

COMPLÉMENT DE COURS

dans le cadre de la thématique

**OUTILS POUR ETUDIER LA STRUCTURE ET LA DYNAMIQUE DES
PEPTIDES ET DES PROTEINES**

par

Marc BAADEN

(INSTITUT DE BIOLOGIE PHYSICO-CHIMIQUE, PARIS)

CNRS UPR9080

Dynamique Moléculaire in silico :

**Fondements théoriques et liste détaillée
de références bibliographiques**

Copyright (c) 2003 **Marc Baaden**

Permission vous est donnée de copier, distribuer et/ou modifier ce document selon les termes de la Licence GNU Free Documentation License, Version 1.1 ou ultérieure publiée par la Free Software Foundation ; avec les sections inaltérables suivantes :

Préface

Avec le texte de première page de couverture suivant :

- pas de texte de première page de couverture -

Avec le texte de dernière page de couverture suivant :

- pas de texte de dernière page de couverture -

Une copie de cette Licence est incluse dans la section appelée **GNU Free Documentation License** de ce document.

Préface

Ce chapitre consacré essentiellement aux méthodes de simulation et à leurs fondements théoriques est extrait de ma thèse doctorale «Etudes de molécules extractantes en solution et aux interfaces liquide-liquide : aspects structuraux et mécanistiques des effets de synergie ».

L'intégralité de la thèse est disponible sur Internet à partir du site <http://ww.marc-baaden.de>.

Ce document peut à la fois servir comme simple guide dans l'exploration de la dynamique moléculaire, mais fournit également une bonne base de références bibliographiques.

La reproduction et distribution de ce document est fortement encouragée. Voir les conditions à la fin.

Pour toute autre question et renseignement, contactez moi par courrier électronique à l'adresse baaden@smplinux.de. Bonne lecture !

Marc Baaden, Avril 2003

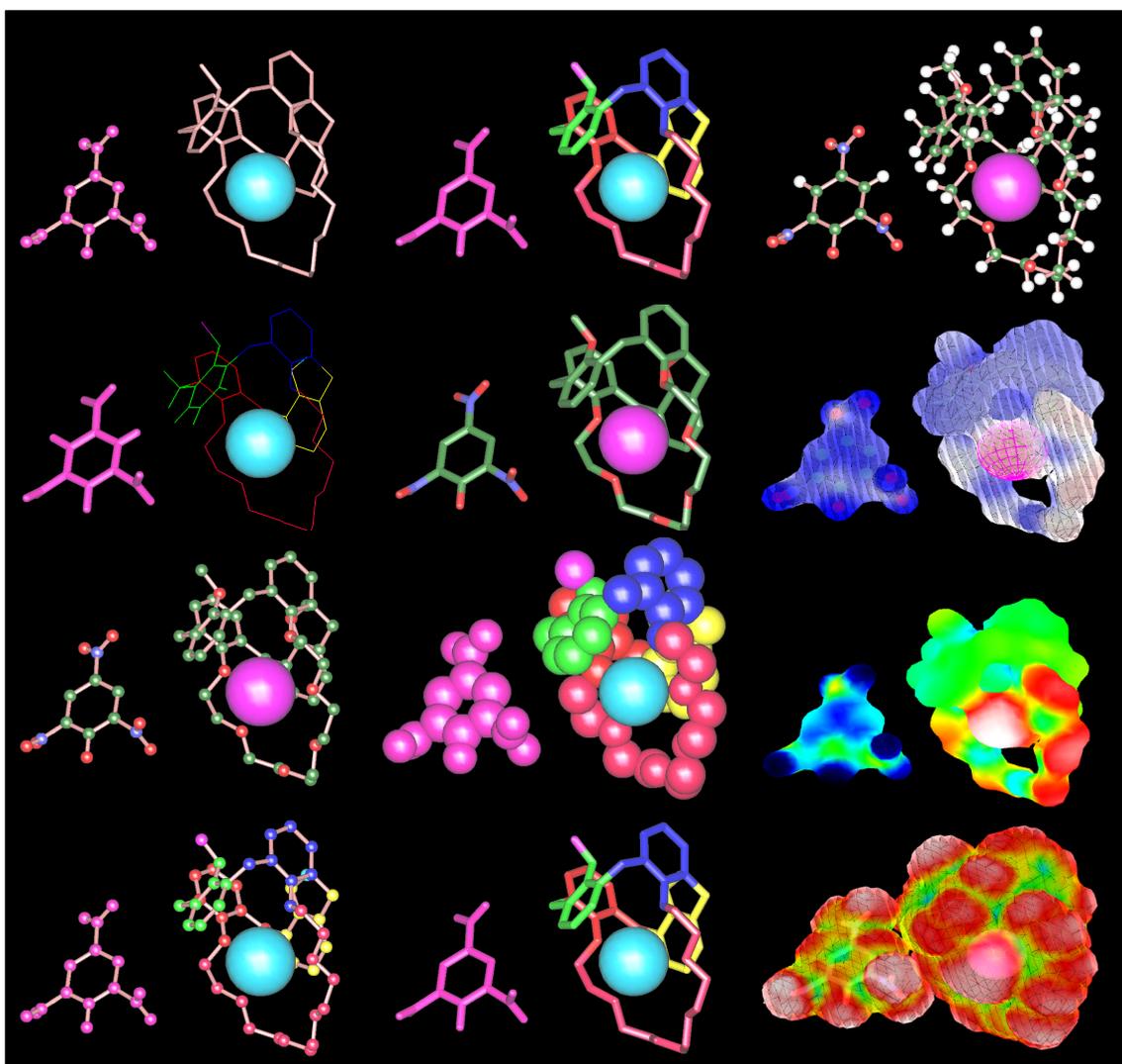
Table des matières · *Table of contents*

Chapitre 1	Méthodes de simulation et fondements théoriques	7
1	Introduction	7
2	Définitions des symboles, notations et de l'espace	8
3	Degrés de liberté, fonction d'interaction et équations de mouvement	9
4	HAMILTONIEN et fonction d'interaction : champs de forces utilisés	11
4.1	Forme de la fonction d'interaction	11
4.1.1	Interactions entre atomes liés	12
4.1.1.1	Elongation des liaisons covalentes entre atomes	12
4.1.1.2	Déformation des angles entre atomes liés	12
4.1.1.3	Termes de torsion, angles dièdres et impropres	13
4.1.2	Interactions non-liées	14
4.1.2.1	Interactions de LENNARD-JONES	14
4.1.2.2	Interactions électrostatiques	15
4.1.3	Forces spéciales dues aux contraintes	16
4.1.4	Inclusion de la polarisabilité	17
4.2	Paramètres s utilisés	17
4.2.1	Atomes, ions et molécules simulés	18
4.2.2	Choix des charges	19
4.2.3	Choix des autres paramètres	23
4.2.4	Modèles des solvants utilisés	28
4.2.4.1	Acétonitrile - MeCN	28
4.2.4.2	Eau - H ₂ O	28
4.2.4.3	Chloroforme - CHCl ₃ et MCl ₃	29
4.3	Améliorations, simplifications, artefacts et pièges	29

4.3.1	Traitement des interactions entre atomes "non-liés"	29
4.3.1.1	Exclusion ou modification de certaines interactions	29
4.3.1.2	Distance de troncature des interactions non-liées	30
4.3.1.3	Interactions électrostatiques	31
4.3.2	Traitement des bords	32
4.3.2.1	Systèmes dans le vide	32
4.3.2.2	Conditions périodiques	32
5	Génération de la topologie, des coordonnées atomiques et mise en solution	34
5.1	Topologie du système	34
5.2	Configuration du système	34
5.3	Mise en solution dans une boîte de liquide pure	34
5.4	Interface liquide/liquide	35
6	Mécanique et dynamique moléculaire	36
6.1	Minimisation de l'énergie	36
6.2	Dynamique moléculaire	37
6.2.1	Principe	37
6.2.2	Algorithmes de VERLET et du "saut de mouton"	38
6.2.3	Contrôle de la température	39
6.2.4	Contrôle de la pression	40
6.2.5	Algorithme SHAKE	41
6.2.6	Dynamique sous contraintes : SHAKE et BELLY	42
6.2.7	Protocole de simulation · Simulation protocol	42
6.3	Calculs de variation d'énergie libre	43
6.3.1	Application à la variation de l'énergie libre de complexation $\Delta\Delta F_c$	44
7	Analyse des résultats · Analysis of the results	45
7.1	Analyse infographique	46
7.2	Analyses géométriques et énergétiques	47
7.2.1	Fonctions de distribution radiale	47
7.2.2	Surface de séparation de GIBBS	47
7.2.3	Tension de surface	48
8	Temps de calcul et ressources informatiques	49
	Références bibliographiques du chapitre 1	51

Chapitre 1 × *Chapter 1*

Méthodes de simulation et fondements théoriques
Simulation methods and theoretical background



Chapitre 1 × Chapter 1

Ⓢ

Méthodes de simulation et fondements théoriques *Simulation methods and theoretical background*

"L'attribution d'une réalité physique quelconque aux calculs de l'histoire passée de l'électron est pure affaire de goût."

*Werner Heisenberg, 1930, Les Principes
physiques de la Théorie des Quantas*

1 Introduction × Introduction

Le comportement d'un système de particules composé de noyaux, d'atomes, d'électrons et de photons est en principe décrit par les lois de la mécanique quantique relativiste. Toute autre description à un niveau supérieur, comme les atomes et molécules au lieu des noyaux et électrons ou la dynamique classique au lieu de la dynamique quantique, représente une approximation. Ces approximations peuvent être classées selon le niveau de détail en allant du détail atomique au comportement à l'échelle macroscopique. [14] Chaque niveau d'approximation fait perdre en détail, mais permet d'accéder à des propriétés nouvelles du système par un gain d'efficacité (en temps de calcul) et de faisabilité (en traitement de données).

Dans ce qui suit nous faisons trois approximations importantes : 1) les vitesses des particules sont faibles par rapport à la vitesse de la lumière, 2) le mouvement des électrons est beaucoup plus rapide que celui des noyaux (approximation de BORN-OPPENHEIMER) et 3) le mouvement des atomes peut être décrit par la mécanique classique. A partir de principes simples de la physique, notamment les lois de NEWTON, on simule ainsi sur ordinateur le comportement des molécules au niveau atomique. Notre modèle utilise une représentation de la molécule en boules-et-bâtonnets, où les atomes sont représentés par les boules, ayant

un rayon, une masse et une charge atomique, et les liaisons par les bâtonnets qui sont des ressorts ramenant les liaisons à leur longer la plus stable après une déformation. Le jeu de boules et de bâtonnets qu'on utilise est caractérisé par un *champ de force*. Les interactions physiques entre les atomes (boules) – telles que la répulsion stérique, les liaisons hydrogène, *etc.* – sont prises en compte et introduites dans l'équation de l'énergie potentielle. Du point de vue thermodynamique cette énergie est une fonction d'état. A partir du modèle physique défini par les paramètres du champ de force et la représentation de l'énergie potentielle choisie, nous avons utilisé les lois de NEWTON pour propager les molécules dans l'*espace des phases*, défini par les vecteurs de position des atomes et les vecteurs associés aux quantités de mouvement. La dynamique moléculaire ("DM") [15] nous permet ainsi de générer des ensembles statistiques des configurations d'atomes et molécules en phase cristalline, en solution ou en phase gazeuse. [16]

Pour effectuer ces simulations, nous avons utilisé les logiciels AMBER ("Assisted Model Building and Energy Refinement") version 4.1 et 5 [17] développé par KOLLMAN *et al.* et GROMOS 96 ("GRONingen Molecular Simulation package") [18] développé par VAN GUNSTEREN *et al.* Les détails concernant les champs de forces, la représentation de l'énergie potentielle et les paramètres de simulation sont propres aux deux logiciels utilisés. De manière générale on retrouve ces caractéristiques ou des paramètres similaires dans d'autres programmes de simulations moléculaires, car les lois physiques utilisées sont les mêmes dans tous les cas.

Dans ce qui suit on décrira d'abord l'espace cartésien en 3 dimensions dans lequel nous effectuons nos simulations et la notation utilisée. Ensuite on discutera les trois choix de base de la modélisation : les degrés de liberté, la fonction d'interaction et les équations du mouvement. La suite est consacrée à la description de l'HAMILTONIEN du système, l'énergie cinétique, les termes physiques de l'énergie potentielle et les termes non-physiques représentant des forces "spéciales". Ensuite on discutera l'application de contraintes, par exemple de distance. Les conditions de traitement des bords de l'espace de simulation sont décrites par la suite. Nous discuterons ensuite les procédures qui permettent de minimiser l'énergie, d'effectuer la dynamique moléculaire et de calculer des différences d'énergie libre entre différents systèmes ou différents états de systèmes moléculaires.

2 Définitions des symboles, notations et de l'espace CARTESIEN × *Definitions of symbols, notations and the CARTESIAN space*

Les matrices et vecteurs sont notés en gras comme par exemple \mathbf{r} pour les positions ou \mathbf{f} pour les forces.

Nous utilisons un système de coordonnées CARTESIENNES avec les trois vecteurs de base orthonormés suivants:

$$\mathbf{e}^{3D} \equiv (\mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z) \quad (1)$$

Le vecteur de position d'un atome i dans cet espace 3D peut être représenté par ses 3 composantes (x_i, y_i, z_i)

$$\mathbf{r}_i^{3D} = x_i \mathbf{e}_x + y_i \mathbf{e}_y + z_i \mathbf{e}_z \quad (2)$$

Dans ce qui suit nous utiliserons la notation (incorrecte) abrégée

$$\mathbf{r}_i^{3D} = (x_i, y_i, z_i) \quad (3)$$

en ignorant la différence entre vecteurs et composants. Le produit scalaire entre les deux vecteurs \mathbf{r}_i^{3D} et \mathbf{r}_j^{3D} est défini par

$$\mathbf{r}_i^{3D} \cdot \mathbf{r}_j^{3D} = x_i x_j + y_i y_j + z_i z_j \quad (4)$$

La longueur d'un vecteur

$$r_{ij}^{3D} = |\mathbf{r}_i^{3D} - \mathbf{r}_j^{3D}| \quad (5)$$

entre les atomes i et j est définie par

$$r_{ij}^{3D} = \left[(x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2 \right]^{1/2} \quad (6)$$

Toutes ces définitions peuvent être généralisées à la 4^{ème} dimension. Les simulations de DM dans l'espace CARTESIEN en quatre dimensions sont décrites dans la référence [19]. Cette technique permet de surmonter des barrières énergétiques dans l'hypersurface de l'énergie potentielle. Dans cette thèse nous nous limitons aux simulations en 3 dimensions.

3 Degrés de liberté, fonction d'interaction et équations de mouvement × *Degrees of freedom, interaction function and equations of motion*

Lors de toute modélisation d'un système, trois choix fondamentaux doivent être faits. [20] D'abord quels *degrés de liberté* seront explicitement traités ou simulés, et lesquels seront uniquement traités implicitement et incorporés dans la fonction qui décrit l'interaction entre les degrés de libertés traités explicitement. Les degrés de liberté explicites peuvent être traités de façon classique ou par la mécanique quantique. Ensuite il faut choisir une *fonction d'interaction* [21] pour décrire l'énergie d'interaction du système moléculaire en fonction des degrés de liberté explicites du modèle. Enfin, il faut choisir une méthode pour *échantillonner* les configurations ou *simuler le mouvement* selon les degrés de libertés explicites, ou pour

rechercher les parties de l'espace configurationnel dans lequel le système moléculaire a une basse énergie.

Dans le cadre de cette thèse nous nous intéressons uniquement aux degrés de libertés de position des coordonnées CARTESIENNES des atomes. Dans les simulations classiques, l'HAMILTONIEN du système moléculaire prend la forme

$$H(\mathbf{p}, \mathbf{r}; m, s) = K(\mathbf{p}; m) + V(\mathbf{r}; s) \quad (7)$$

Le premier terme est le terme d'énergie cinétique

$$K(\mathbf{p}; m) = \sum_{i=1}^N \frac{\mathbf{p}_i^2}{2m_i} = \sum_{i=1}^N \frac{1}{2} m_i \mathbf{v}_i^2 \quad (8)$$

qui dépend uniquement des masses atomiques m , et est indépendant des coordonnées \mathbf{r} de la particule. Le deuxième terme correspond à l'énergie potentielle aussi appelée la fonction d'interaction, qui décrit l'énergie d'interaction en fonction des coordonnées \mathbf{r} des particules

$$V(\mathbf{r}; s) = V(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N; s_1, s_2, \dots, s_M) \quad (9)$$

En général $V(\mathbf{r}; s)$ dépend de M paramètres aussi appelés paramètres du champ de force, indiqués par $s^o (s_1, s_2, \dots, s_M)$. En pratique, la fonction d'interaction est une somme de termes qui représentent différents types d'interactions.

La force \mathbf{f}_i qui agit sur une particule i est donnée par l'équation

$$\mathbf{f}_i = - \frac{\partial}{\partial \mathbf{r}_i} V(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) \quad (10)$$

qui peut également servir à calculer l'énergie d'interaction qui correspond à une force donnée. La trajectoire obtenue lors des simulations de DM dépend uniquement des forces qui agissent sur les atomes, pas explicitement des énergies.

Nous utilisons les équations du mouvement classiques de NEWTON pour décrire l'évolution des atomes dans l'espace des phases. D'autres choix existent tels que les équations de LAGRANGE ou de HAMILTON. Les équations de NEWTON sont utilisées sous la forme

$$\frac{d\mathbf{r}_i(t)}{dt} = \mathbf{v}_i(t) \quad (11)$$

et

$$\frac{d\mathbf{v}_i(t)}{dt} = \frac{\mathbf{f}_i(t)}{m_i} \quad (12)$$

avec

$$\mathbf{f}_i(t) = - \frac{\partial}{\partial \mathbf{r}_i} V(\mathbf{r}_1, \mathbf{r}_1, \dots, \mathbf{r}_N) \quad (13)$$

Les vitesses atomiques sont indiquées par \mathbf{v}_i , la force qui agit sur l'atome i par \mathbf{f}_i et le temps par t . Les équations de NEWTON sont uniquement valables pour les coordonnées

CARTESIENNES \mathbf{r}_i d'une particule de masse m_i . Dans les simulations de dynamique moléculaire, les équations 11 à 13 sont "intégrées" numériquement dans le temps.

Le système moléculaire peut être couplé à des grandeurs externes au système de différentes manières. On peut inclure un terme supplémentaire $V^{restr}(\mathbf{r})$, une *fonction de pénalité*, dans la fonction d'interaction du système qui limite le mouvement des particules de façon à ce que la valeur simulée de la quantité externe se rapproche de la valeur choisie. Cette valeur peut également être imposée comme *contrainte* au système de façon à être vérifiée pour chaque configuration ou pour un ensemble de configurations. Une équation au premier ordre peut être ajoutée aux équations de mouvement des particules pour faire évoluer la valeur simulée de la quantité externe vers une valeur donnée, il s'agit de la *méthode de faible couplage*.

Les deux dernières méthodes - contrainte et faible couplage - entraînent une modification des équations de mouvement classiques du système moléculaire. L'énergie totale sera uniquement conservée lors d'une simulation de DM si l'énergie potentielle $V(r)$ ne dépend que de la configuration instantanée. Cette condition n'est pas vérifiée quand on applique une contrainte ou la méthode de faible couplage avec des paramètres moyennés dans le temps.

4 HAMILTONIEN et fonction d'interaction : champs de forces utilisés × *HAMILTONIAN and interaction function : force fields used*

L'HAMILTONIEN du système moléculaire, donné dans l'équation 7 consiste en deux termes. L'énergie cinétique est entièrement décrite par l'équation 8, et ce qu'on appelle communément un champ de force correspond au deuxième terme, la fonction d'interaction, et aux paramètres s associés. Nous décrivons d'abord les termes utilisés dans les champs de force des logiciels AMBER [22] et GROMOS [23] et ensuite les paramètres que nous avons utilisés pour nos simulations.

4.1 Forme de la fonction d'interaction × *Functional form of the interaction function*

De manière générale, l'énergie potentielle dans les champs de force prend la forme suivante

$$V(\mathbf{r}; s) = V^{bon}(\mathbf{r}; s) + V^{nonb}(\mathbf{r}; s) + V^{spec}(\mathbf{r}; s) \quad (14)$$

où le premier terme correspond aux interactions entre atomes liés par des liaisons covalentes, le deuxième aux interactions dites "non-liées" (VAN DER WAALS et électrostatique) et le troisième à des termes spéciaux, par exemple des contraintes de position.

4.1.1 Interactions entre atomes liés · *Bonded interactions*

Quatre termes interviennent dans l'expression de $V^{bon}(\mathbf{r}; S)$

$$V^{bon} = V^{bond}(\mathbf{r}; S) + V^{angle}(\mathbf{r}; S) + V^{dihed}(\mathbf{r}; S) + V^{improp}(\mathbf{r}; S) \quad (15)$$

Ils décrivent l'élongation des liaisons, la déformation des angles et les torsions pour les angles dièdres périodiques et "impropres".

4.1.1.1 Elongation des liaisons covalentes entre atomes · *Bond-stretching term*

Dans le champ de force d'AMBER on utilise un terme quadratique pour les interactions d'élongation des liaisons, avec une sommation sur toutes les liaisons covalentes $n = 1, \dots, N_b$.

$$V_{bond} = \sum_{n=1}^{N_b} K_{b_n}^{harm} (b_n - b_{0_n})^2 \quad (16)$$

b_n est la longueur instantanée de la liaison n , b_{0_n} la valeur à l'équilibre et $K_{b_n}^{harm}$ la constante de force associée. Les unités couramment utilisées pour b_n, b_{0_n} sont en Å et pour $K_{b_n}^{harm}$ en $kcal\ mol^{-1}\ \text{Å}^{-1}$. Dans GROMOS une forme spéciale de V_{bond} est utilisée pour éviter une opération de racine carrée

$$V_{bond} = \sum_{n=1}^{N_b} \frac{1}{4} K_{b_n} (b_n^2 - b_{0_n}^2)^2 \quad (17)$$

Les unités pour b_n, b_{0_n} sont en nm et pour K_{b_n} en $kJ\ mol^{-1}\ nm^{-4}$. Pour convertir des paramètres provenant de AMBER en format GROMOS, nous suivons les conseils donnés dans le manuel de GROMOS, [23a] page II-15 en appliquant l'équation 18

$$K_{b_n} = \frac{1}{2} K_{b_n}^{harm} b_{0_n}^{-2} \quad (18)$$

4.1.1.2 Déformation des angles entre atomes liés · *Bond-angle bending*

Les équations 19 et 20 donnent les termes de déformation des angles entre atomes liés par une liaison covalente dans le cas de AMBER et GROMOS respectivement, ainsi que la formule de conversion valide pour $K_{\theta_n}^{harm} > 0$ (équation 21).

$$V_{angle} = \sum_{n=1}^{N_\theta} K_{\theta_n}^{harm} (\theta_n - \theta_{0_n})^2 \quad (19)$$

$$V_{angle} = \sum_{n=1}^{N_q} \frac{1}{2} K_{q_n} (\cos q_n - \cos q_{0_n})^2 \quad (20)$$

$$K_{q_n} = \frac{2E_{k_B T}}{\left[\cos \left\{ q_{0_n} + \sqrt{\frac{E_{k_B T}}{K_{q_n}^{harm}}} \right\} - \cos q_{0_n} \right]^2 + \left[\cos \left\{ q_{0_n} - \sqrt{\frac{E_{k_B T}}{K_{q_n}^{harm}}} \right\} - \cos q_{0_n} \right]^2} \quad (21)$$

Les angles sont en général donnés en degrés, $K_{q_n}^{harm}$ en $kcal\ mol^{-1}\ rad^{-2}$ et K_{q_n} en $kJ\ mol^{-1}$. $E_{k_B T}$ est le facteur de BOLTZMANN qui est de $E_{k_B T} = 0.5961575\ kcal\ mol^{-1}$ à 300 K. La sommation dans 19 et 20 concerne tous les angles de liaison $n = 1, \dots, N_q$ du système moléculaire.

