il existe diverses approches

permettant d'améliorer la représentation au travers de multipôles d'ordre plus élevé telles que AMOEBA [48, 84], SIBFA [121] ou le modèle de la molécule d'eau de Millot et Stone [228]. Une troncation à un ordre plus élevé que de simples charges peut, par exemple, être motivée par l'importance du moment quadrupolaire, comme pour les atomes de soufre ou d'oxygène [228]. Il est également possible de prendre en compte des effets d'ordre supérieur dans le potentiel d'interaction électrostatique. En particulier, à courte distance, les effets de pénétration électroniques sont présents. Il existe des fonctions pour reproduire ces effets, comme cela a été décrit par Stone [9] ou Freitag *et al.* [301]. Pour cela, l'introduction de gaussiennes en lieu et place de charges ponctuelles est envisageable [223, 262–265].

L'ajout de multipôles, de potentiels à courte distance et la modification de l'interaction électrostatique peuvent être envisagés pour obtenir un potentiel électrostatique plus précis. En contrepartie, dans un souci de simplicité du modèle, le choix de charges ponctuelles sur des sites atomiques est défendable. Il existe diverses méthodes [28, 302, 303] pour dériver les charges et il a souvent été montré que l'emploi de simples charges suffisait à décrire le potentiel électrostatique de la molécule isolée. Dans le cas d'interactions de molécules, plusieurs articles [30, 163, 304, 305] indiquent qu'un potentiel utilisant uniquement des charges suffit pour reproduire le potentiel intermoléculaire. Ceci est aussi l'approche utilisée par le champ de force AMBER [18] dans laquelle le potentiel électrostatique est « l'empreinte génétique » de la molécule.

La dérivation de charges atomiques à partir de la mécanique quantique peut se faire grâce à plusieurs méthodes. En particulier, citons les méthodes de Mulliken [29] et la méthode **resp** [18, 30, 31, 306]... Ces méthodes sont directement accessibles dans certains programmes de chimie quantique comme GAUSSIAN [307] ou GAMESS [308]. À Nancy, une méthode développée depuis 1993 [28, 32, 178, 179] permet de déterminer des multipôles sur une grille de points. Un des avantages est la non restriction à un modèle de charges ponctuelles, ce qui permet de comparer la reproduction du potentiel électrostatique donné par un simple modèle de charges ponctuelles avec un modèle plus complexe (jusqu'à l'hexadécapôle). Cette approche est disponible dans le logiciel OPEP [33] qui reprend l'approche des DMA (« *distributed multipole analysis* ») introduite par Stone [190] à partir du potentiel électrostatique. Dans la suite de cette thèse, nous utiliserons la méthode disponible dans OPEP pour obtenir les multipôles.

La procédure de détermination des charges (voir article de Chipot et Ángyán [309]) commence par la génération d'une grille de points sur laquelle le potentiel électrostatique est évalué en mécanique quantique. Lors de la génération de la grille, trois paramètres interviennent parmi lesquels deux sont figés et le troisième en est déduit : le nombre de points de grille ( $N_p$ ), la distance entre chaque points de grille ( $\Delta r$ , en Å) et un facteur multiplicatif par rapport au rayon de van der Waals des atomes de la molécule ( $\xi$ ). Pour le calcul du potentiel électrostatique, les valeurs optimales [309] sont  $\Delta r = 0.5$  Å et  $\xi = 3.0$ . Le fait de choisir ces deux valeurs permet d'obtenir une grille ni trop dense, ni trop large, afin de ne pas donner un poids trop important à la contribution dipolaire. De plus, par construction, on supprime les points de grille compris dans une enveloppe où les rayons de van der Waals sont doublés pour éliminer les effets de pénétration. Les grilles ainsi générées comportent généralement quelques milliers de points.

Une fois le potentiel quantique obtenu, la reproduction de celui-ci sur la grille sera effectuée grâce à une minimisation des moindres carrés entre la référence et un développement multipolaire du potentiel électrostatique suivant :

$$V_k^{QM} \sim \tilde{V}_k = \sum_a \sum_{l=0}^{N_t(a)} \sum_{\kappa} T_{l\kappa}^{ka} Q_{l\kappa}^a \quad (3.2)$$

Dans cette équation,  $V_k^{QM}$  représente le potentiel calculé en mécanique quantique sur un des  $k$  points de la grille.  $\tilde{V}_k$  est le potentiel approché écrit comme un développement multipolaire en coordonnées sphériques ( $Q_{l\kappa}^a$ , voir la Table E.2 du livre de Stone [9]).  $l$  et  $\kappa$  sont les ordres du développement multipolaire et  $a$  le site sur lequel la détermination des multipôles est effectuée.  $T_{l\kappa}^{ka}$  est un élément du tenseur d'interaction, qui, dans le cas de charges ponctuelles, se limite à  $1/r_{ak}$  où  $r_{ak}$  est la distance entre le point  $k$  de la grille et  $a$ . Dans ce cas là, l'équation 3.3 revient à :

$$V_k^{QM} \sim \tilde{V}_k = \sum_a \frac{1}{r_{ak}} q^a \quad (3.3)$$

Une fois le modèle généré, deux critères d'erreur sur la reproduction du potentiel sont considérés :

$$\left\{ \begin{array}{l} \text{RMSD} = \left[ \frac{1}{N_p} \sum_k (V_k - \tilde{V}_k)^2 \right]^{1/2} \\ \Delta\varepsilon = \frac{1}{N_p} \sum_k \left| \frac{V_k - \tilde{V}_k}{V_k} \right| \end{array} \right. \quad (3.4)$$

où RMSD est l'écart quadratique moyen (*root mean square deviation*) et,  $\Delta\varepsilon$ , l'erreur moyenne sur le potentiel évalué en  $N_p$  points de grille.

Une fois les charges ponctuelles obtenues, leur utilisation dans le potentiel intermoléculaire est immédiate. Ici, aucune modification des charges n'est apportée, contrairement à d'autres approches comme, par exemple, dans les oscillateurs de Drude [283]. Puisqu'aucun point de grille proche des atomes n'est pris en compte, les charges calculées reproduisent la partie multipolaire exacte du potentiel et n'incluent donc pas les effets de pénétration. De plus, le potentiel en  $1/r$  ne permet pas de reproduire le potentiel à courte distance. Pour cette raison, il existe diverses fonctions *ad hoc* pour modéliser ces effets.

### 3.2.1 Effets de pénétration

Pour prendre en compte les effets à courte distance, plusieurs fonctions ont été développées [9, 301, 310, 311]. Citons, en particulier, la formulation dédiée au cas de charges ponctuelles [311], celle d'un développement multipolaire de la distribution électronique [310] ou celle de Stone *et al.* [312]. Ces méthodes permettent de modéliser ces effets avec plus ou moins de simplicité. Par la suite, seules les fonctions de Piquemal *et al.* [310] et de Cisneros *et al.* [311] seront envisagées.

La formulation de Piquemal *et al.* [310] a été développée de façon à être utilisable pour tout ordre du développement multipolaire. Cette fonction ( $f_{\text{damp}}^{\text{piquemal}}$ ), appliquée au potentiel électrostatique d'interaction des atomes  $i$  et  $j$ , s'écrit sous la forme :

$$\begin{aligned}
 \tilde{U}_{ele}^{MM} &= f_{\text{damp}}^{\text{piquemal}} \cdot U_{ele}^{MM} \\
 &= \frac{1}{r} [Z_i Z_j - Z_i (Z_j - q_j) (1 - \exp(-\alpha_j r)) \\
 &\quad + Z_j (Z_i - q_i) (1 - \exp(-\alpha_i r)) \\
 &\quad + (Z_i - q_i) (Z_j - q_j) (1 - \exp(-\beta_i r)) (1 - \exp(-\beta_j r))]
 \end{aligned} \tag{3.5}$$

Dans cette écriture,  $Z$  représente le nombre d'électrons de valence de l'atome. Il faudra donc considérer que pour l'hydrogène, ce nombre est égal à 1. Dans le cas des cations, ce nombre correspond à la dernière couche électronique occupée (exemple :  $\text{Na}^+$ ,  $Z = 8$ ). Les autres paramètres,  $\alpha$  et  $\beta$ , sont définis comme suit :

$$\begin{cases} \alpha_i = \frac{4.42}{r_{vdw}(i)} \\ \beta_i = \frac{4.12}{r_{vdw}(i)} \end{cases} \tag{3.6}$$

Les constantes de 4.42 et 4.12 ont été déterminées afin de reproduire au mieux les interactions dans des complexes neutres et ioniques possédant une liaison hydrogène, des systèmes incluant des métaux ioniques et des complexes empilés (*stacked*) [310]. Ces paramètres sont considérés comme transférables [310]. Quand aux valeurs de  $r_{vdw}$ , elles sont définies pour un type d'atome donné. Elles n'ont donc pas besoin d'être définies pour chaque interaction. Un des avantages de cette formulation est qu'elle est complètement définie pour tous types d'interactions (en supposant que les constantes 4.42 et 4.12 sont transférables). Le potentiel d'interaction de paires est très simplement défini, mais l'utilisation de quatre exponentielles augmente le temps de calcul du potentiel de Coulomb.

Puisque ce dernier potentiel possède un grand nombre d'exponentielles, une simplification a été proposée quelques années plus tard par Piquemal *et al.* [311]. Par la suite, cette approche sera dénommée Cisneros pour qu'il n'y ait pas d'équivoque entre les deux formulations. Cette réécriture simplifie grandement la somme des exponentielles en modifiant la charge portée par l'atome considéré :

$$\begin{cases} q_i^* = 2q_i - \{Z_i - (Z_i - q_i) [1 - \exp(\Omega_{ij})]\} \\ \Omega_{ij} = \frac{\lambda_{AB} r}{(Z_i - q_i)/Z_i} \end{cases} \quad (3.7)$$

Dans cette expression,  $q_i^*$  est la charge portée par l'atome et  $\lambda_{AB}$  est un paramètre d'interaction de paire déterminé pour chaque type d'interaction. Les paramètres déterminés sont donc spécifiques à chaque interaction et non plus des paramètres atomiques comme dans la formulation précédente. Ainsi, un travail supplémentaire est nécessaire pour déterminer le paramètre  $\lambda_{AB}$  pour chaque type d'interaction de molécules ou de fragments de molécules. Par contre, une seule exponentielle est utilisée ici, ce qui simplifie le potentiel par rapport à la formulation originale de Piquemal.

Dans la suite du chapitre, nous allons voir comment la dérivation d'un modèle électrostatique à partir d'une molécule isolée est envisageable. Puis nous verrons l'apport des effets de pénétration dans l'interaction d'espèces chimiques, tout en observant la reproduction de l'énergie d'interaction par le modèle dérivé pour des molécules isolées.

### 3.2.2 Construction du modèle électrostatique

La construction du modèle électrostatique va être détaillée dans le cas de la molécule d'eau. Nous avons fait ce choix d'après l'importance de cette molécule dans les simulations biologiques mais aussi parce que, au regard de la littérature sur la molécule d'eau, les diverses propriétés moléculaires en phase condensée ou non de cette molécule sont difficilement reproduites par un modèle empirique simple. La difficulté à reproduire toutes les propriétés de la molécule d'eau est évidente si l'on considère les nombreux modèles développés à ce jour et répertoriés, par exemple, dans la revue de Guillot [43].

Tout d'abord, le modèle électrostatique de la molécule d'eau est souvent décrit au travers de charges ponctuelles sur les sites atomiques. Cette construction est nommée modèle à trois points (3-p). Avec la stratégie définie au début de ce chapitre, un modèle est généré sur une grille de 2106 points autour de la molécule d'eau isolée et est reporté dans le tableau 3.2.2.

TABLE 3.1: Modèle de charges atomiques et multipôles régénérés pour les molécules d'eau, de benzène, de formaldéhyde et l'ion formate au niveau de théorie MP2(Full)/Sadlej//MP2(Full)/6-311++G(2d,2p) <sup>a</sup>.

	modèle		multipôles moléculaires		
			régénéré		MP2(Full)/Sadlej
Eau (3-p)	$Q_{00}^O$	-0.672	$Q_{10}$	-0.749	-0.732
	$Q_{00}^H$	0.336	$Q_{20}$	-0.189	-0.231
	RMSD	0.990			
	$\Delta\varepsilon$	48.573			
Eau (4-p)	$Q_{00}^X$	-1.237	$Q_{10}$	-0.730	-0.732
	$Q_{00}^O$	0.000	$Q_{20}$	-0.200	-0.231
	$Q_{00}^H$	-0.6185			
	RMSD	0.231			
	$\Delta\varepsilon$	6.258			
Benzène (12-p)	$Q_{00}^C$	-0.124	$Q_{20}$	-5.553	-5.655
	$Q_{00}^H$	0.124	$Q_{40}$	120.200	163.700
	RMSD	0.219			

TABLE 3.1: (suite)

	$\Delta\varepsilon$	20.849			
Benzène (13-p)	$Q_{00}^X$	0.645	$Q_{20}$	-5.651	-5.655
	$Q_{00}^C$	-0.284	$Q_{40}$	159.2	163.7
	$Q_{00}^H$	0.176			
	RMSD	0.050			
	$\Delta\varepsilon$	6.928			
Benzène (12-p/Dipole)	$Q_{00}^C$	-0.200	$Q_{20}$	-5.648	-5.655
	$Q_{00}^H$	0.200	$Q_{40}$	159.1	163.7
	$Q_{10}^C$	-0.209			
	RMSD	0.055			
	$\Delta\varepsilon$	8.450			
Formaldéhyde	$Q_{00}^C$	0.429	$Q_{10}$	-0.949	-0.946
	$Q_{00}^H$	-0.005	$Q_{20}$	-0.284	-0.488
	$Q_{00}^O$	-0.419	$Q_{30}$	-1.332	0.086
	RMSD	0.432			
	$\Delta\varepsilon$	7.568			
Formate	$Q_{00}^C$	0.864	$Q_{20}$	4.456	4.563
	$Q_{00}^H$	-0.210	$Q_{40}$	-17.507	-13.313
	$Q_{00}^O$	-0.827			
	RMSD	0.468			
	$\Delta\varepsilon$	0.283			

<sup>a</sup> Tous les multipôles sont exprimés en unité atomique. L'écart quadratique moyen (RMSD) entre le potentiel électrostatique déterminé en mécanique quantique et régénéré à partir du modèle est exprimé en  $10^{-3}$  a.u. L'erreur relative correspondante,  $\Delta\varepsilon$ , est donnée en pourcentage.

Avec le modèle à trois points, la reproduction du potentiel électrostatique n'est pas très précise ( $\Delta\varepsilon \simeq 49$  %) du fait de la non reproduction du quadrupôle moléculaire de la molécule d'eau [228]. Afin de mieux reproduire le potentiel de la molécule d'eau, une approche dans l'esprit de TIP4P [313] est envisagée. Ce modèle, nommé modèle à quatre points (4-p), se base sur une charge nulle sur l'oxygène et un atome fictif, délocalisé le long de la bissectrice, porteur de la charge. Un schéma est proposé à la figure 3.1. Pour chaque



position de l'atome fictif le long de l'axe  $C_2$  de la molécule, un nouveau jeu de charges est déterminé afin de reproduire au mieux le potentiel (résultats en figure 3.2). Les variations de position de cette particule sont effectuées entre  $-0.4$  et  $0.2$  Å. La direction positive de l'axe est définie telle que celui-ci est à l'extérieur de l'angle  $\widehat{HOH}$  formé par les atomes de la molécule.

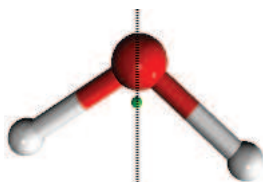


FIGURE 3.1 – Modèle quatre-points dans l'esprit de TIP4P [313]. L'atome fictif (en vert) est placé suivant l'axe bissecteur de l'angle défini par les atomes d'hydrogène et d'oxygène.

La courbe caractérisant l'erreur commise par rapport à la reproduction du potentiel en fonction de la séparation du site fictif de l'oxygène montre un minimum à une distance de  $0.276$  Å. Ce résultat est compatible avec le modèle TIP4P [313] qui place la charge à  $0.15$  Å de l'oxygène. En se situant à ce minimum, la reproduction du potentiel électrostatique est telle que l'erreur relative entre le calcul de mécanique quantique et le modèle est d'environ  $6$  % (voir le tableau 3.2.2). En comparaison avec le modèle à trois points, pour lequel l'erreur est d'environ  $49$  %, le gain en reproduction du potentiel quantique n'est pas négligeable.

Les potentiels électrostatiques du formaldéhyde, de l'ion formate et du benzène [162–164] ont aussi été reproduits à partir de charges ponctuelles. Pour chaque molécule isolée, le potentiel électrostatique a été calculé sur une grille de points. Le nombre de points de la grille est, respectivement, de 3079, 3056 et 8070. Dans le cas du benzène, la reproduction du potentiel électrostatique avec un modèle de charges atomiques n'est pas optimale (environ  $21$  % d'erreur, voir tableau 3.2.2). Comme pour le modèle d'eau, la majeure partie de l'erreur semble être due à une mauvaise reproduction du quadrupôle moléculaire. Pour cette molécule, deux modèles supplémentaires ont donc été dérivés afin de mieux reproduire le quadrupôle moléculaire. Le premier modèle est un modèle à treize points (13-p), dans lequel un site fictif a été ajouté au centre du noyau aromatique. Le second modèle est constitué de charges atomiques et de dipôles atomiques localisés

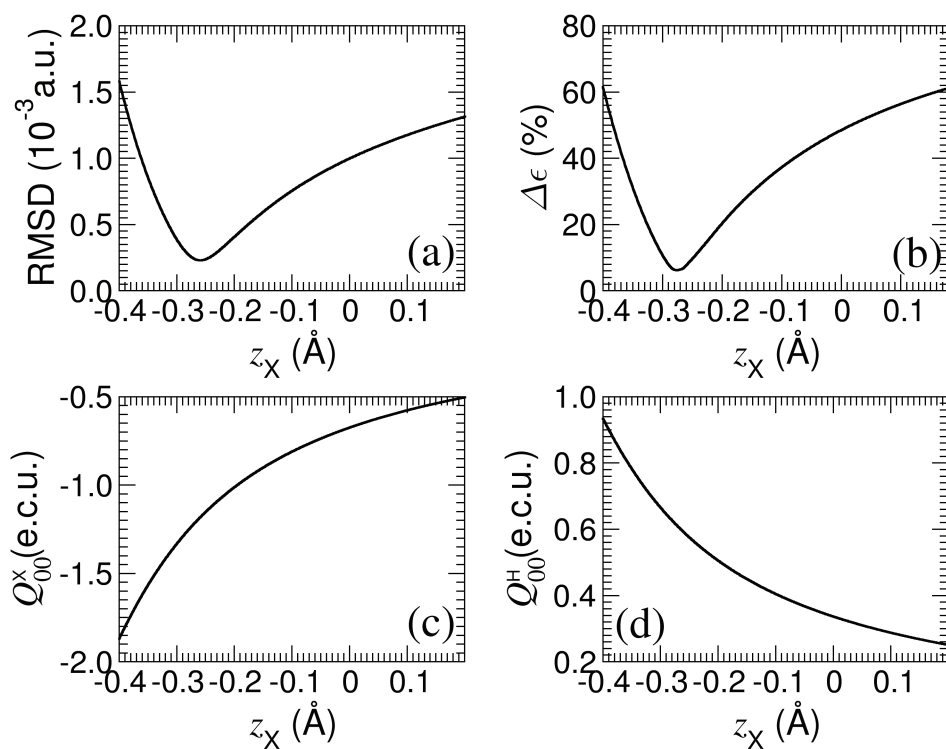


FIGURE 3.2 – (a) Écart quadratique moyen (RMSD) entre le potentiel électrostatique déterminé en mécanique quantique et le modèle de charges ponctuelles exprimé en fonction de la position de l’atome fictif (X) pour la molécule d’eau le long de l’axe  $C_2$ . (b) Erreur correspondante ( $\Delta\epsilon$ ) exprimée comme une fonction de la position de l’atome fictif (X) pour la molécule d’eau le long de l’axe  $C_2$ . (c) Valeur de la charge atomique de l’atome fictif (X) exprimée comme une fonction de la position de l’atome fictif (X) pour la molécule d’eau le long de l’axe  $C_2$ . (d) Valeur de la charge atomique de l’atome d’hydrogène exprimée comme une fonction de la position de l’atome fictif (X) pour la molécule d’eau le long de l’axe  $C_2$ .

sur les atomes carbone (12-p/Dipole). Ces deux modèles reproduisent mieux le potentiel électrostatique du benzène avec, respectivement, 6 et 8 % d'erreur. Ils seront, dans la suite de ce chapitre, comparés au modèle à douze points (12-p) pour la reproduction des énergies d'interaction électrostatique avec divers ions.

### 3.2.3 Interactions électrostatiques

Après avoir défini les modèles pour des molécules isolées, nous allons comparer les énergies d'interaction du modèle électrostatique des molécules (eau, benzène, formaldéhyde, formate) avec des ions et un dimère d'eau en utilisant SAPT ( $\Delta\mathcal{U}_{\text{ele}}^{\text{SAPT}}$ ) comme référence quantique.

Le but de cette étude est de reproduire le potentiel total des molécules en interaction. Pour cela, il faut considérer la gamme de distances sur laquelle le potentiel doit être décrit. Nous avons fait le choix d'un critère énergétique permettant de fixer les distances d'interactions minimales et maximales. Dans chaque cas considéré, sur la courbe d'énergie d'interaction totale ( $\Delta\mathcal{U}_{\text{tot}}$ ), un minimum énergétique ( $U_{\text{min}}$ ) est observé. La distance à laquelle l'énergie calculée en mécanique quantique est  $U_{\text{min}} + 10$  kcal/mol est prise comme distance minimale pour l'interaction. La distance maximale d'interaction est telle que  $\Delta\mathcal{U}_{\text{tot}}$  tende vers zéro ou soit nulle. Cette sélection fournit un modèle où les interactions à courtes distances n'ont pas un trop grand poids étant donné la surface d'énergie potentielle parcourue à 300 K.

Dans un premier temps, je vais décrire les interactions de la molécule d'eau avec des cations, des anions et enfin avec une autre molécule d'eau. Ensuite, je présenterai l'interaction de la molécule de benzène et des cations et finalement, l'interaction du formaldéhyde et de l'ion formate avec le cation calcium. Toutes ces interactions sont schématisées sur la figure 3.3.

#### Interaction molécule d'eau–cations

Dans les articles [163, 164], les interactions de la molécule d'eau et de cations sont présentées. Les cations sont  $\text{Li}^+$ ,  $\text{Na}^+$ ,  $\text{Mg}^{2+}$ ,  $\text{K}^+$  et  $\text{Ca}^{2+}$ . Le positionnement de ces cations a lieu le long de l'axe  $\text{C}_2$  de la molécule d'eau, qui correspond au scénario le plus

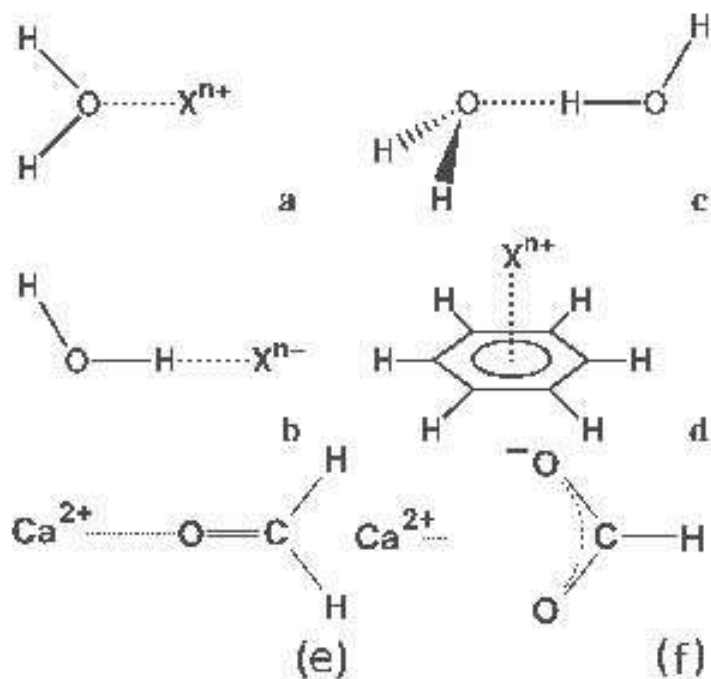


FIGURE 3.3 – (a) Interaction d’une molécule d’eau avec un cation le long de l’axe  $C_2$ . (b) Interaction d’une molécule d’eau avec un anion le long de la liaison O–H. (c) Interaction d’une molécule d’eau avec une autre molécule d’eau formant ainsi une liaison hydrogène. (d) Interaction du benzène avec un cation le long de l’axe  $C_6$ . (e) Interaction du formaldéhyde avec le cation calcium (image issue de [163]). (f) Interaction de l’ion formate avec le cation calcium (image issue de [163]).

TABLE 3.2 – Valeurs du paramètre de pénétration ( $\lambda_{AB}$ ) dans la formulation définie par Cisneros *et al.* (voir équation 3.7). Chaque valeur correspond à une interaction de paire.

	Ca <sup>2+</sup>	K <sup>+</sup>	Li <sup>+</sup>	Mg <sup>2+</sup>	Na <sup>+</sup>	Br <sup>-</sup>	Cl <sup>-</sup>	F <sup>-</sup>	I <sup>-</sup>	H <sub>2</sub> O
H <sub>2</sub> O	-2.90	-2.54	-6.00	-3.60	-3.00	-2.55	-2.82	-4.00	-2.08	-3.34
Benzene	-2.78	-2.50	-2.95	-3.45	-3.34					

favorable énergétiquement [314].

L'interaction de la molécule d'eau et du cation calcium a été étudiée dans les articles [162, 163] et une généralisation aux autres cations a été ensuite apportée [164]. Les interactions de ces cations avec la molécule d'eau sont représentées à la figure 3.4.

Pour chaque interaction, la différence énergétique sur toute la gamme de distances considérée entre le modèle à trois points et le modèle à quatre points est minime. Ceci est visible sur les courbes de la figure 3.4 avec un modèle à trois points et à quatre points qui se superposent. Afin de comprendre cette faible différence, le potentiel électrostatique quantique le long de l'axe C<sub>2</sub> a été comparé au potentiel généré par le modèle de charges ponctuelles. Dans le cas du modèle à trois points l'erreur moyenne pour l'interaction des cations à une distance arbitraire de 2.2 Å est de environ 5 %. Pour le modèle à quatre points l'erreur moyenne est de environ 7 %. Cette légère différence explique en partie que dans le cas de l'interaction du cation magnésium et de l'eau, le modèle à trois points est meilleur que le modèle à quatre points. Le potentiel électrostatique du modèle à trois points est donc bien reproduit le long de cet axe. Par contre, en prenant la liaison chimique O–H, à une distance de 2.2 Å d'un hydrogène, le modèle à trois points conduit à une erreur moyenne d'environ 16 %, alors que celle-ci reste limitée à environ 5 % dans le cas du modèle à quatre points.

L'utilisation des fonctions de pénétration permet de reproduire la courbe d'interaction de la molécule d'eau avec les cations de manière très satisfaisante sur la gamme de distances utilisée. Chaque paramètre de la fonction d'atténuation est ajusté afin de reproduire au mieux le potentiel. Les valeurs sont données dans les tableaux 3.2 et 3.3.

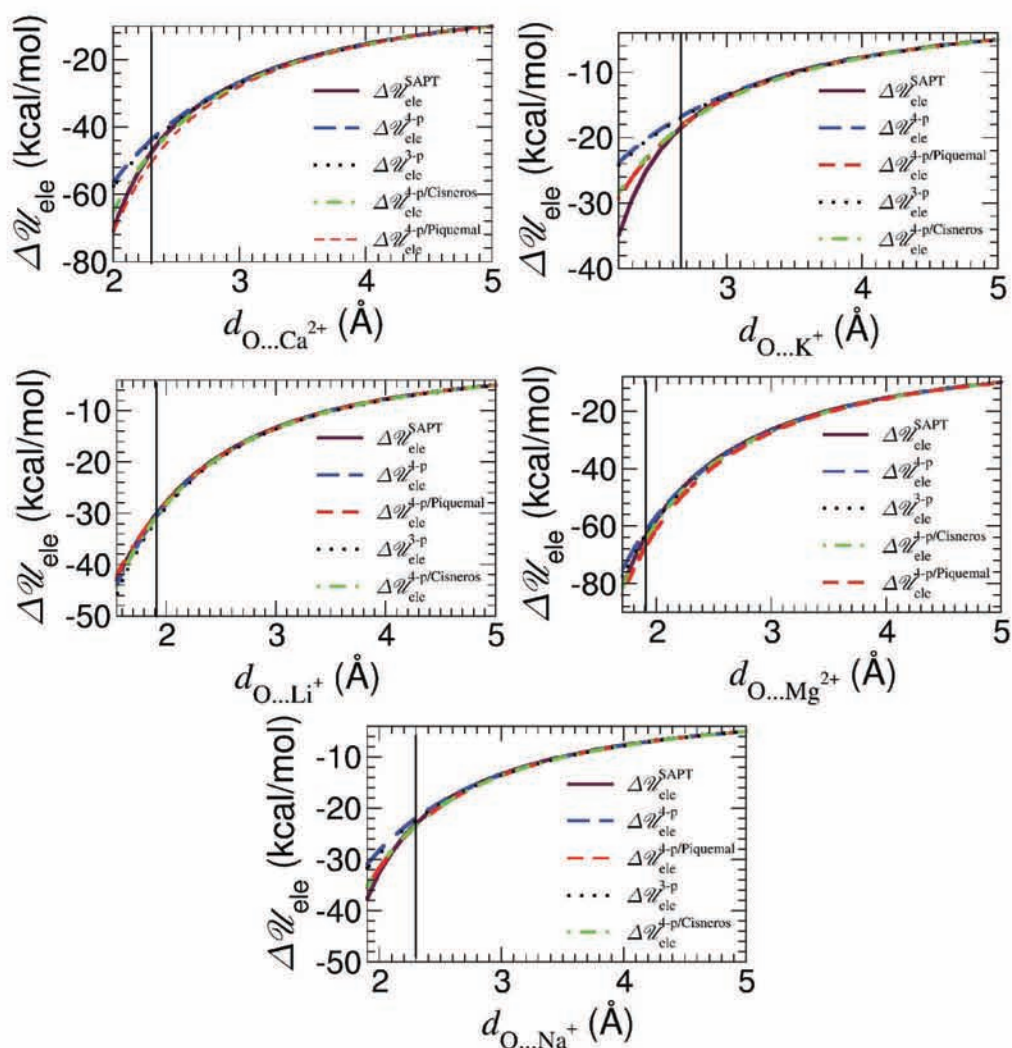


FIGURE 3.4 – Interactions du cation calcium ( $\text{Ca}^{2+}$ ), du cation potassium ( $\text{K}^+$ ), du cation lithium ( $\text{Li}^+$ ), du cation magnésium ( $\text{Mg}^{2+}$ ) et du cation sodium ( $\text{Na}^+$ ) avec la molécule d'eau le long de l'axe  $\text{C}_2$  de celle-ci. La référence (SAPT) est la courbe en trait plein. Le modèle à trois points et à quatre points pour la molécule d'eau sont en pointillés et en ligne discontinue, respectivement. La correction des effets de pénétration est en ligne discontinue courte et ligne discontinue pointillée lors de l'application des corrections de Piquemal *et al.* (équation 3.5) et de Cisneros *et al.* (équation 3.7), respectivement. La barre verticale correspond au minimum de l'énergie d'interaction.

TABLE 3.3 – Valeurs du paramètre de pénétration ( $r_{vdW}$ ) pour chaque type d'atome dans la formulation définie par Piquemal *et al.* (voir équations 3.6 et 3.5).