4.1.1.3 Termes de torsion, angles dièdres et impropres ·*Proper and improper torsions*

Dans GROMOS, la partie de la fonction d'interaction qui représente les termes d'interactions harmoniques impropres de déformation des angles dièdres (hors du plan, hors de la configuration tétraédrique) est décrite par

$$V^{improp}(\mathbf{r}; s) = \sum_{n=1}^{N_\xi} \frac{1}{2} K_{\xi_n} (\xi_n - \xi_{0_n})^2 \quad (22)$$

La somme comprend tous les ensembles d'angles dièdres $n = 1, \dots, N_\xi$ définis pour maintenir des groupes d'atomes proches d'une configuration spatiale donnée. Par exemple pour maintenir un groupe d'atomes plan comme dans des cycles aromatiques, ou des groupement fonctionnels tels que des groupes carboxyliques, ou encore pour maintenir quatre atomes dans une configuration tétraédrique et empêcher une racémisation s'il s'agit de carbones asymétriques représentés dans le modèle de l'atome uni. Une convention IUPAC existe pour la définition exacte de ces angles dièdres [24] et une description qualitative est possible avec les termes définis dans la réf. [25].

Les impropres dans AMBER prennent la même forme que l'équation 24 utilisée pour les autres torsions, avec une multiplicité de $m_n=2$ pour un plan ou $m_n=3$ pour un tétraèdre, γ est égal à 180 dans les deux cas. [22b]

Dans le champ de force de GROMOS, le terme de torsion des angles dièdres dans la fonction d'interaction prend la forme

$$V^{dihed}(\mathbf{r}; s) = \sum_{n=1}^{N_\phi} K_{\phi_n} [1 + \cos(\delta_n) \cos(m_n \phi_n)] \quad (23)$$

avec la constante de force K_{ϕ_n} , l'angle dièdre ϕ_n et les valeurs de δ_n et m_n limités à $\delta_n = 0, \pi$ et $m_n = 1, 2, 3, 4, 5, 6$ ce qui mène à $\cos(\delta_n) = \pm 1.0$. La sommation se fait sur tous les angles dièdres $n = 1, \dots, N_\phi$ sélectionnés lors de la définition de la topologie de la molécule.

Dans AMBER le terme de torsion (propre et impropre) est défini par une série de FOURIER de la manière suivante

$$V^{improp/dihed}(\mathbf{r}; s) = \sum_{n=1}^{N_\phi} \sum_{m=1}^{P_n} K_{\phi_{n,m}} [1 + \cos(m_n \phi_n - \gamma_n)] \quad (24)$$

avec les angles de phase des dièdres γ_n , les termes de la série de Fourier de 1 à P_n et les constantes de forces associées $K_{\phi_{n,m}}$. En général la série ne comporte qu'un seul terme, mais quelques cas sont connus où il en faut plus (e.g. butane, OS-CH-CH-OS ou H-N-C-O). [22b,26]

4.1.2 Interactions non-liées · *Non-bonded interactions*

Les interactions entre atomes non-liés sont décrites par deux termes, V^{LJ} pour les interactions de type LENNARD-JONES et V^{CRF} pour les interactions électrostatiques.

$$V^{nonb}(\mathbf{r}; s) = \sum_{\substack{\text{paires}(i,j) \\ \text{non liés}}} \left\{ V^{LJ}[\mathbf{r}_{ij}, C_{12}(i, j), C_6(i, j)] + V^{CRF}[\mathbf{r}_{ij}, q_i, q_j, \epsilon_0, \epsilon_1, C_{rf}, R_{rf}] \right\} \quad (25)$$

La somme s'étend en principe à toutes les paires d'atomes (i,j), mais un certain nombre d'interactions est en général exclu, voir partie 4.3.1.

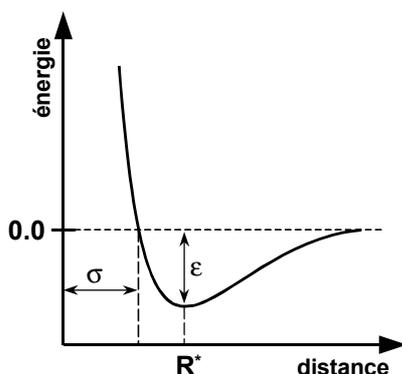


Figure 2 : Modèles (\mathbf{e}, R^*) et (\mathbf{e}, \mathbf{s}) des paramètres de VAN DER WAALS pour une molécule diatomique de type X-X. Énergie d'interaction en fonction de la distance de séparation entre les deux atomes.

4.1.2.1 Interactions de LENNARD-JONES · *LENNARD-JONES interactions*

Différents ensembles de paramètres sont utilisés pour décrire les interactions de type LENNARD-JONES, aussi appelées interactions de VAN DER WAALS (Figure 2). Dans GROMOS, on utilise les paramètres C_6 (en $\text{kJ mol}^{-1} \text{nm}^6$) et C_{12} (en $\text{kJ mol}^{-1} \text{nm}^{12}$) pour l'interaction entre les atomes i et j à une distance r_{ij} (en nm)

$$V^{LJ} = \left[\frac{C_{12}(i, j)}{(r_{ij})^{12}} - \frac{C_6(i, j)}{(r_{ij})^6} \right] \quad (26)$$

dans AMBER nous avons utilisé les paramètres ϵ (en $kcal\ mol^{-1}$) et R^* (en Å) pour ce même potentiel avec r_{ij} en Å également

$$V^{LJ} = \epsilon_{ij} \left(\frac{R_{ij}^*}{r_{ij}} \right)^{12} - 2\epsilon_{ij} \left(\frac{R_{ij}^*}{r_{ij}} \right)^6 \quad (27)$$

Le premier terme en r_{ij}^{-12} correspond à la répulsion entre deux atomes due à l'exclusion de PAULI, le deuxième terme en r_{ij}^{-6} représente la dispersion attractive de LONDON. Pour convertir les paramètres utilisés avec AMBER dans le modèle utilisé avec GROMOS ou vice versa, nous utilisons les équations 28 et 29

$$C_{12}(i, j) = \sqrt{\epsilon(i, i) \cdot \epsilon(j, j)} (R^*(i, i) + R^*(j, j))^{12} \quad (28)$$

$$C_6(i, j) = 2\sqrt{\epsilon(i, i) \cdot \epsilon(j, j)} (R^*(i, i) + R^*(j, j))^6 \quad (29)$$

Ces équations tiennent compte des règles utilisées pour déterminer des paramètres mixtes entre deux atomes i et j de type différent. Dans GROMOS nous utilisons

$$C_{12}(i, j) = \sqrt{C_{12}(i, i) \cdot C_{12}(j, j)} \quad (30)$$

$$C_6(i, j) = \sqrt{C_6(i, i) \cdot C_6(j, j)} \quad (31)$$

Dans AMBER nous utilisons les règles de LORENTZ-BERTHELOT

$$\epsilon(i, j) = \sqrt{\epsilon(i, i) \cdot \epsilon(j, j)} \quad (32)$$

$$R^*(i, j) = R^*(i, i) + R^*(j, j) \quad (33)$$

4.1.2.2 Interactions électrostatiques · *Electrostatic interactions*

La première partie concerne les interactions COULOMBIENNES et les deux autres termes le champ de réaction. [27] L'équation 34 donne l'expression utilisée dans GROMOS et AMBER †

$$V^{CRF} = \frac{q_i q_j}{4\pi\epsilon_0\epsilon_1} \left[\frac{1}{r_{ij}} - \frac{1/2 C_{rf}(r_{ij})^2}{R_{rf}^3} - \frac{1 - 1/2 C_{rf}}{R_{rf}} \right] \quad (34)$$

La correction du champ de réaction de type POISSON-BOLTZMANN est une généralisation du modèle d'ONSAGER [28] qui prend en compte les effets de force ionique. On part d'un modèle où l'extérieur d'une sphère de rayon R_{rf} autour du soluté est modélisé par un milieu homogène polarisable (Figure 3). Plus précisément le terme représente la force sur l'atome i due au champ de réaction induit par l'atome j . On part de l'hypothèse d'un continuum

† Il s'agit d'une version modifiée du programme AMBER, dans laquelle l'algorithme du champ de réaction [27] a été incorporé par E. ENGLER.

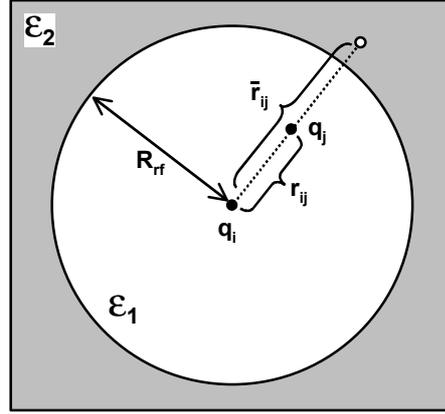


Figure 3 : Représentation de la méthode du champ de réaction. [27] L'atome i au centre de la sphère de rayon R_{rf} est entouré de charges q_j situées à \mathbf{r}_{ij} . Le champ de réaction est créé par les charges images à $\bar{\mathbf{r}}_{ij}$.

électrostatique caractérisé par une permittivité diélectrique relative ϵ_2 et une longueur d'écrantage de DEBYE inverse κ à l'extérieur de la distance de troncature du champ de réaction R_{rf} . Pour chaque atome j , on introduit une charge image hors de la sphère R_{rf} à une distance $\bar{\mathbf{r}}_{ij} = R_{rf}^2 / \mathbf{r}_{ij}$, qui participe au potentiel électrostatique à l'intérieur de la sphère. Le coefficient C_{rf} est déterminé par

$$C_{rf} = \frac{(2\epsilon_1 - 2\epsilon_2)(1 + \kappa R_{rf}) - \epsilon_2(\kappa R_{rf})^2}{(\epsilon_1 + 2\epsilon_2)(1 + \kappa R_{rf}) + \epsilon_2(\kappa R_{rf})^2} \quad (35)$$

et $\epsilon_1 = 1$. Le dernier terme dans 34 est constant et sert à rendre l'interaction nulle à $r_{ij} = R_{rf}$. Pour $\epsilon_2 = 1$ et $\kappa = 0$ la correction du champ de réaction est nulle.

4.1.3 Forces spéciales dues aux contraintes · *Special forces: restraints and constraints*

La fonction d'interaction donnée dans l'équation 14 peut contenir des termes qui ne sont utilisés que dans des cas particuliers. Dans cette thèse nous avons fait appel aux *contraintes harmoniques* sur la position des atomes, sur les distances de liaison, les angles et les torsions et aux *contraintes* de position (option BELLY dans AMBER) et de distances de liaison (Algorithme SHAKE). En guise d'exemple nous donnons l'expression du terme de contrainte harmonique de position utilisé dans GROMOS:

$$V^{pr}(\mathbf{r}; s) = \sum_{n=1}^{N_{pr}} \frac{1}{2} K_n^{pr} (\mathbf{r}_n - \mathbf{r}_n^0)^2 \quad (36)$$

La somme s'étend à tous les N_{pr} atomes sélectionnés. Les positions de référence sont notées \mathbf{r}_n^0 . K_n^{pr} est la constante de force de l'oscillateur harmonique. Les expressions des autres termes sont données dans [23a]. L'option BELLY et l'algorithme SHAKE sont décrits dans 6.2.6 et 6.2.5 respectivement.

4.1.4 Inclusion de la polarisabilité · *Inclusion of polarisability*

Dans les termes additifs par paires de l'équation 34 on néglige la contribution de la polarisation, qui par sa nature est non-additive. Pour l'inclure dans la fonction d'interaction, on se sert de l'approche des polarisabilités atomiques, où on attribue une polarisabilité α_i à chaque atome i . [29] L'énergie de polarisation V^{pol} dépend du champ électrique permanent \mathbf{E}_i^P qui agit sur l'atome i et du moment dipolaire induit \mathbf{m}_i :

$$V^{pol} = -\frac{1}{2} \sum_{i=1}^{N_{at}} \mathbf{E}_i^P \cdot \mathbf{m}_i \quad (37)$$

Pour rendre compte de la polarisation, il faut déterminer \mathbf{E}_i^P par une procédure itérative. A partir d'un champ initial $\mathbf{E}_i^{k=0} = \mathbf{E}_i^P$, on se sert du tenseur dipolaire $\bar{\mathbf{T}}_{ij}$ et de ce champ pour calculer le champ électrique du pas suivant \mathbf{E}_i^{k+1} :

$$\mathbf{E}_i^{k+1} = \mathbf{E}_i^k + \sum_{j \neq i} \langle \bar{\mathbf{T}}_{ij} \rangle \alpha_j \mathbf{E}_j^k \quad (38)$$

La procédure itérative mène à un champ auto-cohérent qui suffit à un critère de tolérance défini. Ensuite on calcule une contribution de ce champ à la force d'interaction entre les atomes i et j :

$$\begin{aligned} \mathbf{F}_{ij} = & \left(\frac{q_j}{r_{ij}^3} - \frac{3\mathbf{m}_j \cdot \mathbf{r}_{ij}}{r_{ij}^5} \right) \mathbf{m}_i - \left(\frac{q_i}{r_{ij}^3} - \frac{3\mathbf{m}_i \cdot \mathbf{r}_{ij}}{r_{ij}^5} \right) \mathbf{m}_j \\ & - \frac{3}{r_{ij}^5} \left(q_j \mathbf{r}_{ij} \cdot \mathbf{m}_i - q_i \mathbf{r}_{ij} \cdot \mathbf{m}_j - \frac{5}{r_{ij}^2} (\mathbf{r}_{ij} \cdot \mathbf{m}_i)(\mathbf{r}_{ij} \cdot \mathbf{m}_j) + (\mathbf{m}_i \cdot \mathbf{m}_j) \right) \mathbf{r}_{ij} \end{aligned} \quad (39)$$

Ce terme s'ajoute comme force spéciale à la fonction d'interaction.

4.2 Paramètres \mathcal{S} utilisés · *Employed parameters \mathcal{S}*

Les logiciels AMBER et GROMOS ainsi que les champs de forces associés ont été développés surtout pour la simulation de (bio)molécules, acides aminés, peptides, protéines et acides nucléiques. Les jeux de paramètres \mathcal{S} fournis avec ces logiciels – parm91 et parm94 dans le cas de AMBER et les paramètres 43A1 et 43B1 de GROMOS – ont été ajustés pour décrire ces classes de molécules. Plus d'information sur la façon d'obtenir ces paramètres se trouve dans la littérature. [21-23] Les systèmes simulés dans cette thèse sont étroitement liés à l'extraction liquide-liquide, et souvent des paramètres adéquats n'existaient pas encore dans les champs de forces d'AMBER et de GROMOS. Dans ces cas, nous avons eu recours aux données publiées dans la littérature scientifique, aux travaux antérieurs effectués dans le laboratoire MSM, ou nous avons développé de nouveaux paramètres.

4.2.1 Atomes, ions et molécules simulés · *Simulated atoms, ions and molecules*

La figure 4 montre les molécules neutres simulées dans cette thèse, la figure 5 les molécules de solvant et la figure 6 les espèces ioniques. La nomenclature de ces composés est donnée en annexe. A chaque atome est attribué un type d'atome qui sert à définir les paramètres du champ de force sauf les charges atomiques. Deux atomes dans des situations chimiquement proches sont du même type. Dans ce qui suit nous donnons le jeu de paramètres s : le choix des charges, les paramètres pour chaque type d'atome et les modèles de solvant utilisés.

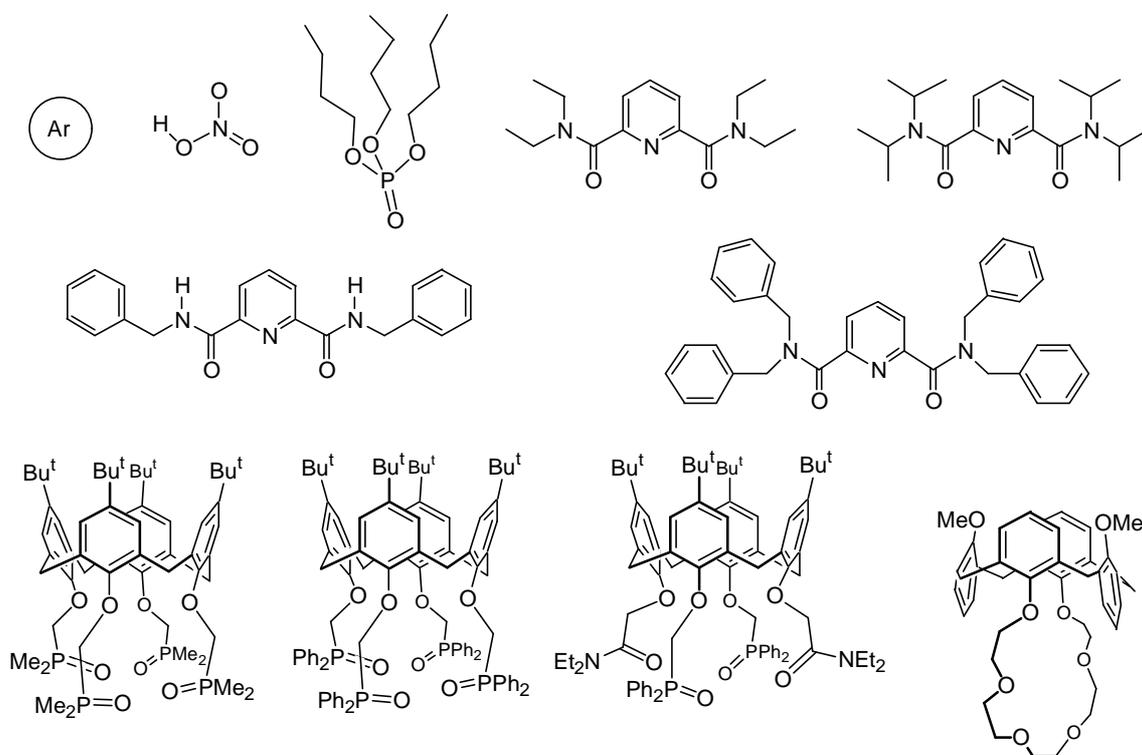


Figure 4 : Espèces neutres simulées. La nomenclature de ces composés est donnée en annexe.

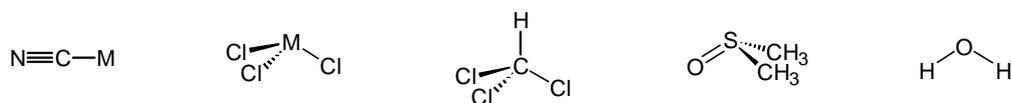


Figure 5 : Solvants simulés. M représente un atome uni, $M=CH_3$ pour l'acétonitrile et $M=CH$ pour le chloroforme. Deux modèles de chloroforme sont représentés : un modèle en 4 points (atome uni) et un modèle en 5 points.

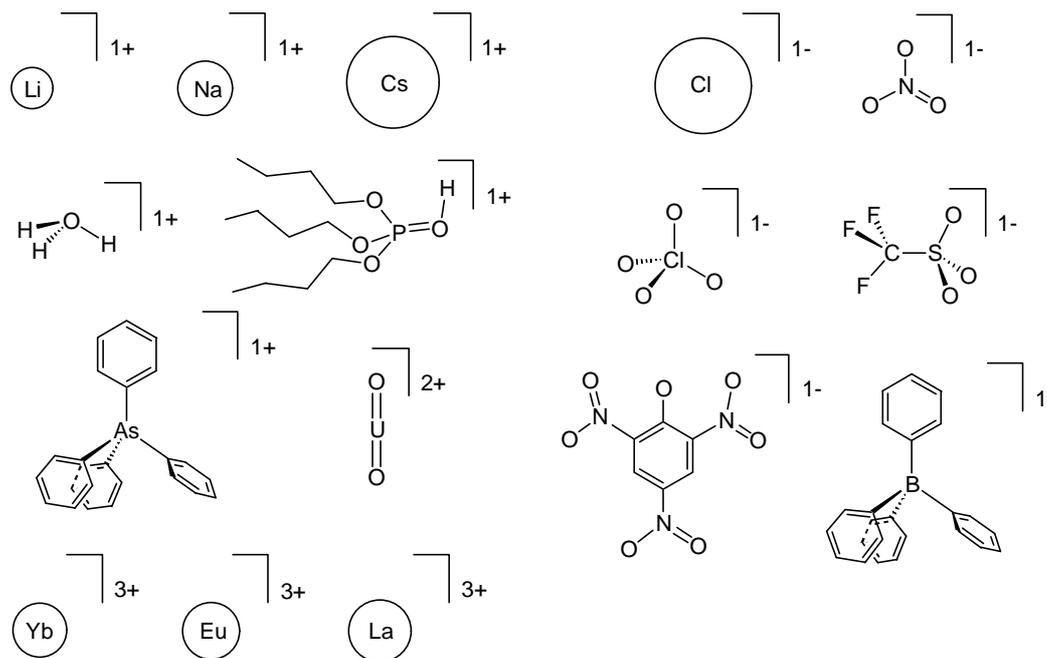


Figure 6 : Espèces ioniques simulées, cations à gauche et anions à droite. Voir la nomenclature donnée en annexe.

4.2.2 Choix des charges · Choice of the charges

Plusieurs approches de dérivation de charges existent, [30] par exemple celle de MULLIKEN, [31] celle des "atomes dans les molécules", [32] celle de CRAMER et TRUHLAR, [33] de GASTEIGER-MARSILI [34] et celle des charges à partir du potentiel électrostatique ("Electrostatic Potential", noté ESP). [35] Nous utiliserons surtout ce dernier type de charges qui a été discuté récemment. [36] Ces charges ESP sont également utilisées dans AMBER. [37] Les charges et types d'atomes de Ar et des ions monoatomiques sont données dans le tableau 1, [23a, 38] ceux des molécules simples et ions moléculaires dans la figure 7, ceux des dérivés du ligand pyridine-dicarboxamide dans la figure 8 et ceux des fragments moléculaires utilisés pour construire les calixarènes simulés sont données dans la figure 9. Toutes les espèces ioniques sont représentées par une charge totale entière +/-n e. Les distributions de charges de l'acide nitrique HNO₃, de l'anion triflate F₃CSO₃⁻ et du cation H₃O⁺ proviennent de l'ajustement de potentiels électrostatiques obtenus par des calculs au niveau HARTREE-FOCK 6-31G*. Cette même approche a été utilisée pour les dérivés de la pyridine-dicarboxamide (Figure 8), mais en considérant plusieurs fragments plutôt que l'intégralité de la molécule. Pour la partie pyridine nous nous référons aux calculs effectués par JORGENSEN *et al.* pour décrire le solvant pur. [39] Nous avons calculé les charges de la partie amide. La représentation finale est obtenue par une répartition symétrique des charges atomiques sur les résidus de même type. Les charges partielles atomiques de l'anion picrate (Figure 7) ont été ajustées au potentiel électrostatique d'un calcul MNDO par TROXLER, [40] ainsi que celles de l'anion perchlorate ClO₄⁻ (BERNY), [41] les charges de l'anion tétraphénylborate et du

Tableau 1: Charges et types d'atomes de Ar et des ions monoatomiques.

Elément	Type d'atome ^a	Charge [e]	Référence
Ar	AR(28)	0	[23a]
Li	LI	+1	[38a]
Na	IP / NA+(34)	+1	[38a] / [23a]
Cs	CS	+1	[38a]
Cl	CL(30)	-1	[23a]
Yb	YB	+3	[38b]
Eu	EU	+3	[38b]
La	LA	+3	[38b]

^a Pour AMBER nous donnons le code à deux lettres du type d'atome, pour GROMOS nous donnons en plus le code numérique associé ("*integer atom code*") entre parenthèses.

cation tétraphénylarsonium au potentiel électrostatique d'un calcul HF 3-21G* par SCHURHAMMER. [42] La distribution de charges de l'anion nitrate est prise dans la littérature (ESP 6-31G*). [43] GUILBAUD a ajusté les charges du cation uranyle pour rendre compte de l'énergie d'hydratation par rapport à Sr²⁺. [44] La représentation des charges du tri-*n*-butylphosphate (TBP) provient d'un calcul ESP au niveau MNDO. [45] Le TBP protoné a été ajusté à partir des charges du TBP neutre et des charges *ab initio* de (MeO)₃P=OH⁺ et (MeO)₃P=O. [46] Pour la distribution de charges des calixarènes simulés (Figure 9), nous appliquons à nouveau le principe des fragments moléculaires. Pour les calixarènes avec 4 bras indépendants nous avons utilisé des charges provenant d'un calcul *ab initio* (ESP 6-31G*) pour les fragment —CH₂(Ph)₂P=O et —(O)CH₂(C=O)N(Et)₂ et des charges de CHARMM pour le fragment anisole. [47] Pour le calixarène couronne, nous avons conservé les charges précédemment utilisées, [48] dans le but d'obtenir des résultats directement comparables aux travaux de MUZET [49] et de LAUTERBACH. [50] Les types d'atomes ont été choisi en analogie avec ceux donnés dans la référence [51].

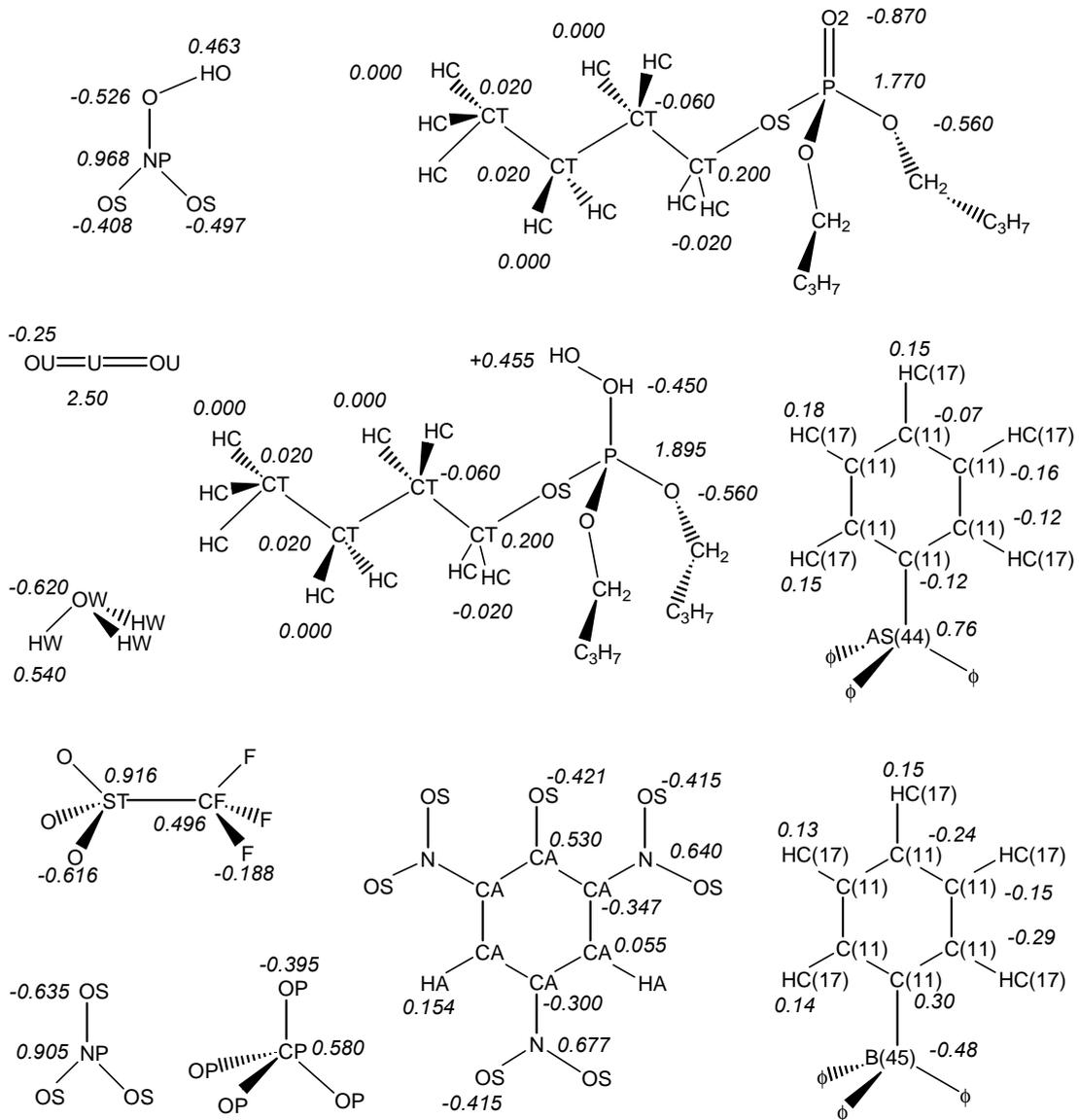


Figure 7 : Charges et types d'atomes des molécules simples et ions moléculaires. En haut les molécules neutres, au milieu les cations et en bas les anions. Pour AMBER nous donnons le code à deux lettres du type d'atome, pour GROMOS nous donnons en plus le code numérique associé ("integer atom code") entre parenthèses.

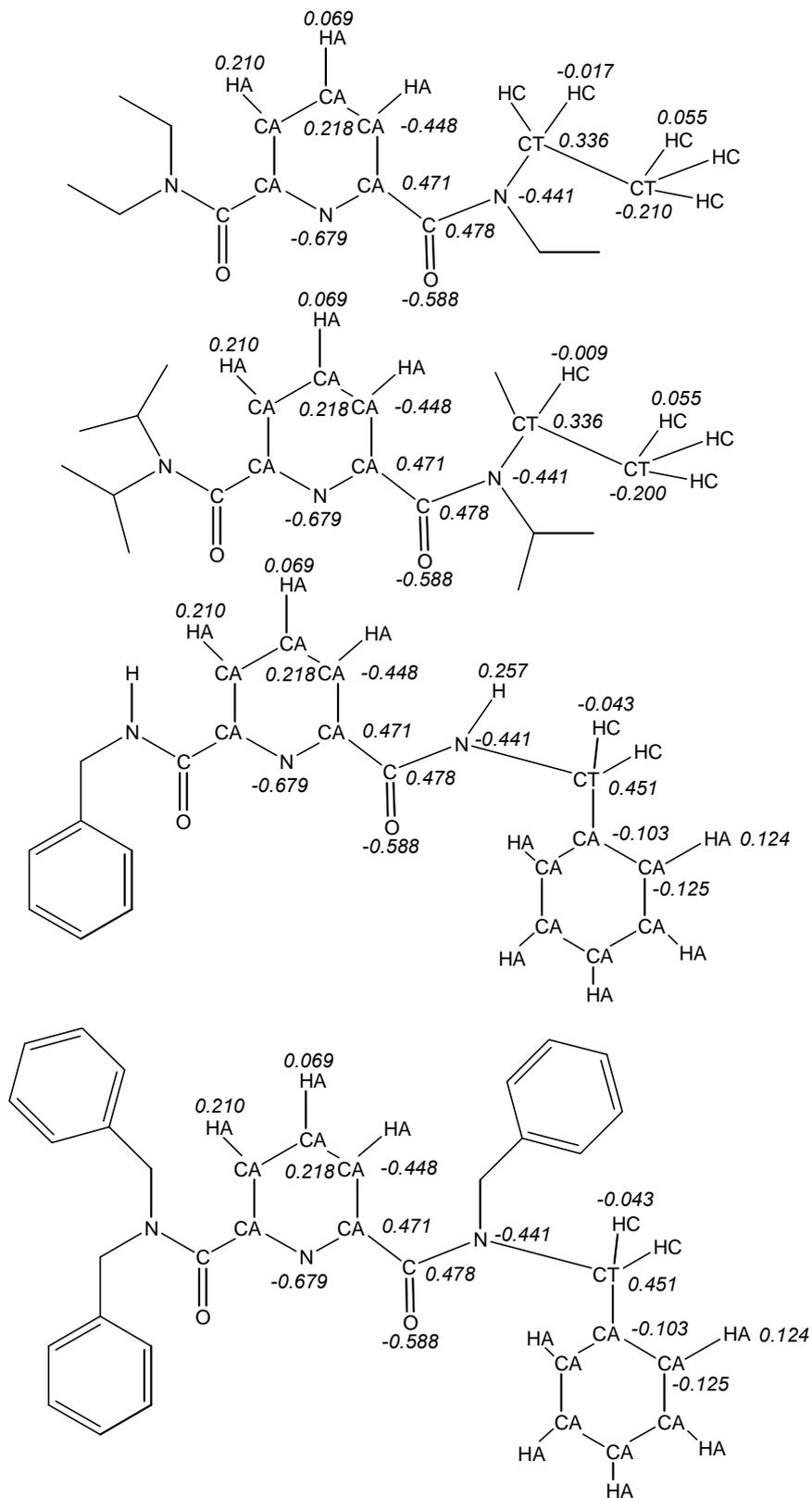


Figure 8 : Charges et types d'atomes pour les dérivés du ligand pyridine-dicarboxamide. Nous donnons le code à deux lettres du type d'atome utilisé dans AMBER.

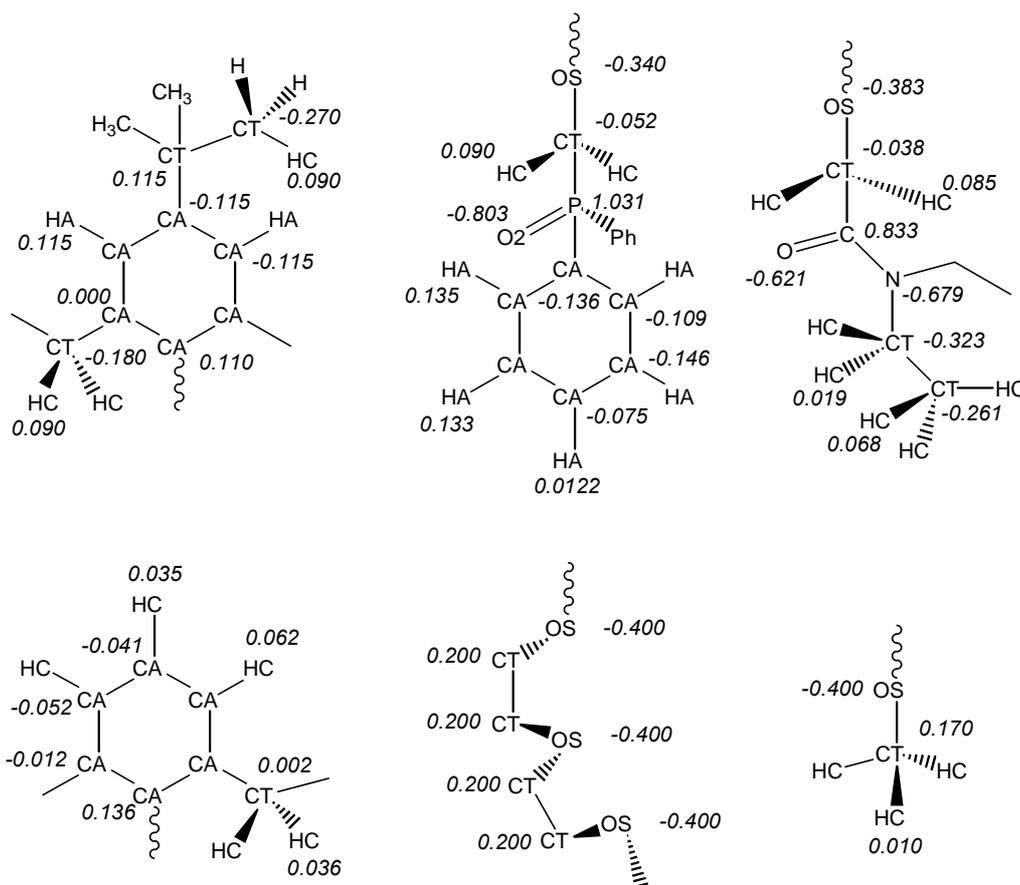


Figure 9 : Charges et types d'atomes pour les fragments de calixarènes, en haut pour les formes tétrasubstituées, en bas pour le calixarène-couronne.

4.2.3 Choix des autres paramètres · Choice of the other parameters

Les paramètres représentant les liaisons, angles, torsions et interactions de VAN DER WAALS proviennent en partie de la littérature, [22,23,38,52] de travaux antérieurs effectués au laboratoire MSM, [40-42,44-46,48-50,53,49] ou ont été mis au point en considérant des données structurales. Une source d'information structurale est la base de données cristallographiques de Cambridge, [54] une autre sont les calculs *ab initio*. † Les polarisabilités atomiques proviennent des travaux de APPLEQUIST et al. et GAO et al. [29]

Une liste des paramètres utilisés avec AMBER est donnée dans les tableaux 2 - paramètres de VAN DER WAALS et polarisabilités, 3 - paramètres des liaisons, 4 - paramètres des angles et 5 - paramètres des torsions. Pour les cations La^{3+} , Eu^{3+} et Yb^{3+} , deux sets de paramètres sont indiqués dans le tableau 2. Ceux pris de la littérature (*set2*), [38b] paramétrisés à partir de calculs d'énergie libre dans l'eau et un deuxième (*set1*) avec des

† Le protocole standard est une optimisation de géométrie au niveau HF avec la base 6-31G* ou au niveau DFT avec la fonctionnelle B3LYP. [55]

Tableau 2 : Paramètres de VAN DER WAALS et polarisabilités utilisés avec AMBER (voir formule 27).

Elément	Type(s) d'atome ^a	Masse molaire [g/mol]	R^* ^b [Å]	ϵ [kcal/mol]	α_i [Å ³]	Référence
C	C	12.010	1.908	0.0860	0.616	[22c]
C	CA	12.010	1.908	0.0860	0.878	[22c]
"	CT	"	1.908	0.1094	0.878	[22c]
"	CF (= "C3")	"	2.000	0.150	0.878	[22b]
"	CX	"	2.048	0.150	0.360	[56]
"	CZ	"	1.914	0.137		[62]
"	MX	15.03	2.119	0.207	0.878	[56]
"	MZ	13.02	2.133	0.080		[52]
Cl	CP (= "LZ"), LZ	35.450	1.947	0.300	1.91	[52]
	FZ	"	1.936	0.275	1.91	[62]
Cs	CS	132.9054	3.397	0.0001		[38a]
Eu	EU	151.960	1.650 (set1) 1.852 (set2)	0.050		[38b]
F	F	19.000	1.750	0.061	0.320	[22c]
H	H	1.008	0.600	0.0157		[22c]
"	HC	"	1.487	0.0157	0.135	[22c]
"	HA	"	1.459	0.0150	0.135	[22c]
"	HO, HW	"	0.000	0.0000		[22c]
"	HZ	"	1.577	0.020		[62]
La	LA	138.910	1.875 (set1) 2.105 (set2)	0.0600		[38b]
Li	LI	6.940	1.137	0.0183		[38a]
Na	IP	22.990	1.868	0.0028		[38a]
N	N, NP (= "N")	14.010	1.824	0.170	0.530	[22c]
"	NX	"	1.796	0.170	0.520	[56]
O	O	16.000	1.6612	0.210	0.434	[22c]
"	OP (= "O")	"	1.600	0.200	0.465	[22b]
"	OS	"	1.6837	0.1700		[22c]
"	O2	"	1.6612	0.210		[22c]
"	OU	"	1.750	0.200		[44]
"	OH	"	1.721	0.2104		[22c]
"	OW	"	1.7683	0.1520		[22c]
P	P	30.970	2.100	0.200		[22c]
S	ST	32.000	2.000	0.200	1.70	[22b]
U	U		1.580	0.400		[44]
Yb	YB	173.040	1.475 (set1) 1.656 (set2)	0.040		[38b]

^a Les types d'atomes entre parenthèses correspondent aux types indiqués dans la référence correspondante. ^b Pour les cations La^{3+} , Eu^{3+} et Yb^{3+} , deux sets de paramètres testés sont indiqués.

valeurs de R^* plus petites, en prenant la valeur de S comme R^* (Figure 2). Le rapport est de $2^{1/6} = 1.12$.

Les paramètres utilisés avec GROMOS sont ceux décrits dans le manuel [23a] pour Ar, Na^+ , Cl⁻, C et H. Nous avons introduit deux nouveaux types d'atomes pour As et B dont les paramètres de VAN DER WAALS sont ceux de SCHURHAMMER [42,46b] convertis en se servant des équations 28 et 29. On utilise $C_6(\text{As},\text{C})=5.7208 \cdot 10^{-3}$, $C_6(\text{B},\text{C})=2.5764 \cdot 10^{-3}$,

Tableau 3 : Paramètres d'élongation des liaisons covalentes utilisés avec AMBER (voir formule 16).

Liaison	K_b^{harm} [$kcal\ mol^{-1}\ \text{\AA}^{-1}$]	b_0 [\AA]	Liaison	K_b^{harm} [$kcal\ mol^{-1}\ \text{\AA}^{-1}$]	b_0 [\AA]
C -CA	469	1.409	CZ-HZ	331	1.1
C -CT	317	1.522	FZ-FZ	100	2.903
C -N	490	1.335	FZ-HZ	100	2.34
C -O	570	1.229	H -N	434	1.01
C -O2	656	1.25	HO-OH	553	0.96
C -OS	425	1.338	HO-OS	553	0.96
CA-CA	469	1.4	HW-HW	553	1.514
CA-CT	317	1.51	MX-CX	500	1.458
CA-HA	367	1.08	MX-NX	500	2.615
CA-N	337	1.41	N -OS	340	1.21
CA-NA	427	1.381	NP-OH	300	1.26
CA-OS	300	1.4	NP-OS	300	1.26
CF-F	400	1.38	O2-P	525	1.48
CP-OP	216	1.758	OH-P	230	1.61
CT-CT	310	1.526	OP-OP	100	2.903
CT-F	367	1.38	OS-P	230	1.61
CT-HC	340	1.09	OW-HW	553	0.957
CT-N	337	1.449	P -CA	331	1.81
CT-OH	320	1.41	P -CT	331	1.82
CT-OS	320	1.41	ST-CF	400	1.82
CX-NX	500	1.157	ST-O	400	1.57
CZ-FZ	216	1.758	U -OU	300	1.8

C_6 (As,H)= $1.0242 \cdot 10^{-3}$ et C_6 (B,H)= $0.4400 \cdot 10^{-3}$ $kJ\ mol^{-1}nm^6$ ainsi que C_{12} (As,C)= $0.1262 \cdot 10^{-6}$, C_{12} (B,C)= $3.3052 \cdot 10^{-6}$, C_{12} (As,H)= $1.4014 \cdot 10^{-6}$ et C_{12} (B,H)= $0.3339 \cdot 10^{-6}$ $kJ\ mol^{-1}nm^{12}$. Les paramètres de liaison sont b_0 (As...C)= 0.1906 , b_0 (B...C)= $0.1656\ nm$ et K_b (As...C)= $1.34175 \cdot 10^6$, K_b (B...C)= $1.77745 \cdot 10^6$ $kJ\ mol^{-1}nm^{-4}$. Les paramètres d'angle sont θ_0 (C...X...C)= 109.5° , θ_0 (X...C...C)= 120° , K_θ (C...X...C)= 165.65 et K_θ (X...C...C)= 223.98 $kJ\ mol^{-1}$ pour X=As ou B. Les paramètres de torsion sont nuls pour permettre une libre rotation des groupes aromatiques. [42]

Tableau 4 : Paramètres de déformation des angles entre atomes liés utilisés avec AMBER (voir formule 19).

Angle	K_{θ}^{harm} [kcal mol ⁻¹ rad ⁻²]	θ_0 [°]	Angle	K_{θ}^{harm} [kcal mol ⁻¹ rad ⁻²]	θ_0 [°]
C -CA-CA	63	120	CT-OS-CT	60	109.5
C -CA-HA	35	120	CT-OS-P	100	120.5
C -CA-N	70	120	CT-P -CA	60	106
C -CT-CT	63	111.1	CT-P -CT	60	105
C -CT-HC	50	109.5	CT-P -O2	60	114.4
C -CT-N	63	110.1	F -CF-F	80	109.5
C -CT-OS	50	109.5	F -CT-F	77	109.1
C -N -CT	50	121.9	H -N -H	35	120
C -N -H	30	120	HC-CT-HC	35	109.5
CA-C -CA	63	120	HC-CT-N	50	109.5
CA-C -N	70	120	HC-CT-OS	35	109.5
CA-C -O	70	120	HC-CT-P	35	109.5
CA-CA-CA	63	120	HO-OH-P	45	108.5
CA-CA-CT	70	120	HW-HW-OW	0	127.74
CA-CA-HA	35	120	HW-OW-HW	100	104.52
CA-CA-N	70	120	MX-CX-NX	100	180
CA-CA-OS	70	120	N -C -O	80	122.9
CA-CT-CA	70	109.5	N -CT-CA	70	120
CA-CT-CT	63	114	NP-OH-HO	45	105.3
CA-CT-HC	50	109.5	O -C -O	80	126
CA-N -C	40	112	O -ST-O	80	109.5
CA-N -CA	70	120	O2-C -O2	80	126
CA-N -H	40	107	O2-P -O2	140	119.9
CA-N -OS	70	118	O2-P -OH	45	108.23
CA-OS-CT	46.5	113	O2-P -OS	100	108.23
CA-P -CA	60	107.2	OH-P -OS	45	102.6
CA-P -O2	60	111.8	OP-CP-OP	70	109.47
CF-ST-O	80	109.5	OS-CT-P	50	114
CT-C -N	70	116.6	OS-N -OS	80	124
CT-C -O	80	120.4	OS-NP-OH	150	120
CT-C -O2	70	117	OS-NP-OS	150	120
CT-CT-CT	40	109.5	OS-P -OS	45	102.6
CT-CT-HC	50	109.5	OU-U -OU	150	180
CT-CT-N	80	109.7	P -CA-CA	40	120
CT-CT-OH	50	109.5	P -CT-C	60	112.2
CT-CT-OS	50	109.5	P -OS-P	100	120.5
CT-N -CT	50	118	ST-CF-F	80	109.5
CT-N -H	30	118.04	X -CF-X	80	109.5
CT-OH-HO	55	108.5	X -ST-X	80	109.5

Tableau 5 : Paramètres de torsion utilisés avec AMBER (voir formule 24).