	O(H <sub>2</sub> O)	H(H <sub>2</sub> O)	Ca <sup>2+</sup>	K <sup>+</sup>	Li <sup>+</sup>	Mg <sup>2+</sup>	Na <sup>+</sup>
$r_{vdW}$ (Å)	1.20	1.56	1.40	1.50	1.10	1.00	0.95

### Interaction molécule d'eau–anion et dimère d'eau

L'interaction de la molécule d'eau suivant la liaison O–H avec des anions (F<sup>−</sup>, Cl<sup>−</sup>, Br<sup>−</sup>, I<sup>−</sup>) et avec une autre molécule d'eau a été étudiée. La géométrie du dimère d'eau est prise par rapport à la référence [315]. Dans cet article, l'énergie d'interaction de nombreuses géométries du dimère est évaluée en mécanique quantique à un haut niveau de théorie. Parmi les géométries, j'ai choisi de prendre la géométrie du dimère au minimum. Les courbes d'interaction sont disponibles aux figures 3.5. Pour ces interactions, le modèle à trois points ne reproduit pas fidèlement l'énergie d'interaction électrostatique de référence. Ce désaccord est lié à l'erreur ( $\Delta\varepsilon \simeq 49\%$ ) observée sur la reproduction du potentiel total (voir tableau 3.2.2). Comme indiqué précédemment, la différence entre le modèle à trois points (erreur d'environ 16 %) et le modèle à quatre points (erreur d'environ 5 %) pour la reproduction du potentiel électrostatique le long de l'axe O–H est d'environ 11 %. Cette différence énergétique entre les deux modèles est tout à fait reproduite au niveau des courbes d'interaction entre la molécule d'eau et des anions et explique l'écart observé sur la figure 3.4 entre la référence quantique, le modèle à trois points et le modèle à quatre points. Pour ce dernier, malgré la bonne reproduction de l'énergie d'interaction le long de l'axe C<sub>2</sub>, la reproduction de l'énergie d'interaction le long de la liaison O–H est catastrophique, même à des distances telles que le minimum de l'interaction (voir la ligne verticale noire dans les graphes). En comparaison, le modèle à quatre points décrit de manière fidèle les interactions des diverses espèces. Ceci n'est vrai que dans les gammes de distances où les effets de pénétration électronique ne sont pas importants car, pour les espèces considérées dans ce cas, la couche externe de valence est totalement remplie. L'extension spatiale de leurs électrons est plus étendue que dans le cas de cations. L'interaction non-multipolaire due aux électrons de ces espèces avec la molécule d'eau est ainsi plus importante que dans le cas de l'interaction avec des cations. Ces effets à courte portée ne peuvent donc pas être reproduits par la seule partie multipolaire exacte du potentiel. Malgré cela, le modèle à

quatre points est satisfaisant sur la gamme de distances proches du minimum d'équilibre et ne conduit pas à des erreurs de la même magnitude que pour le modèle à trois points.

Dans les cas présentés ici, la contribution des effets de pénétration peut être importante (jusque 50 % de l'énergie d'interaction). Pour reproduire très fidèlement l'énergie d'interaction sur toute la gamme de distances considérée, il est possible d'utiliser les deux fonctions présentées ici (voir équations 3.5 et 3.7). Il s'avère que la flexibilité du modèle de Cisneros *et al.* permet de mieux reproduire l'énergie d'interaction à courte distance (voir tableaux 3.2 et 3.3 pour les valeurs des paramètres).

### Interaction benzène–cations

Une étude préliminaire menée par Dehez *et al.* [162] traite l'interaction de la molécule de benzène avec le cation ammonium ( $\text{NH}_4^+$ ). L'article révèle une très bonne reproduction de l'énergie d'interaction électrostatique de ces deux molécules. Puis, l'interaction de cations avec le benzène le long de l'axe  $\text{C}_6$  de celui-ci est proposée [164]. Dans ces dimères, l'énergie électrostatique n'est pas bien reproduite par un simple modèle de charges atomiques. Le premier facteur qui pourrait être à l'origine de cette déficience est la reproduction sous-optimale du potentiel électrostatique (voir tableau 3.2.2). L'écart au potentiel calculé en mécanique quantique est assez significatif et peut engendrer les différences énergétiques observées entre le modèle à douze points et la référence quantique (voir figure 3.6). Un deuxième facteur tient à la localisation plus importante de la charge dans le cas des cations métalliques par rapport à l'ion ammonium. De par la taille des cations, les interactions électrostatiques sont plus localisées et donc l'importance de cette contribution induit un écart dans la reproduction de l'énergie d'interaction électrostatique. Une troisième explication réside dans les possibles effets de pénétration dans le plan de la molécule de benzène, dus aux électrons  $\pi$ . Ces effets de pénétration ne peuvent, par définition, pas être reproduits par un modèle décrivant la partie multipolaire exacte du potentiel électrostatique.

Les trois points soulevés ici sont à l'origine de deux modèles plus complexes développés pour palier ces écarts. Je vais, dans la suite, les décrire et expliquer le choix final pour le modèle électrostatique du benzène. Le premier modèle définit des charges sur chaque atome et des dipôles sur les carbones (modèle 12-p/Dipole). Le deuxième modèle permet



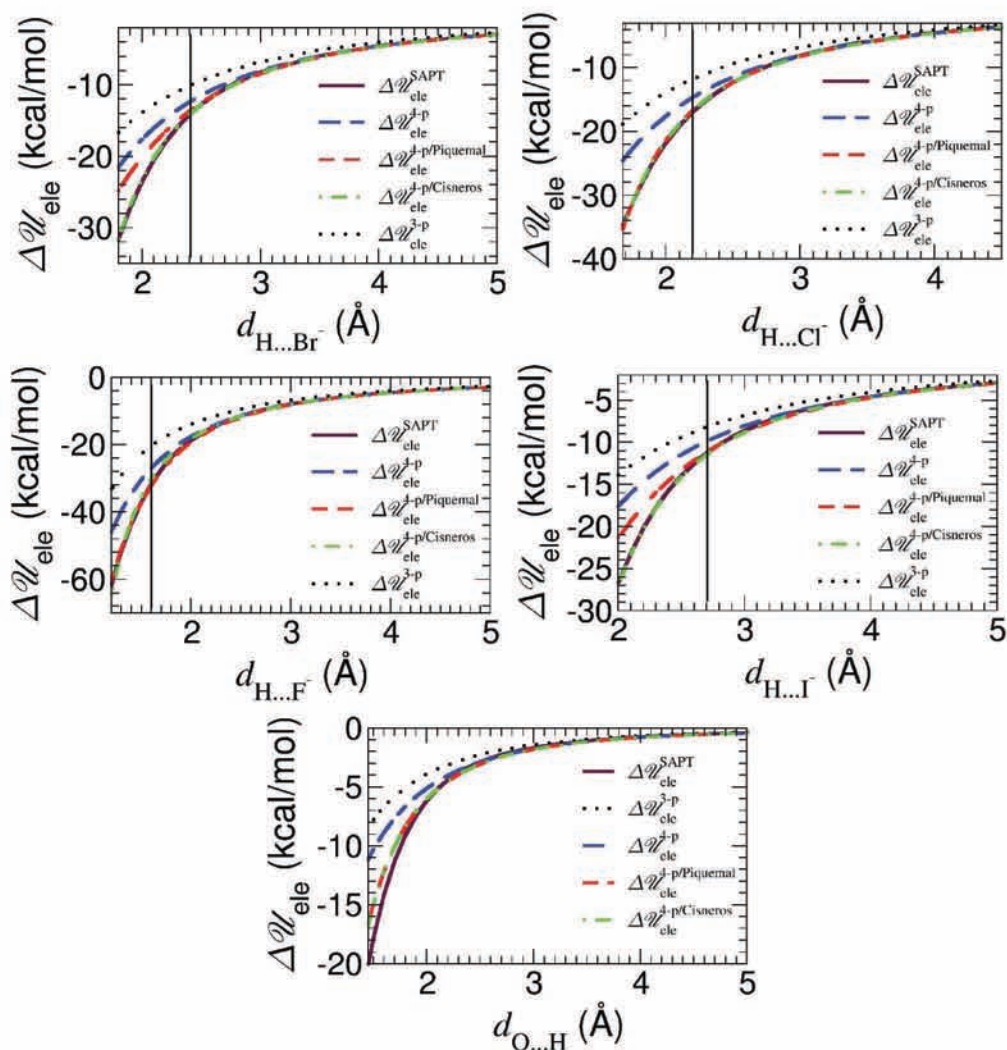


FIGURE 3.5 – Interactions de l’anion bromure ( $\text{Br}^-$ ), de l’anion chlorure ( $\text{Cl}^-$ ), de l’anion fluorure ( $\text{F}^-$ ), de l’anion iodure ( $\text{I}^-$ ) et de l’eau avec la molécule d’eau le long d’une liaison O–H de celle-ci. La référence (SAPT) est la courbe en trait plein. Les modèles à trois points et à quatre points pour la molécule d’eau sont en pointillés et en ligne discontinue, respectivement. La correction des effets de pénétration est en ligne discontinue courte et ligne discontinue pointillée lors de l’application des corrections de Piquemal *et al.* (équation 3.5) et de Cisneros *et al.* (équation 3.7), respectivement. La barre verticale correspond au minimum de l’énergie d’interaction.

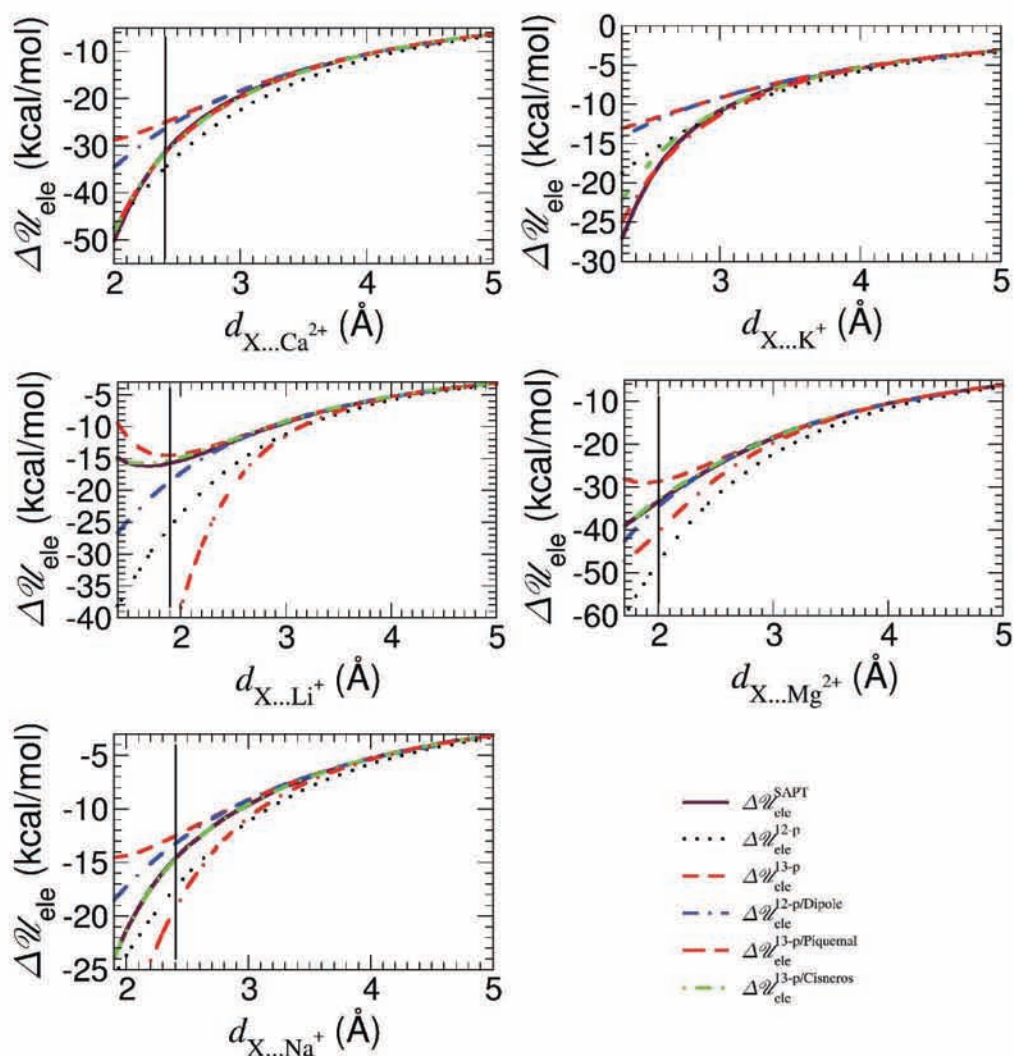


FIGURE 3.6 – Interactions du cation calcium ( $\text{Ca}^{2+}$ ), du cation potassium ( $\text{K}^+$ ), du cation lithium ( $\text{Li}^+$ ), du cation magnésium ( $\text{Mg}^{2+}$ ) et du cation sodium ( $\text{Na}^+$ ) avec la molécule de benzène le long de l'axe  $\text{C}_6$  de celle-ci. La référence (SAPT) est la courbe en trait plein. Les modèles à douze points et à treize points pour la molécule d'eau sont en pointillés et en ligne discontinue, respectivement. La correction des effets de pénétration est en ligne discontinue courte et en ligne discontinue pointillée lors de l'application des corrections de Piquemal *et al.* (équation 3.5) et de Cisneros *et al.* (équation 3.7), respectivement. La barre verticale correspond au minimum de l'énergie d'interaction.

de conserver la symétrie imposée par le benzène et de ne rajouter qu'une seule charge. Dans la littérature [303], des modèles de charge plus complexes ont défini celles-ci en dehors du plan du benzène. Du fait de la symétrie  $D_{6h}$  du benzène, un grand nombre de sites est à considérer (18 ou 24). Pour ne pas utiliser autant de charges, le modèle que j'ai proposé n'a qu'une simple charge fictive au centre de la molécule de benzène (modèle à treize points, 13-p). Pour ces deux modèles, les énergies d'interaction électrostatiques sont mieux reproduites sur une plus grande gamme de distances. De plus, comme présenté dans le cas des énergies d'interactions électrostatiques de l'eau et des cations, les effets de pénétration sont des effets stabilisants. Sans cette contribution, l'énergie d'interaction électrostatique d'un modèle purement multipolaire ne peut être plus stable que la référence quantique. Or, pour le modèle à douze points, nous observons que l'énergie d'interaction électrostatique du modèle est plus stable que la référence. Ceci montre que le nuage électronique du benzène est mal décrit. Pour améliorer la description du nuage électronique et de ses variations en fonction de l'environnement, Hunter *et al.* [316] ont proposé de modifier les charges portées par les atomes de manière empirique. En revanche, aucune fonction explicite ne semble disponible dans la littérature pour décrire précisément cet effet. De plus, il a été montré qu'un modèle de benzène, dont le potentiel électrostatique est décrit par un simple jeu de charges, était capable de reproduire un bon nombre de conformations d'interaction de dimères [317]. Il est donc possible, comme dans le cas des champs de force utilisés en dynamique moléculaire, de compenser l'erreur au travers des autres fonctions du potentiel intermoléculaire, en particulier grâce au potentiel de van der Waals.

Il est intéressant de regarder plus en détail le cas de l'interaction de la molécule de benzène avec le cation lithium. Un comportement à courte distance, différent par rapport aux autres cas, est observé avec un changement de pente de la courbe d'interaction électrostatique vers la distance d'équilibre. L'énergie d'interaction électrostatique devient défavorable pour des distances proches. À très courte distance (inférieure à 2 Å), ce comportement est aussi visible dans le cas de l'interaction du cation sodium avec la molécule de benzène. Cet effet à courte distance s'explique par le fait que les électrons du système  $\pi$  du benzène ne sont pas confinés dans le plan de la molécule. Ils peuvent ainsi interagir avec les électrons du cation lorsque celui-ci s'approche du cycle benzénique. Parmi les modèles présentés ici, seul le modèle à treize points est à même de reproduire ce comportement, mais sa validité physique peut être mise en question. En particulier, la

charge déterminée au centre du noyau aromatique est positive alors que celle-ci devrait reproduire la contribution des électrons  $\pi$ . Comme parfois dans les modèles limités à une distribution de monopôles, la description est mathématiquement correcte, au sens moindres carrés du terme, mais physiquement peu satisfaisante [28, 32, 178, 179].

La comparaison de tous les modèles présentés montre que le plus utilisé, dans le cas qui nous intéresse (la dynamique moléculaire), ne reproduit pas forcément le calcul quantique avec la plus grande précision. De plus, si le modèle à treize points ne peut être réellement pris en considération au regard de la remarque précédente, l'ajout de dipôles sur les atomes de carbone ne convient pas non plus du fait de l'hypothèse choisie d'un modèle électrostatique ne faisant intervenir que des charges. Au final, malgré la reproduction un peu faussée de l'énergie d'interaction électrostatique dans le cas de l'interaction de la molécule de benzène avec des cations, ce modèle est le plus simple et le plus compact disponible pour cette molécule. Les effets tels que la pénétration du nuage électronique  $\pi$  qui interviennent fortement peuvent alors être intégrés directement dans un potentiel adéquat, comme un potentiel de van der Waals. De plus, comme le montre l'étude de l'interaction avec le cation ammonium [162], si la charge est délocalisée, alors le modèle de charges atomiques est tout à fait satisfaisant. On peut donc s'attendre à avoir aussi un bon accord en considérant, par exemple, le cation guanidinium.

### Interaction de l'ion calcium avec le formaldéhyde, le formate et l'eau

Les molécules d'eau, de formaldéhyde et de formate peuvent servir de modèles aux espèces en interaction au site catalytique d'une enzyme comme la phospholipase A2. Au site actif de celle-ci, un calcium est en interaction forte avec les acides aminés (glycine, tyrosine et aspartate) et l'eau. La modélisation précise de cette interaction est nécessaire afin d'obtenir une structure stable au cours de la dynamique moléculaire de l'enzyme (travaux non publiés). Des paramètres spécifiques ont été développés par Shiratori et Nakagawa [318] pour simuler cette enzyme. Toutefois, du fait de la forte proximité du calcium avec l'enzyme et l'eau, un fort champ polarisant est généré et une reparamétrisation est nécessaire pour ajouter de la polarisation explicite. Pour décrire cela plus précisément, la reproduction du potentiel total a été étudiée dans les articles de Dehez *et al.* [162, 163]. Les courbes d'interaction sont consignées dans les figures 3.7. Dans ces figures, il est montré que les charges atomiques obtenues par la mécanique

quantique reproduisent parfaitement l'énergie d'interaction électrostatique, tant que les effets de pénétration sont négligeables.

### 3.3 Induction

La seconde contribution à l'énergie d'interaction de notre modèle est l'énergie d'induction ou de polarisation. Les différentes méthodes qui existent à l'heure actuelle pour prendre en compte ces effets explicitement ont été détaillées dans la section 2.2. Je vais expliquer comment dériver les paramètres de polarisabilité à partir d'une molécule isolée, puis comment calculer l'énergie d'induction dans le cas de dimères en utilisant des fonctions d'atténuation spécifiques. Enfin, je commenterai le comportement de l'énergie d'induction modélisée par rapport à la référence quantique.

#### 3.3.1 Construction du modèle de polarisabilité

Le modèle de polarisabilités est dérivé à partir d'un développement multipolaire. Ceci s'apparente à la méthode décrite dans la partie 2.2.1. La grande différence entre le modèle proposé ici et les modèles de polarisabilités utilisés dans les champs de forces polarisables « usuels » est l'ordre du développement multipolaire. Bien souvent, les modèles de polarisabilités distribuées sont restreints à des polarisabilités dipolaires isotropes (ordre 1). Dans le cas où le nuage électronique d'une molécule ou d'un fragment moléculaire peut se déformer différemment dans les trois directions de l'espace, les polarisabilités dipolaires isotropes peuvent s'avérer incapables de décrire correctement la polarisabilité moléculaire. Un exemple convaincant est celui du benzène. Un simple modèle de polarisabilités dipolaires isotropes situées sur les atomes ne peut pas reproduire chaque élément du tenseur de polarisabilité moléculaire. Prenons les valeurs de polarisabilité obtenues en mécanique quantique dans le tableau 3.4 : la polarisabilité moléculaire du benzène hors du plan ( $\alpha_{10,10}$ ) est d'une magnitude deux fois inférieure à celle dans le plan. Le modèle isotrope reproduira au mieux la trace du tenseur. Pour décrire plus précisément ce tenseur, le modèle doit être anisotrope, par exemple en ajoutant des polarisabilités d'ordre zéro au modèle. Cette polarisabilité, non locale, est appelée polarisabilité de flux de charge (voir page 132 du livre de Stone [9]). Du fait de la non-localité des flux de charge, la polarisabilité moléculaire est anisotrope car une composante « le long des atomes »

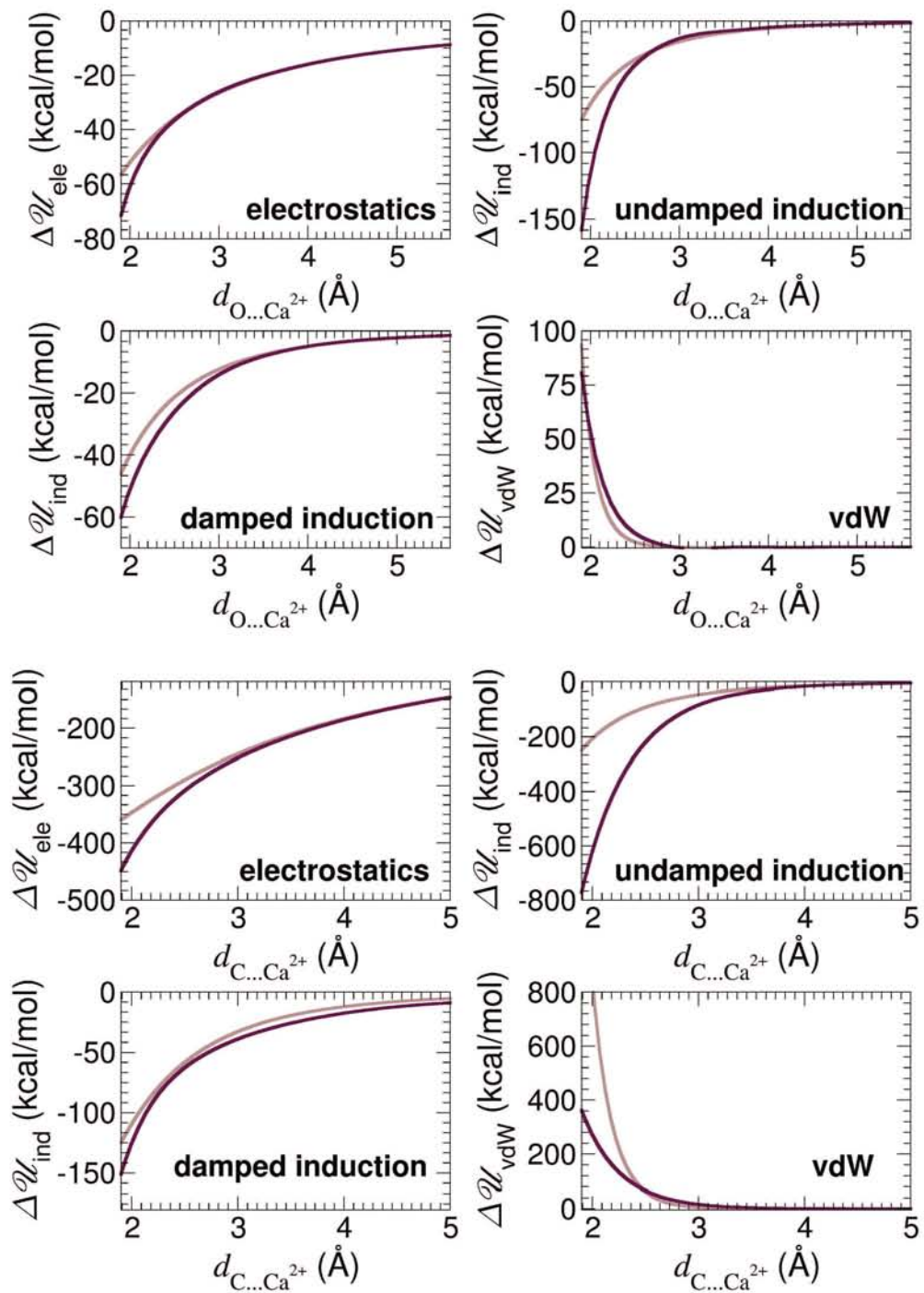


FIGURE 3.7 – Figure issue de l'article [163]. Les interactions représentées sont entre le cation calcium et l'anion formate ( $\text{O}\dots\text{Ca}^{2+}$ ) ou le formaldéhyde ( $\text{C}\dots\text{Ca}^{2+}$ ).

est générée. Il faut noter que cette polarisabilité n'est pas équivalente à l'approche des charges fluctuantes décrites dans la section 2.2.2. Pour prendre en compte l'anisotropie de la polarisabilité moléculaire, nous proposons un modèle composé de polarisabilités distribuées sur les atomes d'ordre zéro (flux de charges) entre atomes liés chimiquement et d'ordre un (dipolaire) isotrope sur les atomes lourds [162–164, 319].

À présent que l'ordre exact du modèle est choisi, une procédure fiable doit être employée. Je vais pour cela utiliser le logiciel OPEP [33]. La section de ce code permettant la dérivation des paramètres a été développée dans les années 2000 [195, 320–322]. Dans ces articles, la stratégie qui permet de déterminer les paramètres pour chaque ordre de polarisabilité est décrite. L'expérience de chaque article a permis de considérer le modèle présenté dans cette thèse comme le plus à même à reproduire chaque partie du tenseur de polarisabilités moléculaires.

Pour générer les polarisabilités distribuées, comme dans le cas de la dérivation des paramètres électrostatiques, il faut dans un premier temps, générer une grille de points autour de la molécule. Une charge polarisante est ensuite placée successivement en chaque point de la grille et l'énergie induite par cette charge sur la molécule est déterminée par un calcul quantique (équation 3.8). Les caractéristiques de cette grille diffèrent sensiblement de celles des grilles utilisées pour dériver des multipôles. Des études ont montrées [33] que la taille de grille optimale était une séparation entre les points ( $\Delta r$ ) de 0.75 Å (0.5 Å pour l'obtention des multipôles). Le facteur multiplicatif par rapport au rayon de van der Waals des atomes ( $\xi$ ) est quant à lui de 5.0 (3.0 les multipôles). Pour les polarisabilités distribuées, la grille optimale nécessaire à une bonne reproduction de la polarisabilité est moins dense que pour les multipôles. La densité de la grille n'est pas si importante dans le cas des multipôles (voir section 3.2) car le calcul de  $\langle \Psi | \frac{1}{R} | \Psi \rangle$  évalué sur  $N$  ou  $2N$  points ne demande pas beaucoup plus de temps. Par contre, dans le cas de la détermination de polarisabilités distribuées, une évaluation de la fonction d'onde à chaque position de la charge polarisante est nécessaire. De plus, la grille d'induction a une plus grande extension spatiale étant donné que les effets de polarisation convergent à grande distance. Ainsi, par construction, le nombre d'évaluation de points de grille est plus important dans le cas d'une grille d'induction et il est donc nécessaire de diminuer la densité de celle-ci pour que le nombre de calculs quantiques soit raisonnable (plusieurs dizaines de milliers de points maximum). Pour des raisons identiques au cas du modèle

électrostatique, aucun point en deçà de deux fois le rayon de van der Waals des atomes n'est considéré pour s'affranchir des effets d'hyperpolarisabilité.

Pour déterminer l'énergie d'induction,  $\mathcal{U}_{ind,k}$ , pour chaque point de grille  $k$ , il faut évaluer le potentiel électrostatique,  $\mathcal{V}_k$ , créé par la molécule au point  $k$ . Il faut aussi obtenir l'énergie de la molécule isolée ( $\mathcal{E}^0$ ) au travers d'un calcul quantique. Enfin, pour chaque point de grille, la molécule considérée va être perturbée par une charge localisée sur un des  $k$  points de grille. Typiquement, une charge  $q_k = +1e$  est utilisée. Pour chaque position de la charge perturbatrice, l'énergie ( $\mathcal{E}_{total,k}$ ) du système charge+molécule est évaluée au niveau de théorie considéré. Ce dernier calcul, à répéter pour chaque position de la charge, est la plus consommateur en terme de ressources de calcul et limite donc la possibilité d'élaborer des grilles pour de très grands systèmes moléculaires. Par exemple, pour la simple molécule d'eau, 3905 points de grille sont nécessaires. Au final, l'énergie d'induction calculée en mécanique quantique est :

$$\mathcal{U}_{ind,k} = \mathcal{E}_{total,k} - \mathcal{E}^0 - q_k \mathcal{V}_k \quad (3.8)$$

Une fois l'énergie d'induction calculée quantiquement, il est possible de l'approximer par un développement multipolaire de la polarisabilité à un ordre donné :

$$\tilde{\mathcal{U}}_{ind,k} = -\frac{1}{2} \sum_{a,b} \sum_{l=0}^{N_l(a)} \sum_{\kappa} \sum_{l'=0}^{N_{l'}(b)} \sum_{\kappa'} q_k T_{l\kappa}^{ka} \alpha_{l\kappa,l'\kappa'}^{ab} T_{l'\kappa'}^{bk} q_k \quad \forall k = 1, \dots, N_p \quad (3.9)$$

Dans cette expression,  $\alpha_{l\kappa,l'\kappa'}^{ab}$  correspond au tenseur de polarisabilité d'ordre  $N_l$  et  $N_{l'}$  relatif aux sites  $a$  et  $b$ , respectivement.  $N_p$  est le nombre de points de grille d'induction.

Le modèle ici présenté se limite à un ordre un isotrope auquel est ajouté l'ordre zéro. En terme chimique, cela se traduit par une polarisabilité de flux de charges pour des atomes connectés par une liaison chimique et une polarisabilité dipolaire isotrope située sur les atomes lourds. Le système linéaire d'équations surdéterminé peut être résolu en minimisant la fonctionnelle :

$$g(\{\alpha_{l\kappa,l'\kappa'}^{ab}\}) = \sum_{k=1}^{N_p} (\mathcal{U}_{ind,k} - \tilde{\mathcal{U}}_{ind,k})^2 \quad (3.10)$$

La résolution d'un tel système d'équations permet d'obtenir des paramètres reproduisant l'énergie d'induction au « sens des moindres carrés ». Afin d'évaluer la précision par rapport à la cible, des critères de fiabilité sont calculés. Ces critères sont les mêmes



que dans le cas du modèle électrostatique : le RMSD et  $\Delta\varepsilon$  (voir équation 3.4). Dans ces équations, il faut remplacer le potentiel électrostatique calculé quantiquement,  $V_k$ , par le potentiel d'induction ( $\mathcal{U}_{ind,k}$ ) et faire de même pour les potentiels approchés ( $\tilde{V}_k$  et  $\tilde{\mathcal{U}}_{ind,k}$ ).

### 3.3.2 Énergie d'interaction d'induction

Une fois le modèle de polarisabilités distribuées dérivé pour des molécules isolées, il faut l'inclure dans un potentiel d'interaction. La contribution de polarisation à l'énergie d'interaction ( $U_{ind}^{MM}$ ) peut se calculer à partir des multipôles et des polarisabilités distribuées (voir aussi le chapitre 8 de l'ouvrage de Stone [9]) :

$$\begin{cases} U_{ind}^{MM} = \frac{1}{2} \sum_A \sum_{B \neq A} \Delta Q_t^a T_{tu}^{ab} Q_u^b \\ \Delta Q_{t'}^{a'} = - \sum_{B \neq A} \alpha_{t't}^{a'a} T_{tu}^{ab} (Q_u^b + \Delta Q_u^b) \end{cases} \quad (3.11)$$

Dans cette équation,  $Q_u^b$  et  $\Delta Q_t^a$  représentent, respectivement, le multipôle d'ordre  $u$  situé sur un site  $b$  et le multipôle induit d'ordre  $t$  situé sur un site  $a$ . Dans le cadre de cette thèse, les multipôles sont d'ordre zéro, c'est-à-dire qu'ils correspondent à des charges ponctuelles.  $T_{tu}^{ab}$  est le tenseur d'interaction d'ordre  $tu$  entre les sites  $a$  et  $b$ .  $\alpha_{t't}^{a'a}$  est la polarisabilité d'ordre  $t't$  entre les sites  $a'$  et  $a$ .

Le modèle de polarisabilités étant limité à des polarisabilités de flux de charges et dipolaires isotropes, le tenseur d'interaction varie, entre un dipôle induit et une charge, en  $1/r^2$ . Le tenseur d'interaction entre une charge induite (le flux de charge) et une charge est en  $1/r$ .

### 3.3.3 Fonctions d'atténuation

Les énergies d'induction, calculées par l'intermédiaire de l'équation 3.11, sont surestimées à courte distance. La correction à l'énergie d'induction d'interaction passe par l'utilisation de fonctions d'atténuation qui rendent compte des contributions d'échange électronique à l'énergie d'induction. Dans la section 2.2.1, trois familles de fonctions d'atténuation ont été introduites. Nous utiliserons ici les fonctions de Jensen et de Tang et Toennies [163, 164].

Dans l'approche de Jensen, la distance intervenant dans le tenseur d'interaction  $T_{tu}^{ab}$  est modifiée suivant l'équation 2.21. Pour chaque type d'interaction, il faut donc déterminer le paramètre  $a$  permettant de reproduire au mieux l'énergie de polarisation à courte distance.