Torsion	n	$K_{\theta} \cdot n$ [$kcal/mol$]	γ [$^{\circ}$]	m	Impropre	K_{θ} [$kcal/mol$]	γ [$^{\circ}$]	m
C -N -CT-C	1	0.2	180	2	C -CA-N -CA	20	180	2
CA-CA-P -CA	6	0	0	2	CA-C -O -N	20	180	2
CT-CT-C -N	1	0.1	0	4	CA-CA-CA-CT	1.1	180	2
		0.07	0	2	CA-CA-P -CA	20	180	2
CT-CT-N -C	1	0.5	180	4	CA-N -CA-CA	20	180	2
		0.15	180	3	CA-N -OS-OS	70	180	2
		0.53	0	1	CT-C -O -N	20	180	2
CT-CT-OS-CT	1	0.38	0	3	N -C -CA-O	20	180	2
		0.1	180	2	N -C -O -CA	20	180	2
CT-P -CA-CA	6	0	0	2	O -C -N -CA	20	180	2
H -N -C -O	1	2.5	180	2	O -CA-C -N	20	180	2
		2	0	1	OS-OH-NP-OS	50	180	2
N -CT-C -N	1	0.4	180	4	OS-OS-NP-OS	50	180	2
		1.35	180	2	X -CA-CA-X	20	180	2
		0.75	180	1	X -CT-N -CT	1	180	2
O2-P -CA-CA	6	0	0	2	X -O2-C -O2	10.5	180	2
OH-CT-CT-OH	1	0.14	0	3	X -X -C -O	10.5	180	2
		1	0	2	X -X -CA-HA	1.1	180	2
OH-P -OS-CT	1	0.25	0	3	X -X -N -H	1	180	2
		1.2	0	2				
OS-CT-CT-OH	1	0.14	0	3				
		1	0	2				
OS-CT-CT-OS	1	0.14	0	3				
		1	0	2				
OS-NP-OH-HO	1	10.2	180	2				
OS-P -OS-CT	1	0.25	0	3				
		1.2	0	2				
X -C -CA-X	4	14.5	180	2				
X -C -CT-X	4	0	0	2				
X -C -N -X	4	10	180	2				
X -CA-CA-X	4	14.5	180	2				
X -CA-CT-X	6	0	0	2				
X -CA-N -X	4	6.8	180	2				
X -CA-NA-X	4	6	180	2				
X -CA-OS-X	2	2	180	2				
X -CT-CT-X	9	1.4	0	3				
X -CT-N -X	6	0	0	2				
X -CT-OH-X	3	0.5	0	3				
X -CT-OS-X	3	1.15	0	3				
X -CT-P -X	9	0	0	2				
X -OH-P -X	3	0.75	0	3				
X -OS-P -X	3	0.75	0	3				
X -ST-CF-X	3	0	0	3				

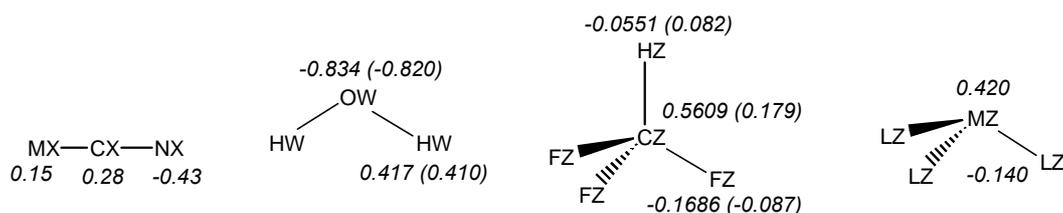


Figure 10 : Charges et types d'atomes des solvants simulés. Pour AMBER nous donnons le code à deux lettres du type d'atome et les charges. Pour GROMOS uniquement les charges des modèles correspondants sont données entre parenthèses.

4.2.4 Modèles des solvants utilisés · *Solvent models*

Les charges et types d'atomes de tous les solvants utilisés dans nos simulations sont donnés dans la figure 10 ci-dessus. Les paramètres correspondants pour AMBER, se trouvent dans les tableaux 2 à 5.

4.2.4.1 Acétonitrile – MeCN · *Acetonitrile*

La représentation du solvant est celle développée par JORGENSEN [56] pour décrire le liquide pur par des potentiels additifs par paires en 1-6-12 (OPLS). Dans ce modèle, les molécules d'acétonitrile sont linéaires et modélisées en 3 "atomes". CH₃ est donc représenté dans l'approximation de "l'atome uni". Ce modèle rend compte implicitement de la polarisation moyenne de l'acétonitrile par lui-même. La non prise en compte explicite de l'énergie de polarisation est une approximation sévère dans le cas de solutés très "durs" comme les ions lanthanides(III) qui polarisent aussi le solvant. Cet aspect sera discuté dans le chapitre suivant.

4.2.4.2 Eau – H₂O · *Water*

Avec AMBER, nous avons utilisé le modèle en 3 points TIP3P (deux atomes d'hydrogène et un oxygène) développé par JORGENSEN *et al.* pour le liquide pur. [57] Pour gagner du temps de calcul, la molécule est maintenue rigide par 3 liaisons, deux entre O et H et une pseudo-liaison H...H. Le modèle d'eau fourni avec GROMOS est appelé SPC ("Simple Point Charge") et diffère du modèle TIP3P ("Transferable Intermolecular Potential 3 Points") dans les charges atomiques utilisées (Figure 10), les paramètres de VAN DER WAALS et la géométrie. [58] Une comparaison des paramètres de ces deux modèles se trouve dans [50], page 28 et une étude systématique des modèles SPC, SPC/E, TIP3P et TIP4P dans [59]. Le nouveau modèle TIP5P qui incorpore une représentation explicite des deux paires libres sur l'oxygène de l'eau a été mis au point récemment, [60] mais n'a pas encore pu être testé dans cette thèse. Une revue des simulations de l'eau sur ordinateur a été donnée par BRODSKY. [61]

4.2.4.3 Chloroforme - CHCl_3 et MCl_3 · *Chloroform*

Dans nos simulations avec AMBER nous utilisons le modèle en 5 points de CHANG *et al.* avec polarisation, [62] mais utilisé sans polarisation. La géométrie a été définie à partir des valeurs expérimentales [63] et les charges § (Figure 10) ont été ajustées de manière à reproduire le moment dipolaire expérimental de 1.01 Debye. [64] Des polarisabilités atomiques extraites du travail d'APPLEQUIST *et al.* reproduisent correctement la polarisabilité expérimentale de la molécule. [65] Ce modèle permet ainsi la prise en compte de la polarisabilité. Les paramètres de LENNARD-JONES (Tableau 2) reproduisent les propriétés dynamiques et thermodynamiques du chloroforme liquide pur et à l'interface liquide/vapeur (enthalpie de vaporisation, densité, RDFs, coefficient de diffusion, tension de surface). Dans ce modèle il y a une charge négative sur l'hydrogène (Figure 10), ce qui est chimiquement surprenant. L'avantage d'un modèle en cinq points est de pouvoir mieux reproduire les interactions de type liaison hydrogène entre molécules de chloroforme ou avec les molécules d'eau. La molécule de chloroforme est maintenue rigide par l'application de contraintes ce qui permet de gagner du temps de calcul. Avec GROMOS nous utilisons également un autre modèle en 5 points élaboré par DIETZ et HEINZINGER. [66]

Des calculs avec AMBER effectués au laboratoire MSM, dont quelques uns seront analysés dans cette thèse, font appel au modèle OPLS du chloroforme, [52] dans lequel le groupement CH est représenté par un atome uni noté M. Il s'agit d'un modèle en 4 points qui permet de réduire de 25 à 16 le nombre de paires d'interactions à calculer par dimère. Le modèle OPLS est maintenu rigide par trois liaisons C-Cl et trois pseudo-liaisons Cl...Cl. Une comparaison des paramètres des deux derniers modèles se trouve dans [50], page 28. Les propriétés de plusieurs modèles de chloroforme ont également été décrites par MUZET. [49]

4.3 Améliorations, simplifications, artefacts et pièges · *Improvements, simplifications, artefacts and caveats*

4.3.1 Traitement des interactions entre atomes "non-liés" · *Treatment of the non-bonded interactions*

4.3.1.1 Exclusion ou modification de certaines interactions · *Exclusion or modification of certain interactions*

On appelle premiers voisins (position 1-2, Figure 11) les atomes *i* et *j* liés par une liaison covalente, et deuxième voisins les atomes *i* et *k* si *k* est lié à *j* (position 1-3, Figure 11).

§ Les charges atomiques ont été calculées à partir du potentiel électrostatique au niveau MP2 sur une géométrie optimisée au niveau Hartree-Fock dans une base 6-31G*.

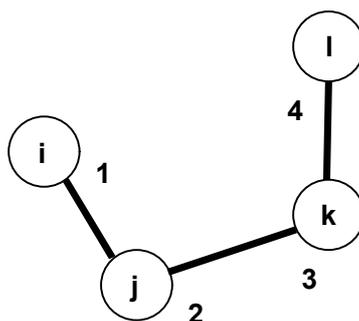


Figure 11 : Interactions entre premiers (1-2), deuxièmes (1-3) et troisièmes voisins (1-4).

La distance entre deux atomes premiers ou deuxième voisins étant courte, les interactions "non-liées" entre ces atomes voisins sera très élevée. En conséquence, *les interactions entre premiers et deuxième voisins ne sont pas pris en compte* dans la sommation de l'équation 25. Leurs interactions sont pris en compte par les termes liés. Les interactions 1-4 entre troisième voisins (Figure 11) peuvent être exclues dans certains cas (*e.g.* atomes dans un cycle aromatique avec GROMOS) ou leur valeur peut être réduite pour compenser une interaction exagérée. [67] Dans ce cas GROMOS fait appel à un set de paramètres de VAN DER WAALS plus petits [68] et AMBER permet d'attribuer un facteur multiplicatif réduisant les interactions électrostatiques 1-4 et/ou de VAN DER WAALS 1-4. [22b] Dans nos simulations nous avons traité les interactions 1-4 de la même manière que les autres interactions non-liées, sauf contre indication dans le texte. Un certain nombre d'interactions "à longue distance" sont exclues (voir distance de troncature, 4.3.1.2 ci-dessous).

4.3.1.2 Distance de troncature des interactions non-liées · *Cutoff and non-bonded interactions*

La somme de l'équation 25 s'étend à toutes les paires d'atomes sauf ceux exclus par les considérations du paragraphe précédent. Le nombre de ces paires (le nombre d'interactions à calculer) augmente avec la taille du système étudié. Pour une boîte d'eau d'environ 20 Å de côté, $0.2 \cdot 10^5$ paires d'atomes interviennent dans le calcul contre $4.4 \cdot 10^6$ pour une boîte d'eau de 30 Å de côté. La majeure partie du temps de calcul est utilisée pour rechercher ces paires d'atomes et pour évaluer leurs interactions. Il existe des méthodes pour augmenter l'efficacité de cette tâche. [69] Afin de limiter les temps de calcul, nous utilisons une double distance de troncature ("twin cutoff") au-delà de laquelle, les interactions "non-liées" (électrostatiques et VAN DER WAALS) sont ignorées. [70] La méthode est illustrée dans la figure 12. L'interaction (i,j) des atomes j à l'intérieur de la zone I, délimitée par un rayon R_c^1 à partir de l'atome i , est calculée à partir des positions instantanées de j pour chaque pas de la dynamique. Si j se trouve à $R_c^1 < r_{ij} < R_c^2$, son interaction avec l'atome i est calculée à partir d'une position sauvegardée et fixe pendant un nombre de pas N_c qui est à définir.

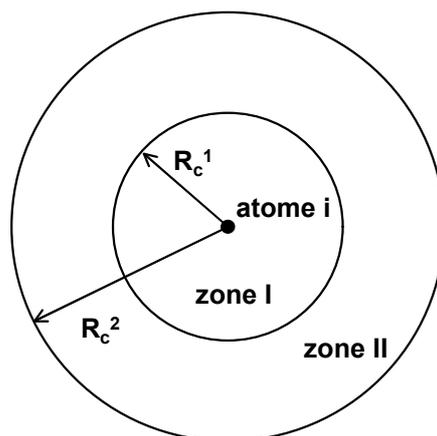


Figure 12 : Méthode de la double distance de troncature. [70]

Cette position est mise à jour après N_c (10 à 100) pas. On néglige ainsi les contributions de haute fréquence des forces à longue portée.

En général cette méthode n'est pas appliquée à des atomes seuls, mais à des groupes d'atomes. Dans ce cas, l'espace délimité par le "cutoff" est égal à l'union des sphères autour de chaque atome du groupe. Dans AMBER on parle de résidus, et dans GROMOS de groupes de charges ("charge groups"). Les résidus dans AMBER correspondent généralement aux molécules ou lorsque une molécule est constituée de plusieurs parties similaires, à chacune de ces parties. Dans GROMOS il s'agit de groupes d'atomes dont la somme des charges est neutre (ou entière pour les ions). Ceci s'impose pour éviter des artefacts (voir ci-dessous).

D'autres problèmes se posent dans des systèmes complexes contenant des espèces chargées (ions) en solution par rapport à l'échantillonnage de leurs mouvements et conformations [71] ou du calcul des énergies d'interactions aux alentours de la distance de troncature. [72] Il peut alors être nécessaire de considérer d'autres méthodes pour le traitement des interactions à longue distance telle que la sommation d'EWALD, la méthode P3M ("particle-particle-particle-mesh") ou des techniques permettant d'éviter cette discontinuité brutale, comme le champ de réaction déjà discuté ci-dessus. [73]

4.3.1.3 Interactions électrostatiques · *Limits on the electrostatic interactions*

Les interactions COULOMBIENNES de l'équation 25 sont inversement proportionnelles à la distance r_{ij} entre les atomes i et j . En conséquence cette interaction est de longue portée et ne peut pas être omise même pour des charges atomiques partielles de l'ordre de 0.1 à 0.4 e à des distances de troncature typiques de 8 à 10 Å. Deux méthodes permettant de réduire les artefacts dus au traitement des interactions "non-liées" sont présentés. Premièrement la non prise en compte des composantes à haute fréquence de la partie à longue distance de l'interaction COULOMBIENNE par l'utilisation d'une double distance de troncature décrite ci-dessus. Deuxièmement la définition d'ensembles à charge neutre, qui réduit l'interaction

électrostatique à une interaction dipolaire. Celle-ci dépend en $1 / r_{ij}^3$ de la distance. Cette dernière méthode consiste à définir des groupes de charges ("charge groups") dans GROMOS, et à calculer les interactions électrostatiques entre résidus neutres avec AMBER.

4.3.2 Traitement des bords · *Treatment of boundaries*

Dans la simulation de systèmes de taille finie il faut se soucier du traitement des bords du système considéré. La situation la plus simple est d'effectuer un calcul *in vacuo* décrit ci-dessous. Dans le cas d'un liquide ou d'une solution, il est d'usage de réduire les effets de bord par une application de *conditions périodiques*, qui seront discutées dans 4.3.2.2.

4.3.2.1 Systèmes dans le vide · *Systems in the gas phase*

La simulation d'un système *in vacuo* correspond à la phase gazeuse à pression zéro. Quand ce protocole est utilisé pour une molécule "en solution", les propriétés des atomes proches de la surface du système sont altérées, [74] et l'effet d'écrantage du solvant sur les interactions électriques entre charges et dipôles de la molécule est négligé. La molécule elle-même peut également être déformée à cause de sa tendance à minimiser son aire de surface. Les quantités de mouvement associés à la translation et à la rotation sont conservées, et il est d'usage de supprimer le mouvement de translation du centre de masse et la rotation autour de ce centre au début d'une telle simulation.

La température d'un système *in vacuo* peut être calculée par l'équation de CLAUSIUS

$$\sum_{n=1}^{N_{at}} \frac{1}{2} m_i \mathbf{v}_i^2 = \frac{1}{2} N_{df} k_B T \quad (40)$$

avec la constante de BOLTZMANN $k_B = 8.31441 \cdot 10^{-3} \text{ kJ mol}^{-1} \text{ K}^{-1}$ et le nombre de degrés de liberté du système noté par

$$N_{df} = 3N_{at} - N_c - N_{cond} \quad (41)$$

N_{at} est le nombre d'atomes, N_c le nombre de contraintes et N_{cond} une variable qui dépend des conditions de simulation. Pour un système dans le vide avec suppression du mouvement et de la rotation du centre de masse, $N_{cond} = 6$.

4.3.2.2 Conditions périodiques · *Periodic boundary conditions*

Pour réduire les effets de bord dans un système fini on utilise souvent les conditions périodiques. Les atomes du système simulé sont introduits dans une boîte cubique** entourée d'images identiques translattées de R_{box} (Figure 13).

** Ou n'importe quelle autre forme permettant de remplir périodiquement l'espace. Par exemple un rectangle, une maille monoclinique ou encore un octaèdre tronqué. [75]

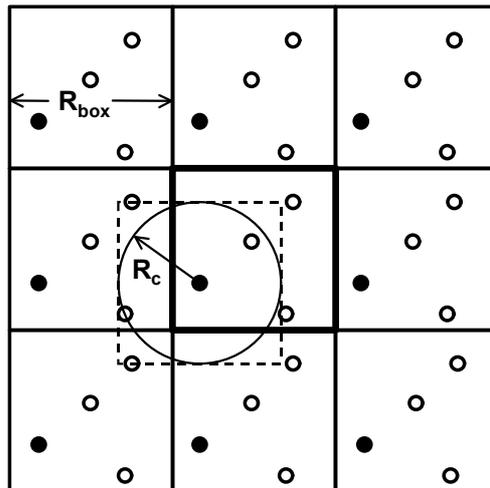


Figure 13 : Conditions périodiques illustrés en 2 dimensions avec une boîte cubique.

Le calcul des forces sur l'atome noir dans la boîte centrale s'effectue à partir des contributions des autres atomes de cette boîte ou des boîtes avoisinantes qui se trouvent à l'intérieur de la distance de troncature R_c . Quand l'atome quitte la boîte d'un côté, il entre en même temps du côté opposé, après translation de R_{box} et avec la même vitesse. L'application des conditions périodiques ressemble ainsi à l'ordre à longue distance dans un cristal. Pour une molécule en solution il s'agit d'un artefact et il faut éviter qu'un atome puisse interagir simultanément avec un autre atome et son image. C'est pourquoi dans les calculs avec AMBER et GROMOS nous ne considérons que les images les plus proches de l'atome central. En conséquence la plus petite dimension de la boîte R_{box} doit excéder deux fois la distance de troncature:

$$R_{box} > 2R_c \quad (42)$$

et quand les interactions sont calculées entre résidus/groupes de charge il faut tenir compte des deux plus grandes dimensions de ces groupes d_1^{max} et d_2^{max} :

$$R_{box} > 2R_c + d_1^{max} + d_2^{max} \quad (43)$$

Pour vérifier que les calculs sont indépendants des conditions périodiques il est utile de simuler un système dans des boîtes de différentes taille ou forme. [76]

5 Génération de la topologie, des coordonnées atomiques et mise en solution · *Generation of the topology, the atomic coordinates and addition of solvent*

5.1 Topologie du système · *Molecular topology*

Pour que les programmes de simulation reconnaissent les molécules en tant que telles il faut leur fournir des informations sur les liaisons entre atomes, la constitution de chaque molécule, les masses et charges à utiliser, *etc.* Il s'agit de la topologie moléculaire. AMBER et GROMOS ont des modules qui permettent de générer cette topologie, soit à partir d'une bibliothèque de fragments déjà existants, soit à partir de coordonnées dont on dispose, ou en créant une nouvelle structure. Une fois établie, cette information est une caractéristique du système et ne change pas pendant la simulation. Pour faciliter la mise en œuvre de systèmes complexes (ion, ligand, molécule extractante, ...) en solution, il existe des modules spécifiques se servant de "boîtes" de solvants préétablies (voir ci-dessous). Comme les molécules de solvant sont identiques entre elles, il suffit de se servir de la topologie d'une molécule comme référence.

5.2 Configuration du système · *System configuration*

Un deuxième type de données, les coordonnées atomiques et des informations dépendantes de celles-ci (les vitesses des atomes, la taille de la boîte de simulation, *etc.*) complète la description du système simulé. Ces informations relatives à la configuration du système changent constamment lors de la simulation. Au début de la simulation il faut générer une configuration initiale du système. Pour le soluté on dispose souvent de coordonnées expérimentales provenant d'une base de données [54, 77] ou de la littérature, ou encore de structures calculées à partir de géométries standard (méthodes *ab initio*, semi-empiriques, mécanique moléculaire). Les boîtes de solvant introduites dans 5.1 sont une source de coordonnées de référence déjà pré-équilibrées par des méthodes de dynamique moléculaire ou MONTE-CARLO. [30c,78]

5.3 Mise en solution dans une boîte de liquide pur · *Putting the solute in a pure solvent box*

Dans les calculs en solution, le système simulé est constitué de boîtes "cubiques" contenant le solvant et le soluté, reproduites périodiquement dans les trois dimensions. Ci-

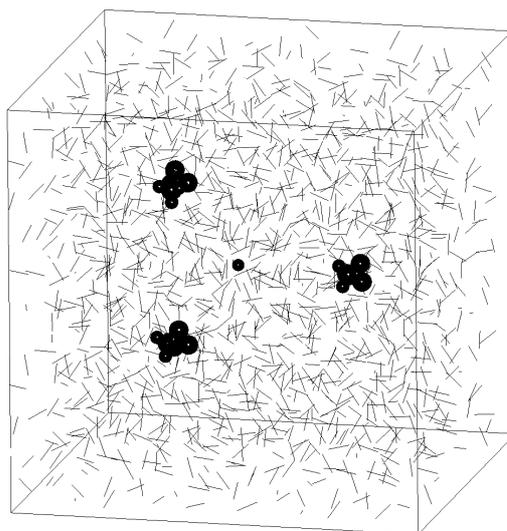


Figure 14 : Le sel triflate de l'euporium (paire dissociée) après mise en solution dans l'acétonitrile.

dessous une telle boîte est donnée en guise d'exemple (Figure 14). La taille de la boîte R_{box} doit être choisie en fonction des dimensions du soluté pour éviter des artefacts dans le calcul (équation 43). Pour construire cette boîte de simulation, le soluté est placé au centre d'un grand cube constitué d'une concaténation de n^3 boîtes de référence du solvant. On découpe un volume adapté en retirant toutes les molécules du solvant dont l'un des atomes est situé à moins de 2 Å du soluté et le système est réduit à une seule boîte dont les dimensions sont définies en fonction du "cutoff" utilisé.

5.4 Interface liquide/liquide · *Liquid-liquid interface*

Pour construire une boîte de simulation contenant une interface liquide-liquide entre deux solvants, nous partons des boîtes d'eau et de chloroforme. Elles ont été équilibrées séparément avant d'être assemblées (Figure 15). Pour éviter une superposition entre les molécules des deux solvants sur la face commune en début de calcul, toutes les molécules du second solvant dont l'un des atomes est situé à moins de 2 Å d'un atome du premier solvant, sont retirées. Le soluté est placé au centre de la boîte de la même manière que pour les boîtes

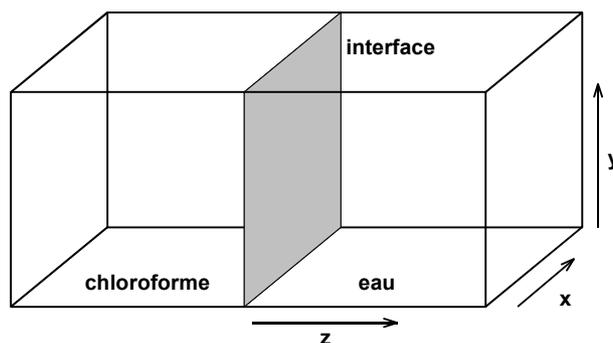


Figure 15 : Schéma d'une boîte de simulation avec interface liquide-liquide.

de liquides purs.

Dans le cas de boîtes possédant une interface, plusieurs manières de traiter la périodicité ont été discutées dans la littérature. Une périodicité en trois dimensions a été utilisée par LINSE et KESSLER. [79] BENJAMIN *et al.* ont appliqué la périodicité uniquement suivant x et y . [80] On peut en plus appliquer des contraintes au bords de boîte sur l'axe z . [23a,50] Dans cette thèse nous avons utilisé la périodicité en 2 dimensions^{††} et la périodicité en 3 dimensions. Dans ce dernier cas - compte tenu de la périodicité - on obtient deux interfaces dans la boîte de simulation. Pour éviter que le soluté interagisse simultanément avec ces deux interfaces, il faut choisir $R_{\text{box}}(z)$ suffisamment grand.

6 Mécanique et dynamique moléculaires

Molecular mechanics and molecular dynamics

6.1 Minimisation de l'énergie · *Energy minimisation*

La mécanique moléculaire nous permet de minimiser l'énergie calculée à partir de l'équation 14 afin d'obtenir des configurations à basse énergie de notre système moléculaire et de réduire des forces initiales trop grandes qui aboutiront à une trajectoire aberrante. Nous avons utilisé deux méthodes de gradient faisant appel à l'énergie potentielle $V(\mathbf{r}; s)$ et sa dérivée suivant \mathbf{r} . L'inconvénient de ces méthodes est qu'on suit à peu de choses près le gradient de l'énergie potentielle en se déplaçant de $\Delta\mathbf{r}_i$ dans l'espace configurationnel.

$$\Delta\mathbf{r}_i \approx - \frac{\partial}{\partial\mathbf{r}_i} V(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) \quad (44)$$

De cette façon on se déplace presque exclusivement vers le bas sur l'hypersurface énergétique, et le minimum atteint est en général le *minimum local* le plus proche de la configuration initiale (Figure 16). Il n'est pas possible de surmonter des barrières énergétiques. Dans ce cas la dynamique moléculaire est un moyen plus efficace de recherche de configurations de basses énergies, elle permet de surmonter des barrières de l'ordre de $k_B T$. [81]

La première méthode de minimisation utilisée dans nos calculs est la *méthode des gradients de plus grande pente* ("steepest descent") de COURANT. [82] Le $n+1$ ème pas de minimisation est constitué du calcul des forces $\mathbf{f}(t_n)$ suivant l'équation 10 pour la configuration $\mathbf{r}(t_n)$ donnée, et du calcul de la configuration suivante $\mathbf{r}(t_{n+1})$, avec un pas de Δx .

$$\mathbf{r}(t_{n+1}) = \mathbf{r}(t_n) + N_n \Delta x \mathbf{f}(t_n) \quad (45)$$

^{††} Avec un potentiel harmonique qui repousse les molécules vers l'intérieur de la boîte, aux extrémités en z

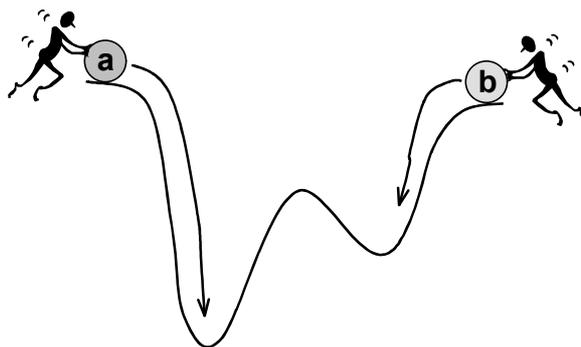


Figure 16 : Inconvénient des méthodes de gradient : le minimum local atteint par les deux boules a et b est différent et dépend de la position (configuration) initiale.

N_n est un facteur de normalisation. Au départ une valeur initiale de Δx est donnée. Aussi longtemps que l'énergie potentielle décroît, on augmente Δx . Si $V(\mathbf{r})$ croît, on diminue Δx . Cette méthode permet de descendre rapidement vers un minimum local, mais elle ne converge pas.

Dans la *méthode des gradients conjugués*, [83] la force (le gradient) $\mathbf{f}(t_{n+1})$ calculé à chaque itération est conjugué au précédent pour déterminer le vecteur unitaire $\mathbf{p}(t_{n+1})$ de la direction de descente sur l'hypersurface énergétique.

$$\mathbf{p}(t_{n+1}) = \mathbf{f}(t_{n+1}) + \beta_n \mathbf{p}(t_n) \quad (46)$$

$$\beta_n = \frac{\langle \mathbf{f}(t_{n+1}) | \mathbf{f}(t_{n+1}) \rangle}{\langle \mathbf{f}(t_n) | \mathbf{f}(t_n) \rangle} \quad (47)$$

Le pas de la minimisation, s_{min} , est déterminé en prenant en compte le gradient de l'itération précédente, ce qui permet de l'optimiser par rapport à la première méthode décrite. La nouvelle configuration est obtenue par

$$\mathbf{r}(t_{n+1}) = \mathbf{r}(t_n) + s_{min} \mathbf{p}(t_n) \quad (48)$$

L'arrêt de la minimisation se fait sur un critère de convergence ou sur un nombre maximal d'itérations. Dans nos calculs nous avons choisi $1.0 \cdot 10^{-4} \text{ kcal mol}^{-1} \text{ \AA}^{-1}$ comme critère et généralement 100-5000 itérations selon la complexité des systèmes. Environ la moitié des pas de minimisation sont faits par la méthode de plus grande pente, la suite par les gradients conjugués.

Les méthodes de minimisation ont été décrites et discutées dans la littérature. [84]

6.2 Dynamique moléculaire · *Molecular Dynamics*

6.2.1 Principe · *Principle*

Les simulations de dynamique moléculaire à partir d'une fonction d'énergie empirique comme celle donnée dans l'équation 14 sont largement utilisées pour étudier la

structure, la dynamique et certains aspects thermodynamiques de systèmes moléculaires à l'équilibre et hors équilibre. Les équations de mouvement des atomes issues de la mécanique classique sont intégrées dans le temps. Dans un système de coordonnées CARTESIENNES, ces équations sont celles de NEWTON (11-13), avec les forces \mathbf{f}_i égales au gradient de l'énergie potentielle mais de signe opposé.

Dans AMBER on utilise l'algorithme de VERLET [85] pour l'intégration des équations de mouvement, dans GROMOS l'algorithme du "saut de mouton" ("*leap-frog*"). Les deux procédures sont décrites ci-dessous. Dans une simulation de DM standard, l'énergie totale E du système moléculaire est une constante. En général on garde également le nombre d'atomes N et le volume de la boîte de simulation fixe. Ces simulations se déroulent alors dans l'ensemble microcanonique (N,V,E) . Souvent il est préférable de maintenir la température constante plutôt que l'énergie, il s'agit alors de simulations (N,V,T) . Dans d'autres cas il peut être avantageux de simuler à pression constante plutôt qu'à volume constant, ce sont des simulations (N,P,T) . Pour les simulations (N,V,T) ou (N,P,T) il est nécessaire de coupler le système moléculaire à un bain thermique ou un bain de pression, respectivement. Ceci est décrit dans 6.2.3 et 6.2.4. A la fin de cette section nous décrivons l'application de contraintes au système moléculaire par les algorithmes SHAKE et BELL, et les conditions initiales utilisées pour nos calculs.

6.2.2 Algorithmes de VERLET et du "saut de mouton" · *VERLET and leap-frog algorithm*

Les deux algorithmes sont équivalents. Dans ce qui suit nous décrivons l'algorithme du saut de mouton. A partir de la différence des deux séries de TAYLOR des vitesses $\mathbf{v}_i(t_n - \Delta t/2)$ à $t = t_n$ et $\mathbf{v}_i(t_n + \Delta t/2)$ à $t = t_n$ on obtient

$$\mathbf{v}_i(t_n + \Delta t/2) = \mathbf{v}_i(t_n - \Delta t/2) + m_i^{-1} \mathbf{f}_i(t_n) \Delta t \quad (49)$$

en négligeant les termes de 3^{ème} ou plus grand ordre. De même, la différence entre les séries de Taylor des positions $\mathbf{r}_i(t_n)$ à $t = t_n + \Delta t/2$ et $\mathbf{r}_i(t_n + \Delta t)$ à $t = t_n + \Delta t/2$ donne

$$\mathbf{r}_i(t_n + \Delta t) = \mathbf{r}_i(t_n) + \mathbf{v}_i(t_n + \Delta t/2) \Delta t \quad (50)$$

Ces deux équations permettent l'intégration dans le temps suivant le schéma montré dans la figure 17.

Pour en déduire l'algorithme de VERLET, il faut éliminer les vitesses \mathbf{v}_i dans les équations 49 et 50 et remplacer t_n par $t_n - \Delta t$ dans l'équation 50. On obtient

$$\mathbf{r}_i(t_n + \Delta t) = 2 \mathbf{r}_i(t_n) - \mathbf{r}_i(t_n - \Delta t) \Delta t + m_i^{-1} \mathbf{f}_i(t_n) (\Delta t)^2 \quad (51)$$

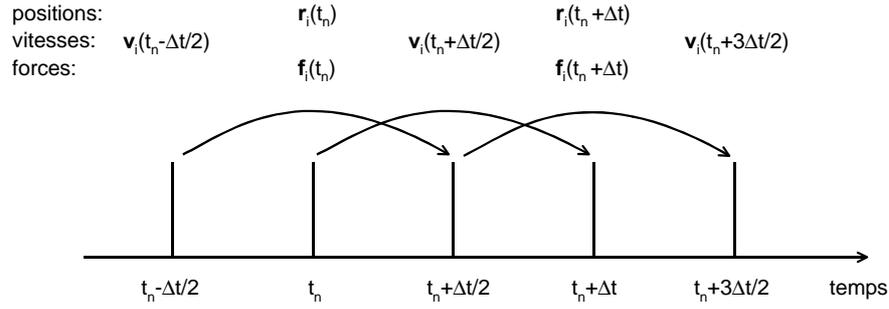


Figure 17 : Intégration des équations de mouvement suivant l'algorithme du saut de mouton.

La particularité de ce dernier algorithme est que la vitesse des atomes n'apparaît pas explicitement dans l'équation, et ainsi un couplage avec un bain thermique par ajustement des vitesses devient impossible.

Le temps de calcul nécessaire est directement proportionnel au pas d'intégration Δt . Le pas Δt doit rester petit devant la période de vibration la plus courte dans le système pour pouvoir représenter ce mouvement. En général ce sont les vibrations X-H qui limitent Δt à 10^{-15} s = 1 fs.

D'autres algorithmes sont ceux de RUNGE-KUTTA, [86] GEAR [87] ou BEEMAN. [88] Au niveau des performances, l'algorithme leap-frog semble le plus avantageux. [89,93b]

6.2.3 Contrôle de la température · *Temperature coupling*

Plusieurs méthodes pour le contrôle de température dans les simulations de DM ont été proposées. [90] Une méthode simple, que nous utilisons dans nos simulations, est la méthode de faible couplage ("weak coupling method"), dans laquelle l'équation de mouvement des atomes est modifiée afin d'obtenir une relaxation du premier ordre de la température T vers la température de référence T_0 . [91]

$$\frac{dT(t)}{dt} = \tau_T^{-1} [T_0 - T(t)] \quad (52)$$

La température d'un ensemble de N_{dF} degrés de liberté est défini par l'équation de CLAUSIUS (40) en termes de l'énergie cinétique de ces degrés de liberté. Le contrôle de température peut ainsi s'effectuer par une modification des vitesses des atomes par un facteur de correction $\lambda(t)$.

$$\lambda(t) = \left[1 + \frac{\Delta t}{\tau_T} \left[\frac{T_0}{T(t)} - 1 \right] \right]^{\frac{1}{2}} \quad (53)$$

La vitesse de relaxation de la température est contrôlée par le *temps de relaxation de la température* τ_T . Ce paramètre peut être ajusté en fonction du système. [16h] Il faut que τ_T soit

suffisamment petit (couplage fort) pour maintenir la température moyenne à T_0 , mais suffisamment grand (couplage faible) pour éviter de perturber le système.

Différentes parties du système moléculaire peuvent présenter des temps de chauffe différents, par exemple par l'utilisation d'une distance de troncature sur les interactions non-liés. Il est alors préférable de coupler les différentes parties du système séparément à des bains thermiques. Dans nos simulations à l'interface liquide-liquide, nous avons systématiquement séparé les deux solvants. Sauf exceptions précisées dans le texte, nous avons utilisé une température de référence de 300 K et un temps de relaxation de 0.1 ps.

6.2.4 Contrôle de la pression ·*Pressure coupling*

Le contrôle de la pression s'effectue également par une méthode de faible couplage, comme le contrôle de la température. La pression prend le rôle de la température et les positions atomiques celui des vitesses. On modifie les équations de mouvement pour obtenir une relaxation de la pression P au premier ordre vers la valeur de référence P_0 .

$$\frac{dP(t)}{dt} = \tau_p^{-1} [P_0 - P(t)] \quad (54)$$

La pression peut être définie en se servant du théorème du viriel,

$$P(t) = \frac{2}{3} \frac{E_{kin}(t) - W(t)}{V(t)} \quad (55)$$

où $W(t)$ est le viriel et $V(t)$ le volume de la boîte de simulation. Le *viriel moléculaire* est défini par

$$W(t) = -\frac{1}{2} \sum_{\alpha < \beta}^{N_M} \mathbf{R}_{\alpha\beta}(t) \mathbf{F}_{\alpha\beta}(t) \quad (56)$$

avec $\mathbf{R}_{\alpha\beta}(t)$ la distance entre les centres de masses des molécules α et β à l'instant t et $\mathbf{F}_{\alpha\beta}(t)$ la force sur le centre de masse de α exercée par la molécule β . Comme la pression à température constante est reliée au volume par la compressibilité isotherme κ_T , le couplage s'effectue par l'ajustement des coordonnées atomiques et de la taille de la boîte de simulation par un facteur de correction μ . Dans le cas d'un ajustement isotropique on obtient

$$\mu(t) = \left[1 - \kappa_T \frac{\Delta t}{\tau_p} [P_0 - P(t)] \right]^{\frac{1}{3}} \quad (57)$$

La valeur de κ_T pour l'eau, utilisée par défaut, est de $44.6 \cdot 10^{-6} \text{ bar}^{-1}$. [92] Comme le *temps de relaxation de la pression* τ_p est un paramètre adaptable, il n'est pas nécessaire de connaître κ_T du système avec certitude. Comme la définition de la pression dépend de l'énergie cinétique, le couplage de la pression ne doit pas être plus fort que celui de la température.

$$\tau_p \geq \tau_T \quad (58)$$

Pour nos calculs à pression constante, nous avons utilisé une valeur de 1.0 ps et une pression de référence de 1 atm.

6.2.5 Algorithme SHAKE · *SHAKE algorithm*

La méthode SHAKE est utilisée pour appliquer des *contraintes de distance* au système moléculaire. [93] Typiquement des contraintes sur les longueurs de liaisons ou les angles peuvent être mis sous la forme d'une contrainte de distance entre les atomes $k1$ et $k2$,

$$\sigma_k(\mathbf{r}) \equiv r_{k_1 k_2}^2 - d_{k_1 k_2}^2 = 0 \quad (59)$$

qui sont contraint à la distance de référence $d_{k_1 k_2}$ et séparés par le vecteur instantané $r_{k_1 k_2}$. Schématiquement l'algorithme est illustré dans la figure 18. La correction sur les positions s'effectue après un pas de dynamique sans contrainte. Mathématiquement on assure la validité de l'équation 59 par l'application de la méthode des multiplicateurs de LAGRANGE. Comme les atomes peuvent être impliqués dans plusieurs contraintes, les équations sont résolues par itération. Ces itérations s'arrêtent soit quand $\sigma_k(\mathbf{r})$ est inférieur à un seuil de tolérance (typiquement 0.0005), le nombre d'itérations est trop grand (> 1000), ou le déplacement lors du pas de dynamique sans contrainte est trop grand pour que SHAKE puisse satisfaire les contraintes (Figure 18).

Nous avons appliqué SHAKE sur toutes les liaisons faisant intervenir un atome d'hydrogène, pour éliminer les plus hautes fréquences de vibration présentes dans nos systèmes. La période d'oscillation de ces vibrations est la borne supérieure du pas d'intégration Δt . En limitant ces vibrations, souvent sans importance pour les mouvements globaux, il est possible d'augmenter Δt sans affecter le résultat de la DM. Toutes les liaisons des solvants sont également contraintes, pour gagner en temps de calcul. [94]

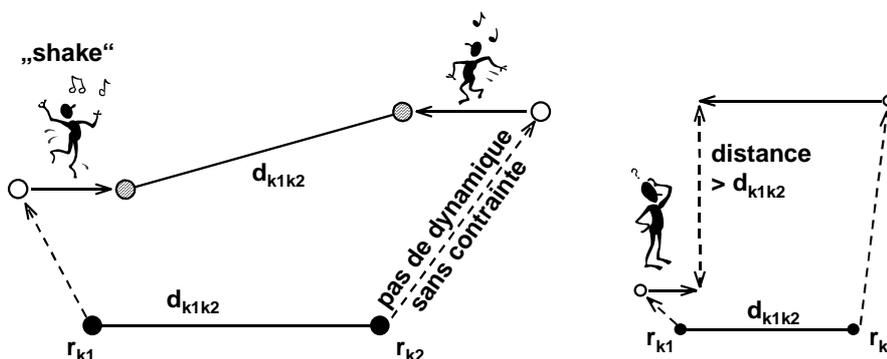


Figure 18 : Méthode SHAKE : à gauche sont illustrées les corrections sur les positions des atomes initiales r_{k1} et r_{k2} (cercles pleins) après un pas de dynamique sans contraintes (cercles vides). Après SHAKE (cercles gris) la distance est celle fixée au départ. A droite une situation dans laquelle SHAKE échoue, à cause d'un déplacement trop grand lors du pas de dynamique sans contrainte.

6.2.6 Dynamique sous contraintes : SHAKE et BELLY · *Constraint dynamics : SHAKE and BELLY*

L'application de contraintes dans les simulations de DM peut entraîner des difficultés dans les calculs. L'algorithme SHAKE a été présenté ci-dessus. Il peut être appliqué à toute sorte de liaison et aux angles. Une forme de contraintes sont les contraintes sur les positions des atomes. L'option BELLY dans AMBER permet de définir un groupe d'atomes dans le système moléculaire qui est contraint à sa position initiale et ainsi maintenu immobile pendant la durée de la simulation. Ceci peut être utile pour l'équilibration d'une partie du système, par exemple pour relaxer les molécules de solvant uniquement, sans affecter la configuration du soluté.

Pour que la dynamique sous contraintes soit physiquement raisonnable et valable il faut considérer plusieurs aspects. Les fréquences des degrés de liberté "gelés" doivent être (considérablement) plus grandes que les autres pour permettre une augmentation notable de Δt . Il faut également s'assurer que ces degrés de liberté sont faiblement couplés aux autres uniquement pour ne pas affecter le mouvement des molécules par l'application des contraintes. Les effets métriques de tenseurs sont également à considérer. [95]

6.2.7 Protocole de simulation · *Simulation protocol*

Dans nos simulations de DM avec AMBER nous avons suivi un protocole semblable pour tous les systèmes traités. Les structures de départ, provenant de l'expérience ou d'une construction par infographie, ont été équilibrées en plusieurs étapes. D'abord par minimisation d'énergie (méthode de plus grande pente, suivie par les gradients conjugués) suivi d'une dynamique courte (50 ps) et d'une deuxième minimisation dans le vide. Après la mise en solution on minimise à nouveau. Si le soluté est composé de plusieurs espèces séparées par le solvant (paires d'ions, complexes, etc.), une dynamique moléculaire est effectuée en relaxant le solvant uniquement. Pour garantir une bonne densité du système, on effectue une seconde dynamique d'équilibration à pression constante (1 atm). Cela permet d'ajuster le volume et de corriger la densité. Ensuite les simulations de DM proprement dites ont été commencées avec des vitesses aléatoires (distribution de MAXWELL) à une température de 300 K et à volume constant. Le pas de temps pour l'intégration des équations du mouvement est de 1 femtoseconde. Les mouvements d'élongation de haute fréquence des hydrogènes et du solvant sont gelés par l'application de l'algorithme SHAKE. La durée des simulations de DM varie en fonction du système. Les interactions en 1-6-12 ont été calculées dans la limite d'un double "cutoff" de 12 / 15 Å, avec mise à jour de la liste des paires toutes les 10 fs, évaluées entre résidus.

Dans les calculs avec GROMOS, nous avons suivi un protocole semblable. Les différences concernent la double distance de troncature qui est de 8/14 Å et l'application systématique de la correction du champ de réaction au-delà de cette limite.

6.3 Calculs de variation d'énergie libre · *Free energy calculations*

Il existe plusieurs méthodes pour calculer la différence d'énergie libre entre deux états A et B d'un système moléculaire ou entre deux systèmes moléculaires A et B, fondées sur la mécanique statistique. [96] Dans l'approche du paramètre de couplage ("*coupling parameter approach*"), l'HAMILTONIEN devient une fonction analytique du paramètre de couplage λ ,

$$H(\mathbf{p}, \mathbf{r}; \lambda) = K(\mathbf{p}; \lambda) + V(\mathbf{r}; \lambda) \quad (60)$$

avec les deux HAMILTONIENS qui caractérisent les états ou systèmes A et B donnés par

$$H(\mathbf{p}, \mathbf{r}; \lambda_A) = H_A(\mathbf{p}, \mathbf{r}) \quad (61)$$

$$H(\mathbf{p}, \mathbf{r}; \lambda_B) = H_B(\mathbf{p}, \mathbf{r}) \quad (62)$$

En utilisant l'HAMILTONIEN sous la forme qui dépend de λ , l'énergie libre de HELMHOLTZ F devient également une fonction de λ ,

$$F(\lambda) = -k_B T \ln Z(\lambda) \quad (63)$$

où Z est la fonction de partition canonique normalisée du système de N atomes donnée par

$$Z(\lambda) = N_{norm}^{-1} \iint e^{-H(\mathbf{p}, \mathbf{r}; \lambda) / k_B T} d\mathbf{p} d\mathbf{r} \quad (64)$$

Ce qui permet de calculer la différence d'énergie libre ΔF_{BA} par

$$\Delta F_{BA} = F(\lambda_B) - F(\lambda_A) = -k_B T \ln \langle e^{-[H(\lambda_B) - H(\lambda_A)] / k_B T} \rangle_{\lambda_A} \quad (65)$$

qui est généralement appelée la *formule de perturbation de l'énergie libre*, initialement établie par ZWANZIG. [97] Dans cette expression, $\langle \dots \rangle_{\lambda_A}$ désigne la moyenne sur l'ensemble statistique généré avec l'HAMILTONIEN $H(\lambda_A)$. Cette moyenne ne converge que si la différence entre les deux états est petite. A partir de la dérivée de $F(\lambda)$ par rapport à λ on obtient la *formule de l'intégration thermodynamique*, équation 67, qu'on doit à KIRKWOOD. [98]

$$F'(\lambda) \equiv \frac{dF(\lambda)}{d\lambda} = \left\langle \frac{\partial H(\lambda)}{\partial \lambda} \right\rangle_{\lambda} \quad (66)$$

$$\Delta F_{BA} = \int_{\lambda_A}^{\lambda_B} F'(\lambda) d\lambda = \int_{\lambda_A}^{\lambda_B} \left\langle \frac{\partial H(\lambda)}{\partial \lambda} \right\rangle d\lambda \quad (67)$$

Ces formules sont strictement applicables aux calculs (N,V,T) dans l'ensemble canonique et mènent aux énergies libres de HELMHOLTZ. Les formules correspondantes pour les simulations (N,P,T) mènent aux énergies libres de GIBBS et sont données dans [16c].