Dans le cas de la fonction d'atténuation de Tang et Toennies, l'énergie d'interaction est multipliée par la fonction  $f_n(r)$  (voir équation 2.22). Le paramètre  $a$  du développement est, comme dans le cas de l'approche de Jensen, déterminé pour chaque interaction. Stone et collaborateurs [91, 92] ont proposé de déterminer ce paramètre à partir des énergies d'ionisation des fragments en interaction. L'application de la fonction d'atténuation à l'énergie d'interaction d'induction (équations 3.11) donne :

$$\begin{cases} U_{ind}^{MM} = \frac{1}{2} \sum_A \sum_{B \neq A} \Delta Q_t^a T_{tu}^{ab} f_n(r_{ab})^{1/2} Q_u^b \\ \Delta Q_{t'}^{a'} = - \sum_{B \neq A} \alpha_{t't}^{a'a} T_{tu}^{ab} f_n(r_{a'b})^{1/2} (Q_u^b + \Delta Q_u^b) \\ f_{n'}(r) = 1 - \exp(-ar) \sum_{k=0}^{n'} \frac{(ar)^k}{k!} \end{cases} \quad (3.12)$$

Le facteur d'atténuation considéré dans cette approche revient à multiplier le tenseur d'interaction  $T_{tu}^{ab}$  par  $\sqrt{f_n(r_{ab})}$ . Stone et Millot [228] ont montré que la fonction d'atténuation de Tang and Toennies pouvait être raisonnablement tronquée à l'ordre  $n' = 6^2$ . Un moyen pour éviter le calcul de la racine carrée de la fonction d'atténuation a été proposé dans un article de Meredith et Stone [229]. La version simplifiée de la fonction d'atténuation tronquée à l'ordre 6,  $f_6(r)$ , est :

$$[f_3(0.725b; r)]^2 \approx f_6(b; r) \quad (3.13)$$

En conclusion, l'énergie d'induction modèle devient :

$$\begin{aligned} U_{ind}^{MM} &= \frac{1}{2} \sum_A \sum_{B \neq A} \Delta Q_t^a T_{tu}^{ab} f_3(\beta, r_{ab}) Q_u^b \\ \Delta Q_{t'}^{a'} &= - \sum_{B \neq A} \alpha_{t't}^{a'a} T_{tu}^{ab} f_3(\beta, r_{a'b}) (Q_u^b + \Delta Q_u^b) \end{aligned} \quad (3.14)$$

où  $\beta = b/0.725$  est le coefficient d'atténuation modifié de Tang et Toennies. Le terme  $f_3(\beta, r_{ab})$  s'écrit explicitement :

$$f_3(\beta, r_{ab}) = 1 - \exp(-\beta r_{ab}) \left[ 1 + \beta r_{ab} + \frac{(\beta r_{ab})^2}{2} + \frac{(\beta r_{ab})^3}{6} \right] \quad (3.15)$$

---

2. De ce fait, seuls les calculs de puissances inférieures ou égales à 6 doivent être effectués. Il est remarquable que celles-ci soient déjà évaluées dans des fonctions telles que le Lennard-Jones 12-6 ou le Buckingham exp-6.

### 3.3.4 Énergie d'interaction d'induction entre molécules

#### Molécule d'eau-cation

Comme présenté dans la partie précédente, les interactions avec les cations ont d'abord été effectuées en prenant le cas de la molécule d'eau avec le cation calcium [162]. Dans ce cas, un très bon accord entre l'énergie de référence SAPT et l'énergie du modèle d'induction est obtenu. Le modèle de polarisabilités distribuées mis au point avec une grille de 3905 points est donné dans le tableau 3.4. Le paramètre de la fonction d'atténuation de Jensen vaut  $a = 0.085$  et est obtenu de manière à ce que la référence quantique soit reproduite sur toute la gamme de distances considérée dans l'article [162].

Les mêmes cas de figure que dans la section 3.2, c'est-à-dire l'interaction de différents cations (calcium, potassium, lithium, magnésium et sodium) avec la molécule d'eau, sont repris pour l'énergie d'interaction d'induction. Les courbes sont rassemblées dans la figure 3.8. Dans le modèle, j'ai choisi d'ajouter une polarisabilité dipolaire isotrope sur le cation potassium car, dans le cas de l'interaction du cation potassium avec la molécule d'eau, la courbe quantique résultante le long de l'axe  $C_2$  de la molécule d'eau ne peut pas être reproduite par un modèle où seule la molécule d'eau est polarisable. Tout se passerait comme s'il n'y avait pas de terme d'échange-induction et l'ajout de cette polarisabilité dipolaire isotrope permet d'obtenir une énergie d'induction multipolaire plus stabilisante que l'énergie d'interaction d'induction (incluant forcément un terme d'échange-induction). De ce fait, la polarisabilité dipolaire isotrope de tous les cations a été calculée quantiquement avec GAUSSIAN [307]. Seule la contribution du cation potassium apparaît comme non négligeable (voir tableau 3.5). Il faut noter que cette polarisabilité est assez importante, elle vaut environ 50 % de la polarisabilité moléculaire de la molécule d'eau. Pour les autres cations, la valeur de la polarisabilité calculée est inférieure à  $2 \text{ bohr}^3$ . Ajouter une polarisabilité dipolaire isotrope sur le cation potassium est aussi mis en avant dans certains articles traitant de la sélectivité du cation potassium par rapport au cation sodium dans les canaux ioniques [99, 101] du fait des interactions à l'extérieur et à l'intérieur du canal.

Le comportement observé pour une interaction entre les multipôles et les polarisabilités distribuées sur-stabilise, à courte distance, l'interaction entre les espèces (voir les pointillés courts sur la figure 3.8). Ce comportement est celui attendu pour une

TABLE 3.4 – Modèles de polarisabilités distribuées et polarisabilités moléculaires régénérées pour l'eau, le benzène, l'anion formate et le formaldéhyde au niveau de théorie MP2(Full)/Sadlej//MP2/6-311++G(2d,2p) <sup>a</sup>.

	distributed polarizabilities		molecular polarizabilities		
			regenerated	MP2(Full)/Sadlej	
Eau	$\alpha_{00,00}^{OH}$	-0.808	$\alpha_{10,10}$	10.177	9.747
	$\alpha_{1\kappa,1\kappa}^{OO}$	8.1808	$\alpha_{11c,11c}$	11.483	10.058
			$\alpha_{11s,11s}$	8.180	9.539
	RMSD	0.127			
	$\Delta\varepsilon$	7.000			
Benzène	$\alpha_{00,00}^{CC}$	-1.822	$\alpha_{10,10}$	47.537	45.121
	$\alpha_{00,00}^{CH}$	-0.280	$\alpha_{11c,11c}$	89.089	81.333
	$\alpha_{1\kappa,1\kappa}^{CC}$	7.953	$\alpha_{11s,11s}$	89.089	81.333
	RMSD	0.025			
	$\Delta\varepsilon$	2.737			
Formate	$\alpha_{00,00}^{CH}$	-2.201	$\alpha_{10,10}$	38.900	36.551
	$\alpha_{00,00}^{CO}$	-2.2104	$\alpha_{11c,11c}$	24.732	24.684
	$\alpha_{1\kappa,1\kappa}^{CC}$	10.459	$\alpha_{11s,11s}$	44.321	42.080
	$\alpha_{1\kappa,1\kappa}^{OO}$	7.134			
	RMSD	0.265			
	$\Delta\varepsilon$	2.700			
Formaldéhyde	$\alpha_{00,00}^{CH}$	-1.984	$\alpha_{10,10}$	27.934	22.419
	$\alpha_{00,00}^{CO}$	-2.070	$\alpha_{11c,11c}$	12.220	12.844
	$\alpha_{1\kappa,1\kappa}^{CC}$	8.133	$\alpha_{11s,11s}$	24.680	17.966
	$\alpha_{1\kappa,1\kappa}^{OO}$	4.087			
	RMSD	0.136			
	$\Delta\varepsilon$	14.228			

<sup>a</sup> L'écart quadratique moyen (RMSD) entre l'énergie d'induction sur une grille de points et le modèle de polarisabilités distribuées en  $10^{-3}$  u.a.. L'erreur moyenne correspondante,  $\Delta\varepsilon$ , est donnée en pourcentage.

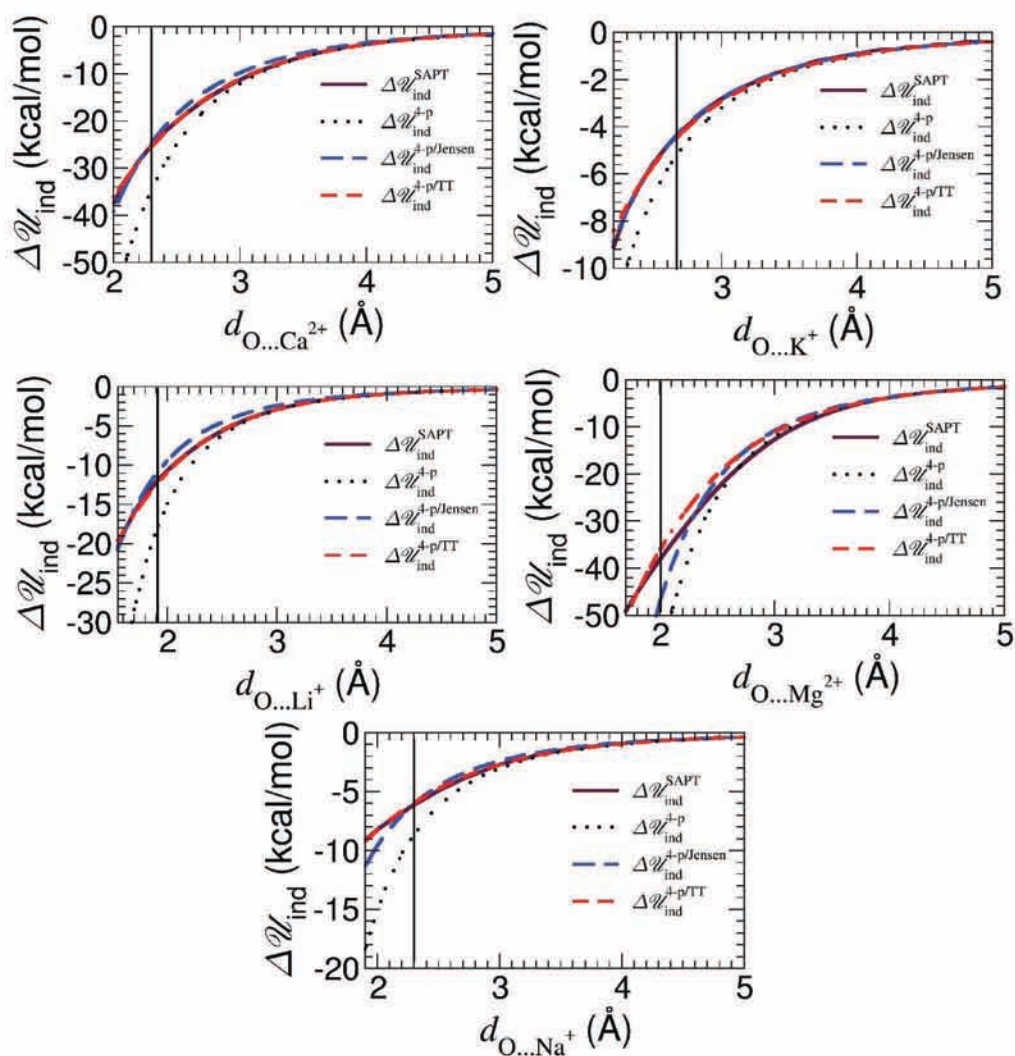


FIGURE 3.8 – Interactions du cation calcium ( $\text{Ca}^{2+}$ ), du cation potassium ( $\text{K}^+$ ), du cation lithium ( $\text{Li}^+$ ), du cation magnésium ( $\text{Mg}^{2+}$ ) et du cation sodium ( $\text{Na}^+$ ) avec la molécule d'eau le long de l'axe  $\text{C}_2$  de celle-ci. La référence (SAPT) est la courbe en trait plein. L'énergie d'induction du modèle non atténué est en pointillés. Le modèle de polarisabilités distribuées atténué par une fonction de Jensen et de Tang et Toennies sont en ligne discontinue et en ligne discontinue pointillée, respectivement. La barre verticale correspond au minimum de l'énergie d'interaction.

TABLE 3.5 – Polarisabilités dipolaires isotropes des ions calculées au niveau de théorie MP2(Full)/Sadlej.

	K <sup>+</sup>	Br <sup>-</sup>	Cl <sup>-</sup>	F <sup>-</sup>	I <sup>-</sup>
$\alpha_{1\kappa,1\kappa}$ (bohr <sup>3</sup> )	5.174	39.074	30.316	10.513	60.333
$\alpha_{1\kappa,1\kappa}^{exp.}$ (bohr <sup>3</sup> ) [10]	5.736	35.226	26.588	9.313	52.704
$\alpha_{1\kappa,1\kappa}^{exp.}$ (bohr <sup>3</sup> ) [323]	5.473	39.410	29.760	10.527	60.128

TABLE 3.6 – Valeurs du paramètre d’atténuation de Jensen *et al.* (a) pour chaque type d’atomes (une règle de combinaison de Lorentz-Berthelot est utilisée dans l’équation 2.21).

	H <sub>2</sub> O	H <sub>2</sub> OX	H <sub>2</sub> O	Ca <sup>2+</sup>	K <sup>+</sup>	Li <sup>+</sup>	Mg <sup>2+</sup>	Na <sup>+</sup>
$r_{vdW}$ (Å)	4.103	4.103	0.358	0.105	0.190	0.095	0.200	0.085

interaction n’utilisant que des polarisabilités distribuées qui ne prennent donc pas en compte l’échange-induction. En atténuant l’énergie avec les fonctions de Jensen *et al.* ou de Tang et Toennies le potentiel quantique de référence est reproduit de manière très fidèle (voir tableaux 3.6 et 3.7 pour les paramètres d’atténuation).

La fonction de Jensen a été, dans un premier temps [162, 163], considérée car elle permet de définir une interaction entre des types d’atomes. De plus, ses paramètres pourraient ainsi être transférables. Dans un second temps, nous avons montré que la fonction de Jensen reproduisait de manière moins fidèle les énergies d’interaction d’induction par rapport à la fonction de Tang et Toennies [164].

TABLE 3.7 – Valeurs du paramètre d’atténuation (b) de la fonction de Tang et Toennies. Chaque valeur correspond à une interaction de paire.

	Ca <sup>2+</sup>	K <sup>+</sup>	Li <sup>+</sup>	Mg <sup>2+</sup>	Na <sup>+</sup>	Br <sup>-</sup>	Cl <sup>-</sup>	F <sup>-</sup>	I <sup>-</sup>	H <sub>2</sub> O
H <sub>2</sub> O	2.10	2.15	2.38	2.10	2.00	5.00	5.00	4.50	5.00	5.00
Benzène	1.85	1.75	2.10	1.85	1.70					

### Molécule d'eau-anion et dimère d'eau

Les courbes quantiques et des énergies d'induction obtenues par le modèle d'induction sur les figures 3.9 sont différentes du cas précédent. Dans ce cas, l'énergie d'induction non-atténuée du modèle empirique sous-estime l'énergie d'induction pour les systèmes incluant un dimère d'eau et de l'eau en interaction avec des anions. L'écart entre l'énergie quantique et l'énergie du modèle peut faire penser au cas du cation potassium pour lequel il a fallu ajouter une polarisabilité dipolaire isotrope au modèle. Or, dans ce cas-là, une forte polarisabilité dipolaire isotrope est évaluée pour chaque anion (voir tableau 3.5). Cette différence entre la référence et le modèle peut provenir du modèle en lui-même, qui ne prend pas en compte assez d'effets, ou de la référence quantique qui inclut des effets d'échange-induction non considérés pour l'instant. Nous détaillerons les contributions à l'énergie d'induction ainsi que le modèle pour expliquer cette différence.

Regardons les contributions au potentiel d'interaction de référence SAPT :  $\Delta\mathcal{U}_{ind}^{MM} = \Delta\mathcal{U}_{ind}^{SAPT} + \Delta\mathcal{U}_{exc-ind}^{SAPT} + \delta Hf$ . Les contributions de cette somme sont le terme d'induction pure,  $\Delta\mathcal{U}_{ind}^{SAPT}$ , le terme d'échange-induction pure,  $\Delta\mathcal{U}_{exc-ind}^{SAPT}$ , et les ordres supérieurs, qui sont inclus dans  $\delta Hf$ . Le signe de  $\delta Hf$  dans le cas des interactions avec les anions n'est pas le même que dans les cas précédents (voir tableau 3.8). Il est négatif donc il contribue à stabiliser l'énergie d'induction. De plus, la contribution des ordres plus élevés,  $\delta Hf$ , est plus grande dans ces systèmes (jusque 50 % de l'énergie d'induction totale). Comme il a déjà été remarqué dans l'article de Dehez *et al.* [163], une comparaison directe de l'énergie d'induction pure (d'échange-induction) n'est pas possible avec l'énergie d'induction d'un modèle utilisant des polarisabilités distribuées (d'un potentiel d'atténuation). Il n'est donc pas possible de comparer le modèle en décomposant terme à terme. Seule la somme des termes est pertinente. De plus, dans la partie électrostatique, j'ai parlé de l'importance des effets de pénétration à courte distance dans ces systèmes. Pour prendre les effets de pénétration en compte, Torheyden et Jansen [324] ont documenté un effet non multipolaire à l'énergie d'induction qui ne peut être décomposé dans le développement actuel SAPT. Ce qui fait que la connaissance précise, pour l'instant, des effets d'induction et d'échange-induction ne sont pas connus. Dans le cas des systèmes où les effets non-multipolaires sont importants, il se peut donc que la référence SAPT ne puisse pas être comparée directement avec la référence quantique puisque la fonction d'échange-induction est déterminée à partir de la référence quantique.

Concernant le modèle, celui-ci reproduit de manière très satisfaisante le potentiel

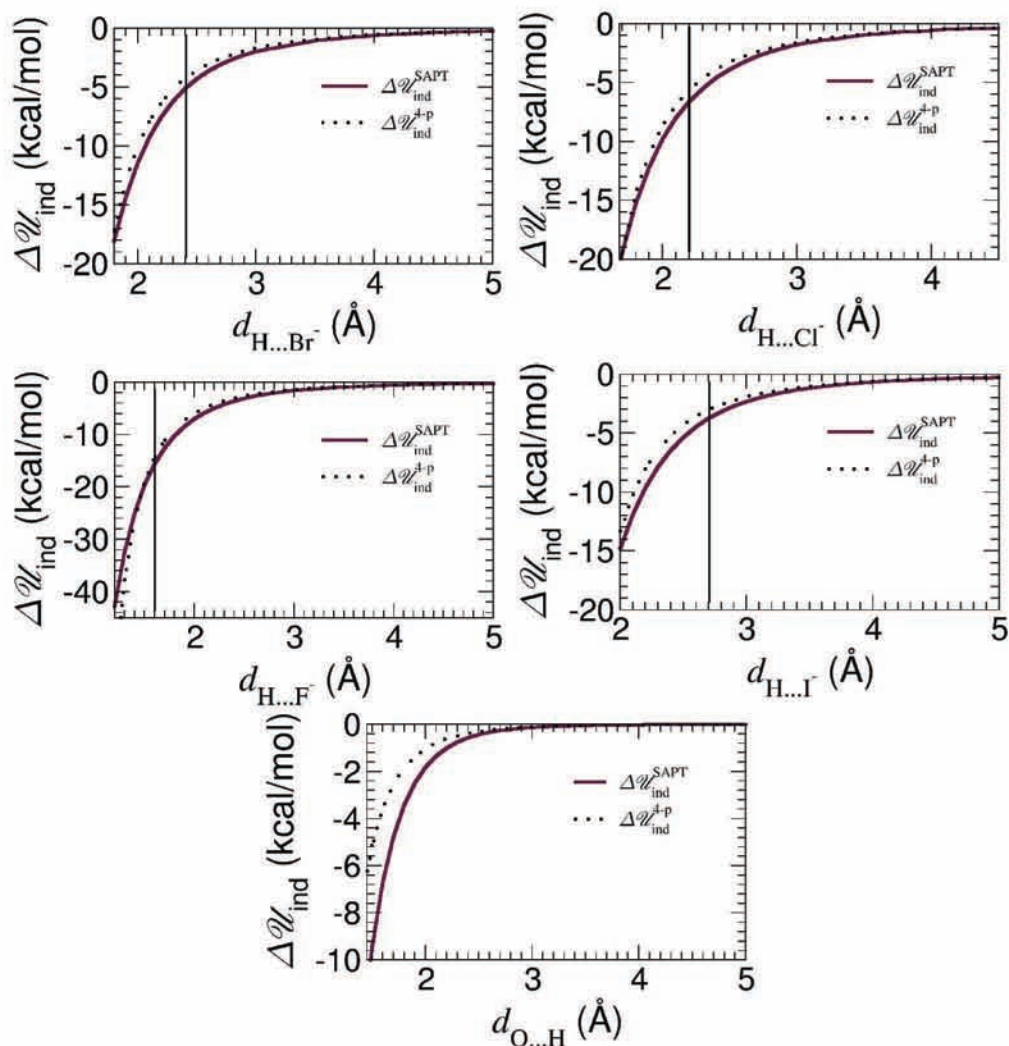


FIGURE 3.9 – Interactions de l’anion bromure ( $\text{Br}^-$ ), de l’anion chlorure ( $\text{Cl}^-$ ), de l’anion fluorure ( $\text{F}^-$ ), de l’anion iodure ( $\text{I}^-$ ) et de l’eau avec la molécule d’eau le long d’une liaison O–H de celle-ci. La référence (SAPT) est la courbe en trait plein. L’énergie d’induction du modèle non atténuée est en pointillés. La barre verticale correspond au minimum de l’énergie d’interaction.

TABLE 3.8 – Contribution des termes d’ordre supérieur à l’énergie d’induction au minimum de l’énergie d’interaction.

$\delta\text{HF}$ (kcal/mol)	$\text{Li}^+$	$\text{Na}^+$	$\text{K}^+$	$\text{Mg}^{2+}$	$\text{Ca}^{2+}$	$\text{F}^-$	$\text{Cl}^-$	$\text{Br}^-$	$\text{I}^-$	$\text{H}_2\text{O}$
$\text{H}_2\text{O}$	1.24	0.76	-0.70	-2.05	3.76	-4.14	-2.10	-1.67	-1.29	-0.46



d'induction quantique étant donné les faibles pourcentages d'erreur du modèle par rapport à la référence (voir molécule d'eau dans le tableau 3.4). De ce fait, le modèle d'induction proposé ici est cohérent.

Des essais afin de prendre en compte la pénétration directement dans le calcul de l'énergie d'induction ont été menés. Pour cela, l'équation 3.11 a été réécrite de la manière suivante :

$$\begin{aligned}
 U_{ind}^{MM} &= \frac{1}{2} \sum_A \sum_{B \neq A} \Delta Q_t^a T_{tu}^{ab} Q_u^{*b} \\
 \Delta Q_{t'}^{*a'} &= - \sum_{B \neq A} \alpha_{t't}^{a'a} T_{tu}^{ab} (Q_u^{*b} + \Delta Q_u^b)
 \end{aligned}
 \tag{3.16}$$

où  $Q_u^{*b}$  est la charge portée par le site polarisant modifiée pour prendre en compte les effets de pénétration comme décrit dans la fonction de Cisneros *et al.* (équation 3.7). Les paramètres déterminés dans le cas de l'électrostatique ont ainsi été utilisés pour modifier cette charge à courte distance. Les énergies d'induction ainsi obtenues n'ont été que très marginalement modifiées et il est apparu qu'en ce qui concerne notre modèle, le traitement explicite des effets de pénétration n'était pas essentiel.

Une approche alternative, mais qui n'a pas été satisfaisante dans les cas présents, a été étudiée et insérée dans l'annexe B. Cette approche permet de déterminer a priori la valeur des paramètres des fonctions de Tang et Toennies. Mais, du fait de la non nécessité d'ajouter une fonction d'atténuation dans le cas des interactions de la molécule d'eau avec des anions, cette approche a été abandonnée car elle prédisait un fort coefficient d'atténuation. Puisqu'un modèle sans potentiel d'échange-induction peut reproduire l'énergie d'induction obtenue en mécanique quantique, j'ai préféré conserver l'approche actuelle qui est plus cohérente avec la référence.

### Molécule de benzène-cations

Pour dériver le modèle d'induction du benzène, une grille de 3192 points est considérée. La faible taille de cette grille vient des propriétés de symétrie de la molécule de benzène ( $D_{6h}$ ). La grille est ainsi découpée en quarts de plan pour lesquels les énergies sont identiques<sup>3</sup>. Cette manipulation permet d'avoir une grille complète de 12768 points pour

<sup>3</sup>. Dans l'absolu, une division par 6 est possible mais le nombre de points évalués ici est facilement accessible.

un effort en terme de calculs quantiques divisé par quatre.

Les énergies d'interaction d'induction dans le cas du benzène avec des cations (voir figures 3.10) est comparable au cas présenté pour l'interaction avec la molécule d'eau, c'est-à-dire que l'énergie d'induction du modèle de polarisabilités distribuées sur-estime fortement l'énergie d'induction évaluée en mécanique quantique. Pour décrire les effets d'échange induction, il est alors nécessaire d'utiliser une fonction d'atténuation.

Il a été montré dans la partie électrostatique que de forts effets de pénétration interviennent pour cette interaction. Or, pour l'énergie d'induction, ces effets ne semblent pas influencer l'énergie d'induction comme vu dans l'interaction entre les anions et la molécule d'eau. Dans le cas de l'interaction avec le cation magnésium, la courbe est un peu différente. Divers articles [121, 267, 271, 325–327] reportent des études de systèmes en interaction avec le cation magnésium et montrent notamment que les effets de transfert de charges peuvent être importants. Ceci peut expliquer la différence dans la courbe de l'énergie d'induction du cation magnésium à la figure 3.10 étant donné que le transfert de charges n'est pas pris explicitement en compte dans le modèle mais est inclut au niveau du terme de polarisation alors que le terme de SAPT contient aussi les effets de transfert de charge<sup>4</sup>. Dans le cas de l'interaction avec le magnésium, il serait peut-être intéressant de vérifier la validité du modèle a posteriori. Malgré cet écart au niveau du magnésium, une fois le paramètre d'atténuation bien ajusté, la référence quantique est très bien reproduite par le modèle.

### Formaldéhyde, formate et eau – calcium

L'étude avait été menée avant l'utilisation de la fonction d'atténuation de Tang et Toennies mais pour ces systèmes [163], le modèle reproduit fidèlement l'énergie d'induction de référence (voir figures 3.7 et 3.8). Les phénomènes d'induction pour ces systèmes sont donc simplement décrits par le modèle de polarisabilités dipolaires isotropes et de flux de charges le long de l'axe décrit dans la section 3.2.3.

---

4. Pour obtenir la contribution de transfert de charge Stone et Misquitta [172] ont décrit une méthode.

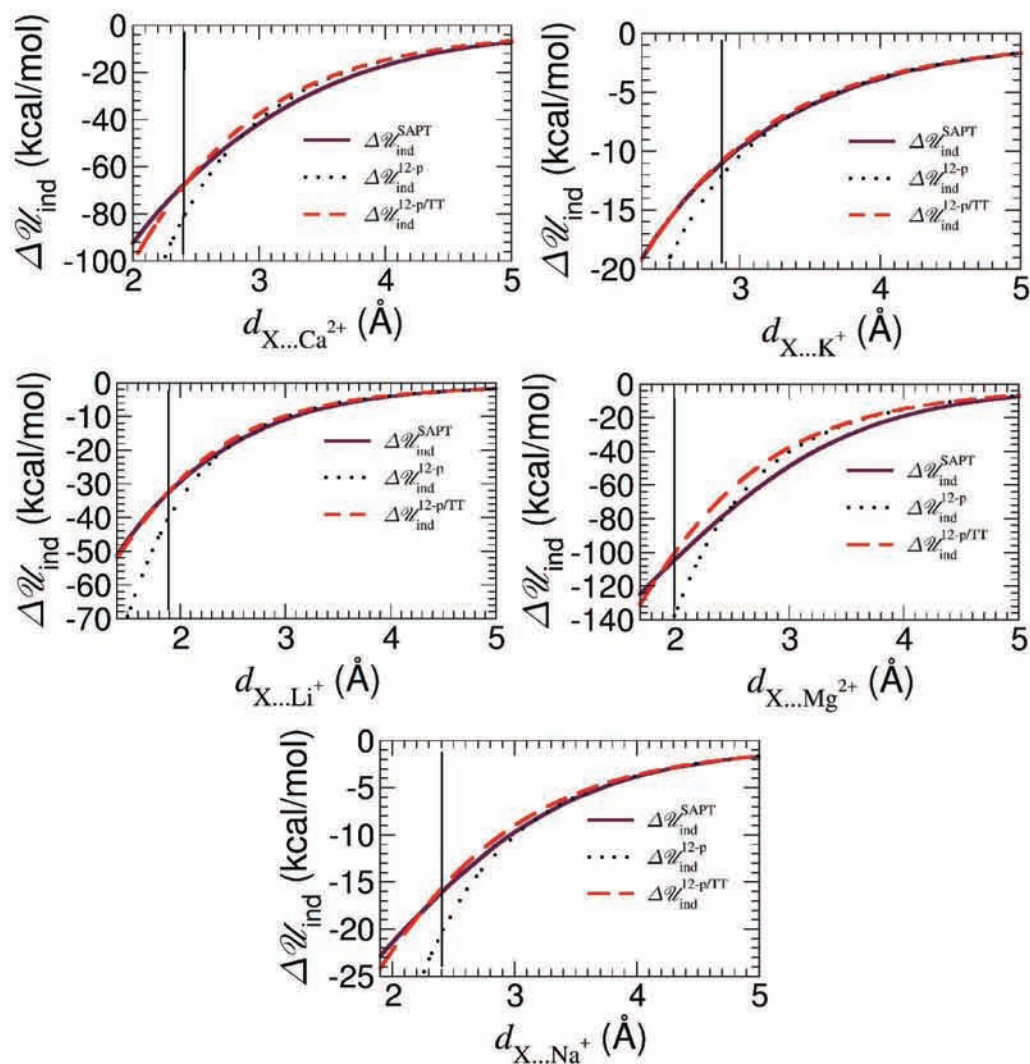


FIGURE 3.10 – Interactions du cation calcium ( $\text{Ca}^{2+}$ ), du cation potassium ( $\text{K}^+$ ), du cation lithium ( $\text{Li}^+$ ), du cation magnésium ( $\text{Mg}^{2+}$ ) et du cation sodium ( $\text{Na}^+$ ) avec la molécule benzène le long de l'axe  $\text{C}_6$  de celle-ci. La référence (SAPT) est la courbe en trait plein. L'énergie d'induction du modèle non atténuée est en pointillés. Le modèle de polarisabilité distribuées atténué par une fonction de Jensen et de Tang et Toennies sont en ligne discontinue et en ligne discontinue pointillée, respectivement. La barre verticale correspond au minimum de l'énergie d'interaction.

### 3.4 van der Waals

Un potentiel de van der Waals est une fonction mathématique qui permet de reproduire les effets d'attraction et de répulsion entre molécules. Dans le cas de deux atomes  $A$  et  $B$ , la forme du potentiel est une courbe qui est répulsive à courte distance, a un minimum énergétique de profondeur  $\varepsilon_{AB}$  à une distance caractéristique,  $R_{ij}^*$ , et tend vers zéro à longue distance. Différentes expressions pour le potentiel de van der Waals sont possibles et, ici, j'ai étudié en particulier trois formulations.

#### 3.4.1 Différentes expressions du potentiel de van der Waals

Le potentiel le plus employé pour des simulations de systèmes biologiques (CHARMM, AMBER, GROMACS...) est celui de Lennard-Jones. L'expression du potentiel est, par exemple celui du champ de force CHARMM décrit dans la section 1 à l'équation 1.4 :

$$\mathcal{U}_{vdW}^{LJ} = \varepsilon_{ij} \left[ \left( \frac{R_{ij}^*}{r_{ij}} \right)^{12} - 2 \left( \frac{R_{ij}^*}{r_{ij}} \right)^6 \right] \quad (3.17)$$

Une autre forme du potentiel de van der Waals a été proposée par Halgren [328] (voir équation 3.18) et est utilisée dans le champ de force polarisable AMOEBA [48, 85, 143, 326, 329–331].

$$\begin{aligned} \mathcal{U}_{vdW}^{Halgren} &= \sum_{i < j} \varepsilon_{ij} \left( \frac{1.07 R_{ij}^*}{r_{ij} + 0.07 R_{ij}^*} \right)^7 \left( \frac{1.12 R_{ij}^{*7}}{r_{ij}^7 + 0.12 R_{ij}^{*7}} - 2 \right) \\ \varepsilon_{ij} &= \frac{2\varepsilon_{ii}\varepsilon_{jj}}{\varepsilon_{ii} + \varepsilon_{jj}} \\ R_{ij}^* &= \frac{R_{ii}^{*3} + R_{jj}^{*3}}{R_{ii}^{*2} + R_{jj}^{*2}} \end{aligned} \quad (3.18)$$

Dans cette expression,  $r$  est la distance entre les sites,  $\varepsilon_{ij}$  est la profondeur du puits et  $R^*$  est le rayon de van der Waals pour la paire considérée. Les valeurs 1.07 et 1.12 ont été déterminées à partir d'interactions calculées quantiquement pour des gaz rares (voir paragraphe 2.3 de l'article [328]). Pour obtenir les puissances dans le potentiel de Halgren, des interactions entre des gaz rares ont été considérées. En faisant varier la partie répulsive de 12 à 15 et la partie attractive de 6 à 8, la meilleure reproduction du potentiel a été déterminée pour le jeu 14 et 7, respectivement, pour la partie répulsive et attractive.