Dans les calculs d'énergie libre standard, les masses atomiques $m(\lambda)$ et les paramètres du champ de force $s(\lambda)$ deviennent des fonctions du paramètre de couplage λ . Dans le cas de la méthode de fenêtrage que nous utilisons avec AMBER, les équations qui relient la fonction d'interaction à λ sont données dans [99]. Le calcul de ΔF_{BA} est fragmenté en "fenêtres" correspondant chacune à une valeur de λ qui varie entre 0 ($H(\lambda = 0) = H_A$) et 1 ($H(\lambda = 1) = H_B$). On obtient la différence d'énergie libre par

$$\Delta F_{BA} = F(\lambda_B) - F(\lambda_A) = -\sum_{\lambda=0}^1 k_B T \ln \langle e^{-[H(\lambda+\delta\lambda)-H(\lambda)]/k_B T} \rangle_{\lambda} \quad (68)$$

Le pas $\delta\lambda$ est pris suffisamment petit ($\delta\lambda=0.02$) de manière à vérifier l'approximation d'une perturbation. En général, on effectue à chacune des 51 fenêtres une période d'équilibration de 2 ps, et une période de collection de données de 3 ps. Dans ces calculs nous utilisons un simple "cutoff" de 12 Å, parce que la version 4.1 d'AMBER ne permet pas d'effectuer ce type de calcul avec un double "cutoff".

D'autres méthodes et approches pour calculer des différences d'énergie libre, dans le contexte du calcul d'un potentiel de forces moyennes, seront discutées plus en détail dans le chapitre 5 de cette thèse. Il s'agit entre autre de l'échantillonnage sous contraintes ("*umbrella sampling*"), de la méthode de croissance lente ("*slow growth*") et de l'intégration thermodynamique. Ci-dessous nous illustrons l'application des calculs de variation d'énergie libre aux systèmes décrites dans les chapitres 2 à 4 de cette thèse.

6.3.1 Application à la variation de l'énergie libre de complexation

DDF_c · *Application to the calculation of binding free energies*

Dans ce travail nous nous intéressons à l'extraction liquide-liquide et aux phénomènes de complexation. Pour quantifier la reconnaissance moléculaire d'une molécule extractante pour un substrat, on se sert de la différence d'énergie libre de complexation relative $\Delta\Delta F_c = \Delta F_1 - \Delta F_2$ entre deux substrats 1 et 2 par un même complexant L. La prédiction par les calculs de cette quantité thermodynamique a subi d'énormes progrès. Dans les systèmes étudiés, les substrats sont en général des cations M_1^{n+} et M_2^{n+} . Comme la simulation directe du processus de complexation est souvent difficile, l'approche introduite par WIPFF *et al.* [100] consiste à utiliser un "chemin alchimique" via des états intermédiaires non physiques, dans notre cas en "mutant" le cation M_1^{n+} en M_2^{n+} d'abord libre puis complexé.

Les variations d'énergie libre ΔF_3 et ΔF_4 correspondent respectivement à la mutation des ions libres et des complexes en solution qui peut s'effectuer avec une faible incertitude statistique. Comme F est une fonction d'état, la différence d'énergie libre ΔF est

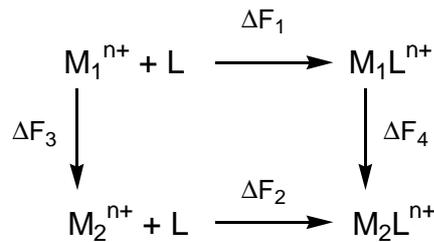


Schéma 1 : Cycle thermodynamique de complexation de deux cations M_1^{n+} et M_2^{n+} par un ligand L .

indépendante du chemin qui relie les états finals, aussi longtemps que le système est en équilibre et que la mutation s'effectue de façon réversible. [101] Dans le cas du cycle thermodynamique fermé qui est montré dans le schéma 1, on a $\Delta F = 0$. Ainsi $\Delta\Delta F_c$ peut également s'écrire

$$\Delta\Delta F_c = \Delta F_4 - \Delta F_3 \quad (69)$$

Des exemples se trouvent dans la littérature. [16e,96a,102] La variation d'énergie libre ΔF_4 n'est pas accessible expérimentalement, mais les valeurs de ΔF_3 calculées peuvent être comparées aux variations d'énergie libre de solvation expérimentales. [103]

En pratique, on transforme progressivement le cation M_1^{n+} ($\lambda = 0$) en M_2^{n+} ($\lambda = 1$) en couplant les paramètres ϵ et R^* correspondant pour simuler un ion "hybride" ($\lambda \neq 0$) par

$$\epsilon_\lambda = \lambda \epsilon_{M_1^{n+}} + (1 - \lambda) \epsilon_{M_2^{n+}} \quad (70)$$

$$R_\lambda^* = \lambda R_{M_1^{n+}}^* + (1 - \lambda) R_{M_2^{n+}}^* \quad (71)$$

comme décrit ci-dessus (méthode de fenêtrage). A chaque fenêtrage, on calcule la variation ΔF_λ par rapport à l'état précédent $\lambda - \delta\lambda$ et l'état $\lambda + \delta\lambda$. Les deux valeurs de ΔF calculées, dans le sens direct ("forward") et dans le sens inverse ("backward"), doivent en théorie être égales. Leur écart donne une indication sur l'erreur minimum des valeurs de $\Delta\Delta F_c$ calculées.

7 Analyse des résultats · *Analysis of the results*

Les données générées lors de nos simulations de DM sont les trajectoires des atomes dans l'espace configurationnel $\mathbf{r}(t)$ et leurs vitesses instantanées $\mathbf{v}(t)$. Ces trajectoires ont été stockées toutes les 0.2 ps dans le cas des dynamiques et toutes les 0.5 ps pour les calculs de variation d'énergie libre. Des fichiers contenant les caractéristiques du calcul (température, énergies cinétiques et potentielles du système, ...) ont également été créés pour permettre une première analyse rapide de chaque simulation, bien que ces grandeurs soient directement reproductibles à partir des trajectoires stockées (\mathbf{r} , \mathbf{v}).

7.1 Analyse infographique · *Visual analysis*

La première étude du système est une observation visuelle, infographique. [104] Nous avons utilisé le logiciel DRAW, [105] mis au point au sein du laboratoire MSM, qui permet de suivre l'évolution globale du système en visualisant la dynamique pas à pas. On peut ne visualiser qu'une partie du système, par exemple le soluté sans le solvant, pour mieux déceler son comportement. DRAW permet également d'effectuer un certain nombre de mesures ponctuelles d'énergie ou de géométrie en temps réel. Une grande partie des images de systèmes moléculaires qui illustrent cette thèse ont été créées avec ce logiciel, et la figure couleur 1-1 dans le tome II en donne un aperçu. D'autres vues ont été créées avec les logiciels VMD [106] et SCHAKAL. [107]

Plus d'informations sur l'analyse infographique se trouvent dans [108].

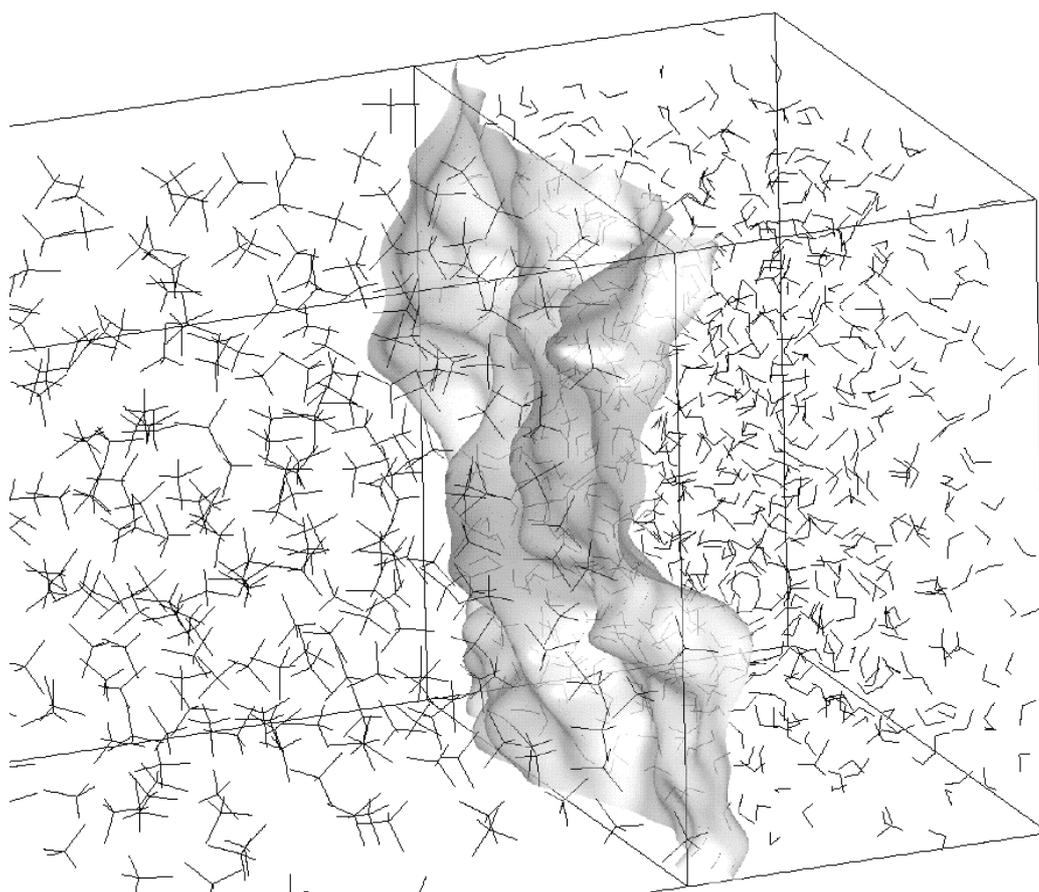


Figure 19 : Visualisation de l'interface eau/chloroforme avec le logiciel DRAW. [105]

Dans le cas particulier de systèmes présentant une interface liquide-liquide, la visualisation de la surface de contact entre les deux solvants nous a permis de décrire la rugosité de cette interface qualitativement. En guise d'exemple la figure 19 montre l'interface eau-chloroforme qui n'est pas complètement plane comme on pourrait le croire.

7.2 Analyses géométriques et énergétiques · *Geometry and energy analysis*

Nous utilisons le programme MDS [109] développé au laboratoire pour effectuer des analyses statistiques sur les distances entre atomes, centres de masses ou groupes d'atomes, les angles, dièdres ou distances à des plans. On peut calculer des moyennes/médianes et les fluctuations au cours du temps ou suivre l'évolution des observables choisies. Il est également possible de déterminer les densités des systèmes - solutés ou solvants - ou de mesurer la distribution statistique du solvant autour d'un soluté.

L'analyse des composantes énergétiques s'effectue avec le module ANAL de AMBER par groupes d'atomes ou de molécules au cours du temps. Les différentes composantes de la fonction d'interaction, comme les interactions électrostatiques ou de VAN DER WAALS peuvent être calculées séparément. Mais ces analyses ne permettent pas de calculer l'entropie ni l'énergie libre du système.

7.2.1 Fonctions de distribution radiale · *Radial distribution functions*

Pour décrire la position moyenne des particules en solution, il est possible d'utiliser les fonctions de distribution radiale. Dans ce qui suit nous nous intéressons uniquement au calcul des fonctions normalisées par paires ("*pair distribution function*", RDF), abrégées par $g(R)$. [110] On choisit d'abord le centre i de la fonction ($R=0$), souvent un atome ou le barycentre d'une molécule dont on veut caractériser la solvation et l'espèce j dont on veut calculer la distribution autour de i . Le calcul de $g_{ij}(R)$ fait intervenir la densité ρ_j des atomes j et le nombre moyen $N(R)$ d'atomes j dans une sphère de rayon R autour de l'atome i .

$$g_{ij}(R) = \frac{N(R + dR) - N(R)}{\frac{4}{3} \pi \rho_j (R + dR)^3 - R^3} \quad (72)$$

Les pics de RDF décrivent la structure des couches de solvation ou de coordination autour de l'atome i . Leur intégration permet d'accéder au nombre d'atomes correspondant qui représente le nombre de coordination ou d'hydratation.

Les fonctions $g(R)$ sont accessibles expérimentalement par la diffraction des rayons X à l'état solide ou en solution, voir [111] pour le cas de l'eau.

7.2.2 Surface de séparation de GIBBS · *GIBBS surface*

Plusieurs définitions thermodynamiques de l'interface liquide-liquide existent. Nous suivons le modèle de GIBBS, qui a considéré l'interface comme une surface de division sans épaisseur, et non celui de VAN DER WAALS qui suppose une épaisseur faible mais finie. On

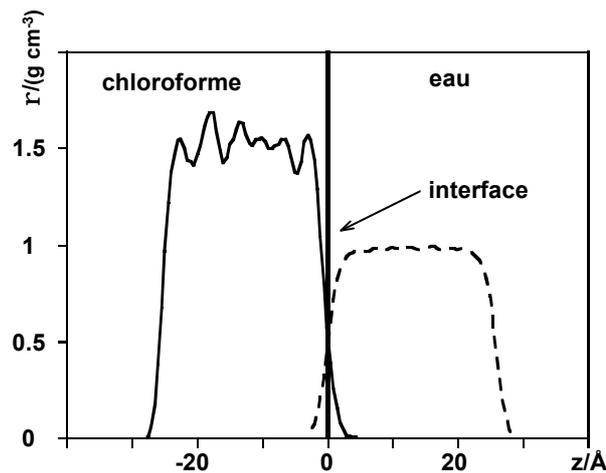


Figure 20 : Définition de la surface de séparation de GIBBS.

peut se servir des courbes de densité des deux liquides pour définir la position de la *surface de séparation de GIBBS* comme leur intersection (Figure 20). Plus précisément, dans nos analyses nous considérons comme interface le plan $z=0$ qui passe par l'intersection des courbes de densités moyennes des deux solvants suivant z .

7.2.3 Tension de surface ·Surface tension

La notion de tension superficielle introduite par LAPLACE, permet de caractériser le comportement d'un système comportant une (des) interface(s) liquide-liquide. Les interactions d'un atome dans le liquide massif et à l'interface sont différentes, car l'environnement microscopique de l'atome n'est pas le même. La différence entre les énergies d'interaction dans ces deux positions est à l'origine de la tension interfaciale. Elle peut être réduite par la présence de molécules dites tensioactives, qui s'y adsorbent et jouent un rôle de tampon. Dans le cas contraire d'un soluté repoussé par l'interface, la tension interfaciale augmente.

Le programme MDS que nous utilisons pour les analyses des trajectoires permet de calculer la tension de surface en se servant de la formule de KIRKWOOD et BUFF, [112] généralisée aux molécules polyatomiques.

$$\gamma = \frac{1}{2} \int_{-\infty}^{\infty} [P_N(z) - P_T(z)] dz \quad (73)$$

Dans cette équation nous intégrons les pressions tangentielles $P_T(z)$ et normales $P_N(z)$ à l'interface, calculées dans des tranches de largeur Δz . Le facteur $\frac{1}{2}$ s'applique uniquement dans le cas d'un système périodique avec 2 interfaces. Ces calculs souffrent de fluctuations qui peuvent être de l'ordre du double de la valeur calculée. [113]

8 Temps de calcul et ressources informatiques

Computer time and resources

La puissance de calcul progresse fortement depuis de nombreuses années, ce qui a été démontré de façon spectaculaire par la première simulation par DM classique d'une durée de 1 μ s. [114] Cette évolution a pu se sentir au cours de cette thèse, et les derniers calculs ont été effectués sur des ordinateurs de type PC (architecture x86) plutôt que sur des supercalculateurs. [115] Les deux logiciels utilisés, AMBER et GROMOS, sont écrits en FORTRAN77, un langage facilement portable sur la majorité des architectures existantes. Le code de AMBER est parallélisé, ce qui permet de tirer avantage de machines comportant plusieurs processeurs. Dans le cadre des systèmes étudiés, un nombre de 8 processeurs semble un bon compromis entre vitesse de calcul et efficacité. Il existe également une version parallèle à mémoire partagée de GROMOS et une version parallèle pour PCs biprocesseurs.

Nous avons effectué des calculs sur des machines de type HP 735 et 871 (2 processeurs), Grappe, Cray C64, T3E (256 proc.), Origin2000 (32 proc.), Origin 200 (4 proc.) et PC (686, 2 proc.). Dans cette liste, les ordinateurs à base de PC sont les plus puissants par rapport au nombre de processeurs. De nos jours, une simulation de 100 ps sur 10000 atomes ne prend guère plus que 24 heures sur une telle machine.

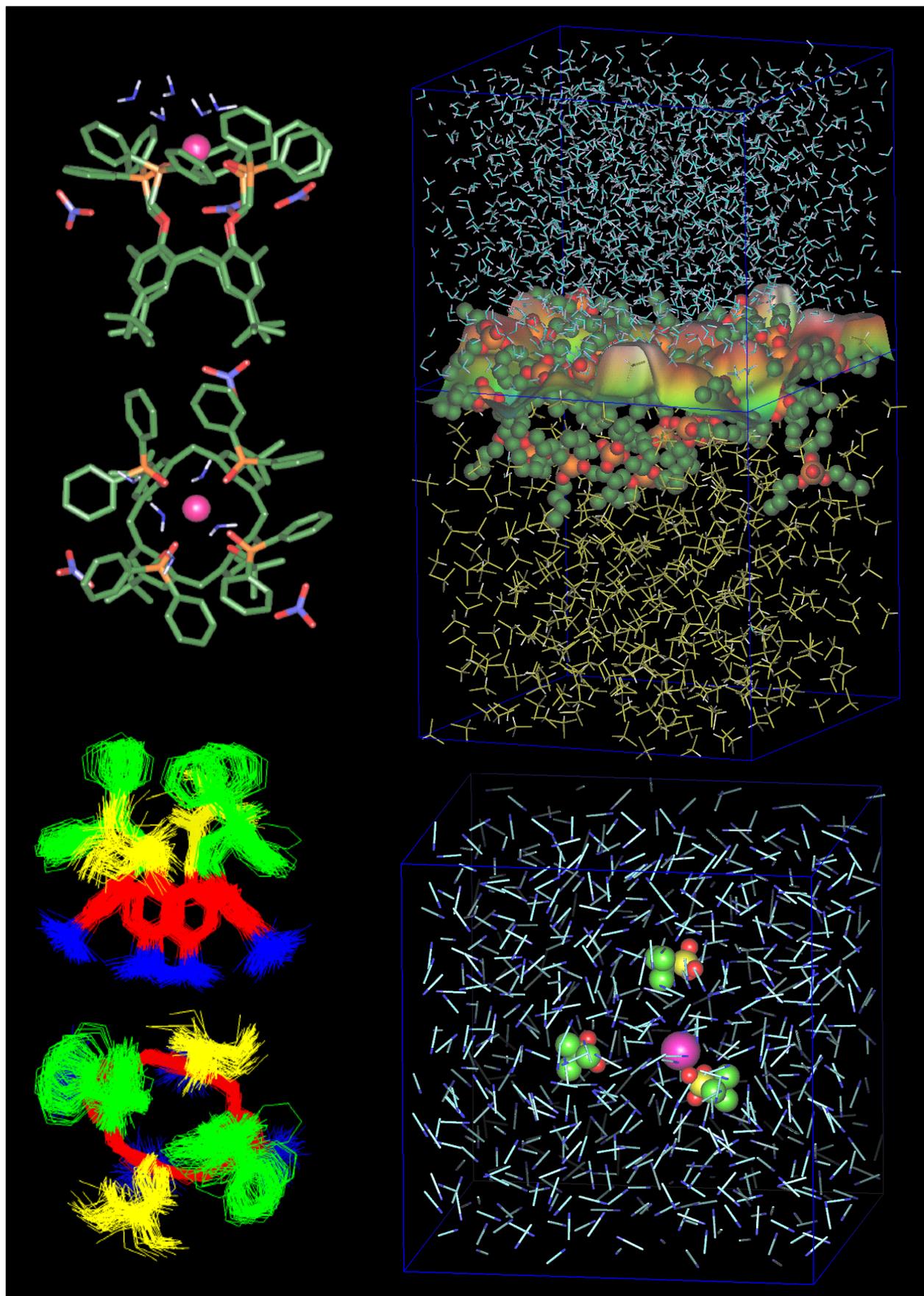


Figure 1-1: Several examples of visualisation results with the MD-DRAW software. On top: orthogonal views of a calixarene with selected solvent molecules (left) and biphasic water/chloroform system (right) with TBP molecules at the interface. The water (blue)/chloroform (yellow) interface is shown. Bottom: cumulated orthogonal views of a calixarene (left) and a $\text{La}(\text{F}_3\text{CSO}_3)_3$ salt in an acetonitrile simulation box (right).

Références bibliographiques du chapitre 1

Bibliographic references for chapter 1

- [14] BERENDSEN, H. J. C., 2000, *Hierarchical Modeling*, manuscrit d'un cours donné à l'ETH Zürich, Suisse.
- [15] a) BOYD, D. B. et LIPKOWITZ, K. B., 1982, *J. Chem. Educ.*, **59**, 269-274; b) VAN GUNSTEREN, W. F. et BERENDSEN, H. J. C., 1990, *Angew. Chem. Int. Ed. Engl.*, **29**, 992-1023; c) LEACH, A., 1996, *Molecular Modelling: Principles and Applications* (Addison Wesley Longman Limited, Essex); d) ALLINGER, N. L. et BURKERT, U., 1982, *Molecular Mechanics* (Oxford University Press); e) HAILE, J. M., 1992, *Molecular Dynamics Simulation: Elementary Methods* (Wiley, New York); f) RAPPE, A. K. et CASEWIT, C. J., 1997, *Molecular Mechanics Across Chemistry*, illustré par John Choi (University Science Books); g) RAPAPORT, D. C., 1997, *The Art of Molecular Dynamics Simulation* (Cambridge University Press); h) ALLEN, M. P. et TILDSELEY, D. J., 1987, *Computer Simulations of Liquids* (Clarendon, Oxford); i) Divers auteurs, 1996, *Understanding Molecular Simulation*, édité par D. Frenkel et B. Smit (Academic Press).
- [16] a) KOLLMAN, P. A., 1985, *Acc. Chem. Res.*, **18**, 105-111; b) KOLLMAN, P. A., 1987, *Ann. Rev. Phys. Chem.*, **38**, 303-316; c) VAN GUNSTEREN, W. F., 1988, *Prot. Eng.*, **2**, 5-13; d) BERENDSEN, H. J. C., 1996, *Science*, **271**, 954-955; e) MEZEL, M. et BEVERIDGE, D. L., 1986, *Computer Simulation of Chemical and Biomolecular Systems*, édité par W. L. Jorgensen (Annals of the New York Academy of Sciences, **482**), pp.1-26; f) AUFFINGER, P. et WESTHOF, E., 1998, *Curr. Opin. Struct. Biol.*, **8**, 227-236; g) Divers auteurs, 1996, *Computer Modelling in Inorganic Crystallography*, édité par C. R. A. Catlow (Academic Press, London); h) MCCAMMON, J. A. et HARVEY, S. C., 1987, *Dynamics of Proteins and Nucleic Acids* (Cambridge University Press); i) BROOKS III, C. L., KARPLUS, M. et PETTIT, B. M., 1988, *Proteins: A Theoretical Perspective of Dynamics, Structure and Thermodynamics* (Wiley, New York); j) MCCAMMON, J. A. et KARPLUS, M., 1983, *Acc. Chem. Res.*, **16**, 187-193. k) GARCIA, A. E. et KRUMHANS, J. A., 1987, *J. Biomol. Struct. & Dynamics*, **5**, 513-535.
- [17] a) WEINER, P. K. et KOLLMAN, P. A., 1981, *J. Comp. Chem.*, **2**, 287-303; b) PEARLMAN, D. A., CASE, D. A., CADWELL, J. W., ROSS, W. S., CHEATHAM III, T. E., DEBOLT, S., FERGUSON, D., SEIBEL, G. et KOLLMAN, P. A., 1995, *Comp. Phys. Commun.*, **91**, 1-41; c) PEARLMAN, D. A., CASE, D. A., CADWELL, J. W., ROSS, W. S., CHEATHAM III, T. E., FERGUSON, D. M., SEIBEL, G. L., SINGH, U. C., WEINER, P. K. et KOLLMAN, P. A., 1995, *AMBER4.1* (University of California, San Francisco); d) CASE, D. A., PEARLMAN, D. A., CADWELL, J. W., CHEATHAM III, T. E., ROSS, W. S., SIMMERLING, C. L., DARDEN, T. A., MERZ, K. M., STANTON, R. V., CHENG, A. L., VINCENT, J. J., CROWLEY, M., FERGUSON, D. M., RADMER, R. J., SEIBEL, G. L., SINGH, U. C., WEINER, P. K. et KOLLMAN, P. A., 1997, *AMBER5* (University of California, San Francisco).
- [18] a) SCOTT, W. R. P., HÜNENBERGER, P. H., TIRONI, I. G., MARK, A. E., BILLETER, S. R., FENNEN, J., TORDA, A. E., HUBER, T., KRÜGER, P. et VAN GUNSTEREN, W. F., 1999, *J. Phys. Chem. A*, **103**, 3596-3607; b) VAN GUNSTEREN, W. F., BILLETER, S. R., EISING, A. A., HÜNENBERGER, P. H., KRÜGER, P., MARK, A. E., SCOTT, W. R. P. et TIRONI, I. G., 1996, *GROMOS96* (BIOMOS b. v., ETH Zürich).
- [19] a) VAN SCHAİK, R. C., BERENDSEN, H. J. C., TORDA, A. E. et VAN GUNSTEREN, W. F., 1993, *J. Mol. Biol.*, **234**, 751-762; b) BEUTLER, T. C. et VAN GUNSTEREN, W. F., 1994, *J. Chem. Phys.*, **101**, 1417-1422.
- [20] VAN GUNSTEREN, W. F., HUBER, T. et TORDA, A. E., 1995, *ECCC1 Computational Chemistry Conference Proceedings*, **330**, édité par F. Bernardi et J.-L. Rivail (American Institute of Physics), pp. 235-268.
- [21] HÜNENBERGER, P. H. et VAN GUNSTEREN, W. F., 1997, dans *Computer Simulation of Biomolecular Systems, Theoretical and Experimental Applications*, Vol. 3, édité par W. F. van Gunsteren, P. K. Weiner, A. J. Wilkinson (Kluwer Academic Publishers, Dordrecht), pp. 3-82.
- [22] a) WEINER, S. J., KOLLMAN, P. A., NGUYEN, D. T. et CASE, D. A., 1986, *J. Comput. Chem.*, **7**, 230-252; b) WEINER, S. J., KOLLMAN, P. A., CASE, D. A., SINGH, U. C., GHIO, C., ALAGONA, G., PROFETA JR., S. et WEINER, P., 1984, *J. Am. Chem. Soc.*, **106**, 765-784; c) CORNELL, W. D., CIEPLAK, P., BAYLY, C. I., GOULD, I. R., MERZ, K. M., FERGUSON, D. M., SPELLMEYER, D. C., FOX, T., CALDWELL, J. W. et KOLLMAN, P. A., 1995, *J. Am. Chem. Soc.*, **117**, 5179-5197; d) CORNELL, W. D., CIEPLAK, P., BAYLY, C. I., GOULD, I. R., MERZ, K. M., FERGUSON, D. M., SPELLMEYER, D. C., FOX, T., CALDWELL, J. W. et KOLLMAN, P. A., 1996, *J. Am. Chem. Soc.*, **118**, 2309.
- [23] a) VAN GUNSTEREN, W. F., BILLETER, S. R., EISING, A. A., HÜNENBERGER, P. H., KRÜGER, P., MARK, A. E., SCOTT, W. R. P. et TIRONI, I. G., 1996, *Biomolecular Simulation: The GROMOS96 Manual and User Guide*

- (Hochschulverlag, ETH Zürich); b) DAURA, X., MARK, A. E. et VAN GUNSTEREN, W. F., 1998, *J. Comp. Chem.*, **19**, 535-547; c) VAN GUNSTEREN, W. F., DAURA, X. et MARK, A. E., 1998, dans *Encyclopedia of Computational Chemistry*, **2**, édité par P. v. R. Schleyer, N. L. Allinger, T. Clark, J. Gasteiger, P. A. Kollman et H. F. Schaefer III (Wiley & Sons, Chichester), pp. 1211-1216.
- [24] a) IUPAC-IUB, 1970, *J. Mol. Biol.*, **52**, 1-17; b) IUPAC-IUB, 1970, *Biochemistry*, **9**, 3471-3479.
- [25] KLYNE, W. et PRELOG, V., 1960, *Experientia*, **16**, 521-568.
- [26] a) HAYES, D., ROTHENBERG, S. et KOLLMAN, P., 1977, *J. Am. Chem. Soc.*, **99**, 2150; b) MARSH, F., WEINER, P., DOUGLAS, J., KOLLMAN, P., KENYON, G. et GERLT, J., 1980, *J. Am. Chem. Soc.*, **102**, 1660; c) GOVIL, G., 1976, *Biopolymers*, **15**, 2303; d) OLSON, W., 1982, *J. Am. Chem. Soc.*, **104**, 278; e) BARKER, R. et BOUDREAUX, G., 1967, *Spectrochim. Acta, Part A*, **23A**, 727.
- [27] a) TIRONI, I. G., SPERB, R., SMITH, P. E. et VAN GUNSTEREN, W. F., 1995, *J. Chem. Phys.*, **102**, 5451-5459; b) TIRONI, I. G., 1997, *Thèse* (ETH Zürich); c) HÜNENBERGER, P. H., 1997, *Thèse* (ETH Zürich);
- [28] a) ONSAGER, L., 1936, *J. Am. Chem. Soc.*, **58**, 1486; b) BARKER, J. A. et WATTS, R. O., 1973, *Mol. Phys.*, **26**, 789; c) NEUMANN, M., 1983, *Mol. Phys.*, **50**, 841.
- [29] a) APPLEQUIST, J., CARL, J. R. et FUNG, K.-K., 1972, *J. Am. Chem. Soc.*, **94**, 2952-2960; b) BODE, K. A. et APPLEQUIST, J., 1996, *J. Phys. Chem.*, **100**, 17820-17824; c) GAO, J., PAVELITES, J. J. et HABIBOLLAZADEH, D., 1996, *J. Phys. Chem.*, **100**, 2689-2697; d) BAKOWIES, D., 1994, *Thèse* (Hartrung-Gorre Verlag, Zürich), pp. 54-61;
- [30] a) BACHRACH, S. M., 1994, *Rev. Comp. Chem.*, **5**, 171; b) BROOKS, B. R., BRUCCOLERI, R. E., OLAFSON, B. D., STATES, D. J., SWAMINATHAN, S. et KARPLUS, M., 1983, *J. Comp. Chem.*, **4**, 187-217; c) JORGENSEN, W. L. et TIRADO-RIVES, J., 1988, *J. Am. Chem. Soc.*, **110**, 1657-1671; d) OROZCO, M. et LUQUE, F. J., 1990, *J. Comp. Chem.*, **11**, 909-923; e) CIEPLAK, P. et KOLLMAN, P. A., 1991, *J. Comp. Chem.*, **12**, 1232-1236; f) MERZ JR., K. M., 1992, *J. Comp. Chem.*, **13**, 749-767; g) ALAGONA, G. et GHIO, C., 1992, *J. Mol. Struct.(THEOCHEM)*, **88**, 187-216; h) WIBERG, K. B. et RABLEN, P. R., 1993, *J. Comp. Chem.*, **14**, 1504-1518; i) REYNOLDS, C. A., ESSEX, J. W. et RICHARDS, W. G., 1992, *J. Am. Chem. Soc.*, **114**, 9075-9079.
- [31] MULLIKEN, R. S., 1962, *J. Chem. Phys.*, **36**, 3428.
- [32] BADER, R. W. F., 1990, *Atoms in Molecules. A Quantum Theory* (Clarendon Press, Oxford).
- [33] a) CRAMER, C. J. et TRUHLAR, D. G., 1992, *J. Comput.-Aided Mol. Des.*, **6**, 629-666; b) STORER, J. W., GIESEN, D. J., CRAMER, C. J. et TRUHLAR, D. G., 1995, *J. Comput.-Aided Mol. Des.*, **9**, 87-110; c) KAMINSKI, G. A. et JORGENSEN, W. L., 1998, *J. Phys. Chem. B*, **102**, 1787-1796.
- [34] a) GASTEIGER, J. et MARSILI, M., 1980, *Tetrahedron*, **36**, 3219-3228; b) GASTEIGER, J. et MARSILI, M., 1980, *Croat. Chem. Acta*, **53**, 601-614; c) GASTEIGER, J. et MARSILI, M., 1980, *Z. Analyt. Chem.*, **304**, 258-259.
- [35] a) CHIRLAIN, L. E. et FRANCL, M. M., 1987, *J. Comp. Chem.*, **6**, 894; b) BRENEMAN, C. M. et WIBERG, K. B., 1990, *J. Comp. Chem.*, **11**, 361; c) SINGH, U. C. et KOLLMAN, P. A., 1984, *J. Comp. Chem.*, **5**, 129; d) BESLER, B. H., MERZ, K. M. et KOLLMAN, P. A., 1990, *J. Comp. Chem.*, **11**, 431.
- [36] SIGFRIDSSON, E. et RYDE, U., 1998, *J. Comp. Chem.*, **19**, 377-395.
- [37] a) BAYLY, C. I., CIEPLAK, P., CORNELL, W. D. et KOLLMAN, P. A., 1993, *J. Phys. Chem.*, **97**, 10269; b) CORNELL, W. D., CIEPLAK, P., BAYLY, C. I. et KOLLMAN, P. A., 1993, *J. Am. Chem. Soc.*, **115**, 9620-9631; c) CIEPLAK, P., CORNELL, W. D., BAYLY, C. et KOLLMAN, P. A., 1995, *J. Comp. Chem.*, **16**, 1357-1377; d) KOLLMAN, P. A., CORNELL, W. D., ST-AMANT, A. et HALGREN, T. A., 1995, *J. Comp. Chem.*, **16**, 1483-1506.
- [38] a) ÅQVIST, J., 1990, *J. Phys. Chem.*, **94**, 8021-8024; b) VAN VEGGEL, F. C. J. M. et REINHOUDT, D., 1999, *Chem. Eur. J.*, **5**, 90-95.
- [39] CARLSON, H. A., NGUYEN, T. B., OROZCO, M. et JORGENSEN, W. L., 1993, *J. Comput. Chem.*, **14**, 1240-1249.
- [40] TROXLER, L., WIPFF, G. et HARROWFIELD, J., 1998, *J. Phys. Chem.*, **102**, 6821-6830.
- [41] BERNY, F., 2000, *Thèse* (Université Louis Pasteur, Strasbourg).
- [42] SCHURHAMMER, R. et WIPFF, G., 1999, *New J. Chem.*, **23**, 381-391.
- [43] a) HOWELL, J. M. et SAPSE, A. M., 1982, *J. Chem. Phys.*, **86**, 2345; b) RUBIN-CARREZ, C., 1991, *Thèse* (Université de Nancy).
- [44] GUILBAUD, P. et WIPFF, G., 1996, *J. Mol. Struct. (THEOCHEM)*, **366**, 55-63.
- [45] a) BEUDAERT, P., LAMARE, V., DOZOL, J.-F., TROXLER, L. et WIPFF, G., 1998, *Solv. Extr. Ion Exch.*, **16**, 597-618. b) BEUDAERT, P., 1998, *Thèse* (Université Louis Pasteur, Strasbourg).
- [46] a) SCHURHAMMER, R., ERHART, V., TROXLER, L. et WIPFF, G., 1999, *Perkin Trans. 2*, 2423-2431; b) SCHURHAMMER, R., 2000, *communication privée*.

- [47] FISCHER, S., GROOTENHUIS, P. D. J., GROENEN, L. C., VAN HOORN, W. P., VAN VEGGEL, F. C., REINHOUDT, D. N. et KARPLUS, M., 1995, *J. Am. Chem. Soc.*, **117**, 1611-1620.
- [48] WIPFF, G. et LAUTERBACH, M., 1995, *Supramol. Chem.*, **6**, 187-207.
- [49] MUZET, N., 1999, *Thèse* (Université Louis Pasteur, Strasbourg).
- [50] LAUTERBACH, M., 1997, *Thèse* (Université Louis Pasteur, Strasbourg).
- [51] GROOTENHUIS, P. D. J., KOLLMAN, P. A., GROENEN, L. C., REINHOUDT, D. N., VAN HUMMEL, G. J., UGOZZOLI, F. et ANDRETTI, G. D., 1990, *J. Am. Chem. Soc.*, **112**, 4165-4176.
- [52] JORGENSEN, W. L., BRIGGS, J. M. et CONTRERAS, M. L., 1990, *J. Phys. Chem.*, **94**, 1683-1686.
- [53] a) GUILBAUD, P., 1995, *Thèse* (Université Louis Pasteur, Strasbourg); b) TROXLER, L., 1995, *Thèse* (Université Louis Pasteur, Strasbourg).
- [54] ALLEN, F. H. et KENNARD, O., 1993, *Chem. Des. Autom. News*, **8**, 31-37.
- [55] a) PARR, R. G. et YANG, W., 1989, *Density Functional Theory of Atoms and Molecules* (Oxford University Press, New York); b) BECKE, A. D., 1993, *J. Chem. Phys.*, **98**, 5648.
- [56] JORGENSEN, W. L. et BRIGGS, J. M., 1988, *Mol. Phys.*, **63**, 547-558.
- [57] JORGENSEN, W. L., CHANDRASEKHAR, J., MADURA, J. D., IMPEY, R. W. et KLEIN, M. L., 1983, *J. Chem. Phys.*, **79**, 926-936.
- [58] a) BERENDSEN, H. J. C., POSTMA, J. P. M., VAN GUNSTEREN, W. F. et HERMANS, J., 1981, dans *Intermolecular Forces*, édité par B. Pullman (Reidel Publishing Company, Dordrecht), pp. 331-342; b) Postma, J. P. M., 1985, *Thèse* (Université de Groningen); c) SMITH, P. E. et VAN GUNSTEREN, W. F., 1993, *Chem. Phys. Lett.*, **215**, 315-318; d) SMITH, P. E. et VAN GUNSTEREN, W. F., 1994, *J. Chem. Phys.*, **100**, 3169-3174; e) BERWEGER, C. D., VAN GUNSTEREN, W. F. et MÜLLER-PLATHE, F., 1995, *Chem. Phys. Lett.*, **232**, 429-436.
- [59] VAN DER SPOEL, D., VAN MAAREN, P. J. et BERENDSEN, H. J. C., 1998, *J. Chem. Phys.*, **108**, 10220-10226.
- [60] MAHONEY, M. W. et JORGENSEN, W. L., 2000, *J. Chem. Phys.*, **112**, 8910-8922.
- [61] BRODSKY, A., 1996, *Chem. Phys. Lett.*, **261**, 563.
- [62] CHANG, T. M., DANG, L. X. et PETERSON, K. A., 1997, *J. Phys. Chem. B*, **101**, 3413-3419.
- [63] BARTELL, L. S., BROCKWAY, L. O. et SCHWENDEMAN, R. H., 1955, *J. Chem. Phys.*, **23**, 1854.
- [64] *CRC Handbook of Chemistry and Physics*, 1998 (CRC Press, Boca Raton, New York).
- [65] APPLEQUIST, J., CARL, J. R., et FUNG, K.-K., 1972, *J. Am. Chem. Soc.*, **94**, 2952.
- [66] a) DIETZ, W. et HEINZINGER, K., 1984, *Ber. Bunsenges. phys. Chem.*, **88**, 543; b) DIETZ, W. et HEINZINGER, K., 1985, *Ber. Bunsenges. phys. Chem.*, **89**, 968; c) TIRONI, I.G. et VAN GUNSTEREN, W. F., 1994, *Mol. Phys.*, **83**, 381-403.
- [67] a) MOMANY, F., MCGUIRE, R., BURGESS, A. et SCHERAGA, H., 1975, *J. Phys. Chem.*, **79**, 2361; b) DUNFIELD, L., BURGESS, A. et SCHERAGA, H., 1978, *J. Phys. Chem.*, **82**, 2609.
- [68] VAN GUNSTEREN, W. F. et KARPLUS, M., 1982, *Macromolecules*, **15**, 1528-1544.
- [69] VAN GUNSTEREN, W. F., BERENDSEN, H. J. C., COLONNA, F., PERAHIA, D., HOLLENBERG, J. P. et LELLOUCH, D., 1984, *J. Comp. Chem.*, **5**, 272-279.
- [70] BERENDSEN, H. J. C., 1985, dans *Molecular Dynamics and Protein Structure*, édité par J. Hermans (Polycrystal Book Service, Western Springs), pp. 18-22.
- [71] a) BROOKS III, C. L., PETTIT, B. M. et KARPLUS, M., 1985, *J. Chem. Phys.*, **83**, 5897; b) SMITH, P. E. et PETTIT, B. M., 1991, *J. Chem. Phys.*, **95**, 8430; c) SCHREIBER, H. et STEINHAUSER, O., 1992, *J. Mol. Biol.*, **228**, 909.
- [72] a) ASHBAUGH, H. S. et WOOD, R. H., 1997, *J. Chem. Phys.*, **106**, 8135-8139; b) HUMMER, G., PRATT, L. R., GARCIA, A. E., BERNE, G. J. et RICK, S. W., 1997, *J. Phys. Chem. B*, **101**, 3017-3020; c) ÅQVIST, J. et HANSSON, T., 1998, *J. Phys. Chem. B*, **102**, 3387-3840; d) ASHBAUGH, H. S., SAKANE, S. et WOOD, R. H., 1998, *J. Phys. Chem. B*, **102**, 3844-3845; e) HUMMER, G., PRATT, L. R., GARCIA, A. E. et GARDE, S., 1998, *J. Phys. Chem. B*, **102**, 3841-3843.
- [73] a) EWALD, P., 1921, *Ann. Phys.*, **64**, 253; b) DE LEEUW, S. W., PERRAM, J. W. et SMITH, E. R., 1980, *Proc. R. Soc. London*, **A373**, 27; c) HEYES, D. M., 1981, *J. Chem. Phys.*, **74**, 1924; d) DARDEN, T., YORK, D. et PEDERSEN, L., 1993, *J. Chem. Phys.*, **98**, 10089-10092; e) ESSMANN, U., PERERA, L., BERKOWITZ, M. L., DARDEN, T., LEE, H. et PEDERSEN, L. G., 1995, *J. Chem. Phys.*, **103**, 5877-8593; f) HÜNENBERGER, P. H., 1999, dans *Simulation and Theory of Electrostatic Interactions in Solution: Computational Chemistry, Biophysics and Aqueous Solution*, édité par L. R. Pratt et G. Hummer (American Institute of Physics, New York), pp. 17-83; g) BAKER, N. A., HÜNENBERGER, P. H. et MCCAMMON, J. A., 1999, *J. Chem. Phys.*, **110**, 10679-10692;