Le troisième et dernier potentiel considéré ici est le potentiel proposé par Buckingham *et al.* [332]. Ce potentiel est écrit sous la forme d'une exponentielle pour le terme répulsif et d'une fonction en  $1/r^6$  pour le terme attractif :

$$\begin{aligned} \mathcal{U}_{vdW}^{Buckingham} &= \sum_{i < j} \varepsilon_{ij} \exp(-A_{1ij} * r_{rij}) - \left( \frac{A_{2ij}}{r_{ij}} \right)^6 \\ \varepsilon_{ij} &= \sqrt{\varepsilon_{ii} \varepsilon_{jj}} \\ A_{1ij} &= \frac{A_{1ii} + A_{1jj}}{2} \\ A_{2ij} &= \frac{A_{2ii} + A_{2jj}}{2} \end{aligned} \tag{3.19}$$

$\varepsilon_{ij}$ ,  $A_{1ij}$ ,  $A_{2ij}$  sont les paramètres de forme du potentiel de Buckingham. Pour cette fonction, à ma connaissance, il n'existe pas de règles de combinaison explicitement développées. J'ai donc considéré une formulation inspirée des règles de Lorentz-Berthelot utilisées dans le potentiel 6–12 de Lennard-Jones (voir équation 3.19). Par analogie avec le potentiel de Lennard-Jones :  $\varepsilon_{ii}$  est le puits de potentiel de l'atome ou du site considéré et  $A_{2ii}$  est le rayon de van der Waals. Le terme  $A_{1ii}$  est la pente de la fonction de Buckingham qui traduit les effets répulsifs de l'interaction [332].

Pour les potentiels de Lennard-Jones et de Halgren, la partie répulsive est le carré de la partie attractive. Pour le potentiel de Buckingham, apparaît donc un surcoût calculatoire dû à la partie en exponentielle. Depuis plusieurs années, les améliorations des algorithmes mathématiques font que le coût de calcul de ces exponentielles devient de plus en plus faible. En particulier, Andrews *et al.* [333] ont proposé des méthodes intégrées dans les compilateurs comme `ifort`<sup>5</sup> de INTEL qui diminue le surcoût engendré. De plus, l'utilisation d'une exponentielle dans le potentiel de répulsion a un meilleur sens physique<sup>6</sup> que des fonctions en  $1/r^n$ ,  $n \geq 12$ . Ce qui donne au potentiel de Buckingham une validité supérieure pour décrire la répulsion malgré un coût en temps de calcul supérieur.

### 3.4.2 Détermination des paramètres de van der Waals

La détermination des paramètres de van der Waals avec une méthode précise doit être définie dans l'esprit des précédentes méthodes avec OPEP. Avant de commencer, il

5. <http://software.intel.com/en-us/intel-compilers>

6. La répulsion calculée quantiquement est une somme d'exponentielles [14].

est nécessaire de choisir la contribution exacte à reproduire ainsi que la fonction de van der Waals à considérer.

Initialement [162, 163], la détermination des paramètres de van der Waals a été effectuée en utilisant un algorithme de minimisation non-linéaire de type Levenberg-Marquardt [334, 335]. Cet algorithme est robuste et permet d’avoir les paramètres d’une équation non-linéaire à partir d’une série de valeurs. La qualité de l’ajustement s’estime directement à partir du coefficient de corrélation ( $r^2$ ) entre la courbe initiale et la courbe générée. Les paramètres de van der Waals sont déterminés, dans un premier temps [162, 163], pour des valeurs de distance non fixées par des règles précises et, en particulier, à des courtes distances d’interaction. La conclusion principale de ces articles est que les paramètres de Lennard-Jones des champs de forces de type CHARMM ne peuvent pas être transférés directement pour reproduire les énergies et qu’un ajustement des paramètres est nécessaire pour tous les types d’atomes. De plus, la forme du potentiel de Lennard-Jones 6–12 n’est pas à même de reproduire avec précision le potentiel quantique.

À partir de ces conclusions, plusieurs questions sont apparues :

- Sur quelle gamme de distances intermoléculaires le potentiel se doit-il d’être précis ?
- À partir de quelle grandeur les paramètres de van der Waals doivent-ils être ajustés ?
- Ce terme inclut-il d’autres contributions ?
- Quelle fonction serait la plus adaptée ?

Nous avons déjà répondu à la première question (voir section 3.2.3). La gamme de distances est définie telle que la distance minimum choisie correspond à une énergie d’environ +10 kcal/mol par rapport au minimum énergétique de l’interaction. Cette sélection de distances est très importante car, si trop de points sont considérés dans la partie répulsive, alors le potentiel de van der Waals exagérera la répulsion et décrira mal la partie attractive.

Les réponses aux deuxième et troisième questions sont brièvement évoquées dans la section 3.1. D’après le développement SAPT, le potentiel de van der Waals est [162, 163] :  $\Delta\mathcal{U}_{\text{vdW}}^{\text{SAPT}} = \Delta\mathcal{U}_{\text{disp}}^{\text{SAPT}} + \Delta\mathcal{U}_{\text{exc-disp}}^{\text{SAPT}}$ . Dans notre modèle, certaines contributions à courtes distances sont négligées dans les parties électrostatique et d’induction. Le potentiel de van der Waals doit alors compenser ces contributions dans la reproduction du potentiel total. Nous avons donc fait le choix [164] d’ajuster notre potentiel de van der Waals,

$\Delta\mathcal{U}_{vdW}^{MM}$ , sur la quantité suivante :  $\Delta\mathcal{U}_{vdW}^{MM} = \Delta\mathcal{U}_{tot}^{SAPT} - \Delta\mathcal{U}_{ele}^{MM} - \Delta\mathcal{U}_{ind}^{TT}$ .

Pour répondre à la dernière question, il est nécessaire de comparer la capacité des trois fonctions considérées à reproduire globalement les énergies d'interaction.

### Interaction molécule d'eau–cation, anion ou molécule d'eau

Pour ces interactions, les courbes obtenues pour les trois potentiels de van der Waals considérés sont fournies en figures 3.11 et 3.12. Les paramètres pour chaque fonction sont référencés dans le tableau 3.9.

Au travers de ces figures, il apparaît clairement qu'un potentiel de Lennard-Jones 6–12 ne peut pas reproduire simultanément la partie attractive et la partie répulsive et que les potentiels de Halgren et de Buckingham sont en très bon accord avec  $\Delta\mathcal{U}_{vdW}^{MM}$ . Il est aussi à noter que la convergence de l'algorithme de Levenberg-Marquardt dans le cas du potentiel de Lennard-Jones dépend fortement des conditions initiales.

### Interaction benzène–cations

Dans le cas du benzène, en première approximation, la détermination des paramètres est effectuée entre le centre du benzène et le cation. De part la symétrie de la molécule de benzène, les paramètres obtenus au centre du benzène peuvent être redistribués aisément sur les atomes de carbone.

Dans le tableau 3.10, les paramètres de van der Waals pour les divers potentiels sont disponibles sous la forme d'interactions de paire. Les courbes obtenues sont visibles sur les figures 3.19.

L'observation de ces courbes nous permet de dire que le potentiel de van der Waals qui utilise une fonction de Buckingham reproduit mieux le potentiel que nous voulons ajuster (non représenté) par rapport au potentiel de Halgren. Ainsi, le choix de la fonction de Buckingham sera effectué par rapport au potentiel de Halgren.

## 3.5 Énergie totale

La reproduction de l'énergie totale est un des objectifs les plus importants dans cette étude. C'est pour cela que tout le cheminement a amené à un moyen de reproduire au

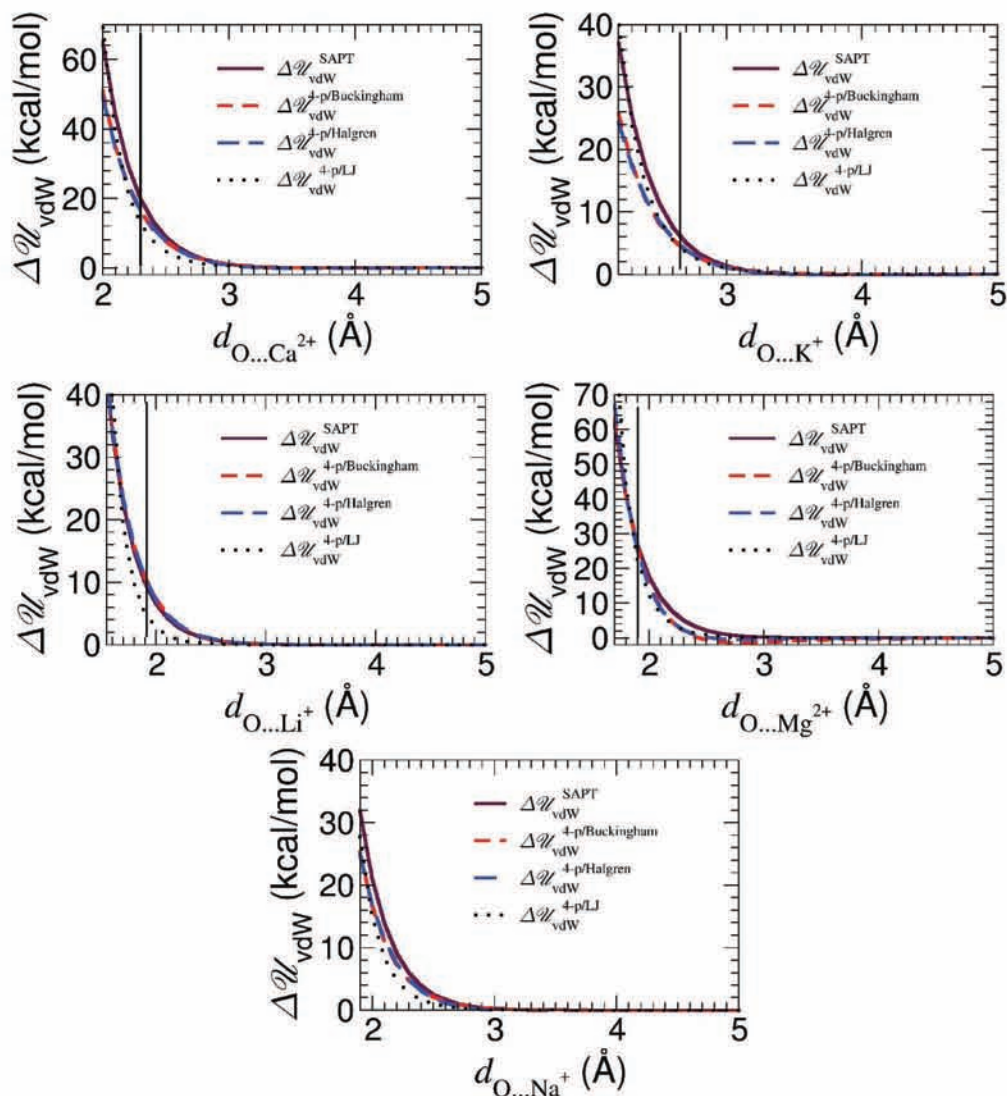


FIGURE 3.11 – Interactions du cation calcium ( $\text{Ca}^{2+}$ ), du cation potassium ( $\text{K}^+$ ), du cation lithium ( $\text{Li}^+$ ), du cation magnésium ( $\text{Mg}^{2+}$ ) et du cation sodium ( $\text{Na}^+$ ) avec la molécule de benzène le long de l'axe  $\text{C}_6$  de celle-ci. La référence (SAPT) est la courbe en trait plein. Le potentiel de Buckingham et de Halgren sont en trait court discontinu et long discontinu, respectivement. Le potentiel de Lennard-Jones est en pointillés. La barre verticale correspond au minimum de l'énergie d'interaction.



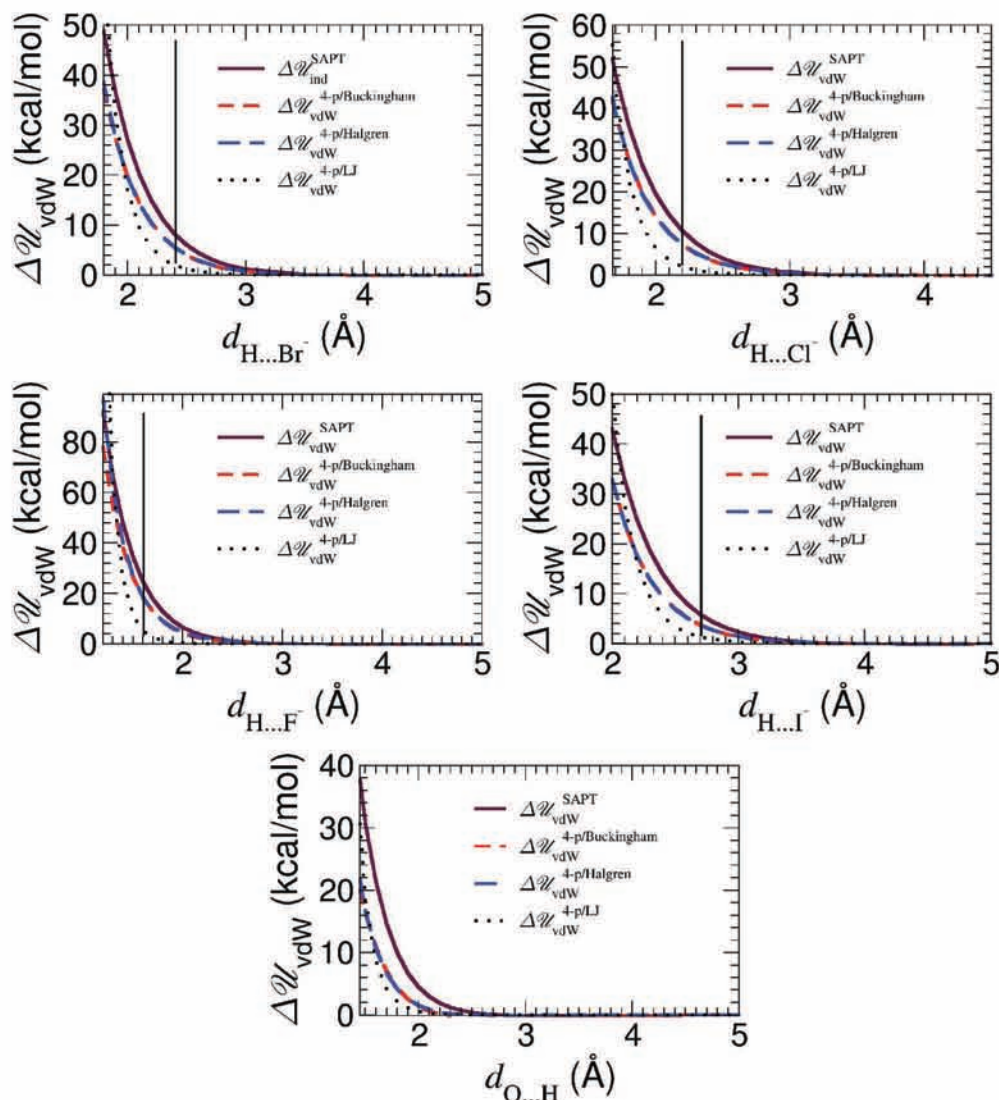


FIGURE 3.12 – Interactions de l’anion bromure ( $\text{Br}^-$ ), de l’anion chlorure ( $\text{Cl}^-$ ), de l’anion fluorure ( $\text{F}^-$ ) et de l’anion iodure ( $\text{I}^-$ ) avec la molécule de benzène le long de l’axe  $\text{C}_6$  de celle-ci. La référence (SAPT) est la courbe en trait plein. Le potentiel de Buckingham et de Halgren sont en trait court discontinu et long discontinu, respectivement. Le potentiel de Lennard-Jones est en pointillés. La barre verticale correspond au minimum de l’énergie d’interaction.

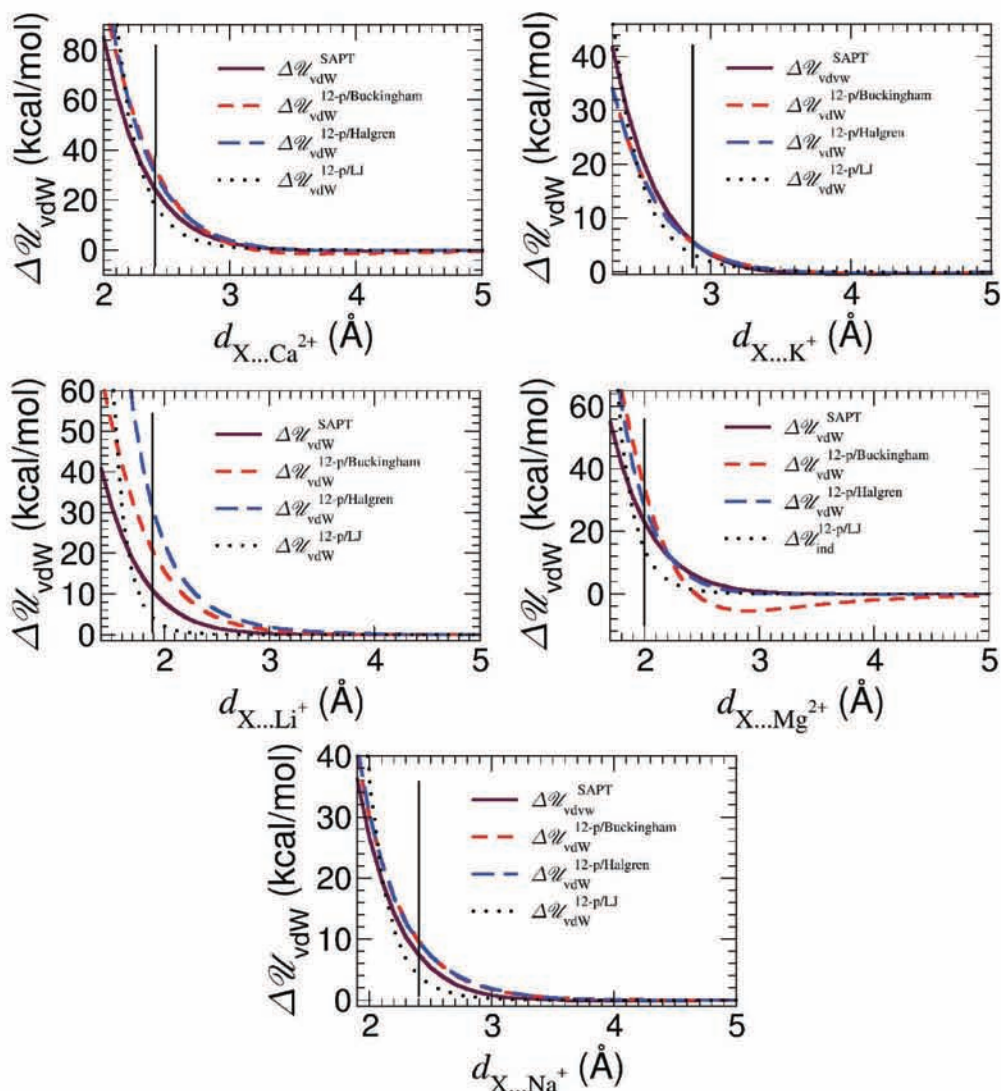


FIGURE 3.13 – Interactions du cation calcium ( $\text{Ca}^{2+}$ ), du cation potassium ( $\text{K}^+$ ), du cation lithium ( $\text{Li}^+$ ), du cation magnésium ( $\text{Mg}^{2+}$ ) et du cation sodium ( $\text{Na}^+$ ) avec la molécule de benzène le long de l’axe  $\text{C}_6$  de celle-ci. La référence (SAPT) est la courbe en trait plein. Le potentiel de Buckingham et de Halgren sont en trait court discontinu et long discontinu, respectivement. Le potentiel de Lennard-Jones est en pointillés. La barre verticale correspond au minimum de l’énergie d’interaction.

mieux l'énergie totale tout en sachant que la simplicité du modèle proposé ne pouvait pas contenir toutes les contributions physiques.

Dans cet esprit, le modèle électrostatique est constitué de charges ponctuelles (monopôles) situées sur les atomes ou, pour certaines molécules (cas de l'eau), sur des sites fictifs afin de reproduire au mieux le potentiel électrostatique de référence. Pour le modèle d'induction, des polarisabilités de flux de charges (ordre zéro) et dipolaires isotropes (ordre un) sont utilisées. L'énergie d'induction du modèle de polarisabilités distribuées est, ou non, atténuée par une série de Tang et Toennies tronquée à l'ordre trois de façon à prendre en compte les effets d'échange-induction et d'éviter la « catastrophe de polarisation » à courte distance. Le reste de l'énergie d'interaction est décrit par un potentiel de van der Waals modélisé par une fonction de Buckingham. Ce potentiel est déterminé à partir de l'énergie d'interaction totale obtenue par la mécanique quantique à laquelle a été retranchée les contributions électrostatique et d'induction de notre modèle empirique. Une fois toute cette méthodologie appliquée, pour les systèmes étudiés, la reproduction de l'énergie totale est reportée sur les figures 3.14, 3.15 et 3.16. Dans ces images sont représentées l'énergie d'interaction totale ( $\Delta\mathcal{U}_{\text{tot}}^{\text{SAPT}}$ ), le potentiel empirique avec un potentiel de Lennard-Jones ( $\Delta\mathcal{U}_{\text{tot}}^{\text{LJ}}$ ) et un potentiel de Buckingham ( $\Delta\mathcal{U}_{\text{tot}}^{\text{Buckingham}}$ ).

Les énergies au niveau du minimum ainsi que les distances minimum du modèle et de la référence quantique pour toutes les interactions considérées sont dans les tableaux 3.11 et 3.12. La différence entre l'énergie minimum ou la distance minimum entre la référence et notre modèle est inférieure à 5 % dans tous les cas. Sur toute la gamme de distances, l'énergie totale quantique est également reproduite par le modèle empirique. L'étape suivante est d'étudier dans quelles mesures ces paramètres obtenus sur la base de dimères peuvent être transférés directement en simulation de phase condensée.

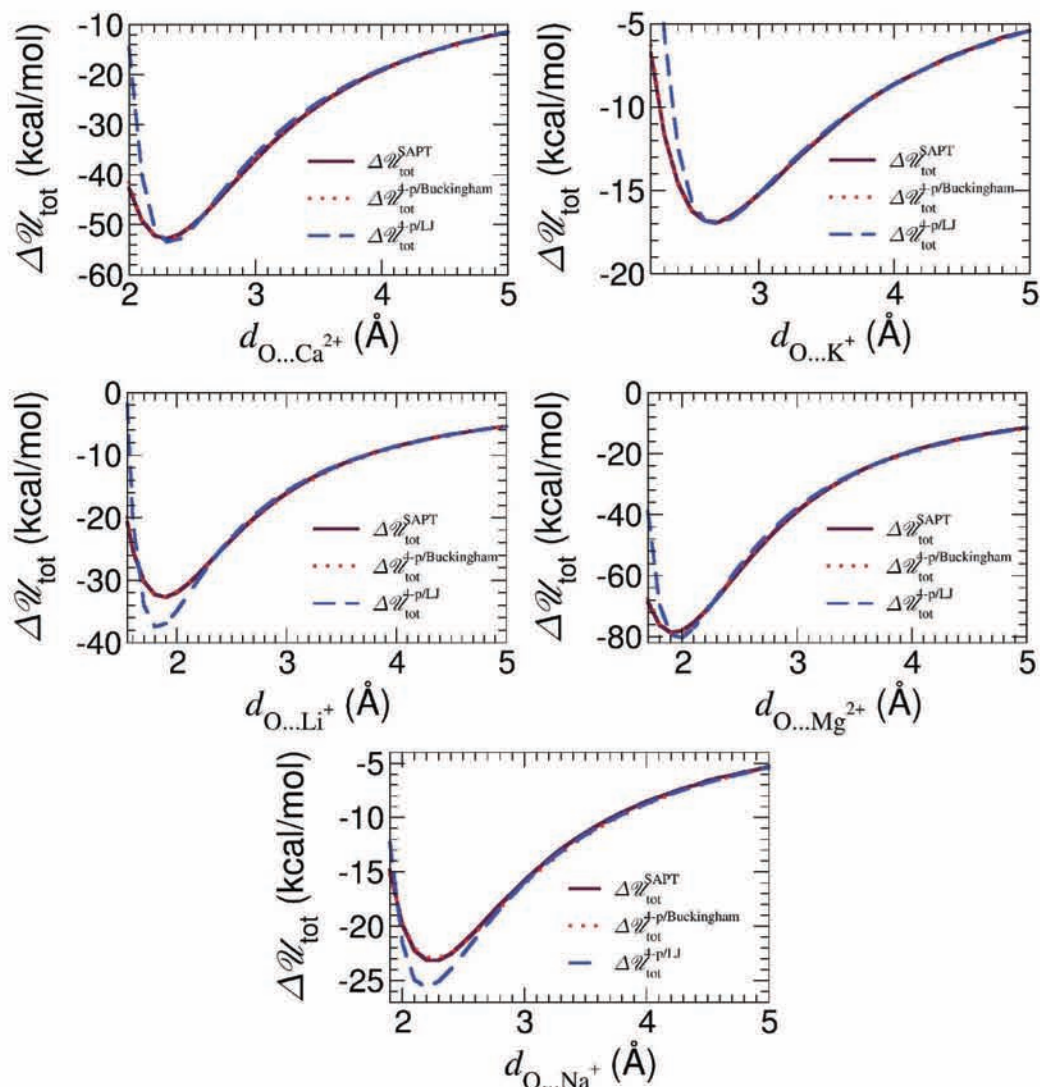


FIGURE 3.14 – Interactions du cation calcium ( $\text{Ca}^{2+}$ ), du cation potassium ( $\text{K}^+$ ), du cation lithium ( $\text{Li}^+$ ), du cation magnésium ( $\text{Mg}^{2+}$ ) et du cation sodium ( $\text{Na}^+$ ) avec la molécule de benzène le long de l'axe  $\text{C}_6$  de celle-ci. La référence (SAPT) est la courbe en trait plein. Le potentiel de Buckingham et de Halgren est en trait court discontinu et long discontinu, respectivement. Le potentiel de Lennard-Jones est en pointillé. La barre verticale correspond au minimum de l'énergie d'interaction.

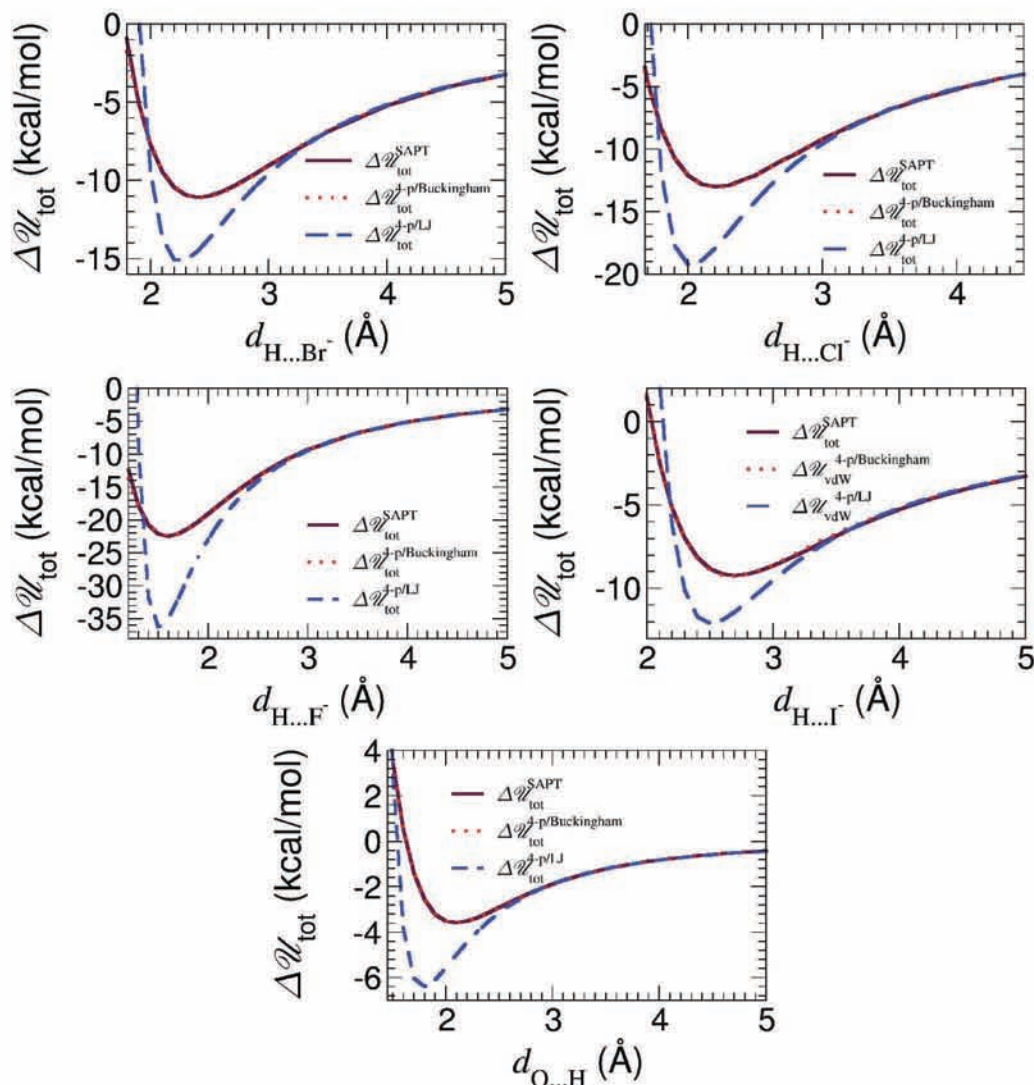


FIGURE 3.15 – Interactions de l’anion bromure ( $\text{Br}^-$ ), de l’anion chlorure ( $\text{Cl}^-$ ), de l’anion fluorure ( $\text{F}^-$ ) et de l’anion iode ( $\text{I}^-$ ) avec la molécule de benzène le long de l’axe  $\text{C}_6$  de celle-ci. La référence (SAPT) est la courbe en trait plein. Le potentiel de Buckingham et de Halgren sont en trait court discontinu et long discontinu, respectivement. Le potentiel de Lennard-Jones est en pointillés. La barre verticale correspond au minimum de l’énergie d’interaction.

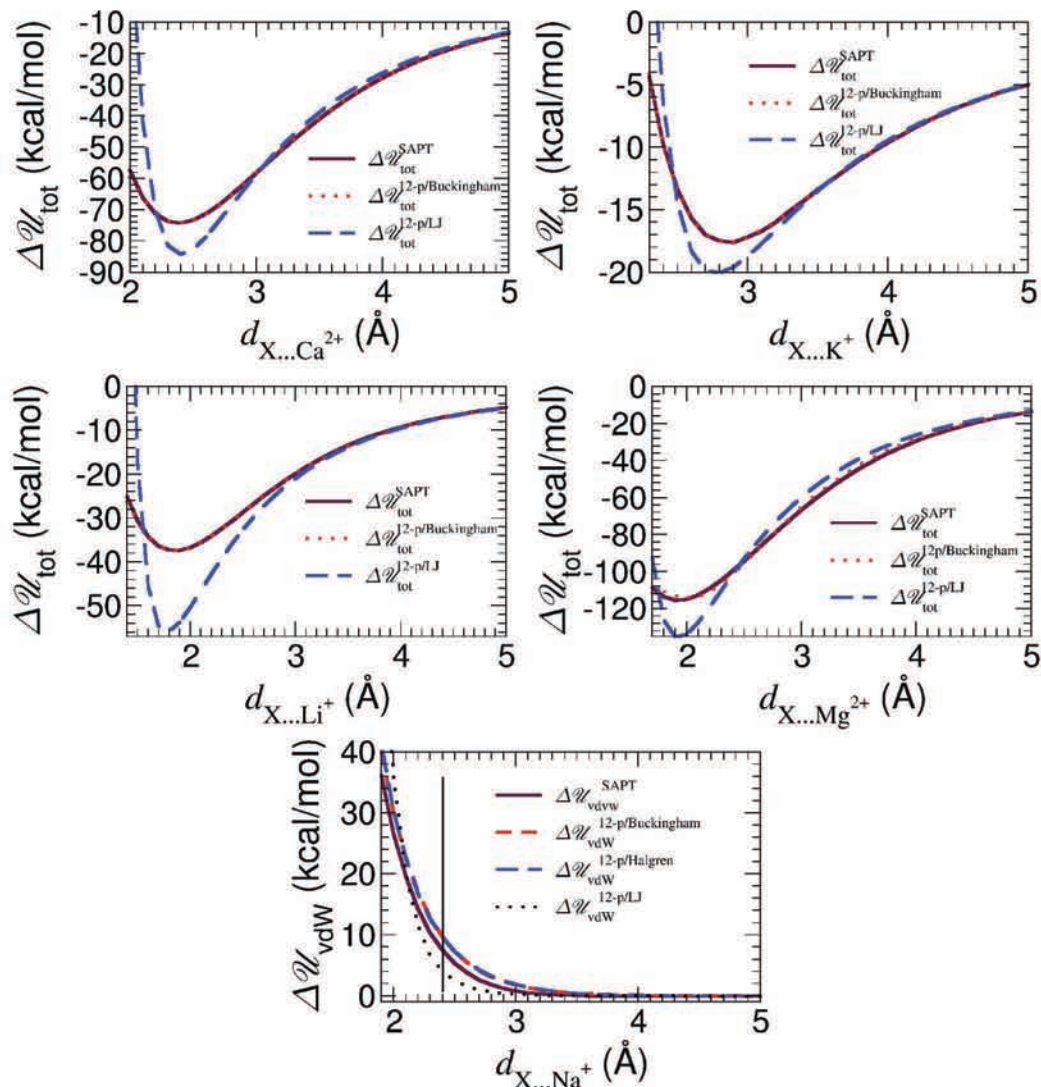


FIGURE 3.16 – Interactions du cation calcium ( $\text{Ca}^{2+}$ ), du cation potassium ( $\text{K}^{+}$ ), du cation lithium ( $\text{Li}^{+}$ ), du cation magnésium ( $\text{Mg}^{2+}$ ) et du cation sodium ( $\text{Na}^{+}$ ) avec la molécule de benzène le long de l'axe  $\text{C}_6$  de celle-ci. La référence (SAPT) est la courbe en trait plein. Le potentiel de Buckingham et de Halgren sont en trait court discontinu et long discontinu, respectivement. Le potentiel de Lennard-Jones est en pointillés. La barre verticale correspond au minimum de l'énergie d'interaction.

TABLE 3.9 – Paramètres de van der Waals pour des systèmes  $\text{H}_2\text{O}-\text{X}$ . Les paramètres sont obtenus en utilisant un modèle électrostatique de charges ponctuelles et un modèle d'induction de polarisabilités distribuées atténué par la fonction de Tang et Toennies. (1)  $\epsilon$  et  $A_n$  ( $n = 1, 2$ ) sont les paramètres de la fonction de Buckingham. (2)  $\epsilon$  et  $R^*$  sont les paramètres de la fonction de Halgren. (3)  $\epsilon$  et  $\sigma$  sont les paramètres de la fonction de Lennard-Jones.  $\epsilon$  est donné en kcal/mol. Tous les autres paramètres sont en Ångström.

		Li <sup>+</sup>	Na <sup>+</sup>	K <sup>+</sup>	Mg <sup>2+</sup>	Ca <sup>2+</sup>	F <sup>-</sup>	Cl <sup>-</sup>	Br <sup>-</sup>	I <sup>-</sup>	H <sub>2</sub> O
(1)	$\epsilon$	19149	66526	120694	94937	94131	5837	11843	12340	17204	14946
	$A_1$	3.944	4.147	3.842	3.900	3.738	3.598	3.352	3.106	3.132	3.949
	$A_2$	0.025	0.032	1.418	3.385	2.332	0.019	1.679	2.600	0.034	2.541
(2)	$\epsilon$	0.041	0.100	0.115	0.433	0.151	0.003	0.021	0.025	0.038	0.041
	$R^*$	3.412	3.390	3.880	2.848	3.700	5.132	4.113	4.205	4.284	2.864
(3)	$\epsilon$	468.3	94.27	11.58	1308	7.547	65.52	93.85	323.2	12.20	61468
	$\sigma$	1.302	1.715	2.395	1.355	2.405	1.301	1.600	1.566	2.248	0.763

TABLE 3.10 – Paramètres de van der Waals pour les systèmes benzène- $\text{X}^{n+}$ . Les paramètres sont obtenus en utilisant un modèle électrostatique de charges ponctuelles et un modèle d'induction de polarisabilités distribuées atténué par la fonction de Tang et Toennies. (1)  $\epsilon$  et  $A_n$  ( $n = 1, 2$ ) sont les paramètres de la fonction de Buckingham. (2)  $\epsilon$  et  $R^*$  sont les paramètres de la fonction de Halgren. (3)  $\epsilon$  et  $\sigma$  sont les paramètres de la fonction de Lennard-Jones.  $\epsilon$  est donné en kcal/mol. Tous les autres paramètres sont en ångström.

		Li <sup>+</sup>	Na <sup>+</sup>	K <sup>+</sup>	Mg <sup>2+</sup>	Ca <sup>2+</sup>
(1)	$\epsilon$	3556	8821	38649	103962	61104
	$A_1$	2.624	2.841	2.924	3.184	2.753
	$A_2$	2.419	0.067	3.502	4.583	4.661
(2)	$\epsilon$	0.002	0.016	0.104	0.138	0.157
	$R^*$	6.331	4.871	4.261	3.471	4.157
(3)	$\epsilon$	8124	539.9	6.103	33.21	9.573
	$\sigma$	2.624	2.841	2.924	1.860	2.753

$d_{A...B}$	Li <sup>+</sup>	Na <sup>+</sup>	K <sup>+</sup>	Mg <sup>2+</sup>	Ca <sup>2+</sup>	F <sup>-</sup>	Cl <sup>-</sup>	Br <sup>-</sup>	I <sup>-</sup>	H <sub>2</sub> O
H <sub>2</sub> O	1.90	2.22	2.70	1.90	2.30	1.60	2.20	2.40	2.70	2.10
	<b>1.90</b>	<b>2.25</b>	<b>2.66</b>	<b>1.90</b>	<b>2.30</b>	<b>1.60</b>	<b>2.20</b>	<b>2.40</b>	<b>2.70</b>	<b>2.10</b>
Benzene	1.90	2.41	2.85	2.02	2.40					
	<b>1.90</b>	<b>2.41</b>	<b>2.90</b>	<b>1.92</b>	<b>2.40</b>					

TABLE 3.11 – Minimum de distance pour l'interaction entre la molécule d'eau et des cations ou des anions ainsi qu'entre la molécule de benzène et des cations. La valeur quantique de référence est en gras et toutes les distances sont exprimées en Ångström.



$\Delta\mathcal{U}_{tot,min}$	Li <sup>+</sup>	Na <sup>+</sup>	K <sup>+</sup>	Mg <sup>2+</sup>	Ca <sup>2+</sup>	F <sup>-</sup>	Cl <sup>-</sup>	Br <sup>-</sup>	I <sup>-</sup>	H <sub>2</sub> O
H <sub>2</sub> O	-32.63	-22.91	-16.85	-78.80	-52.98	-22.38	-13.07	-11.14	-9.25	-3.60
	<b>-32.69</b>	<b>-23.12</b>	<b>-16.91</b>	<b>-79.85</b>	<b>-52.75</b>	<b>-22.46</b>	<b>-13.00</b>	<b>-11.08</b>	<b>-9.22</b>	<b>-3.60</b>
Benzene	-37.45	-23.41	-17.57	-114.17	-74.25					
	<b>-37.43</b>	<b>-23.70</b>	<b>-17.60</b>	<b>-116.41</b>	<b>-74.23</b>					

TABLE 3.12 – Minimum d'énergie potentielle pour l'interaction entre la molécule d'eau et des cations ou des anions ainsi qu'entre la molécule de benzène et des cations. La valeur quantique de référence est en gras et toutes les distances sont exprimées en Ångström.

## 3.6 Vers les simulations de dynamique moléculaire

L'objectif final de cette thèse est de pouvoir effectuer des dynamiques moléculaires de phase condensée, dans la mesure du possible, grâce à des paramètres issus de la mécanique quantique seule. Afin de présenter les résultats obtenus par notre modèle, il faut, tout d'abord, discuter de la stratégie qui permet de résoudre les moments induits en dynamique moléculaire. Puis, je détaillerai l'approche utilisée dans le logiciel TINKER pour effectuer les simulations. Finalement, nous verrons pourquoi nous devons optimiser les paramètres de van der Waals obtenus précédemment.

## 3.7 Traitement des moments induits en dynamique moléculaire

Afin de traiter les moments induits en dynamique moléculaire, il existe deux méthodes usuellement employées : le Lagrangien étendu et le traitement auto-cohérent.

### 3.7.1 Le Lagrangien étendu

Pour calculer les dipôles induits, il est possible d'étendre le Lagrangien du système en considérant ces dipôles comme des variables dynamiques supplémentaires. Le lagrangien est  $L = K - U$  avec  $U$  l'énergie potentielle électrostatique (incluant l'énergie de polarisation) et  $K$  l'énergie cinétique des particules et des dipôles considérés. En définissant la pseudo-masse  $m_\mu$  aux dipôles induits, le Lagrangien du système s'écrit alors :

$$L = \frac{1}{2} \sum_{i=1}^N \frac{p_i^2}{m_i} - (U - U_{pol}) + \frac{1}{2} \sum_{i=1}^N m_\mu \left( \frac{d\mu_i}{dt} \right)^2 - U_{pol} \quad (3.20)$$

Il est alors possible de réécrire les équations lagrangiennes de propagation des moments induits comme :

$$\frac{d}{dt} \frac{\partial L}{\partial \left( \frac{d\mu_i}{dt} \right)} = \frac{\partial L}{\partial \mu_i} \quad (3.21)$$

d'où :

$$m_\mu \frac{d^2 \mu_i}{dt^2} = -\frac{\mu_i}{\alpha_i} + \mathbf{E}_i \quad (3.22)$$

La valeur attribuée au paramètre  $m_\mu$  a une forte influence dans les calculs de type Lagrangien et sa détermination nécessite des calculs préliminaires. En effet,  $m_\mu$  doit être suffisamment petit pour que les moments induits soient dans un régime quasi-SCF, mais pas trop afin d'éviter de rendre les trajectoires instables (voir référence [280] par exemple).

### 3.7.2 Le traitement auto-cohérent

Nous avons déjà vu dans la partie spécifique aux effets de polarisation (chapitre 2) à l'équation 3.11 comment il était possible d'écrire l'énergie d'induction en considérant un traitement auto-cohérent de l'énergie de polarisation. Dans cette approche, le paramètre important est le nombre de pas nécessaires pour converger les moments induits. Il existe diverses approches pour optimiser cette convergence dont une qui est présentée en annexe C.

Le choix fait ici pour effectuer les simulations de dynamique moléculaire s'est porté sur TINKER<sup>7</sup> qui utilise un traitement auto-cohérent. Ce logiciel a l'avantage de posséder déjà un grand nombre de fonctionnalités et de types de tenseurs d'interactions nécessaires pour rapidement effectuer des simulations « benchmark » avec les paramètres obtenus via l'approche détaillée précédemment. En contrepartie, la vitesse de calcul s'en ressent mais permet tout de même de mener à bien des calculs de simulation sur des systèmes qui nous intéresseront ici pour éprouver le modèle.

De plus, l'intégration des équations du mouvement dans les ensembles thermodynamiques NPT ou NVT requiert l'usage d'un seul thermostat afin de maintenir la pression ou la température constante pour les simulations avec un traitement auto-cohérent des moments induits. Par contre, les approches Lagrangiennes, telles que Drude [336], nécessitent plusieurs thermostats afin d'éviter les problèmes dits « de points chauds ». Il faut aussi noter que dans le cas d'un traitement auto-cohérent des moments induits, il est possible d'utiliser un pas de temps supérieur à celui nécessaire dans le cas du Lagrangien étendu du fait de la masse fictive. Typiquement, un pas de 0.5-1 fs est requis [73, 290, 336] pour intégrer les équations du mouvement dans le cas du Lagrangien étendu alors que dans le cas d'un traitement auto-cohérent, un pas de 1-2 fs est utilisable.

---

7. <http://dasher.wustl.edu/tinker>

### 3.8 Code de dynamique moléculaire : Tinker

Avant de pouvoir utiliser TINKER, il a fallu modifier certaines routines et ajouter des options au programme :

— Pour le potentiel électrostatique, les molécules seront considérées comme rigides. Dans le cas de la molécule d'eau, le site fictif ne posera donc aucun problème. Étant donné que nous ne traitons que des monopôles pour les sites atomiques ou les extra-sites, le programme TINKER ne nécessite aucune modification pour ce potentiel.

— Le potentiel de polarisation est un peu plus complexe. TINKER, dans sa version standard, ne gère que les polarisabilités dipolaires isotropes c'est pourquoi je me suis limité, dans un premier temps, à l'étude du modèle de l'eau faisant appel uniquement à ce type de polarisabilité. L'extension de TINKER aux flux de charges requerra la modification du programme sachant que la plupart des éléments du tenseur d'interaction électrostatique sont déjà codés. Au niveau de l'échange induction, le seul schéma d'atténuation fourni par TINKER est le modèle de Thole [199]. J'ai donc ajouté la fonction d'atténuation de Tang et Toennies. Comme exprimé dans la section 3.3.3, cela revient à multiplier le tenseur d'interaction électrostatique par la fonction de Tang et Toennies. La dérivée de la fonction de Tang et Toennies (équation 3.15) est de la forme :

$$\frac{df_3(\beta, r_{ab})}{dr_{ab}} = \frac{\beta^4 r_{ab}^3}{6} \exp(-\beta r_{ab}) \quad (3.23)$$

— Le dernier potentiel qu'il faut manipuler est celui de van der Waals, et en particulier, la forme proposée par Buckingham. Pour l'inclure, il faut les dérivées successives de l'équation 3.19 par rapport à la distance :

$$\begin{aligned} \frac{d\mathcal{U}_{vdW}^{Buckingham}}{dr_{ij}} &= \sum_{i<j} \left[ -\varepsilon_{ij} * A_{1ij} * \exp(-A_{1ij} * r_{rij}) + 6 * \left( \frac{A_{2ij}}{r_{ij}} \right)^6 \frac{1}{r_{ij}} \right] \\ \frac{d^2\mathcal{U}_{vdW}^{Buckingham}}{dr_{ij}^2} &= \sum_{i<j} \left[ \varepsilon_{ij} * A_{1ij}^2 * \exp(-A_{1ij} * r_{rij}) - 42 * \left( \frac{A_{2ij}}{r_{ij}} \right)^6 \frac{1}{r_{ij}^2} \right] \end{aligned} \quad (3.24)$$

Une fois les modifications faites, j'ai pu réaliser les premiers tests à partir d'un modèle d'eau utilisant des polarisabilités dipolaires isotropes.

### 3.9 Cas d'une simulation de l'eau

Le modèle défini dans le chapitre 3 va être utilisé pour des simulations de l'eau. Pour un premier essai, une simulation NPT d'une boîte de 216 molécules est effectuée. Pour cette simulation, la boîte est générée grâce au logiciel PACKMOL<sup>8</sup> à partir d'un monomère. La longueur de liaison oxygène-hydrogène est de 0.958101 Å et la liaison oxygène-atome fictif est de 0.275623 Å. L'angle H-O-H de la molécule d'eau est de 104.26°.

Les paramètres atomiques pour le potentiel de Buckingham, sont déterminés à partir du tableau 3.9. Pour rappel, dans le cas de l'interaction du dimère de l'eau, les paramètres sont obtenus pour l'interaction de deux sites : l'atome d'oxygène pour une molécule et un atome d'hydrogène pour l'autre. À partir des règles de conversion dues à la distance entre les atomes, les paramètres pour l'atome d'oxygène de l'eau sont :

- $\varepsilon = 210057.0$
- $A_{1ii} = 3.7550$
- $A_{2ii} = 2.6495$

Pour cette simulation, il n'y a aucun paramètre d'atténuation pour l'énergie d'induction, comme spécifié dans la partie 3.3. Pour ce qui est de la polarisabilité, la valeur dipolaire isotrope est donnée dans le tableau 3.4 en  $10^{-3}$  a.u.. Pour le logiciel TINKER, les polarisabilités doivent être données en Å<sup>3</sup> donc la valeur utilisée sera de 1.449.

Lorsque la simulation a convergé vers une taille de boîte fixe, c'est-à-dire après une longue étape d'équilibration, il est possible de regarder les propriétés obtenues. La simulation utilisée est de 100 ps, j'ai sauvegardé les états toutes les 0.1 ps avec un pas de temps de 1 fs. Dans ce cas, la densité est largement sous-estimée et aux alentours de 0.82 g.cm<sup>3</sup> environ. De plus, la fonction de distribution radiale (*rdf*, figure 3.17) entre les atomes d'oxygène des molécules d'eau montre une mauvaise structuration des molécules d'eau dans la boîte ce qui se caractérise principalement par l'absence de liaisons hydrogène entre les molécules d'eau.

La fonction distribution obtenue n'est donc pas du tout représentative de celle obtenue expérimentalement (voir figure 3.18 partie (a)).

---

8. <http://www.ime.unicamp.br/~martinez/packmol>

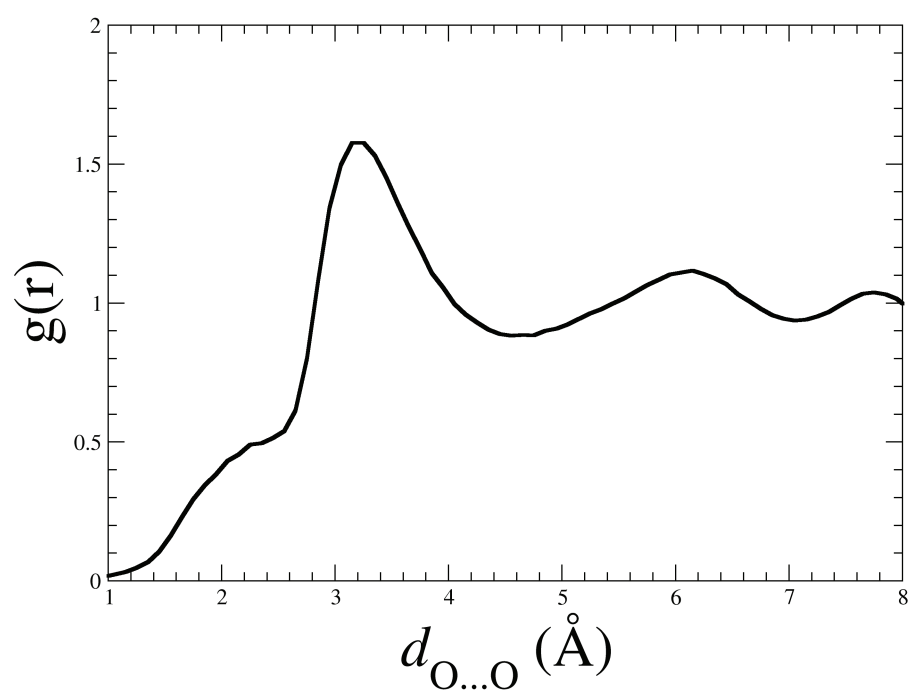


FIGURE 3.17 – Fonction de distribution radiale OO pour le modèle polarisable obtenu dans l'article [164]. En abscisse, les distances sont en ångström.

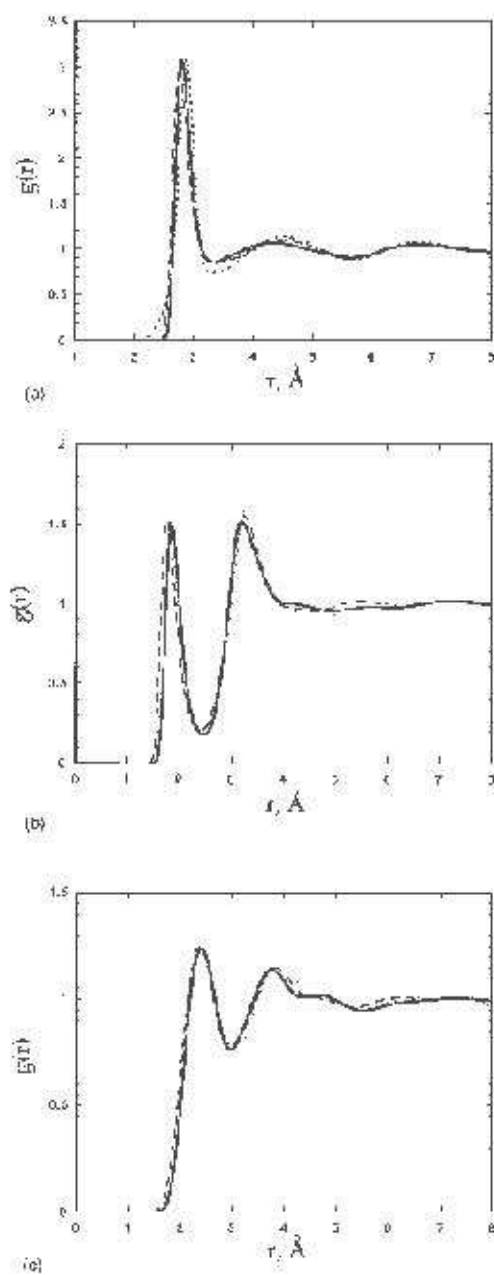


FIGURE 3.18 – Comparaison entre les valeurs mesurées expérimentalement et calculées pour la fonction de distribution radiale à 298 K (a) pour OO, train plein (référence [337]), en ligne discontinue (SPC/E) et en pointillés (expérience), (b) comme (a), pour OH, (c) comme (a), pour HH.

La source de ce problème réside principalement dans la stratégie de mise au point du potentiel de van der Waals où, pour déterminer les paramètres, un simple axe d'interaction a été considéré. Cela ne permet pas d'avoir une statistique suffisante pour reproduire les grandeurs spatiales d'interaction des molécules d'eau. Pour améliorer la qualité de notre modèle, une stratégie consiste à affiner les paramètres afin de reproduire certaines propriétés de la phase condensée (c'est ce qui est fait dans la plupart des champs de force actuels polarisables ou non). Ici, sur la base de plusieurs simulations successives nous avons obtenu un jeu de paramètres de van der Waals permettant de bien reproduire la structure et la densité. Les paramètres obtenus sont :

- $\varepsilon = 67000.0$
- $A_{1ii} = 4.7371$
- $A_{2ii} = 2.7606$

Toutefois, cette approche ne permet pas de connaître précisément les sources d'erreurs du modèle et comment celles-ci se distribuent dans les différentes composantes. Nous avons donc préféré travailler en amont pour proposer une meilleure dérivation *ab initio* des paramètres de notre modèle pour s'affranchir le plus possible d'une optimisation sur les propriétés de la phase condensée.

Dans un premier temps, j'ai cherché à identifier quelle(s) était(ent) le (les) termes du potentiel intermoléculaire responsable(s) de l'échec de la première simulation d'eau. Pour ce faire, nous avons regardé le comportement des trois contributions (électrostatique, induction et van der Waals) sur des multimères (jusqu'à l'hexamère) qui ont été obtenus en utilisant la décomposition de Kitaura et Morokuma (KM) (puisque SAPT ne permet de calculer qu'un dimère). Nous avons validé l'utilisation de KM en comparant l'énergie du dimère obtenu avec SAPT et aussi, pour chaque multimère, un calcul GAUSSIAN au niveau de théorie X3BLYP/Sadlej [338] est également effectué. L'énergie totale obtenue est alors comparée au résultat fourni par GAMESS. Dans tous les cas, la décomposition KM donne des résultats proches du calcul de plus haut niveau. Les énergies des multimères sont dans le tableau 3.13.

Avec ce dernier tableau, il est intéressant de s'apercevoir que même un modèle isotrope de polarisabilité reproduit vraiment fidèlement le comportement de l'énergie d'induction. Par contre, les deux autres quantités ne sont pas aussi bien reproduites. Pour la contribution électrostatique, nous avons montré dans la section 3.2.3 que les effets de pénétration dans le cas de la molécule d'eau sont importants et ne sont pas pris



	trimère		tétramère		pentamère		hexamère	
	KM	Modèle	KM	Modèle	KM	Modèle	KM	Modèle
Électrostatique	-27.51	-19.24	-51.74	-33.54	-67.41	-43.03	-85.27	-53.56
Induction	-4.83	-4.36	-11.41	-11.62	-16.20	-18.01	-21.62	-25.39
van der Waals	21.03	15.23	42.49	26.73	55.81	34.68	72.36	45.62
Totale	-11.31	-8.37	-20.66	-18.42	-27.80	-26.36	-34.53	-33.33

TABLE 3.13 – Multimères d'eau obtenus à la référence [338]. L'énergie est donnée pour le modèle et la référence quantique (KM avec le programme GAMESS) au niveau de théorie HF/Sadlej en kcal/mol.

en compte explicitement dans notre modèle. Par contre, nous avons défini un potentiel de van der Waals à même de compenser cette approximation. Or, plus le nombre de multimères augmente, moins la compensation est importante. Cela veut donc dire que le modèle de van der Waals ne reproduit pas fidèlement les interactions selon plusieurs directions comme c'est le cas avec les multimères.

### 3.9.1 Nouveaux paramètres de van der Waals

Après avoir fait ces premières analyses sur des multimères, nous avons voulu les utiliser pour obtenir des paramètres de van der Waals. J'ai donc extrait les 35 dimères possibles à partir des multimères de la référence [338]. Pour chaque dimère, un calcul SAPT a été effectué et à partir de cela, un nouveau jeu de paramètres a été obtenu :

- $\varepsilon = 244565.0$
- $A_{1ii} = 3.7604$
- $A_{2ii} = 3.1182$

La densité du modèle est de  $0.84 \text{ g.cm}^3$  ce qui ne convient toujours pas. Cela provient du manque d'interactions à courte distance étant donné que seuls les états d'équilibre sont considérés pour ces dimères mais aussi de la faible statistique avec les 35 dimères considérés.

Afin d'obtenir des paramètres de van der Waals plus représentatifs la méthode employée ici est l'extraction des géométries de dimère issues de simulations de dynamique

moléculaire pour des molécules d'eau. Pour cela, j'ai réalisé une simulation du modèle polarisable de Dang et Chang [337, 339] au cours de laquelle j'ai prélevé 861 dimères à des distances intermoléculaires inférieures à 10 Å. Ce dernier critère est pris car le rayon de coupure (*cut-off*) choisi pour la simulation est de 11 Å.

Pour chaque dimère de la simulation, un calcul quantique (corrigé de la BSSE, *Basis Set Superposition Error*) permet d'obtenir l'énergie d'interaction. On retranche alors à l'énergie totale la contribution de l'énergie d'interaction électrostatique et d'induction de notre modèle empirique. Cela permet d'obtenir un nuage de points pour différentes conformations. Un programme<sup>9</sup> nous permet, en utilisant un algorithme de Levenberg-Marquardt, de trouver de nouveaux paramètres pour la molécule d'eau. Les nouveaux paramètres sont :

- $\varepsilon = 2845438.0$
- $A_{1ii} = 4.7532$
- $A_{2ii} = 2.6936$

Une simulation de dynamique moléculaire de l'eau réalisée avec ce nouveau modèle conduit à une densité de 0.98 g.cm<sup>3</sup> et une distribution radiale (voir figure 3.19) très proche de l'expérience, excepté une intensité du premier pic trop forte. Ces premiers résultats, très encourageants, du modèle pour les simulations de dynamique moléculaire montrent qu'une approche toute *ab initio* est plausible pour obtenir des paramètres non biaisés par une reproduction de quelques valeurs obtenues expérimentalement.

---

9. Ce programme, `fitmultisite.py`, a été développé en python spécifiquement au laboratoire pour ajuster les paramètres de van der Waals. Celui-ci permettra aussi de déterminer des paramètres pour divers types d'atomes pour la suite du projet.

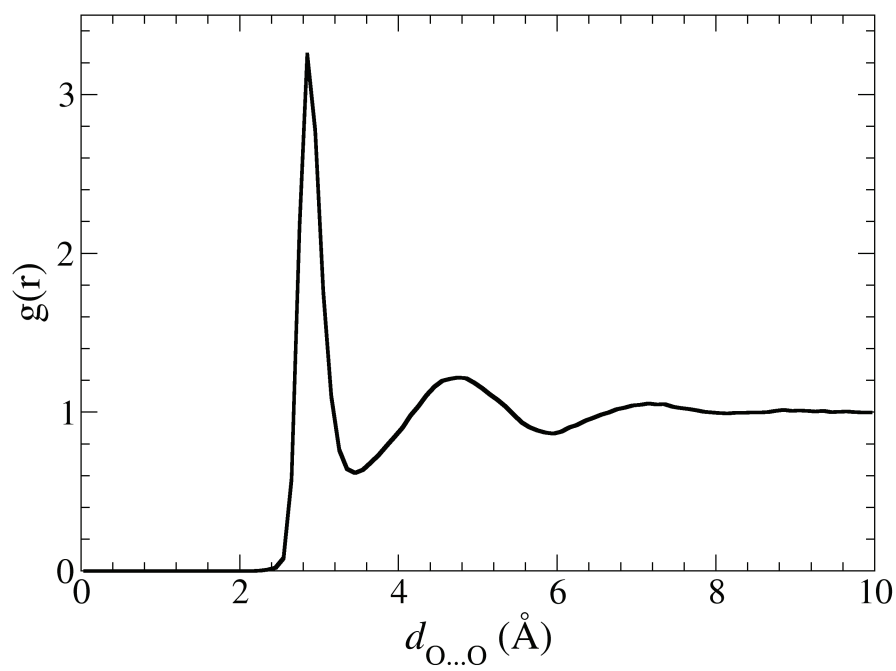


FIGURE 3.19 – Fonction de distribution radiale OO pour le « nouveau » modèle polarisable de l'eau. En abscisse, les distances sont en ångström.

# Conclusion et perspectives

---

Le développement des champs de force polarisables a été extensif cette dernière décennie. Cependant, il n'existe que peu de cas (exemple : les métallo-protéases) pour lesquels les champs de force polarisables démontrent une meilleure reproduction des systèmes biologiques par rapport aux champs de force additifs de paire. Nous avons vu que trois grandes familles de modèles polarisables pour des systèmes biologiques se distinguaient. Les charges fluctuantes se basent sur un principe de monopôles variant au fur et à mesure de la simulation par rapport à l'environnement, ce qui permet un calcul rapide du terme de polarisation explicite (inclus dans le potentiel électrostatique). Les oscillateurs de Drude ont également un modèle de monopôles mais ceux-ci sont fixes tout au long de la simulation. L'ajout de charges délocalisées reliées par un ressort permet de faire varier le moment dipolaire total par rapport à l'influence du champ électrique externe. Ce modèle est assez aboutit car il permet d'effectuer des simulations longues sur des systèmes d'intérêt biologique comme montré dans les références [72, 73, 143, 214, 218, 289, 290]. Le dernier modèle, qui est aussi le modèle utilisé dans cette thèse, a pour principe l'ajout de polarisabilités distribuées afin de définir un terme d'énergie d'induction explicite à l'énergie d'interaction.

Le modèle présenté ici contient un terme électrostatique tronqué à l'ordre zéro du développement multipolaire (charges), un terme d'induction tronqué à l'ordre un (flux de charges et polarisabilités dipolaires isotropes) et un terme de van der Waals utilisant un potentiel de Buckingham ( $\exp-r^{-6}$ ). Afin de considérer l'énergie d'échange-induction, un potentiel d'atténuation (Tang et Toennies) tronqué à l'ordre trois du développement peut être utilisé. La méthodologie pour dériver les paramètres de chacun des potentiels a été décrite par Archambault *et al.* [164]. Elle consiste en une dérivation de charges dérivées du potentiel pour la molécule isolée pour la partie électrostatique. Les polarisabilités distribuées sont également obtenues à partir de plusieurs calculs quantiques pour la molécule isolée. Les paramètres pour la fonction de Tang et Toennies sont obtenus en

comparant l'énergie d'induction du modèle et celle de la référence quantique (SAPT). Le potentiel de Buckingham est ajusté afin de reproduire au mieux l'énergie d'interaction totale. Ainsi, il embrasse toutes les contributions d'échange qui n'auraient pas été prises en compte explicitement dans les précédents potentiels (par exemple les effets de pénétration dans le potentiel électrostatique).

Le modèle ainsi défini nous a permis de reproduire de façon très fidèle un grand nombre d'énergies d'interaction entre des molécules (eau et benzène) et des cations ou anions. Cet accord suivant une direction spécifique de l'interaction a besoin de faire ses preuves sur diverses orientations présentes dans des simulations de systèmes biologiques. Les premiers essais sur des agrégats d'eau, en simulation de dynamique moléculaire, ont montré que les paramètres de van der Waals obtenus n'étaient pas nécessairement optimaux. Ceux-ci avaient la tendance d'exagérer certaines configurations ou de perdre toute structure dans l'enchaînement des molécules d'eau au sein de la boîte. L'objectif est maintenant d'améliorer et d'étendre la méthode d'ajustement pour obtenir des paramètres de van der Waals adaptés à des simulations en phase condensée sans avoir à reparamétrer le potentiel. Un début de réponse a été apporté grâce à des géométries représentatives de la phase aqueuse.

Les polarisabilités distribuées pour les vingt acides aminés naturels permettraient d'avoir un champ de force plus complet et utilisable pour simuler des systèmes biologiques réalistes. La carte d'induction de fragments d'acides aminés nécessite un temps de calcul important et n'a donc pas pu être menée à bout au cours de cette thèse. Une fois les assemblages d'acides aminés constitués, il restera à trouver une méthode pour joindre les polarisabilités obtenues sur des fragments isolés.

# Fichiers d'entrée et de sortie

## SAPT

---

### Sommaire

---

<b>A.1 Fichiers d'entrée</b> . . . . .	<b>101</b>
<b>A.2 Fichier de sortie</b> . . . . .	<b>104</b>

---

Afin de présenter les résultats des calculs, une explication des fichiers d'entrée et de sortie et leur définition va être explicité dans le cas de l'interaction de l'ion formate ( $\text{HCOO}^-$ ) avec le cation calcium ( $\text{Ca}^{2+}$ ).

### A.1 Fichiers d'entrée de la suite SAPT

Afin de démarrer un calcul perturbatif, la préparation de fichiers d'entrée est nécessaire au bon fonctionnement de la suite de logiciel SAPT2008. Comme nous avons pu le voir dans la partie théorique, le calcul est en trois parties : le dimère, le monomère A dans la base de B et le monomère B dans la base de A.

Afin d'explicitier les fichiers nécessaires, le nom de base du fichier sera `nom_de_fichierX.extension` où X correspondra à « rien » pour le dimère, A le monomère A dans la base de B, B le monomère B dans la base de A et l'extension sera définie suivant le type de fichier que nous allons détailler par la suite.

La suite SAPT doit être couplée à un logiciel de chimie et en particulier, ici, `Atmol1024` [340] est employé parce que une version modifiée spécialement pour SAPT est fourni. Pour pouvoir être compris par `Atmol1024`, un fichier `*.scfinp` est généré (voir tableau A.1).

SCF	116	21	0	1	ED3
TITLE					
HCOO- Ca.					
MFILE					
MT3					
1					
0					
DIIS	20	1	0		
AUTO	0				
MAXCYC	60				
ACCURACY	1	10			
START					
ENTER	1				

TABLE A.1 – Exemple de fichier `scfinp` dans le cas de l'interaction de l'ion formate avec le cation calcium.

La première ligne correspond à des valeurs nécessaires pour le *scf*, respectivement, le nombre de fonctions de base (116), le nombre d'électrons sur 2 du système considéré (21), la charge globale du système (0) et le spin (1). Les autres mots-clefs sont définis pour le programme et une liste plus exhaustive peut être trouvée sur le site de `Atmol`<sup>1</sup>. La partie la plus importante de ce fichier d'entrée est le `MAXCYC` qui permet d'indiquer le nombre maximum de cycles pour converger la fonction d'onde (une discussion sera effectuée en A.2 sur cette partie).

Le second type de fichier à considérer est nommé `*.intinp`. Dans celui-ci, la géométrie de l'ensemble du dimère est considérée. La structure de fichier est définie dans le tableau A.2.

Les deux premières lignes du fichier correspondent au titre pour le calcul à effectuer puis nous spécifions la géométrie du système en spécifiant bien **tous** les atomes considérés dans le dimère. Afin de définir le monomère dans la base de l'autre monomère, il suffit

1. <http://tc5.chem.uu.nl/ATMOL/scf.html>

```

TITLE
ca2+ hcoo-
GEOMETRY ANGSTROM
0.000000      0.306183   0.000000   6.0   C1
1.120462      -0.204599   0.000000   8.0   O1
-1.120339     -0.204197   0.000000   8.0   O2
0.000000      1.433262   0.000000   1.0   H1
0.0           0.306183   2.400000   20.0  CA1
END
GTOS
...
END
MAXBLOCK 3000000
ACCURACY 12      13      12
MAINFILE MT3
DUMPFIL  ED3
IBLOCK   1
NOPRINT  GROU    GTOS    BASI
BYPASS   PROPERTY
ENTER    1

```

TABLE A.2 – Exemple de fichier `intinp` dans le cas de l'interaction de l'ion formate avec le cation calcium.



```
CaHCOO-
&TRN
      ISITANEW=T,
      OUT=F, TOLER=15, DIMER=T, MEMTRAN=100000000
&END
&CCINP
      CCPRINT=F, VCRIT=1.0D-10, TOLITER=1.0D-5
&END
&INPUTCOR
      SAPT2=T,
      CONVAMP=F,
      MEMSAPT=150000000,
      FROZEN=F
&END
```

TABLE A.3 – Exemple de fichier `data` dans le cas de l'interaction de l'ion formate avec le cation calcium.

de définir le numéro atomique (4<sup>ème</sup> colonne) comme étant égal à 0.0. La base<sup>2</sup> à utiliser est définie entre les balises `GTOS`, celle-ci est écrite comme pour le programme de chimie quantique `GAUSSIAN`.

Un dernier fichier (`*P.data`) est nécessaire, il permet de donner les directives au programme `SAPT`. Il ressemble au tableau A.3.

Une fois ces fichiers générés, le calcul peut alors être démarré à partir de la commande :

```
SAPT nom_de_fichier scfcp >& fichier.out
```

## A.2 Fichier de sortie de la suite SAPT

Des fichiers sont générés au cours des diverses étapes de calcul, mais le fichier d'importance est celui où est écrit la décomposition de l'énergie d'interaction que nous

---

2. <http://gnode2.pnl.gov/bse/portal>

avons écrit dans `fichier.out`.

Dans ce fichier, nous devons avant toute analyse de la sortie regarder si notre fonction d'onde calculée avec le programme `Atmol1024` est convergée parce que le programme ne s'arrête pas si la convergence n'est pas atteinte. Pour effectuer cette vérification il suffit de chercher dans le fichier si l'occurrence **NO CONVERGENCE** est présente. Si jamais la convergence n'est pas atteinte, il faudra augmenter le `MAXCYC`.

Une fois cette précaution prise, la décomposition de l'énergie peut alors être analysée. Celle-ci est résumée dans un tableau fournit vers la fin du fichier et un exemple est donné dans le tableau A.2. Chaque contribution dans cette sortie est explicitée dans la documentation fournie avec le logiciel SAPT.

TABLE A.4: Exemple de tableau extrait d'un fichier de sortie dans le cas de l'interaction de l'ion formate avec le cation calcium.

```

=====
                          Summary Table
=====
          Mono A :   12 occupied   104 virtual   116 total
          Mono B :    9 occupied   107 virtual   116 total
----- Molecule A      24 Electron(s)
      ATOM      XX          YY          ZZ          Charge
      1  0.000000000   0.578601972   0.000000000        6.0
      2  2.117366161  -0.386636048   0.000000000        8.0
      3 -2.117133724  -0.385876378   0.000000000        8.0
      4  0.000000000   2.708472450   0.000000000        1.0

----- Molecule B      18 Electron(s)
      5  0.000000000   0.578601972   4.535342373       20.0
E^{HF}_{AB}          -864.7695040419170027   hartrees
E^{HF}_{A}           -188.2576368834700133   hartrees
E^{HF}_{B}           -676.1366197731699685   hartrees

```

TABLE A.4: (suite)

Correction	mHartree	Kcal/mol	1/cm
— SCF (SAPT_super) —			
$E^{\{HF\}}_{\{int\}}$	-375.247385277	-235.47148674	-82357.2810
$E^{\{(10)\}}_{\{elst\}}$	-350.675179957	-220.05218217	-76964.3054
$E^{\{(10)\}}_{\{exch\}}$	36.786169702	23.08368935	8073.6310
$E^{\{(10)\}}_{\{exch\}}\{S^2\}$	36.434734171	22.86316004	7996.4998
$E^{\{(10)\}}_{\{exch\}}-S^2$	0.351435531	0.22052931	77.1312
$E^{\{(20)\}}_{\{ind,resp\}}$	-121.039988166	-75.95380297	-26565.2066
$E^{\{(20)\}}_{\{ex-ind\}}$	52.658250965	33.04357906	11557.1501
$E^{\{(20)\}}_{\{ex-ind,r\}}$	67.294093942	42.22771689	14769.3464
SAPT SCF_{resp}^b	-367.634904479	-230.69457891	-80686.5346
$\delta^{\{HF\}}_{\{int,r\}}$	-7.612480798	-4.77690783	-1670.7464
— CORRELATION —			
$E^{\{(12)\}}_{\{elst,resp\}}$	-6.537266568	-4.10220014	-1434.7642
$\epsilon^{\{(1)\}}_{\{elst,r\}}(k)$	-6.537266568	-4.10220014	-1434.7642
$E^{\{(11)\}}_{\{exch\}}$	-0.532589937	-0.33420551	-116.8900
$E^{\{(12)\}}_{\{exch\}}$	5.206817843	3.26733026	1142.7644
$\epsilon^{\{(1)\}}_{\{exch\}}(k)$	4.674227907	2.93312475	1025.8744
$^tE^{\{(22)\}}_{\{ind\}}$	-21.570583101	-13.53575660	-4734.1957
$^tE^{\{(22)\}}_{\{ex-ind\}}^*$	11.992506506	7.52541776	2632.0509
$E^{\{(20)\}}_{\{disp\}}$	-5.449471941	-3.41959814	-1196.0208
$E^{\{(2)\}}_{\{disp\}}(k)$	-5.449471941	-3.41959814	-1196.0208
$E^{\{(20)\}}_{\{exch-disp\}}$	0.933510474	0.58578716	204.8819
SAPT_{corr}	-9.419810156	-5.91102507	-2067.4093
SAPT_{corr,resp}	-15.957076724	-10.01322521	-3502.1735
— TOTAL (hybrid) —			
SCF+SAPT_{corr}	-384.667195433	-241.38251181	-84424.6904
SCF+SAPT_{corr,resp}	-391.204462001	-245.48471195	-85859.4546
=====			

TABLE A.4: (suite)

---

Conversion Factors :	1000.0	627.51	219474.63
----------------------	--------	--------	-----------

---

\* denotes ESTIMATED corrections

$$^a \text{SAPT SCF} = E^{\{(10)\}}_{\text{elst}} + E^{\{(10)\}}_{\text{exch}} \\ + E^{\{(20)\}}_{\text{ind}} + E^{\{(20)\}}_{\text{exch-ind}}$$

$$^b \text{SAPT SCF}_{\text{resp}} = E^{\{(10)\}}_{\text{elst}} + E^{\{(10)\}}_{\text{exch}} \\ + E^{\{(20)\}}_{\text{ind,resp}} + E^{\{(20)\}}_{\text{exch-ind,resp}}$$


---

Les grandeurs énergétiques qui nous intéressent sont alors lues dans ce tableau et correspondent à :

$$\Delta \mathcal{U}_{ele}^{SAPT} = E^{\{(10)\}}_{\text{elst}} + E^{\{(12)\}}_{\text{elst,resp}}$$

$$\Delta \mathcal{U}_{ind}^{SAPT} = E^{\{(20)\}}_{\text{ind,resp}} + {}^t E^{\{(22)\}}_{\text{ind}}$$

$$\Delta \mathcal{U}_{exch}^{SAPT} = E^{\{(10)\}}_{\text{exch}} + \epsilon^{\{(1)\}}_{\text{exch}}(k)$$

$$\Delta \mathcal{U}_{disp}^{SAPT} = E^{\{(20)\}}_{\text{disp}}$$

$$\Delta \mathcal{U}_{exc-ind}^{SAPT} = E^{\{(20)\}}_{\text{ex-ind,r}} + {}^t E^{\{(22)\}}_{\text{ex-ind}}$$

$$\Delta \mathcal{U}_{exc-disp}^{SAPT} = E^{\{(20)\}}_{\text{exch-disp}}$$

$$\delta Hf = \Delta^{\text{HF}}_{\text{int,r}}$$

$$\Delta \mathcal{U}_{tot}^{SAPT} = \text{SCF+SAPT}_{\text{corr,resp}}$$

À partir des valeurs données dans la sortie et en sommant certaines contributions, il est possible de comparer directement les grandeurs fournies par SAPT avec le modèle empirique.



# Détermination des paramètres de Tang and Toennies en mécanique quantique

---

L'approche pour déterminer le terme  $b$  dans la fonction de Tang et Toennies est proposée par Stone et collaborateurs [91, 92]. Dans cette approche, de simples calculs quantiques permettent d'obtenir la valeur de ce paramètre :

$$b = \sqrt{2I_A} + \sqrt{2I_B} \quad (\text{B.1})$$

où  $I_A$  est le potentiel d'ionisation calculé en mécanique quantique au niveau de théorie PBE0/Sadlej (en *Hartree*). Les valeurs obtenues sont données dans le tableau B.1. Pour les espèces où la valeur du potentiel est disponible expérimentalement, l'écart entre la valeur quantique et la valeur expérimentale est très faible. Ce qui permet de dire que la fonctionnelle PBE0 est adaptée au calcul de cette propriété. L'avantage aussi de cette méthode est la possibilité d'avoir des paramètres pour chaque fragment et pas pour chaque interaction de paire.

Cette méthode semble prometteuse mais étant donné l'état actuel de la décomposition SAPT il n'est pas possible de savoir si le terme est vraiment viable ou si il ne fonctionne que dans certains cas. Par exemple, l'article de Millot et Stone [228] donne une valeur pour le paramètre d'atténuation de 1.92 dans le cas du dimère de l'eau. Le calcul effectué avec les potentiels d'ionisation donne  $2 * \sqrt{2 * 0.464} \approx 1.93$ . Dans le cas de l'interaction entre l'eau et des anions, appliquer le coefficient  $b$  obtenu par le calcul des énergies d'ionisation conduit à une énergie de polarisation presque nulle sur une grande gamme de distance. Pour cela, j'ai considéré que, pour l'instant, l'idée de ne prendre en compte aucun potentiel d'atténuation pour ces interactions était l'approche la plus raisonnable.

	Exp.	PBE0/Sadlej	Erreur (%)
Eau	0.464	0.464	0.00
Formaldéhyde	0.400	0.397	0.76
Calcium	0.225	0.220	1.89
Benzène	0.340	0.339	0.18
Formate	/	0.132	/
Calcium ( $\text{Ca}^{2+}$ )	/	1.874	/
Chlorure ( $\text{Cl}^-$ )	0.133	0.133	0.00
Sodium ( $\text{Na}^+$ )	/	1.771	/
$\text{H}^+$	/	0.500	/
Potassium ( $\text{K}^{2+}$ )	/	1.165	/
Ammonium ( $\text{NH}_4^+$ )	/	0.974	/
Lithium ( $\text{Li}^+$ )	/	2.788	/
Fluorure ( $\text{F}^-$ )	0.120	0.125	4.03
Magnesium ( $\text{Mg}^{2+}$ )	/	2.985	/
Bromure ( $\text{Br}^-$ )	0.124	0.129	4.44
Iodure ( $\text{I}^-$ )	0.112	0.124	10.26

TABLE B.1 – Valeurs des potentiels d’ionisation de différents espèces. La valeur quantique est calculée au niveau de théorie PBE0/Sadlej Les valeurs expérimentales (Exp.) sont obtenues à partir du Handbook de chimie-physique [341]. Toutes les énergies sont données en *Hartree*.

# Intégration des moments induits en dynamique moléculaire

---

Il est souvent connu que les méthodes utilisant des moments induits sont très lentes du fait de la convergence des moments induits. Tout d'abord, pourquoi avoir besoin d'effectuer une convergence des moments induits pour un calcul utilisant des polarisabilités distribuées ?

Pour répondre à cela, observons les équations 3.11. Le calcul de l'énergie d'induction requiert le calcul des moments induits ( $\Delta Q_{\nu'}^a$ , équation 2.17) qui nécessitent la connaissance des moments induits ( $\Delta Q_u^b$ ). En résumé, un site est polarisé par les moments permanents et induits des sites environnants. Il est nécessaire d'itérer sur les moments pour obtenir le moment total (induit et permanent) situé sur le site ou l'atome polarisable. Afin d'obtenir une convergence rapide des moments induits, deux étapes sont nécessaires : le choix de départ des moments induits et des algorithmes itératifs qui permettent une convergence rapide.

Pour obtenir un bon choix de départ, deux moyens sont détaillés par Wang *et al.* [342] : la méthode des moindres carrés et la méthode du polynôme optimum. Étant donné le peu d'utilisation de cette méthode dans cette thèse je ne détaillerai pas ces approches. Afin de converger plus rapidement les moments induits, des méthodes de préconditionneurs sont utilisées mais ne seront pas expliquées ici<sup>1</sup>. Cette approche a l'avantage d'avoir été incluse de manière très concise dans un code nommé MDX. Ce code n'est pas distribué librement mais permet d'effectuer des simulations de dynamique moléculaire sur des boîtes d'eau utilisant trois modèles (POL1 [343], POL3 [337], RPOL [339]) polarisables avec des polarisabilités dipolaires isotropes. Les calculs sur des simulations de 216 molécules d'eau

---

1. Voir les chapitres 3.2 et 3.3 de la thèse de Wang (<http://www.cs.uiuc.edu/research/techreports.php?report=UIUCDCS-R-2005-2522>) pour plus de détails.



polarisables utilisant divers algorithmes d'intégration ou de piston pour la pression ou la température permettent d'obtenir, sur un seul processeur, des simulations de 1 ns en moins de 12 heures.

Par contre, dans les équations développées dans le modèle de Wang, seule une polarisabilité dipolaire isotrope est considérée. Pour inclure notre modèle, les polarisabilités de flux de charges doivent donc être ajoutées. Dans les notations de l'article de Wang et Skeel [196], les équations s'écrivent :

$$E^{el}(\mathbf{r}, \mathbf{d}) = \frac{1}{2} \mathbf{q}^T \mathbf{G}_0(\mathbf{r}) \mathbf{q} + \mathbf{d}^T \mathbf{G}_1(\mathbf{r}) \mathbf{q} + \frac{1}{2} \mathbf{d}^T \mathbf{G}_2(\mathbf{r}) \mathbf{d} + \frac{1}{2} \mathbf{d}^T \mathbf{D}_\alpha^{-1} \mathbf{d} \quad (\text{C.1})$$

L'ajout de polarisabilité de flux de charges s'effectue au niveau du terme induit (i.e.  $\mathbf{d}$ ) où seuls trois termes sont à modifier :  $\mathbf{G}_1$ ,  $\mathbf{G}_2$  et  $\mathbf{D}_\alpha$ .  $\mathbf{d}$  peut donc être réécrit en :

$$\mathbf{d}' = \begin{pmatrix} \Delta Q^1 \\ \Delta Q^2 \\ \vdots \\ \Delta Q^N \\ \mathbf{d} \end{pmatrix}$$

La matrice est de taille  $4N$  avec  $N$  le nombre d'atomes dans la molécule. En comparaison, la taille de la matrice de polarisabilité dipolaire est de taille  $3N$ . Puisque nous avons défini la nouvelle matrice de polarisabilité induite ( $\mathbf{d}'$ ), il est possible de réécrire les opérateurs  $\mathbf{G}_i, i = \{0, 1, 2\}$ . L'opérateur d'ordre zéro ( $\mathbf{G}_0$ ) est inchangé en ajoutant la polarisabilité de flux de charges. Pour l'ordre un, la nouvelle matrice, de taille  $4N * N$ , est :

$$\mathbf{G}'_1 = \begin{pmatrix} \mathbf{G}_0 \\ \mathbf{G}_1 \end{pmatrix}$$

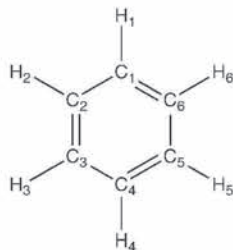
Pour l'ordre deux, la réécriture de la matrice donne une matrice de taille  $4N * 4N$  :

$$\mathbf{G}'_2 = \begin{pmatrix} \mathbf{G}_0 & \mathbf{G}_1^T \\ \mathbf{G}_1 & \mathbf{G}_2 \end{pmatrix}$$

Nous avons ainsi tous les opérateurs  $\mathbf{G}_i$ . Il ne reste plus qu'à réécrire l'opérateur  $\mathbf{D}_\alpha$  :

$$\mathbf{D}'_\alpha = \begin{pmatrix} \mathbf{D}_{\alpha_{00,00}} & 0 \\ 0 & \mathbf{D}_\alpha \end{pmatrix}$$

Ce qui fait que cette matrice est de taille  $4N * 4N$ . Dans cette écriture, la matrice  $\mathbf{D}_{\alpha_{00,00}}$  contient l'opérateur de flux de charges du modèle. Cet opérateur est défini comme expliqué dans la partie 2.2.1 et est détaillé dans le cas du benzène (figure C).



La matrice de polarisation de flux de charges,  $\mathbf{D}_{\alpha_{00,00}}$ , est alors :

$$\begin{pmatrix} \alpha_{00,00}^{C_1 C_1} & \alpha_{00,00}^{C_1 C_2} & 0 & 0 & 0 & \alpha_{00,00}^{C_1 C_6} & \alpha_{00,00}^{C_1 H_1} & 0 & 0 & 0 & 0 & 0 \\ \alpha_{00,00}^{C_2 C_1} & \alpha_{00,00}^{C_2 C_2} & \alpha_{00,00}^{C_2 C_3} & 0 & 0 & 0 & 0 & \alpha_{00,00}^{C_2 H_2} & 0 & 0 & 0 & 0 \\ 0 & \alpha_{00,00}^{C_3 C_2} & \alpha_{00,00}^{C_3 C_3} & \alpha_{00,00}^{C_3 C_4} & 0 & 0 & 0 & 0 & \alpha_{00,00}^{C_3 H_3} & 0 & 0 & 0 \\ 0 & 0 & \alpha_{00,00}^{C_4 C_3} & \alpha_{00,00}^{C_4 C_4} & \alpha_{00,00}^{C_4 C_5} & 0 & 0 & 0 & 0 & \alpha_{00,00}^{C_4 H_4} & 0 & 0 \\ 0 & 0 & 0 & \alpha_{00,00}^{C_5 C_4} & \alpha_{00,00}^{C_5 C_5} & \alpha_{00,00}^{C_5 C_6} & 0 & 0 & 0 & 0 & \alpha_{00,00}^{C_5 H_5} & 0 \\ \alpha_{00,00}^{C_6 C_1} & 0 & 0 & 0 & \alpha_{00,00}^{C_6 C_5} & \alpha_{00,00}^{C_6 C_6} & 0 & 0 & 0 & 0 & 0 & \alpha_{00,00}^{C_6 H_6} \\ \alpha_{00,00}^{H_1 C_1} & 0 & 0 & 0 & 0 & 0 & \alpha_{00,00}^{H_1 H_1} & 0 & 0 & 0 & 0 & 0 \\ 0 & \alpha_{00,00}^{H_2 C_2} & 0 & 0 & 0 & 0 & 0 & \alpha_{00,00}^{H_2 H_2} & 0 & 0 & 0 & 0 \\ 0 & 0 & \alpha_{00,00}^{H_3 C_3} & 0 & 0 & 0 & 0 & 0 & \alpha_{00,00}^{H_3 H_3} & 0 & 0 & 0 \\ 0 & 0 & 0 & \alpha_{00,00}^{H_4 C_4} & 0 & 0 & 0 & 0 & 0 & \alpha_{00,00}^{H_4 H_4} & 0 & 0 \\ 0 & 0 & 0 & 0 & \alpha_{00,00}^{H_5 C_5} & 0 & 0 & 0 & 0 & 0 & \alpha_{00,00}^{H_5 H_5} & 0 \\ 0 & 0 & 0 & 0 & 0 & \alpha_{00,00}^{H_6 C_6} & 0 & 0 & 0 & 0 & 0 & \alpha_{00,00}^{H_6 H_6} \end{pmatrix}$$

L'équation C.1 peut être ainsi définie avec le terme de polarisabilité de flux de charges défini explicitement :

$$E^{el}(\mathbf{r}, \mathbf{d}) = \frac{1}{2} \mathbf{q}^T \mathbf{G}_0(\mathbf{r}) \mathbf{q} + \mathbf{d}'^T \mathbf{G}'_1(\mathbf{r}) \mathbf{q} + \frac{1}{2} \mathbf{d}'^T \mathbf{G}'_2(\mathbf{r}) \mathbf{d}' + \frac{1}{2} \mathbf{d}'^T \mathbf{D}'_{\alpha}^{-1} \mathbf{d}' \quad (\text{C.2})$$

Toute cette réécriture est détaillée pour être intégrée directement dans le code de dynamique moléculaire de Wang et Skeel. Ceci n'est pour l'instant pas effectué étant donné la difficulté au niveau de la programmation de ce code. Afin de faire les premières simulations de dynamique moléculaire, un code plus simple et plus complet (en terme d'intégrateurs, de potentiel...) va être utilisé.



# Bibliographie

1. Thouless, D. J. *The quantum mechanics of many-body systems* (éd. Press, A.) (New York, 1972).
2. Elrod, M. J. et Saykally, R. J. Many-Body Effects in Intermolecular Forces. *Chem. Rev.* **94**, 1975–1997 (1994).
3. Cruzan, J. D., Braly, L. B., Liu, K., Brown, M. G., Loeser, J. G. et Saykally, R. J. Quantifying Hydrogen Bond Cooperativity in Water : VRT Spectroscopy of the Water Tetramer. *Science* **271**, 59–62 (1996).
4. Liu, K., Cruzan, J. D. et Saykally, R. J. Water Clusters. *Science* **271**, 929–933 (1996).
5. Liu, K., Brown, M. G., Carter, C., Saykally, R. J., Gregory, J. K. et Clary, D. C. Characterization of a cage form of the water hexamer. *Nature* **381**, 501–503 (1996).
6. Berendsen, H. J. C., Grigera, J. R. et Straatsma, T. P. The Missing Term in Effective Pair Potentials. *J. Phys. Chem.* **91**, 6269–6271 (1987).
7. Stone, A. J. et Alderton, M. Distributed multipole analysis. *Mol. Phys.* **56**, 1047–1064 (1985).
8. Vigné-Maeder, F. et Claverie, P. The exact multicenter multipolar part of a molecular charge distribution and its simplified representations. *J. Chem. Phys.* **88**, 4934–4948 (1988).
9. Stone, A. J. *The theory of intermolecular forces* (éd. Oxford, C. P.) (International series of monographs on chemistry, 1996).
10. Coker, H. Polarizability changes on ion hydration. *J. Phys. Chem.* **80**, 2084–2091 (1976).
11. Okruss, M., Müller, R. et Hese, A. High-resolution ultraviolet laser spectroscopy on jet-cooled benzene molecules : Ground and excited electronic state polarizabilities determined from static Stark effect measurements. *J. Chem. Phys.* **110**, 10393 (1999).
12. Kitaura, K. et Morokuma, K. A new energy decomposition scheme for molecular interactions within the Hartree-Fock approximation. *Int. J. Quantum Chem.* **10**, 325–340 (1975).

13. Schrödinger, E. *Annalen der Physik. Vierte Folge* **80**, 437 (1926).
14. Jeziorski, B., Moszynski, R. et Szalewicz, K. Perturbation Theory Approach to Intermolecular Potential Energy Surfaces of van der Waals Complexes. *Chem. Rev.* **94**, 1887–1930 (1994).
15. Born, M. et Oppenheimer, R. Zur Quantentheorie der Moleküle. *Annalen der Physik* **84**, 457–484 (1927).
16. Brooks, B. R., Brucoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S. et Karplus, M. CHARMM : A program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.* **4**, 187–217 (1983).
17. Brooks, B. R. *et al.* CHARMM : the biomolecular simulation program. *J. Comput. Chem.* **30**, 1545–1614 (2009).
18. Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M. Jr., Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. C. et Kollman, P. A. A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J. Am. Chem. Soc.* **117**, 5179–5197 (1995).
19. Oostenbrink, C., Villa, A., Mark, A. E. et van Gunsteren, W. F. A biomolecular force field based on the free enthalpy of hydration and solvation : The GROMOS force-field parameter sets 53A5 and 53A6. *J. Comput. Chem.* **25**, 1656–1676 (2004).
20. Kaminski, G. A., Stern, H. A., Berne, B. J. et Friesner, R. A. Development of an Accurate and Robust Polarizable Molecular Mechanics Force Field from ab Initio Quantum Chemistry. *J. Phys. Chem. A* **108**, 621–627 (2004).
21. Becker, O. M., MacKerell Jr., A. D., Roux, B. et Watanabe, M. *Computational Biochemistry and Biophysics* (éd. New York, B.) (Marcel Dekker, Inc., 2001).
22. MacKerell, A. D. Jr. *et al.* All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *J. Phys. Chem. B* **102**, 3586–3616 (1998).
23. Allen, F. H., Bellard, S., Brice, M. D., Cartwright, B. A., Doubleday, A., Higgs, H., Hummelink, T., Hummelink-Peters, B. G., Kennard, O., Motherwell, W. D. S., Rodgers, J. R. et Watson, D. G. The Cambridge Crystallographic Data Centre : computer-based search, retrieval, analysis and display of information. *Acta Cryst.* **B35**, 2331–2339 (1979).

24. Barone, V., Cossi, M. et Tomasi, J. Geometry optimization of molecular structures in solution by the polarizable continuum model. *J. Comput. Chem.* **19**, 404–417 (1998).
25. MacKerell Jr., A. D. et Karplus, M. Importance of attractive van der Waals contribution in empirical energy function models for the heat of vaporization of polar liquids. *J. Phys. Chem.* **95**, 10559–10560 (1991).
26. MacKerell Jr., A. D., Wiorkiewicz-Kuczera, J. et Karplus, M. An all-atom empirical energy function for the simulation of nucleic acids. *J. Am. Chem. Soc.* **117**, 11946–11975 (1995).
27. Jorgensen, W. L. Theoretical studies of medium effects on conformational equilibria. *J. Phys. Chem.* **87**, 5304–5314 (1983).
28. Ángyán, J. G. et Chipot, C. A comprehensive approach to molecular charge density models : From distributed multipoles to fitted atomic charges. *Int. J. Quant. Chem.* **52**, 17–37 (1994).
29. Mulliken, R. S. Electronic Population Analysis on LCAO-MO Molecular Wave Functions. I. *J. Chem. Phys.* **23**, 1833–1840 (1955).
30. Bayly, C. I., Cieplak, P., Cornell, W. et Kollman, P. A. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges : the RESP model. *J. Phys. Chem.* **97**, 10269–10280 (1993).
31. Cornell, W. D., Cieplak, P., Bayly, C. I. et Kollmann, P. A. Application of RESP charges to calculate conformational energies, hydrogen bond energies, and free energies of solvation. *J. Am. Chem. Soc.* **115**, 9620–9631 (1993).
32. Chipot, C., Ángyán, J. G., Maignet, B. et Scheraga, H. A. Modeling amino acid side chains. 2. Determination of point charges from electrostatic properties : toward transferable point charge models. *J. Phys. Chem.* **97**, 9788–9796 (1993).
33. Ángyán, J. G., Chipot, C., Dehez, F., Hättig, C., Jansen, G. et Millot, C. OPEP : A tool for the optimal partitioning of electric properties. *J. Comput. Chem.* **24**, 997–1008 (2003).
34. Yin, D. et MacKerell Jr., A. D. Combined ab initio/empirical approach for optimization of Lennard-Jones parameters. *J. Comput. Chem.* **19**, 334–348 (1998).

35. Wang, Z.-X., Zhang, W., Wu, C., Lei, H., Cieplak, P. et Duan, Y. Strike a balance : Optimization of backbone torsion parameters of AMBER polarizable force field for simulations of proteins and peptides. *J. Comput. Chem.* **27**, 781–790 (2005).
36. Warshel, A. et Russell, S. T. Calculations of electrostatic interactions in biological systems and in solutions. *Q. Rev. Biophys.* **17**, 283–421 (1984).
37. Kollman, P. Free energy calculations : Applications to chemical and biochemical phenomena. *Chem. Rev.* **93**, 2395–2417 (1993).
38. Warshel, A. et Parson, W. W. Dynamics of biochemical and biophysical reactions : insight from computer simulations. *Q. Rev. Biophys.* **34**, 563–670 (2001).
39. Karplus, M. et McCammon, J. A. Molecular dynamics simulations of biomolecules. *Nature Structural Biology* **9**, 646–652 (2002).
40. Dudev, T. et Lim, C. Metal Binding Affinity and Selectivity in Metalloproteins : Insights from Computational Studies. *Annual Review of Biophysics* **37**, 97–116 (2008).
41. Shostak, S. L., Ebenstein, W. L. et Muentner, J. S. The dipole moment of water. I. Dipole moments and hyperfine properties of H<sub>2</sub>O and HDO in the ground and excited vibrational states. *J. Chem. Phys.* **94**, 5875–5882 (1991).
42. Gubskaya, A. V. et Kusalik, P. G. The total molecular dipole moment for liquid water. *J. Chem. Phys.* **117**, 5290–5302 (2002).
43. Guillot, B. A reappraisal of what we have learnt during three decades of computer simulations on water. *J. Mol. Liq.* **101**, 219–260 (2002).
44. Berendsen, H. J. C. dans *CECAM Workshop* (1972).
45. Warshel, A. Calculations of chemical processes in solutions. *J. Phys. Chem.* **83**, 1640–1652 (1979).
46. Åqvist, J. et Hansson, T. On the Validity of Electrostatic Linear Response in Polar Solvents. *J. Phys. Chem.* **100**, 9512–9521 (1996).
47. Rizzo, R. C. et Jorgensen, W. L. OPLS All-Atom Model for Amines : Resolution of the Amine Hydration Problem. *J. Am. Chem. Soc.* **121**, 4827–4836 (1999).
48. Grossfield, A., Ren, P. et Ponder, J. W. Ion Solvation Thermodynamics from Simulation with a Polarizable Force Field. *J. Am. Chem. Soc.* **125**, 15671–15682 (2003).

49. Roux, B. et Karplus, M. Ion transport in the gramicidin channel : free energy of the solvated right-handed dimer in a model membrane. *J. Am. Chem. Soc.* **115**, 3250–3262 (1993).
50. Allen, T. W., Andersen, O. S. et Roux, B. Energetics of ion conduction through the gramicidin channel. *Proc. Natl. Acad. Sci. USA* **101**, 117–122 (2004).
51. Bukowski, R., Szalewicz, K., Groenenboom, G. C. et van der Avoird, A. Predictions of the Properties of Water from First Principles. *Science* **315**, 1249–1252 (2007).
52. Stone, A. J. Water from First Principles. *Science* **315**, 1228–1229 (2007).
53. Stone, A. J. Intermolecular Potentials. *Science* **321**, 787–789 (2008).
54. Dang, L. X. Intermolecular interactions of liquid dichloromethane and equilibrium properties of liquid-vapor and liquid-liquid interfaces : A molecular dynamics study. *J. Chem. Phys.* **110**, 10113–10122 (1999).
55. Chang, T.-M. et Dang, L. X. Molecular dynamics simulations of CCl<sub>4</sub>-H<sub>2</sub>O liquid-liquid interface with polarizable potential models. *J. Chem. Phys.* **104**, 6772–6783 (1996).
56. Martin, M. G. et Siepmann, J. I. Transferable Potentials for Phase Equilibria. 1. United-Atom Description of n-Alkanes. *J. Phys. Chem. B* **102**, 2569–2577 (1998).
57. Martin, M. G. et Siepmann, J. I. Novel Configurational-Bias Monte Carlo Method for Branched Molecules. Transferable Potentials for Phase Equilibria. 2. United-Atom Description of Branched Alkanes. *J. Phys. Chem. B* **103**, 4508–4517 (1999).
58. Chen, B. et Siepmann, J. I. Transferable Potentials for Phase Equilibria. 3. Explicit-Hydrogen Description of Normal Alkanes. *J. Phys. Chem. B* **103**, 5370–5379 (1999).
59. Wick, C. D., Martin, M. G. et Siepmann, J. I. Transferable Potentials for Phase Equilibria. 4. United-Atom Description of Linear and Branched Alkenes and Alkylbenzenes. *J. Phys. Chem. B* **104**, 8008–8016 (2000).
60. Chen, B., Potoff, J. J., et Siepmann, J. I. Monte Carlo Calculations for Alcohols and Their Mixtures with Alkanes. Transferable Potentials for Phase Equilibria. 5. United-Atom Description of Primary, Secondary, and Tertiary Alcohols. *J. Phys. Chem. B* **105**, 3093–3104 (2001).
61. Stubbs, J. M., Potoff, J. J. et Siepmann, J. I. Transferable Potentials for Phase Equilibria. 6. United-Atom Description for Ethers, Glycols, Ketones, and Aldehydes. *J. Phys. Chem. B* **108**, 17596–17605 (2004).



62. Jungwirth, P. et Tobias, D. J. Ions at the Air/Water Interface. *J. Phys. Chem. B* **106**, 6361–6373 (2002).
63. Jungwirth, P. et Tobias, D. J. Specific Ion Effects at the Air/Water Interface. *Chem. Rev.* **106**, 1259–1281 (2006).
64. Dang, L. X. et Chang, T.-M. Molecular Mechanism of Ion Binding to the Liquid/Vapor Interface of Water. *J. Phys. Chem. B* **106**, 235–238 (2002).
65. Hu, J. H., Shi, Q., Davidovits, P., Worsnop, D. R., Zahniser, M. S. et Kolb, C. E. Reactive Uptake of Cl<sub>2</sub>(g) and Br<sub>2</sub>(g) by Aqueous Surfaces as a Function of Br- and I- Ion Concentration : The Effect of Chemical Reaction at the Interface. *J. Phys. Chem.* **99**, 8768–8776 (1995).
66. Liberman, Y. A. et Topaly, V. P. Permeability of bimolecular phospholipid membranes for fat-soluble ions. *Biofizika* **14**, 452–461 (1969).
67. Haydon, D. A. et Myers, V. B. Surface charge, surface dipoles and membrane conductance. *Biochim. Biophys. Acta* **307**, 429–443 (1973).
68. Gawrisch, K., Ruston, D., Zimmerberg, J., Parsegian, V. A., Rand, R. P. et Fuller, N. Membrane dipole potentials, hydration forces, and the ordering of water at membrane surfaces. *Biophys. J.* **61**, 1213–1223 (1992).
69. Schamberger, J. et Clarke, R. J. Hydrophobic ion hydration and the magnitude of the dipole potential. *Biophys. J.* **82**, 3081–3088 (2002).
70. Wang, L., Bose, P. S. et Sigworth, F. J. Using cryo-EM to measure the dipole potential of a lipid membrane. *Proc. Natl. Acad. Sci. USA* **103**, 18528–18533 (2006).
71. Brockman, H. Dipole potential of lipid membranes. *Chem. Phys. Lipids* **73**, 57–79 (1994).
72. Harder, E., MacKerell, A. D. Jr. et Roux, B. Many-Body Polarization Effects and the Membrane Dipole Potential. *J. Am. Chem. Soc.* **131**, 2760–2761 (2009).
73. Yu, H., Whitfield, T. W., Harder, E., Lamoureux, G., Vorobyov, I., Anisimov, V. M., MacKerell Jr., A. D. et Roux, B. Simulating Monovalent and Divalent Ions in Aqueous Solution Using a Drude Polarizable Force Field. *J. Chem. Theory Comput.* **6**, 774–786 (2010).
74. Davis, J. E. et Patel, S. Charge Equilibration Force Fields for Lipid Environments : Applications to Fully Hydrated DPPC Bilayers and DMPC-Embedded Gramicidin A. *J. Phys. Chem. B* **113**, 9183–9196 (2009).

75. Vorobyov, I. et Allen, T. W. The electrostatics of solvent and membrane interfaces and the role of electronic polarizability. *J. Chem. Phys.* **132**, 185101 (2010).
76. Xia, X.-R., Monteiro-Riviere, N. A. et Riviere, J. E. An index for characterization of nanomaterials in biological systems. *Nature Nanotechnology* **5**, 671–675 (2010).
77. Asuria, P., Balea, S. S., Karajanagia, S. S. et Kane, R. S. The protein-nanomaterial interface. *Current Opinion in Biotechnology* **17**, 562–568 (2006).
78. Kane, R. S. et Stroock, A. D. Nanobiotechnology : Protein-Nanomaterial Interactions. *Biotechnol. Prog.* **23**, 316–319 (2007).
79. Leach, A. *Molecular Modelling : Principles and Applications (2nd Edition)* (Prentice Hall, 2001).
80. Schwede, T. et Peitsch, M. C. *Computational Structural Biology : Methods and Applications* (World Scientific Publishing Company, 2008).
81. Muegge, I. et Rarey, M. Small molecule docking and scoring. *Rev. Comput. Chem.* **17**, 1–60 (2001).
82. Jiao, D., Golubkov, P. A., Darden, T. A. et Ren, P. Calculation of protein-ligand binding free energy by using a polarizable potential. *Proc. Natl. Acad. Sci. USA* **105**, 6290–6295 (2008).
83. Ren, P. et Ponder, J. W. Consistent treatment of inter- and intramolecular polarization in molecular mechanics calculations. *J. Comput. Chem.* **23**, 1497–1506 (2002).
84. Ren, P. et Ponder, J. W. Polarizable Atomic Multipole Water Model for Molecular Mechanics Simulation. *J. Phys. Chem. B* **107**, 5933–5947 (2003).
85. Ponder, J. W., Wu, C., Ren, P., Pande, V. S., Chodera, J. D., Schnieders, M. J., Haque, I., Mobley, D. L., Lambrecht, D. S., Robert A. DiStasio, J., Head-Gordon, M., Clark, G. N. I., Johnson, M. E. et Head-Gordon, T. Current Status of the AMOEBA Polarizable Force Field. *J. Phys. Chem. B* **114**, 2549–2564 (2010).
86. Gresh, N., Cisneros, G. A., Darden, T. A. et Piquemal, J.-P. Anisotropic, Polarizable Molecular Mechanics Studies of Inter- and Intramolecular Interactions and Ligand-Macromolecule Complexes. A Bottom-Up Strategy. *J. Chem. Theory Comput.* **3**, 1960–1986 (2007).

87. De Courcy, B., Piquemal, J.-P., Garbay, C. et Gresh, N. Polarizable Water Molecules in Ligand-Macromolecule Recognition. Impact on the Relative Affinities of Competing Pyrrolopyrimidine Inhibitors for FAK Kinase. *J. Am. Chem. Soc.* **132**, 3312–3320 (2010).
88. Gavezzotti, A. *Molecular Aggregation : Structure Analysis and Molecular Simulation of Crystals and Liquids* (éd. of Crystallography Monographs on Crystallography, I. U.) (Oxford University Press, USA, 2007).
89. Welch, G. W. A., Karamertzanis, P. G., Misquitta, A. J., Stone, A. J. et Price, S. L. Is the Induction Energy Important for Modeling Organic Crystals? *J. Chem. Theory Comput.* **4**, 522–532 (2008).
90. Misquitta, A. J., Welch, G. W., Stone, A. J. et Price, S. L. A first principles prediction of the crystal structure of C6Br2ClFH2. *Chem. Phys. Lett.* **456**, 105–109 (2008).
91. Misquitta, A. J. et Stone, A. J. Accurate Induction Energies for Small Organic Molecules : 1. Theory. *J. Chem. Theory Comput.* **4**, 7–18 (2008).
92. Misquitta, A. J., Stone, A. J. et Price, S. L. Accurate Induction Energies for Small Organic Molecules. 2. Development and Testing of Distributed Polarizability Models against SAPT(DFT) Energies. *J. Chem. Theory Comput.* **4**, 19–32 (2008).
93. Gavezzotti, A. Calculation of Intermolecular Interaction Energies by Direct Numerical Integration over Electron Densities. I. Electrostatic and Polarization Energies in Molecular Crystals. *J. Phys. Chem. B* **106**, 4145–4154 (2002).
94. Gavezzotti, A. Calculation of Intermolecular Interaction Energies by Direct Numerical Integration over Electron Densities. 2. An Improved Polarization Model and the Evaluation of Dispersion and Repulsion Energies. *J. Phys. Chem. B* **107**, 2344–2353 (2003).
95. Eckhardt, C. J. et Gavezzotti, A. Computer Simulations and Analysis of Structural and Energetic Features of Some Crystalline Energetic Materials. *J. Phys. Chem. B* **111**, 3430–3437 (2007).
96. Eisenman, G. et Horn, R. Ionic selectivity revisited : The role of kinetic and equilibrium processes in ion permeation through channels. *J. Membr. Biol.* **76**, 197–222 (1983).

97. Eisenman, G. et Alvarez, O. Structure and function of channels and channelogs as studied by computational chemistry. *J. Membr. Biol.* **119**, 109–132 (1991).
98. Luzhkov, V. B. et Åqvist, J. K<sup>+</sup>/Na<sup>+</sup> selectivity of the KcsA potassium channel from microscopic free energy perturbation calculations. *Biochim. Biophys. Acta* **1548**, 194–202 (2001).
99. Dudev, T. et Lim, C. Determinants of K<sup>+</sup> vs Na<sup>+</sup> Selectivity in Potassium Channels. *J. Am. Chem. Soc.* **131**, 8092–8101 (2009).
100. Eisenberg, B. Ion channels allow atomic control of macroscopic transport. *Phys. Stat. Solid.* **5**, 708–713 (2008).
101. Illingworth, C. J. et Domene, C. Many-body effects and simulations of potassium channels. *Proc. R. Soc. A* **465**, 1701–1716 (2009).
102. Patel, S., Davis, J. E. et Bauer, B. A. Exploring Ion Permeation Energetics in Gramicidin A Using Polarizable Charge Equilibration Force Fields. *J. Am. Chem. Soc.* **131**, 13890–13891 (2009).
103. Roux, B. et MacKinnon, R. The Cavity and Pore Helices in the KcsA K<sup>+</sup> Channel : Electrostatic Stabilization of Monovalent Cations. *Science* **285**, 100–102 (1999).
104. Burykin, A., Kato, M. et Warshel, A. Exploring the origin of the ion selectivity of the KcsA potassium channel. *Proteins : Struct., Funct., Genet.* **52**, 412–426 (2003).
105. Kato, M., Pisljakov, A. V. et Warshel, A. The barrier for proton transport in aquaporins as a challenge for electrostatic models : The role of protein relaxation in mutational calculations. *Proteins : Struct., Funct., Genet.* **64**, 829–844 (2006).
106. Chatelain, F. C., Alagem, N., Xu, Q., Pancaroglu, R., Reuveny, E. et Minor, D. L. The Pore Helix Dipole Has a Minor Role in Inward Rectifier Channel Function. *Neuron* **47**, 833–843 (2005).
107. Murata, K., Mitsuoka, K., Hirai, T., Walz, T., Agre, P., Heymann, J. B., Engel, A. et Fujiyoshi, Y. Structural determinants of water permeation through aquaporin-1. *Nature* **407**, 599–605 (2000).
108. Sui, H., Han, B.-G., Lee, J. K., Walian, P. et Jap, B. K. Structural basis of water-specific transport through the AQP1 water channel. *Nature* **414**, 872–878 (2001).
109. De Groot, B. L., Engel, A. et Grubmüller, H. A refined structure of human aquaporin-1. *FEBS Lett.* **504**, 206–211 (2001).

110. Agre, P. et Kozono, D. Aquaporin water channels : molecular mechanisms for human diseases. *FEBS Lett.* **555**, 72–78 (2003).
111. Wang, Y., Schulten, K. et Tajkhorshid, E. What makes an aquaporin a glycerol channel? A comparative study of AqpZ and GlpF. *Structure* **13**, 1107–1118 (2005).
112. Chen, H., Wu, Y. et Voth, G. A. Origins of Proton Transport Behavior from Selectivity Domain Mutations of the Aquaporin-1 Channel. *Biophys. J.* **90**, 73–75 (2006).
113. Hénin, J., Tajkhorshid, E., Schulten, K. et Chipot, C. Diffusion of Glycerol through Escherichia coli Aquaglyceroporin GlpF. *Biophys. J.* **94**, 832–839 (2008).
114. Nico, B. et Ribatti, D. Aquaporins in tumor growth and angiogenesis. *Cancer Lett.* **294**, 135–138 (2010).
115. Chen, L. Free-energy landscape of glycerol permeation through aquaglyceroporin GlpF determined from steered molecular dynamics simulations. *Biophys. Chem.* **151**, 178–180 (2010).
116. Oliva, R., Calamita, G., Thornton, J. M. et Pellegrini-Calace, M. Electrostatics of aquaporin and aquaglyceroporin channels correlates with their transport selectivity. *Proc. Natl. Acad. Sci. USA* **107**, 4135–4140 (2010).
117. Burykin, A. et Warshel, A. On the origin of the electrostatic barrier for proton transport in aquaporin. *FEBS Lett.* **570**, 41–46 (2004).
118. Jensen, M. Ø., Tajkhorshid, E. et Schulten, K. Electrostatic Tuning of Permeation and Selectivity in Aquaporin Water Channels. *Biophys. J.* **85**, 2884–2899 (2003).
119. Ilan, B., Tajkhorshid, E., Schulten, K. et Voth, G. A. The mechanism of proton exclusion in aquaporin channels. *Proteins : Struct., Funct., Bioinform.* **55**, 223–228 (2004).
120. Gresh, N., Stevens, W. J. et Krauss, M. Mono- and poly-ligated complexes of Zn<sup>2+</sup> : An ab initio analysis of the metal-ligand interaction energy. *J. Comput. Chem.* **16**, 843–855 (1995).
121. Gresh, N. et Garmer, D. R. Comparative binding energetics of Mg<sup>2+</sup>, Ca<sup>2+</sup>, Zn<sup>2+</sup>, and Cd<sup>2+</sup> to biologically relevant ligands : Combined ab initio SCF supermolecule and molecular mechanics investigation. *J. Comput. Chem.* **17**, 1481–1495 (1996).

122. Garmer, D. R., Gresh, N. et Roques, B.-P. Modeling of inhibitor-metalloenzyme interactions and selectivity using molecular mechanics grounded in quantum chemistry. *Proteins : Struct., Funct., Bioinform.* **31**, 42–60 (1998).
123. Gresh, N. et Spomer, J. Complexes of Pentahydrated Zn<sup>2+</sup> with Guanine, Adenine, and the Guanine-Cytosine and Adenine-Thymine Base Pairs. Structures and Energies Characterized by Polarizable Molecular Mechanics and ab Initio Calculations. *J. Phys. Chem. B* **103**, 11415–11427 (1999).
124. Tiraboschi, G., Gresh, N., Giessner-Prettre, C., Pedersen, L. G. et Deerfield, D. W. Parallel ab initio and molecular mechanics investigation of polycoordinated Zn(II) complexes with model hard and soft ligands : Variations of binding energy and of its components with number and charges of ligands. *J. Comput. Chem.* **21**, 1011–1039 (2000).
125. Tiraboschi, G., Fournié-Zaluski, M.-C., Roques, B.-P. et Gresh, N. Intramolecular chelation of Zn<sup>2+</sup> by  $\alpha$ - and  $\beta$ -mercaptocarboxamides. A parallel ab initio and polarizable molecular mechanics investigation. Assessment of the role of multipole transferability. *J. Comput. Chem.* **22**, 1038–1047 (2001).
126. Gresh, N. et Derreumaux, P. Generating Conformations for Two Zinc-Binding Sites of HIV-1 Nucleocapsid Protein from Random Conformations by a Hierarchical Procedure and Polarizable Force Field. *J. Phys. Chem. B* **107**, 4862–4870 (2003).
127. Shaw, D. E. *et al.* dans *Conference on High Performance Networking and Computing* (2009).
128. Freddolino, P. L., Liu, F., Gruebele, M. et Schulten, K. Ten-Microsecond Molecular Dynamics Simulation of a Fast-Folding WW Domain. *Biophys. J.* **94**, 75–77 (2008).
129. Freddolino, P. L., Park, S., Roux, B. et Schulten, K. Force field bias in protein folding simulations. *Biophys. J.* **96**, 3772–3780 (2009).
130. Freddolino, P. L., Harrison, C. B., Liu, Y. et Schulten, K. Challenges in protein folding simulations : Force field, timescale, sampling, and analysis. *Nature Physics* **In press** (2010).
131. Nguyen, H., Jäger, M., Kelly, J. W. et Gruebele, M. Engineering a  $\beta$ -Sheet Protein toward the Folding Speed Limit. *J. Phys. Chem. B* **109**, 15182–15186 (2005).
132. Lei, H. et Duan, Y. Two-stage Folding of HP-35 from Ab Initio Simulations. *J. Mol. Biol.* **370**, 196–206 (2007).

133. Xu, Z., Luo, H. H. et Tieleman, D. P. Modifying the OPLS-AA force field to improve hydration free energies for several amino acid side chains using new atomic charges and an off-plane charge model for aromatic residues. *J. Comput. Chem.* **28**, 689–697 (2007).
134. Clementi, C. Coarse-grained models of protein folding : toy models or predictive tools? *Current Opinion in Structural Biology* **18**, 10–15 (2008).
135. Ababneh, A. M. The role of polarization interactions in the wrapping/unwrapping of nucleosomal DNA around the histone octamer : Implications to gene regulation. *J. Theo. Bio.* **258**, 229–239 (2009).
136. Halgren, T. A. et Damm, W. Polarizable force fields. *Current Opinion in Structural Biology* **11**, 236–242 (2001).
137. Rick, S. W. et Stuart, S. J. Potentials and Algorithms for Incorporating Polarizability in Computer Simulations. *Rev. Comput. Chem.* **18**, 89–146 (2003).
138. Masia, M., Probst, M. et Rey, R. On the performance of molecular polarization methods. I. Water and carbon tetrachloride close to a point charge. *J. Chem. Phys.* **121**, 7362 (2004).
139. Yu, H. et van Gunsteren, W. F. Accounting for polarization in molecular simulation. *Comp. Phys. Comm.* **172**, 69–85 (2005).
140. Warshel, A., Kato, M. et Pislakov, A. V. Polarizable Force Fields : History, Test Cases, and Prospects. *J. Chem. Theory Comput.* **3**, 2034–2045 (2007).
141. Jorgensen, W. L. Special issue on polarization. *J. Chem. Theory. Comput.* **3**, 1877 (2007).
142. Cieplak, P., Dupradeau, F.-Y., Duan, Y. et Wang, J. Polarization effects in molecular mechanical force fields. *J. Phys. : Condens. Matter* **21**, 333102–333123 (2009).
143. Lopes, P. E. M., Roux, B. et MacKerell, A. D. Jr. Molecular modeling and dynamics studies with explicit inclusion of electronic polarizability : theory and applications. *Theor. Chem. Acc.* **124**, 11–28 (2009).
144. Oxtoby, D. W. Local polarization theory for field-induced molecular multipoles. *J. Chem. Phys.* **72**, 5171 (1980).
145. Lamoureux, G., MacKerell, A. D. Jr. et Roux, B. A simple polarizable model of water based on classical Drude oscillators. *J. Chem. Phys.* **119**, 5185 (2003).

146. Mortier, W. J., Ghosh, S. K. et Shankar, S. Electronegativity Equalization Method for the Calculation of Atomic Charges in Molecules. *J. Am. Chem. soc.* **108**, 4315–4320 (1986).
147. Rappé, A. K. et Goddard, W. A. Charge equilibration for molecular dynamics simulations. *J. Phys. Chem.* **95**, 3358–3363 (1991).
148. London, F. On the theory and system of molecular forces. *Z. Phys.* **63**, 369–397 (1930).
149. London, F. The general theory of molecular forces. *Trans. Faraday Soc.* **33**, 8–26 (1937).
150. Margeneau, H. Van der Waals forces. *Rev. Mod. Phys.* **11**, 1–35 (1939).
151. Longet-Higgins, H. The electronic state of composite systems. *Proc. R. Soc. A* **235**, 537–543 (1956).
152. Buckingham, A. D. Permanent and induced molecular moments and long-range intermolecular forces. *Adv. Chem. Phys.* **12**, 107–143 (1967).
153. Hirschfelder, J. O. Perturbation theory for exchange forces, I. *Chem. Phys. Lett.* **1**, 325–329 (1967).
154. Hirschfelder, J. O. Perturbation theory for exchange forces, II. *Chem. Phys. Lett.* **1**, 363–368 (1967).
155. Møller, C. et Plesset, M. S. Note on an Approximation Treatment for Many-Electron Systems. *Phys. Rev.* **46**, 618–622 (1934).
156. Patkowski, K., Szalewicz, K. et Jeziorski, B. Third-order interactions in symmetry-adapted perturbation theory. *J. Chem. Phys.* **125**, 154107 (2006).
157. Rybak, S., Szalewicz, K., Jeziorski, B. et Jaszunski, M. Intraatomic correlation effects for the He-He dispersion and exchange-dispersion energies using explicitly correlated Gaussian geminals. *J. Chem. Phys.* **86**, 5652 (1987).
158. Fetter, A. L. et Walecka, J. D. *Quantum Theory of Many-Particle Systems Book Description* (éd. Francisco, M.-H. S.) (Dover Publications, 1971).
159. Onida, G., Reining, L. et Rubio, A. Electronic excitations : density-functional versus many-body Green's-function approaches. *Rev. Mod. Phys.* **74**, 601–659 (2002).



160. Murdachaew, G., Szalewicz, K. et Bukowski, R. Efficient Generation of Flexible-Monomer Intermolecular Potential Energy Surfaces. *Phys. Rev. Lett.* **88**, 123202 (2002).
161. Stone, A. J. et Misquitta, A. J. Atom-atom potentials from ab initio calculations. *Int. Rev. Phys. Chem.* **26**, 193–222 (2007).
162. Dehez, F., Ángyán, J. G., Gutiérrez, I. S., Luque, F. J., Schulten, K. et Chipot, C. Modeling Induction Phenomena in Intermolecular Interactions with an Ab Initio Force Field. *J. Chem. Theory Comput.* **3**, 1914–1926 (2007).
163. Dehez, F., Archambault, F., Gutiérrez, I. S., Luque, F. J. et Chipot, C. An ab initio strategy for handling induction phenomena in metal ion complexes. *Mol. Phys.* **106**, 1685–1696 (2008).
164. Archambault, F., Chipot, C., Soteras, I., Luque, F. J., Schulten, K. et Dehez, F. Polarizable Intermolecular Potentials for Water and Benzene Interacting with Halide and Metal Ions. *J. Chem. Theory Comput.* **5**, 3022–3031 (2009).
165. Kim, D., Tarakeshwar, P. et Kim, K. S. Theoretical Investigations of Anion- $\pi$  Interactions : The Role of Anions and the Nature of  $\pi$  Systems. *J. Phys. Chem. A* **108**, 1250–1258 (2004).
166. Podeszwa, R., Bukowski, R. et Szalewicz, K. Potential Energy Surface for the Benzene Dimer and Perturbational Analysis of  $\pi$  - $\pi$  Interactions. *J. Phys. Chem. A* **110**, 10345–10354 (2006).
167. Smith, T., Slipchenko, L. V. et Gordon, M. S. Modeling  $\pi$  - $\pi$  Interactions with the Effective Fragment Potential Method : The Benzene Dimer and Substituents. *J. Phys. Chem. A* **112**, 5286–5294 (2008).
168. Soteras, I., Orozco, M. et Luque, F. J. Induction effects in metal cation-benzene complexes. *Phys. Chem. Chem. Phys.* **10**, 2616–2624 (2008).
169. Jenness, G. R. et Jordan, K. D. DF-DFT-SAPT Investigation of the Interaction of a Water Molecule to Coronene and Dodecabenzocoronene : Implications for the Water-Graphite Interaction. *J. Phys. Chem. C* **113**, 10242–10248 (2009).
170. Geng, Y., Takatani, T., Hohenstein, E. G. et Sherrill, C. D. Accurately Characterizing the  $\pi$  - $\pi$  Interaction Energies of Indole-Benzene Complexes. *J. Phys. Chem. A* **114**, 3576–3582 (2010).

171. Quinonero, D., Deya, P. M., Carranza, M. P., Rodriguez, A. M., Jalon, F. A. et Manzano, B. R. Experimental and computational study of the interplay between C-H/ $\pi$  and anion- $\pi$  interactions. *Dalton Trans.* **39**, 794–806 (2010).
172. Stone, A. J. et Misquitta, A. J. Charge-transfer in Symmetry-Adapted Perturbation Theory. *Chem. Phys. Lett.* **473**, 201–205 (2009).
173. Piquemal, J.-P., Marquez, A., Parisel, O. et Giessner-Prettre, C. A CSOV study of the difference between HF and DFT intermolecular interaction energy values : The importance of the charge transfer contribution. *J. Comput. Chem.* **Volume 26**, 1052–1062 (2005).
174. Sakharov, D. V. et Lim, C. Zn Protein Simulations Including Charge Transfer and Local Polarization Effects. *J. Am. Chem. Soc.* **127**, 4921–4929 (2005).
175. Lee, Y.-m. et Lim, C. Physical Basis of Structural and Catalytic Zn-Binding Sites in Proteins. *J. Mol. Biol.* **379**, 545–553 (2008).
176. Sakharov, D. V. et Lim, C. Force fields including charge transfer and local polarization effects : Application to proteins containing multi/heavy metal ions. *J. Comput. Chem.* **30**, 191–202 (2009).
177. Ponder, J. W. et Case, D. A. Force Fields for Protein Simulation. *Adv. Prot. Chem.* **66**, 27–85 (2003).
178. Chipot, C., Ángyán, J. G., Maignret, B. et Scheraga, H. A. Modeling amino acid side chains. 3. Influence of intra- and intermolecular environment on point charges. *J. Phys. Chem.* **97**, 9797–9807 (1993).
179. Chipot, C., Maignret, B., Rivail, J. L. et Scheraga, H. A. Modeling amino acid side chains. 1. Determination of net atomic charges from ab initio self-consistent-field molecular electrostatic properties. *J. Chem. Phys.* **96**, 10276–10284 (1993).
180. Cox, S. R. et Williams, D. E. Representation of the molecular electrostatic potential by a net atomic charge model. *J. Comput. Chem.* **2**, 304–323 (1981).
181. Stone, A. Distributed multipole analysis, or how to describe a molecular charge distribution. *Chem. Phys. Lett.* **83**, 233–239 (1981).
182. Bader, R. A quantum theory of molecular structure and its applications. *Chem. Rev.* **91**, 893–928 (1991).
183. Dang, L. X., Rice, J. E., Caldwell, J. et Kollman, P. A. Ion solvation in polarizable water : molecular dynamics simulations. *J. Am. Chem. Soc.* **113**, 2481–2486 (1991).

184. Karim, O. A. Simulation of an anion in water : effect of ion polarizability. *Chem. Phys. Lett.* **184**, 560–565 (1991).
185. Dang, L. X. Development of Nonadditive Intermolecular Potentials Using Molecular Dynamics : Solvation of Li<sup>+</sup> and F<sup>-</sup> Ions in Polarizable Water. *J. Chem. Phys.* **96**, 6970–6977 (1992).
186. Soetens, J.-C., Millot, C., Chipot, C., Jansen, G., Ángyán, J. et Maigret, B. Effect of polarizability on the potential of mean force of two cations. The guanidinium-guanidinium ion pair in water. *J. Phys. Chem. B* **101**, 10910–10917 (1997).
187. Carignano, M. A., Karlström, G. et Linse, P. Polarizable Ions in Polarizable Water : A Molecular Dynamics Study. *J. Phys. Chem. B* **107**, 1142–1147 (1997).
188. Chang, T.-M. et Dang, L. X. Ion Solvation in Polarizable Chloroform : A Molecular Dynamics Study. *J. Phys. Chem. B* **101**, 10518–10526 (1997).
189. Chang, T.-M. et Dang, L. X. Detailed Study of Potassium Solvation Using Molecular Dynamics Techniques. *J. Phys. Chem. B* **103**, 4714–4720 (1999).
190. Stone, A. J. Distributed polarizabilities. *Mol. Phys.* **56**, 1065–1082 (1985).
191. Le Sueur, C. R. et Stone, A. J. Localization methods for distributed polarizabilities. *Mol. Phys.* **83**, 293–307 (1994).
192. Bader, R. F. W. Atoms in molecules in external fields. *J. Chem. Phys.* **91**, 6989 (1989).
193. Bader, R. F. W., Keith, T. A., Gough, K. M. et Laidig, K. E. Properties of atoms in molecules : additivity and transferability of group polarizabilities. *Mol. Phys.* **75**, 1167–1189 (1992).
194. Hättig, C., Hebss, B. A., Jansen, G. et Ángyán, J. G. Topologically partitioned dynamic polarizabilities using the theory of atoms in molecules. *Can. J. Chem.* **74**, 976–987 (1996).
195. Dehez, F., Chipot, C., Millot, C. et Ángyán, J. G. Fast and accurate determination of induction energies : reduction of topologically distributed polarizability models. *Chem. Phys. Lett.* **338**, 180–188 (2001).
196. Wang, W. et Skeel, R. D. Fast evaluation of polarizable forces. *J. Chem. Phys.* **123**, 164107 (2005).

197. Lopes, P. E., Harder, E., Roux, B. et Mackerell, A. D. Formalisms for the Explicit Inclusion of Electronic Polarizability in Molecular Modeling and Dynamics Studies. *Challenges and Advances in Computational Chemistry and Physics* **7**, 219–257 (2009).
198. Applequist, J., Carl, J. R. et Fung, K.-K. An Atom dipole interaction model for molecular polarizability. Application to polyatomic molecules and determination of atom polarizabilities. *J. Am. Chem. Soc.* **94**, 2952–2960 (1972).
199. Thole, B. T. Molecular polarizabilities calculated with a modified dipole interaction. *Chem. Phys.* **59**, 341–350 (1981).
200. Jensen, L., Åstrand, P.-O., Sylvester-Hvid, K. O. et Mikkelsen, K. V. Frequency-Dependent Molecular Polarizability Calculated within an Interaction Model. *J. Phys. Chem. A* **104**, 1563–1569 (2000).
201. Tang, K. T. et Toennies, J. P. A simple theoretical model for the van der Waals potential at intermediate distances. I. Spherically symmetric potentials. *J. Chem. Phys.* **66**, 1496 (1977).
202. Tang, K. T. et Toennies, J. P. Erratum : A simple theoretical model for the van der Waals potential at intermediate distances. I. Spherically symmetric potentials. *J. Chem. Phys.* **67** and **68**, 375 and 786 (1977).
203. Tang, K. T. et Toennies, J. P. An improved simple model for the van der Waals potential based on universal damping functions for the dispersion coefficients. *J. Chem. Phys.* **80**, 3726 (1984).
204. Masia, M., Probst, M. et Rey, R. Polarization damping in halide-water dimers. *Chem. Phys. Lett.* **420**, 267–270 (2006).
205. Elking, D., Darden, T. et Woods, R. J. Gaussian induced dipole polarization model. *J. Comput. Chem.* **28**, 1261–1274 (2007).
206. Masella, M. et Flament, J. P. A pairwise and two many-body models for water : Influence of nonpairwise effects upon the stability and geometry of (H<sub>2</sub>O)<sub>n</sub> cyclic (n = 3-6) and cagelike (n = 6-20) clusters. *J. Chem. Phys.* **107**, 9105 (1997).
207. Masella, M. et Cuniasse, P. A many-body model to study proteins. I. Applications to ML<sub>n</sub><sup>m+</sup> complexes, Mm<sup>+</sup> = Li<sup>+</sup>, Na<sup>+</sup>, K<sup>+</sup>, Mg<sup>2+</sup>, Ca<sup>2+</sup>, and Zn<sup>2+</sup>, L = H<sub>2</sub>O, CH<sub>3</sub>OH, HCONH<sub>2</sub>, n = 1–6, and to small hydrogen bonded systems. *J. Chem. Phys.* **119**, 1866–1873 (2003).

208. Masella, M. The multiple time step r-RESPA procedure and polarizable potentials based on induced dipole moments. *Mol. Phys.* **104**, 415–428 (2006).
209. Masella, M., Borgis, D. et Cuniasse, P. Combining a polarizable force-field and a coarse-grained polarizable solvent model : Application to long dynamics simulations of Bovine Pancreatic Trypsin Inhibitor. *J. Comput. Chem.* (2007).
210. Houriez, C., Ferré, N., Masella, M. et Siri, D. Prediction of nitroxide hyperfine coupling constants in solution from combined nanosecond scale simulations and quantum computations. *J. Chem. Phys.* **128**, 244504 (2008).
211. Masella, M., Borgis, D. et Cuniasse, P. Combining a polarizable force-field and a coarse-grained polarizable solvent model : Application of long dynamics simulations of bovine pancreatic trypsin inhibitor. *J. Comp Chem.* **29**, 1707 (2008).
212. Houriez, C., Ferré, N., Siri, D. et Masella, M. Further Insights into the Environmental Effects on the Computed Hyperfine Coupling Constants of Nitroxides in Aqueous Solution. *J. Phys. Chem. B* **113**, 15047–15056 (2009).
213. Houriez, C., Ferré, N., Masella, M. et Siri, D. Quantitative evaluation of the aqueous dihydronitroxide nitrogen hyperfine coupling constant from QM/MM//MD computations. *J. Mol. Struct. : THEOCHEM* **898**, 49–55 (2009).
214. Baker, C. M. et Jr., A. D. M. Polarizability rescaling and atom-based Thole scaling in the CHARMM Drude polarizable force field for ethers. *J. Mol. Model* **16**, 567–576 (2009).
215. Fanourgakis, G. S., Markland, T. E. et Manolopoulos, D. E. A fast path integral method for polarizable force fields. *J. Chem. Phys.* **131**, 094102–1–094102–10 (2009).
216. Souaille, M., Loirat, H., Borgis, D. et Gageot, M. MDVRY : a polarizable classical molecular dynamics package for biomolecules. *Comp. Phys. Comm.* **180**, 276–301 (2009).
217. Xie, W., Pu, J. et Gao, J. A Coupled Polarization-Matrix Inversion and Iteration Approach for Accelerating the Dipole Convergence in a Polarizable Potential Function. *J. Phys. Chem. A* **12**, 2109–2116 (2009).
218. Baker, C. M. et MacKerell Jr., A. D. Polarizability rescaling and atom-based Thole scaling in the CHARMM Drude polarizable force field for ethers. *J. Mol. Mod.* **16**, 567–576 (2010).

219. Zhu, X. et Mackerell Jr., A. D. Polarizable empirical force field for sulfur-containing compounds based on the classical Drude oscillator model. *J. Comput. Chem.* **31**, 2330–2341 (2010).
220. Kumar, R., Wang, F.-F., Jenness, G. R. et Jordan, K. D. A second generation distributed point polarizable water model. *J. Chem. Phys.* **132**, 014309 (2010).
221. Jensen, L., Åstrand, P.-O., Osted, A., Kongsted, J. et Mikkelsen, K. V. Polarizability of molecular clusters as calculated by a dipole interaction model. *J. Chem. Phys.* **116**, 4001–4010 (2002).
222. Jensen, L., Swart, M., van Duijnen, P. T. et Snijders, J. G. Medium perturbations on the molecular polarizability calculated within a localized dipole interaction model. *J. Chem. Phys.* **117**, 3316 (2002).
223. Jensen, L., van Duijnen, P. T. et Snijders, J. G. A discrete solvent reaction field model within density functional theory. *J. Chem. Phys.* **118**, 514 (2003).
224. Masia, M., Probst, M. et Rey, R. On the performance of molecular polarization methods. II. Water and carbon tetrachloride close to a cation. *J. Chem. Phys.* **123**, 164505 (2005).
225. Holt, A. et Karlström, G. An intramolecular induction correction model of the molecular dipole moment. *J. Comp Chem.* **29**, 1084–1091 (2007).
226. Mayer, A., Gonzalez, A., Aikens, C. et Schatz, G. A charge-dipole interaction model for the frequency-dependent polarizability of silver clusters. *Nanotechnology* **20**, 195204 (2009).
227. Zhao, K., Yu, Y. et Luo, Y. Molecular dynamics and quantum chemistry study on conformations and optical properties of hydrogen bonded dipolar merocyanine dyes. *J. Phys. Chem. B* **113**, 10271–10276 (2009).
228. Millot, C. et Stone, A. J. Towards an accurate intermolecular potential for water. *Mol. Phys.* **77**, 439–462 (1992).
229. Meredith, A. W. et Stone, A. J. An ab Initio and Diffusion Monte Carlo Study of the Potential Energy Surface of the CO Dimer. *J. Phys. Chem. A* **102**, 434–445 (1998).
230. Millot, C., Soetens, J.-C. et Martins Costa, M. T. C. Revised Anisotropic Site Potentials for the Water Dimer and Calculated Properties. *J. Phys. Chem. A* **102**, 754–770 (1998).

231. Hermida-Ramón, J. M. et Ríos, M. A. A New Intermolecular Polarizable Potential for a Formaldehyde Dimer. Application to Liquid Simulations. *J. Phys. Chem. A* **102**, 10818–10827 (1998).
232. Hermida-Ramón, J. M. et Ríos, M. A. An ab initio polarizable intermolecular potential for dimethyl ether : application to liquid simulations. *Chem. Phys.* **262**, 423–436 (2000).
233. Hodges, M. P. et Stone, A. J. A new representation of the dispersion interaction. *Mol. Phys.* **98**, 275–286 (2000).
234. Patil, S. H., Tang, K. T. et Toennies, J. P. Damping functions for the pairwise sum model of the atom-surface potential. *J. Chem. Phys.* **116**, 8118 (2002).
235. Mas, E. M., Bukowski, R. et Szalewicz, K. Ab initio three-body interactions for water. I. Potential and structure of water trimer. *J. Chem. Phys.* **118**, 4386 (2003).
236. Karamertzanis, P. G. et Pantelides, C. C. Optimal Site Charge Models for Molecular Electrostatic Potentials. *Mol. Sim.* **30**, 413–436 (2004).
237. Giese, T. J. et York, D. M. Charge-dependent model for many-body polarization, exchange, and dispersion interactions in hybrid quantum mechanical/molecular mechanical calculations. *J. Chem. Phys.* **127**, 194101 (2007).
238. Karamertzanis, P. G., Day, G. M., Welch, G. W. A., Kendrick, J., Leusen, F. J. J., Neumann, M. A., et Price, S. L. Modeling the interplay of inter- and intramolecular hydrogen bonding in conformational polymorphs. *J. Chem. Phys.* **128**, 244708 (2008).
239. Bukowski, R., Szalewicz, K., Groenenboom, G. C. et van der Avoird, A. Polarizable interaction potential for water from coupled cluster calculations. I. Analysis of dimer potential energy surface. *J. Chem. Phys.* **128**, 094313 (2008).
240. Bukowski, R., Szalewicz, K., Groenenboom, G. C. et van der Avoird, A. Polarizable interaction potential for water from coupled cluster calculations. II. Applications to dimer spectra, virial coefficients, and simulations of liquid water. *J. Chem. Phys.* **128**, 094314 (2008).
241. Cooper, T. G., Hejczyk, K. E., Jones, W. et Day, G. M. Molecular Polarization Effects on the Relative Energies of the Real and Putative Crystal Structures of Valine. *J. Chem. Theory Comput.* **4**, 1795–1805 (2008).

242. Tang, K. T. et Toennies, J. P. The dynamical polarisability and van der Waals dimer potential of mercury. *Mol. Phys.* **106**, 1645–1653 (2008).
243. Rick, S. W., Stuart, S. J. et Berne, B. J. Dynamical fluctuating charge force fields : Application to liquid water. *J. Chem. Phys.* **101**, 6141–6156 (1994).
244. Patel, S. et Brooks III, C. L. CHARMM fluctuating charge force field for proteins : I parameterization and application to bulk organic liquid simulations. *J. Comput. Chem.* **25**, 1–15 (2004).
245. York, D. M. et Yang, W. A chemical potential equalization method for molecular simulations. *J. Chem. Phys.* **104**, 159 (1996).
246. Nalewajski, R. F. Charge sensitivities of the externally interacting open reactants. *Int. J. Quantum Chem.* **78**, 168–178 (2000).
247. Llanta, E., Ando, K. et Rey, R. Fluctuating Charge Study of Polarization Effects in Chlorinated Organic Liquids. *J. Phys. Chem. B* **105**, 7783–7791 (2001).
248. Smith, P. E. Local Chemical Potential Equalization Model for Cosolvent Effects on Biomolecular Equilibria. *J. Phys. Chem. B* **108**, 16271–16278 (2004).
249. Chelli, R., Barducci, A., Bellucci, L., Schettino, V. et Procacci, P. Behavior of polarizable models in presence of strong electric fields. I. Origin of nonlinear effects in water point-charge systems. *J. Chem. Phys.* **123**, 194109 (2005).
250. Medeiros, M. Monte Carlo simulation of polarizable systems : Early rejection scheme for improving the performance of adiabatic nuclear and electronic sampling Monte Carlo simulations. *Theor. Chem. Acc.* **113**, 178–182 (2005).
251. Piquemal, J.-P., Chelli, R., Procacci, P. et Gresh, N. Key Role of the Polarization Anisotropy of Water in Modeling Classical Polarizable Force Fields. *J. Phys. Chem. A* **111**, 8170–8176 (2007).
252. Warren, G. L., Davis, J. E. et Patel, S. Origin and control of superlinear polarizability scaling in chemical potential equalization methods. *J. Chem. Phys.* **128**, 144110 (2008).
253. Zhang, Y. et Lin, H. Flexible-Boundary Quantum-Mechanical/Molecular-Mechanical Calculations : Partial Charge Transfer between the Quantum-Mechanical and Molecular-Mechanical Subsystems. *J. Chem. Theory Comput.* **4**, 414–425 (2008).



254. Sefcik, J., Demiralp, E., Cagin, T. et Goddard III, W. A. Dynamic Charge Equilibration-Morse stretch force field : Application to energetics of pure silica zeolites. *J. Comput. Chem.* **23**, 1507–1514 (2002).
255. Kitao, O. et Ogawa, T. Consistent charge equilibration (CQEq). *Mol. Phys.* **101**, 3–17 (2003).
256. Nistor, R. A., Polihronov, J. G., Möser, M. H. et Mosey, N. J. A generalization of the charge equilibration method for nonmetallic materials. *J. Chem. Phys.* **125**, 094108 (2006).
257. Ogawa, T., Kurita, N., Sekino, H., Kitao, O. et Tanaka, S. Consistent charge equilibration (CQEq) method : application to amino acids and crambin protein. *Chem. Phys. Lett.* **397**, 382–387 (2004).
258. Brodersen, S., Wilke, S., Leusen, F. J. J. et Engel, G. E. Comparison of Static and Fluctuating Charge Models for Force-Field Methods Applied to Organic Crystals. *Cryst. Growth Des.* **5**, 925–933 (2005).
259. Tanaka, M. et Siehl, H.-U. An application of the consistent charge equilibration (CQEq) method to guanidinium ionic liquid systems. *Chem. Phys. Lett.* **457**, 263–266 (2008).
260. Mulliken, R. S. A New Electroaffinity Scale ; Together with Data on Valence States and on Valence Ionization Potentials and Electron Affinities. *J. Chem. Phys.* **2**, 782–793 (1934).
261. Parr, R. G. et Pearson, R. G. Absolute hardness : companion parameter to absolute electronegativity. *J. Am. Chem. Soc.* **105**, 7512–7516 (1983).
262. Garmer, D. R. et Stevens, W. J. Transferability of molecular distributed polarizabilities from a simple localized orbital based method. *J. Chem. Phys.* **93**, 8263 (1989).
263. Nakagawa, S. Polarizable Model Potential Function for Ion-Methanol Systems. *J. Phys. Chem. A* **104**, 5281–5290 (2000).
264. Giese, T. J. et York, D. M. Improvement of semiempirical response properties with charge-dependent response density. *J. Chem. Phys.* **123**, 164108 (2005).
265. Piquemal, J.-P., Cisneros, G. A., Reinhardt, P., Gresh, N. et Darden, T. A. Towards a force field based on density fitting. *J. Chem. Phys.* **124**, 12 (2006).

266. Bagus, P. S., Hermann, K. et Charles W. Bauschlicher, J. A new analysis of charge transfer and polarization for ligand-metal bonding : Model studies of  $\text{Al}_4\text{CO}$  and  $\text{Al}_4\text{NH}_3$ . *J. Chem. Phys.* **80**, 4378 (1984).
267. Cheng, Y.-H., Liu, L., Fu, Y., Chen, R., Li, X.-S. et Guo, Q.-X. Counterion Effects on the Cation- $\pi$  Interaction between Alkaline Earth Cations and Benzene. *J. Phys. Chem. A* **106**, 11215–11220 (2002).
268. Gresh, N., Polcar, C. et Giessner-Prettre, C. Modeling Copper(I) Complexes : SIBFA Molecular Mechanics versus ab Initio Energetics and Geometrical Arrangements. *J. Phys. Chem. A* **106**, 5660–5670 (2002).
269. Karpfen, A. The intermolecular interaction in the charge-transfer complexes between amines and halogens : A theoretical characterization of the trends in halogen bonding. *Theoretical Chemistry Accounts : Theory, Computation, and Modeling (Theoretica Chimica Acta)* **110**, 1–9 (2003).
270. Peraroa, M. D., Spiegel, K., Lamoureux, G., Vivo, M. D., DeGrado, W. F. et Klein, M. L. Modeling the charge distribution at metal sites in proteins for molecular dynamics simulations. *J. Struct. Biol.* **157**, 444–453 (2007).
271. Faralli, C., Pagliai, M., Cardini, G. et Schettino, V. Ab Initio Molecular Dynamics Study of  $\text{Mg}^{2+}$  and  $\text{Ca}^{2+}$  Ions in Liquid Methanol. *J. Chem. Theory Comput.* **4**, 156–163 (2008).
272. Zhao, Y., Schultz, N. E. et Truhlar, D. G. Design of Density Functionals by Combining the Method of Constraint Satisfaction with Parametrization for Thermochemistry, Thermochemical Kinetics, and Noncovalent Interactions. *J. Chem. Theory Comput.* **2**, 364–382 (2006).
273. Chelli, R., Procacci, P., Righini, R. et Califano, S. Electrical response in chemical potential equalization schemes. *J. Chem. Phys.* **111**, 8569 (1999).
274. Drude, P. *The theory of optics* (éds. London et Bombay) (Longmans, Green, et Co., 1902).
275. Yu, H., Hansson, T. et van Gunsteren, W. F. Development of a simple, self-consistent polarizable model for liquid water. *J. Chem. Phys.* **118**, 221–234 (2003).
276. Yu, H., Geerke, D. P., Liu, H. et van Gunsteren, W. F. Molecular dynamics simulations of liquid methanol and methanol-water mixtures with polarizable models. *J. Comput. Chem.* **27**, 1494–1504 (2006).

277. Darden, T., York, D. et Pedersen, L. Particle mesh Ewald : An  $N^*\log(N)$  method for Ewald sums in large systems. *J. Chem. Phys.* **98**, 10089–10092 (1993).
278. Essmann, U., Perera, L., Berkowitz, M. L., Darden, T., Lee, H. et Pedersen, L. G. A smooth particle mesh Ewald method. *J. Chem. Phys.* **103**, 8577–8593 (1995).
279. Smith, W. Point Multipoles in the Ewald Summation (Revisited). *CCP5 Newsletter* **46**, 18–30 (1998).
280. Toukmaji, A., Sagui, C., Board, J. et Darden, T. Efficient particle-mesh Ewald based approach to fixed and induced dipolar interactions. *J. Chem. Phys.* **113**, 10913 (2000).
281. Patel, S., Mackerell, A. D. Jr. et Brooks III, C. L. CHARMM fluctuating charge force field for proteins : II Protein/solvent properties from molecular dynamics simulations using a nonadditive electrostatic model. *J. Comput. Chem.* **25**, 1504–1514 (2004).
282. Lamoureux, G., Harder, E., Vorobyov, I. V., Roux, B. et Alexander D. MacKerell, J. A polarizable model of water for molecular dynamics simulations of biomolecules. *Chem. Phys. Lett.* **418**, 245–249 (2006).
283. Anisimov, V. M., Lamoureux, G., Vorobyov, I. V., Huang, N., Roux, B. et MacKerell Jr., A. D. Determination of Electrostatic Parameters for a Polarizable Force Field Based on the Classical Drude Oscillator. *J. Chem. Theory Comput.* **1**, 153–168 (2005).
284. Harder, E., Anisimov, V. M., Vorobyov, I. V., Lopes, P. E. M., Noskov, S. Y., MacKerell Jr., A. D. et Roux, B. Atomic Level Anisotropy in the Electrostatic Modeling of Lone Pairs for a Polarizable Force Field Based on the Classical Drude Oscillator. *J. Chem. Theory Comput.* **2**, 1587–1597 (2006).
285. Anisimov, V. M., Lamoureux, G., Vorobyov, I. V., Huang, N., Roux, B. et MacKerell Jr., A. D. Determination of Electrostatic Parameters for a Polarizable Force Field Based on the Classical Drude Oscillator. *J. Chem. Theor. Comput.* **1**, 153–168 (2004).
286. Anisimov, V. M., Vorobyov, I. V., Roux, B. et Mackerell Jr., A. D. Polarizable Empirical Force Field for the Primary and Secondary Alcohol Series Based on the Classical Drude Model. *J. Chem. Theory Comput.* **3**, 1927–1946 (2007).

287. Lopes, P. E. M., Lamoureux, G., Roux, B. et MacKerell Jr., A. D. Polarizable Empirical Force Field for Aromatic Compounds Based on the Classical Drude Oscillator. *J. Phys. Chem. B* **111**, 2873–2885 (2007).
288. Warren, G. L. et Patel, S. Electrostatic Properties of Aqueous Salt Solution Interfaces : A Comparison of Polarizable and Nonpolarizable Ion Models. *J. Phys. Chem. B* **112**, 11679–11693 (2008).
289. Lopes, P. E. M., Lamoureux, G. et Mackerell Jr., A. Polarizable empirical force field for nitrogen-containing heteroaromatic compounds based on the classical Drude oscillator. *J. Comput. Chem.* **30**, 1821–1838 (2009).
290. Baker, C. M., Lopes, P. E. M., Zhu, X., Roux, B. et MacKerell Jr., A. D. Accurate Calculation of Hydration Free Energies using Pair-Specific Lennard-Jones Parameters in the CHARMM Drude Polarizable Force Field. *J. Chem. Theory Comput.* **6**, 1181–1198 (2010).
291. Hariharan, P. C. et Pople, J. A. The effect of d-functions on molecular orbital energies for hydrocarbons. *Chem. Phys. Lett.* **16**, 217–219 (1972).
292. Francl, M. M., Pietro, W. J., Hehre, W. J., Binkley, J. S., Gordon, M. S., DeFrees, D. J., et Pople, J. A. Self-consistent molecular orbital methods. XXIII. A polarization-type basis set for second-row elements. *J. Chem. Phys.* **77**, 3654 (1982).
293. Hehre, W. J., Radom, L., Schleyer, P. v. R. et Pople, J. A. *Ab Initio Molecular Orbital Theory* (éd. York, N.) (Wiley, 1986).
294. Sadlej, A. J. Medium-size polarized basis sets for high-level correlated calculations of molecular electric properties. *Collec. Czech. Chem. Commun.* **53**, 1995–2016 (1988).
295. Sadlej, A. J. Medium-size polarized basis sets for high-level-correlated calculations of molecular electric properties. *Theor. Chim. Acta* **79**, 123–140 (1991).
296. Sadlej, A. J. Medium-size polarized basis sets for high-level-correlated calculations of molecular electric properties. *Theor. Chim. Acta* **81**, 45–63 (1991).
297. Sadlej, A. J. Medium-size polarized basis sets for high-level-correlated calculations of molecular electric properties. *Theor. Chim. Acta* **81**, 339–354 (1991).
298. Pereza, J. J. et Sadlej, A. J. Ab initio calculations of the polarizability. *J. Mol. Struct. : THEOCHEM* **371**, 31–36 (1996).
299. Szalewicz, K. *Private communication*

300. Chen, W. et Gordon, M. S. Energy Decomposition Analyses for Many-Body Interaction and Applications to Water Complexes. *J. Phys. Chem.* **100**, 14316–14328 (1996).
301. Freitag, M. A., Gordon, M. S., Jensen, J. H. et Stevens, W. J. Evaluation of charge penetration between distributed multipolar expansions. *J. Chem. Phys.* **112**, 7300 (2000).
302. Swart, M., van Duijnen, P. T. et Snijders, J. G. A charge analysis derived from an atomic multipole expansion. *J. Comput. Chem.* **22**, 79–88 (2000).
303. Dzyabchenko, A. V. A multipole approximation of the electrostatic potential of molecules. *Russ. J. Phys. Chem. A* **82**, 758–766 (2008).
304. Chen, B., Xing, J. et Siepmann, J. I. Development of Polarizable Water Force Fields for Phase Equilibrium Calculations. *J. Phys. Chem. B* **104**, 2391–2401 (2000).
305. Curutchet, C., Bofill, J. M., Hernández, B., Orozco, M. et Luque, F. J. Energy decomposition in molecular complexes : Implications for the treatment of polarization in molecular simulations. *J. Comput. Chem.* **24**, 1263–1275 (2002).
306. Cieplak, P., Caldwell, J. et Kollman, P. Molecular mechanical models for organic and biological systems going beyond the atom centered two body additive approximation : aqueous solution free energies of methanol and N-methyl acetamide, nucleic acid base, and amide hydrogen bonding and chloroform/water partition coefficients of the nucleic acid bases. *J. Comput. Chem.* **22**, 1048–1057 (2001).
307. Frisch, M. J. *et al. Gaussian 03, Revision C.02* Gaussian, Inc., Wallingford, CT, 2004.
308. Schmidt, M. W., Baldridge, K. K., Boatz, J. A., Elbert, S. T., Gordon, M. S., Jensen, J. H., Koseki, S., Matsunaga, N., Nguyen, K. A., Su, S., Windus, T. L., Dupuis, M. et Montgomery, J. A. *J. Comput. Chem.* **14**, 1347 (1993).
309. Chipot, C. et Ángyán, J. G. Continuing challenges in the parametrization of intermolecular force fields. Towards an accurate description of electrostatic and induction terms. *New J. Chem.* **29**, 411–420 (2005).
310. Piquemal, J.-P., Gresh, N. et Giessner-Prettre, C. Improved Formulas for the Calculation of the Electrostatic Contribution to the Intermolecular Interaction Energy from Multipolar Expansion of the Electronic Distribution. *J. Phys. Chem. A* **107**, 10353–10359 (2003).

311. Cisneros, G. A., Tholander, S. N.-I., Parisel, O., Darden, T. A., Elking, D., Perera, L. et Piquemal, J.-P. Simple formulas for improved point-charge electrostatics in classical force fields and hybrid quantum mechanical/molecular mechanical embedding. *Int. J. Quant. Chem.* **108**, 1905–1912 (2008).
312. Stone, A. J. dans *Theoretical models of chemical bonding* (éd. Maksić, H.) 103–131 (Springer–Verlag, Berlin, 1991).
313. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. et Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935 (1983).
314. Cieplak, P., Lybrand, T. P. et Kollman, P. A. Calculation of free energy changes in ion-water clusters using nonadditive potentials and the Monte Carlo method. *J. Chem. Phys.* **86**, 6393–6403 (1987).
315. Mas, E. M. et Szalewicz, K. Effects of monomer geometry and basis set saturation on computed depth of water dimer potential. *J. Chem. Phys.* **104**, 7606 (1996).
316. Hunter, C. A., Lawson, K. R., Perkins, J. et Urch, C. J. Aromatic interactions. *J. Chem. Soc., Perkin Trans.* **2**, 651–669 (2001).
317. Chipot, C., Jaffe, R., Maigret, B., Pearlman, D. A. et Kollman, P. A. Benzene Dimer : A Good Model for  $\pi$ -Interactions in Proteins? A Comparison between the Benzene and the Toluene Dimers in the Gas Phase and in an Aqueous Solution. *J. Am. Chem. Soc.* **118**, 11217–11224 (1996).
318. Shiratori, Y. et Nakagawa, S. Parametrization of calcium binding site in proteins and molecular dynamics simulation on phospholipase A2. *J. Comput. Chem.* **12**, 717–730 (1991).
319. Soteras, I., Curutchet, C., Bidon-Chanal, A., Dehez, F., Ángyán, J. G., Orozco, M., Chipot, C. et Luque, F. J. Derivation of Distributed Models of Atomic Polarizability for Molecular Simulations. *J. Chem. Theory Comput.* **3**, 1901–1913 (2007).
320. Celebi, N., Ángyán, J. G., Dehez, F., Millot, C. et Chipot, C. Distributed polarizabilities derived from induction energies : A finite perturbation approach. *J. Chem. Phys.* **112**, 2709–2717 (2000).
321. Dehez, F., Soetens, J.-C., Chipot, C., Ángyán, J. G. et Millot, C. Determination of Distributed Polarizabilities from a Statistical Analysis of Induction Energies. *J. Phys. Chem. A* **104**, 1293–1303 (2000).

322. Chipot, C., Dehez, F., Ángyán, J. G., Millot, C., Orozco, M. et Luque, F. J. Alternative Approaches for the Calculation of Induction Energies : Characterization, Effectiveness, and Pitfalls. *J. Phys. Chem. A* **105**, 11505–11514 (2001).
323. Fajans, K. *J. Phys. Chem.* **74**, 3407–3410 (1970).
324. Torheyden, M. et Jansen, G. A new potential energy surface for the water dimer obtained from separate fits of ab initio electrostatic, induction, dispersion and exchange energy contributions. *Mol. Phys.* **104**, 2101–2138 (2006).
325. Jiao, D., King, C., Grossfield, A., Darden, T. A. et Ren, P. Simulation of Ca<sup>2+</sup> and Mg<sup>2+</sup> Solvation Using Polarizable Atomic Multipole Potential. *J. Phys. Chem. B* **110**, 18553–18559 (2006).
326. Piquemal, J.-P., Perera, L., Cisneros, G. A., Ren, P., Pedersen, L. G. et Darden, T. A. Towards accurate solvation dynamics of divalent cations in water using the polarizable amoeba force field : From energetics to structure. *J. Chem. Phys.* **125**, 054511 (2006).
327. Picard, M., Jensen, A.-M. L., Sorensen, T. L.-M., Champeil, P., Moller, J. V. et Nissen, P. Ca<sup>2+</sup> versus Mg<sup>2+</sup> Coordination at the Nucleotide-binding site of the Sarcoplasmic Reticulum Ca<sup>2+</sup>-ATPase. *J. Mol. Biol.* **368**, 1–7 (2007).
328. Halgren, T. A. The representation of van der Waals (vdW) interactions in molecular mechanics force fields : potential form, combination rules, and vdW parameters. *J. Am. Chem. soc.* **114**, 7827–7843 (1992).
329. Schnieders, M. J. et Ponder, J. W. Polarizable Atomic Multipole Solutes in a Generalized Kirkwood Continuum. *J. Chem. Theory Comput.* (2007).
330. Zhang, Q. et Yang, Z.-Z. A comparative study of water clusters by classical and polarizable water models. *Wuli Huaxue Xuebao/ Acta Physico - Chimica Sinica* **23**, 1565–1571 (2007).
331. Walsh, T. R. et Liang, T. A multipole-based water potential with implicit polarization for biomolecular simulations. *J. Comput. Chem.* **30**, 893–899 (2008).
332. Buckingham, R. A., Hamilton, J. et Massey, H. S. W. The Low-Temperature Properties of Gaseous Helium. *Proc. R. Soc. A* **179**, 103 (1941).
333. Andrews, M., Jaeger, D., McCormick, S. et Taylor, G. Evaluation of functions on microcomputers : exp(x). *Computers and Mathematics with Applications* **7**, 503–508 (1981).

334. Levenberg, K. A Method for the Solution of Certain Non-Linear Problems in Least Squares. *The Quarterly of Applied Mathematics* **2**, 164–168 (1944).
335. Marquardt, D. An Algorithm for Least-Squares Estimation of Nonlinear Parameters. *SIAM J. Appl. Math.* **11**, 431–441 (1963).
336. Lamoureux, G. et Roux, B. Modeling induced polarization with classical Drude oscillators : Theory and molecular dynamics simulation algorithm. *J. Chem. Phys.* **119**, 3025–3039 (2003).
337. Dang, L. X. et Chang, T.-M. Molecular dynamics study of water clusters, liquid, and liquid-vapor interface of water with many-body potentials. *J. Chem. Phys.* **106**, 8149 (1997).
338. Bryantsev, V. S., Diallo, M. S., van Duin, A. C. T. et Goddard, I. W. A. Evaluation of B3LYP, X3LYP, and M06-Class Density Functionals for Predicting the Binding Energies of Neutral, Protonated, and Deprotonated Water Clusters. *J. Chem. Theory Comput.* **5**, 1016–1026 (2009).
339. Dang, L. X. The nonadditive intermolecular potential for water revised. *J. Chem. Phys.* **97**, 2659–2660 (1992).
340. Saunders, V. et Guest, M. ATMOL Program Package, SERC Daresbury Laboratory, Daresbury, Great Britain
341. Lide, D. R. *CRC Handbook of Chemistry and Physics* (éd. Raton, B.) (CRC Press, 2005).
342. Wang, W. *Fast Polarizable Force Field Computation in Biomolecular Simulations* thèse de doct. (University of Illinois at Urbana-Champaign, 2005). <<http://www.cs.uiuc.edu/research/techreports.php?report=UIUCDCS-R-2005-2522>>.
343. Caldwell, J., Dang, L. X. et Kollman, P. A. Implementation of nonadditive intermolecular potentials by use of molecular dynamics : development of a water-water potential and water-ion cluster interactions. *J. Am. Chem. Soc.* **112**, 9144–9147 (1990).





---

## Vers une modélisation plus réaliste des systèmes biologiques

**Résumé :** La plupart des fonctions énergie potentielle utilisées pour simuler les systèmes biologiques complexes ne traitent qu'implicitement la polarisation électronique et ce, de façon très incomplète. Bien qu'efficaces pour un large éventail d'applications, ces champs de force atteignent rapidement leurs limites dès lors que les effets de polarisation électronique sont importants. Tel est le cas par exemple au site actif des métalloprotéines où l'ion métallique polarise fortement son environnement. Dans cette thèse, j'ai développé une approche basée sur la mécanique quantique pour obtenir des paramètres d'un champ de force polarisable ayant pour composantes des charges, des polarisabilités distribuées d'ordre zéro et un (isotrope) et un potentiel de van der Waals décrit par une fonction de Buckingham. L'énergie d'induction peut être atténuée par une fonction de Tang et Toennies pour décrire l'énergie d'échange-induction. Cette approche a été effectuée avec succès pour l'interaction d'ions avec l'eau et le benzène mais aussi dans le cas d'un dimère d'eau. Une première étude des résultats en dynamique moléculaire montre que les paramètres obtenus en phase gazeuse peuvent se transférer pour les simulations en phase condensée.

**Mots clés :** Champ de force polarisable, polarisabilités distribuées, modèles polarisables, dynamique moléculaire, systèmes biologiques

---