- h) HÜNENBERGER, P. H. et MCCAMMON, J. A., 1999, *Biophys. Chem.*, **78**, 69-88; i) HÜNENBERGER, P. H. et MCCAMMON, J. A., 1999, *J. Chem. Phys.*, **110**, 1856-1872;
- [74] a) VAN GUNSTEREN, W. F. et KARPLUS, M., 1982, *Biochemistry*, **21**, 2259-2274; b) FRATERNALI, F. et VAN GUNSTEREN, W. F., 1996, *J. Mol. Biol.*, **256**, 939-948.
- [75] ADAMS, D. J., 1979, *Chem. Phys. Lett.*, **62**, 329.
- [76] a) REMERIC, K., VAN GUNSTEREN, W. F. et ENGBERTS, J. B. F. N., 1986, *Chem. Phys.*, **101**, 27; b) MANDELL, M. J., 1976, *J. Stat. Phys.*, **15**, 299; c) PRATT, L. R. et HAAN, S. W., 1981, *J. Chem. Phys.*, **74**, 1864-1873; d) GUPTA, S. et HAILE, J. M., 1984, *Chem. Phys.*, **83**, 481.
- [77] a) BERMAN, H. M., WESTBROOK, J., FENG, Z., GILLILAND, G., BHAT, T. N., WEISSIG, H., SHINDYALOV, I. N. et BOURNE, B. E., 2000, *Nucleic Acids Research*, **28**, 235-242; b) GILLILAND, G.L., TUNG, M., BLAKESLEE, D.M. et LADNER, J., 1994, *Acta Cryst.*, **D50**, 408-413; c) COUTINHO, P. M. et HENRISSAT, B., 1999, dans *Recent Advances in Carbohydrate Bioengineering*, édité par H. J. Gilbert, G. Davies, B. Henrissat et B. Svensson (Royal Society of Chemistry, Cambridge), sous presse; d) COUTHINO, P. M. et HENRISSAT, B., 1999, dans *Genetics, Biochemistry and Ecology of Cellulose Degradation*, édité par K. Ohmiya, K. Hayashi, K. Sakka, Y. Kobayashi, S. Karita et T. Kimura (Uni Publishers, Tokyo), pp. 15-23; e) BERGERHOFF, G., HUNDT, R., SIEVERS, R. et BROWN, I. S., 1983, *J. Chem. Inform. & Comp. Sc.*, **23**, 66-69.
- [78] a) JORGENSEN, W. L., MAXWELL, D. S. et TIRADO-RIVES, J., 1996, *J. Am. Chem. Soc.*, **118**, 11225-11236; b) JORGENSEN, W. L., MADURA, J. D. et SWENSON, C. J., 1984, *J. Am. Chem. Soc.*, **106**, 6638-6646; c) JORGENSEN, W. L., 1986, *J. Phys. Chem.*, **90**, 1276-1284; d) JORGENSEN, W. L., 1986, *J. Phys. Chem.*, **90**, 6379-6388; e) BRIGGS, J. M., MATSUI, T. et JORGENSEN, W. L., 1990, *J. Comput. Chem.*, **11**, 958-971; f) BRIGGS, J. M., NGUYEN, T. B. et JORGENSEN, W. L., 1991, *J. Phys. Chem.*, **95**, 3315-3322; g) JORGENSEN, W. L. et SWENSON, C. J., 1985, *J. Am. Chem. Soc.*, **107**, 1489-1496; h) JORGENSEN, W. L. et SEVERANCE, D. L., 1990, *J. Am. Chem. Soc.*, **112**, 4768-4774; i) JORGENSEN, W. L. et NGUYEN, T. B., 1993, *J. Comput. Chem.*, **14**, 195-205; j) PRANATA, J., WIERSCHKE, S. G. et JORGENSEN, W. L., 1991, *J. Am. Chem. Soc.*, **113**, 2810-2819; k) JORGENSEN, W. L. et TIRADO-RIVES, J., 1988, *J. Am. Chem. Soc.*, **110**, 1657-1666; l) JORGENSEN, W. L. et TIRADO-RIVES, J., 1995, *Perspec. Drug. Discovery Des.*, **3**, 123-138.
- [79] a) LINSE, P., 1993, *J. Am. Chem. Soc.*, **115**, 8793-8797; b) GUBA, W., HAESSNER, R., BREIPOHL, G., HENKE, S., KNOLLE, J., SANDAGADA, V. et KESSLER, H., 1994, *J. Am. Chem. Soc.*, **116**, 7532; c) GUBA, W. et KESSLER, H., 1994, *J. Phys. Chem.*, **98**, 23.
- [80] MICHAEL, D. et BENJAMIN, I., 1995, *J. Phys. Chem.*, **99**, 1530.
- [81] a) VAN GUNSTEREN, W. F. et BERENDSEN, H. J. C., 1984, *J. Mol. Biol.*, **176**, 559-564; b) LAUTZ, J., KESSLER, H., KAPTEIN, R. et VAN GUNSTEREN, W. F., 1987, *J. Comp.-Aid. Mol. Des.*, **1**, 219-241.
- [82] COURANT, R., 1943, *Bull. Am. Math. Soc.*, **5**, 1-23.
- [83] FLETCHER, R. et REEVES, C. M., 1964, *Computer J.*, **7**, 149.
- [84] a) VAN GUNSTEREN, W. F. et KARPLUS, M., 1980, *J. Comp. Chem.*, **1**, 266-274; b) JACOBY, S. L. S., KOWALIK, J. et S., PIZZO, J. T., 1972, *Iterative Methods for Non-linear Optimization Problems* (Englewood Cliffs, New Jersey); c) FLETCHER, R., 1980, *Practical Methods of Optimization. Volume 1: Unconstrained Optimization* (New York), 1980.
- [85] VERLET, L., 1967, *Phys. Rev.*, **159**, 98.
- [86] PRESS, W. H., FLANNERY, B. P., TEUKOLSY, S. A. et VETTERLING, W. T., 1986, *Numerical Recipes* (Cambridge University Press, Cambridge).
- [87] GEAR, C. W., 1971, *Numerical Initial Value Problems in Ordinary Differential Equations* (Prentice-Hall, Englewood Cliffs).
- [88] BEEMAN, D., 1976, *J. Comput. Phys.*, **20**, 130.
- [89] a) VAN GUNSTEREN, W. F., 1990, dans *Modelling of Molecular Structures and Properties*, édité par J.-L. Rivail (Elsevier, Amsterdam), pp. 463-478; b) ÅQVIST, J., VAN GUNSTEREN, W.F., LEIJONMARCK, M. et TAPIA, O., 1985, dans *Molecular Dynamics and Protein Structure*, édité par J. Hermans (Polycrystal Book Service, Western Springs), pp. 145-147; c) BERENDSEN, H. J. C. et VAN GUNSTEREN, W. F., 1986, dans *Molecular Dynamics Simulation of Statistical Mechanical Systems*, édité par G. Ciccotti et W. G. Hoover (North-Holland, Amsterdam), pp. 43-65.
- [90] VAN GUNSTEREN, W. F., NANZER, A. P. et TORDA, A. E., 1996, dans *Monte Carlo and Molecular Dynamics of Condensed Matter*, **49**, édité par K. Binder et C. Ciccotti (SIF, Bologna), pp. 777-788.
- [91] BERENDSEN, H. J. C., POSTMA, J. P. M., VAN GUNSTEREN, W. F., DINOLA, A. et HAAK, J. R., 1984, *J. Chem. Phys.*, **81**, 3684-3690.

- [92] KELL, G. S., 1967, *J. Chem. Engineering Data*, **12**, 66-69.
- [93] a) RYCKAERT, J. P., CICCOTTI, G. et BERENDSEN, H. J. C., 1977, *J. Comput. Phys.*, **23**, 327-336; b) VAN GUNSTEREN, W. F. et BERENDSEN, H. J. C., 1977, *Mol. Phys.*, **34**, 1311-1327
- [94] TIRONI, I. G., BRUNNE, R. M. et VAN GUNSTEREN, W. F., 1996, *Chem. Phys. Lett.*, **250**, 19-24.
- [95] a) VAN GUNSTEREN, W. F., 1980, *Mol. Phys.*, **40**, 1015-1019; b) VAN GUNSTEREN, W. F. et KARPLUS, M., 1982, *Macromolecules*, **15**, 1528-1544; c) BERENDSEN, H. J. C. et VAN GUNSTEREN, W. F., 1984, dans *Molecular Liquids – Dynamics and Interactions*, édité par A. J. Barnes *et al.* (Reidel, Dordrecht), pp. 475-500.
- [96] a) JORGENSEN, W. L., 1989, *Acc. Chem. Res.*, **22**, 184-189; b) VAN GUNSTEREN, W.F., 1989, dans *Computer Simulations of Biomolecular Systems, Theoretical and Experimental Applications*, édité par W. F. van Gunsteren et P. K. Weiner (Escom Science Publishers, Leiden), pp. 27-59; c) KING, P. M., 1993, dans *Computer Simulations of Biomolecular Systems, Theoretical and Experimental Applications, Vol. 2*, édité par W. F. van Gunsteren, P. K. Weiner et A. J. Wilkinson (Escom Science Publishers, Leiden), pp. 267-314; d) VAN GUNSTEREN, W. F., BEUTLER, T. C., FRATERNALI, F., KING, P. M., MARK, A. E. et SMITH, P. E., 1993, dans *Computer Simulations of Biomolecular Systems, Theoretical and Experimental Applications, Vol. 2*, édité par W. F. van Gunsteren, P. K. Weiner et A. J. Wilkinson (Escom Science Publishers, Leiden), pp. 315-348; e) VAN GUNSTEREN, W. F., DAURA, X. et MARK, A. E., soumis à publication dans NATO ASI Series.
- [97] ZWANZIG, R. W., 1954, *J. Chem. Phys.*, **22**, 1420-1426.
- [98] KIRKWOOD, J. G., 1935, *J. Chem. Phys.*, **3**, 300-313.
- [99] PEARLMAN, D. A. et KOLLMAN, P. A., *J. Chem. Phys.*, **90**, 2460-2470.
- [100] LYBRAND, T. P., MCCAMMON, J. A. et WIPFF, G., 1986, *Proc. Natl. Acad. Sci.*, **83**, 833-835.
- [101] MARK, A. E., VAN GUNSTEREN, W. F. et BERENDSEN, H. J. C., 1991, *J. Chem. Phys.*, **94**, 3808-3816.
- [102] a) BASH, P. A., SINGH, U. C., BROWN, F. K., LANGRIDGE, R. et KOLLMAN, P. A., 1987, *Science*, **235**, 574-576; b) BASH, P. A., SINGH, U. C., LANGRIDGE, R. et KOLLMAN, P. A., 1987, *Science*, **236**, 564-568; c) STRAATSMA, T. P. et MCCAMMON, J. A., 1992, *Annu. Rev. Phys. Chem.*, **43**, 407-435.
- [103] JORGENSEN, W. L. et RAVIMOHAN, C., 1985, *J. Chem. Phys.*, **83**, 3050-3054.
- [104] BRICKMANN, J., 1984, *Bild der Wissenschaft*, **3**, 38-42.
- [105] ENGLER, E. et WIPFF, G., 1996, dans *Crystallography of Supramolecular Compounds*, édité par G. Tsoucaris (Kluwer, Dordrecht), pp. 471-476.
- [106] HUMPHREY, W., DALKE, A. et SCHULTEN, K., 1996, *J. Mol. Graphics*, **14**, 33-38.
- [107] DÜCK, U., KELLER, E. et KRÄMER, V., 1997, SCHAKAL97 (Universität Freiburg).
- [108] a) DOUCET, J.-P. et WEBER, J., 1996, *Computer-Aided Molecular Design* (Academic Press, London); b) BRICKMANN, J., 1983, *Konstanzer Blätter für Hochschulfragen*, 119-133; c) BRICKMANN, J., 1984, *Int. J. Quantum Chem.: Quant. Chem. Symp.*, **18**, 647-659; d) BRICKMANN, J., 1992, *J. Chim. Phys.*, **89**, 1709-1721; e) BRICKMANN, J. et TESCHNER, M., 1994, *ECCC1 Computational Chemistry Conference Proceedings*, **330**, édité par F. Bernardi et J.-L. Rivail (American Institute of Physics), pp. 427-446; f) BRICKMANN, J., EXNER, T., KEIL, M., MARHÖFER, R. et MOECKEL, G., 1998, dans *The Encyclopedia of Computational Chemistry*, édité par P.v. R. Schleyer, N. L. Allinger, T. Clark, J. Gasteiger, P. A. Kollman, H. F. Schaefer III et P. R. Schreiner (John Wiley & Sons, Chichester), pp.1679-1693; g) BRICKMANN, J., KEIL, M., EXNER, T. et MARHÖFER, R. J., 2000, *J. Mol. Model.*, **6**, 328-340; h) BERGMAN, D. L., LAAKSONEN, L. et LAAKSONEN, A., 1997, *J. Mol. Graph. & Model.*, **15**, 301-306; i) ABAGYAN, R. et ARGOS, P., 1992, *J. Mol. Biol.*, **225**, 519-532; j) PITERA, J. et KOLLMAN, P., 1997, *J. Mol. Graphics Mod.*, **15**, 355-358.
- [109] ENGLER, E. et WIPFF, G., MDS (non publié).
- [110] a) HANSEN, J. P. et McDONALD, I. R., 1986, *Theory of Simple Liquids* (Academic Press, London); b) POWLES, J. G., 1973, *Advan.Phys.*, **22**, 1; c) GRAY, C. G. et GUBBINS, K. E., 1984, *The Theory of Molecular Fluids. I. Fundamentals* (Clarendon Press, Oxford); d) SOPER, A. K. et EGELSTAFF, P. A., 1980, *Mol. Phys.*, **39**, 1202; e) GRAY, C. G. et HENDERSON, R. I., 1978, *Can.J. Phys.*, **56**, 571; f) GRAY, C. G. et HENDERSON, R. I., 1979, *Can.J. Phys.*, **57**, 1605; g) EGELSTAFF, P. A., GRAY, C. G. et GUBBINS, K. E., 1975, *Molecular Structure and Properties* (Butterworth, London).
- [111] NARTEN, A. H., DANFORD, M.D. et LEVY, H. A., 1967, *Discuss. Faraday Soc.*, **43**, 97.
- [112] KIRKWOOD, J. G. et BUFF, F. P., 1949, *J. Chem. Phys.*, **17**, 338-343.
- [113] MIYAZAKI, J., BARKER, J. A. et POUND, G. M., 1976, *J. Chem. Phys.*, **64**, 3364-3369.
- [114] DUAN, Y. et KOLLMAN, P. A., 1998, *Science*, **282**, 740-744.
- [115] NICKLAUS, M. C., WILLIAMS, R. W., BIENFAIT, B., BILLINGS, E. S. et HODOSCEK, M., 1998, *J. Chem. Inf. Comp. Sci.*, **38**, 893-905.

GNU Free Documentation License
Version 1.2, November 2002

Copyright (C) 2000,2001,2002 Free Software Foundation, Inc.
59 Temple Place, Suite 330, Boston, MA 02111-1307 USA
Everyone is permitted to copy and distribute verbatim copies
of this license document, but changing it is not allowed.

0. PREAMBLE

The purpose of this License is to make a manual, textbook, or other functional and useful document "free" in the sense of freedom: to assure everyone the effective freedom to copy and redistribute it, with or without modifying it, either commercially or noncommercially. Secondly, this License preserves for the author and publisher a way to get credit for their work, while not being considered responsible for modifications made by others.

This License is a kind of "copyleft", which means that derivative works of the document must themselves be free in the same sense. It complements the GNU General Public License, which is a copyleft license designed for free software.

We have designed this License in order to use it for manuals for free software, because free software needs free documentation: a free program should come with manuals providing the same freedoms that the software does. But this License is not limited to software manuals; it can be used for any textual work, regardless of subject matter or whether it is published as a printed book. We recommend this License principally for works whose purpose is instruction or reference.

1. APPLICABILITY AND DEFINITIONS

This License applies to any manual or other work, in any medium, that contains a notice placed by the copyright holder saying it can be distributed under the terms of this License. Such a notice grants a world-wide, royalty-free license, unlimited in duration, to use that work under the conditions stated herein. The "Document", below, refers to any such manual or work. Any member of the public is a licensee, and is addressed as "you". You accept the license if you copy, modify or distribute the work in a way requiring permission under copyright law.

A "Modified Version" of the Document means any work containing the Document or a portion of it, either copied verbatim, or with modifications and/or translated into another language.

A "Secondary Section" is a named appendix or a front-matter section of the Document that deals exclusively with the relationship of the publishers or authors of the Document to the Document's overall subject (or to related matters) and contains nothing that could fall directly within that overall subject. (Thus, if the Document is in part a textbook of mathematics, a Secondary Section may not explain any mathematics.) The relationship could be a matter of historical connection with the subject or with related matters, or of legal, commercial, philosophical, ethical or political position regarding

them.

The "Invariant Sections" are certain Secondary Sections whose titles are designated, as being those of Invariant Sections, in the notice that says that the Document is released under this License. If a section does not fit the above definition of Secondary then it is not allowed to be designated as Invariant. The Document may contain zero Invariant Sections. If the Document does not identify any Invariant Sections then there are none.

The "Cover Texts" are certain short passages of text that are listed, as Front-Cover Texts or Back-Cover Texts, in the notice that says that the Document is released under this License. A Front-Cover Text may be at most 5 words, and a Back-Cover Text may be at most 25 words.

A "Transparent" copy of the Document means a machine-readable copy, represented in a format whose specification is available to the general public, that is suitable for revising the document straightforwardly with generic text editors or (for images composed of pixels) generic paint programs or (for drawings) some widely available drawing editor, and that is suitable for input to text formatters or for automatic translation to a variety of formats suitable for input to text formatters. A copy made in an otherwise Transparent file format whose markup, or absence of markup, has been arranged to thwart or discourage subsequent modification by readers is not Transparent. An image format is not Transparent if used for any substantial amount of text. A copy that is not "Transparent" is called "Opaque".

Examples of suitable formats for Transparent copies include plain ASCII without markup, Texinfo input format, LaTeX input format, SGML or XML using a publicly available DTD, and standard-conforming simple HTML, PostScript or PDF designed for human modification. Examples of transparent image formats include PNG, XCF and JPG. Opaque formats include proprietary formats that can be read and edited only by proprietary word processors, SGML or XML for which the DTD and/or processing tools are not generally available, and the machine-generated HTML, PostScript or PDF produced by some word processors for output purposes only.

The "Title Page" means, for a printed book, the title page itself, plus such following pages as are needed to hold, legibly, the material this License requires to appear in the title page. For works in formats which do not have any title page as such, "Title Page" means the text near the most prominent appearance of the work's title, preceding the beginning of the body of the text.

A section "Entitled XYZ" means a named subunit of the Document whose title either is precisely XYZ or contains XYZ in parentheses following text that translates XYZ in another language. (Here XYZ stands for a specific section name mentioned below, such as "Acknowledgements", "Dedications", "Endorsements", or "History".) To "Preserve the Title" of such a section when you modify the Document means that it remains a section "Entitled XYZ" according to this definition.

The Document may include Warranty Disclaimers next to the notice which states that this License applies to the Document. These Warranty Disclaimers are considered to be included by reference in this License, but only as regards disclaiming warranties: any other implication that these Warranty Disclaimers may have is void and has

no effect on the meaning of this License.

2. VERBATIM COPYING

You may copy and distribute the Document in any medium, either commercially or noncommercially, provided that this License, the copyright notices, and the license notice saying this License applies to the Document are reproduced in all copies, and that you add no other conditions whatsoever to those of this License. You may not use technical measures to obstruct or control the reading or further copying of the copies you make or distribute. However, you may accept compensation in exchange for copies. If you distribute a large enough number of copies you must also follow the conditions in section 3.

You may also lend copies, under the same conditions stated above, and you may publicly display copies.

3. COPYING IN QUANTITY

If you publish printed copies (or copies in media that commonly have printed covers) of the Document, numbering more than 100, and the Document's license notice requires Cover Texts, you must enclose the copies in covers that carry, clearly and legibly, all these Cover Texts: Front-Cover Texts on the front cover, and Back-Cover Texts on the back cover. Both covers must also clearly and legibly identify you as the publisher of these copies. The front cover must present the full title with all words of the title equally prominent and visible. You may add other material on the covers in addition. Copying with changes limited to the covers, as long as they preserve the title of the Document and satisfy these conditions, can be treated as verbatim copying in other respects.

If the required texts for either cover are too voluminous to fit legibly, you should put the first ones listed (as many as fit reasonably) on the actual cover, and continue the rest onto adjacent pages.

If you publish or distribute Opaque copies of the Document numbering more than 100, you must either include a machine-readable Transparent copy along with each Opaque copy, or state in or with each Opaque copy a computer-network location from which the general network-using public has access to download using public-standard network protocols a complete Transparent copy of the Document, free of added material. If you use the latter option, you must take reasonably prudent steps, when you begin distribution of Opaque copies in quantity, to ensure that this Transparent copy will remain thus accessible at the stated location until at least one year after the last time you distribute an Opaque copy (directly or through your agents or retailers) of that edition to the public.

It is requested, but not required, that you contact the authors of the Document well before redistributing any large number of copies, to give them a chance to provide you with an updated version of the Document.

4. MODIFICATIONS

You may copy and distribute a Modified Version of the Document under the conditions of sections 2 and 3 above, provided that you release the Modified Version under precisely this License, with the Modified Version filling the role of the Document, thus licensing distribution and modification of the Modified Version to whoever possesses a copy of it. In addition, you must do these things in the Modified Version:

- A. Use in the Title Page (and on the covers, if any) a title distinct from that of the Document, and from those of previous versions (which should, if there were any, be listed in the History section of the Document). You may use the same title as a previous version if the original publisher of that version gives permission.
- B. List on the Title Page, as authors, one or more persons or entities responsible for authorship of the modifications in the Modified Version, together with at least five of the principal authors of the Document (all of its principal authors, if it has fewer than five), unless they release you from this requirement.
- C. State on the Title page the name of the publisher of the Modified Version, as the publisher.
- D. Preserve all the copyright notices of the Document.
- E. Add an appropriate copyright notice for your modifications adjacent to the other copyright notices.
- F. Include, immediately after the copyright notices, a license notice giving the public permission to use the Modified Version under the terms of this License, in the form shown in the Addendum below.
- G. Preserve in that license notice the full lists of Invariant Sections and required Cover Texts given in the Document's license notice.
- H. Include an unaltered copy of this License.
- I. Preserve the section Entitled "History", Preserve its Title, and add to it an item stating at least the title, year, new authors, and publisher of the Modified Version as given on the Title Page. If there is no section Entitled "History" in the Document, create one stating the title, year, authors, and publisher of the Document as given on its Title Page, then add an item describing the Modified Version as stated in the previous sentence.
- J. Preserve the network location, if any, given in the Document for public access to a Transparent copy of the Document, and likewise the network locations given in the Document for previous versions it was based on. These may be placed in the "History" section. You may omit a network location for a work that was published at least four years before the Document itself, or if the original publisher of the version it refers to gives permission.
- K. For any section Entitled "Acknowledgements" or "Dedications", Preserve the Title of the section, and preserve in the section all the substance and tone of each of the contributor acknowledgements and/or dedications given therein.
- L. Preserve all the Invariant Sections of the Document, unaltered in their text and in their titles. Section numbers or the equivalent are not considered part of the section titles.
- M. Delete any section Entitled "Endorsements". Such a section may not be included in the Modified Version.
- N. Do not retitle any existing section to be Entitled "Endorsements" or to conflict in title with any Invariant Section.
- O. Preserve any Warranty Disclaimers.

If the Modified Version includes new front-matter sections or appendices that qualify as Secondary Sections and contain no material copied from the Document, you may at your option designate some or all of these sections as invariant. To do this, add their titles to the

list of Invariant Sections in the Modified Version's license notice. These titles must be distinct from any other section titles.

You may add a section Entitled "Endorsements", provided it contains nothing but endorsements of your Modified Version by various parties--for example, statements of peer review or that the text has been approved by an organization as the authoritative definition of a standard.

You may add a passage of up to five words as a Front-Cover Text, and a passage of up to 25 words as a Back-Cover Text, to the end of the list of Cover Texts in the Modified Version. Only one passage of Front-Cover Text and one of Back-Cover Text may be added by (or through arrangements made by) any one entity. If the Document already includes a cover text for the same cover, previously added by you or by arrangement made by the same entity you are acting on behalf of, you may not add another; but you may replace the old one, on explicit permission from the previous publisher that added the old one.

The author(s) and publisher(s) of the Document do not by this License give permission to use their names for publicity for or to assert or imply endorsement of any Modified Version.

5. COMBINING DOCUMENTS

You may combine the Document with other documents released under this License, under the terms defined in section 4 above for modified versions, provided that you include in the combination all of the Invariant Sections of all of the original documents, unmodified, and list them all as Invariant Sections of your combined work in its license notice, and that you preserve all their Warranty Disclaimers.

The combined work need only contain one copy of this License, and multiple identical Invariant Sections may be replaced with a single copy. If there are multiple Invariant Sections with the same name but different contents, make the title of each such section unique by adding at the end of it, in parentheses, the name of the original author or publisher of that section if known, or else a unique number. Make the same adjustment to the section titles in the list of Invariant Sections in the license notice of the combined work.

In the combination, you must combine any sections Entitled "History" in the various original documents, forming one section Entitled "History"; likewise combine any sections Entitled "Acknowledgements", and any sections Entitled "Dedications". You must delete all sections Entitled "Endorsements".

6. COLLECTIONS OF DOCUMENTS

You may make a collection consisting of the Document and other documents released under this License, and replace the individual copies of this License in the various documents with a single copy that is included in the collection, provided that you follow the rules of this License for verbatim copying of each of the documents in all other respects.

You may extract a single document from such a collection, and distribute it individually under this License, provided you insert a copy of this

License into the extracted document, and follow this License in all other respects regarding verbatim copying of that document.

7. AGGREGATION WITH INDEPENDENT WORKS

A compilation of the Document or its derivatives with other separate and independent documents or works, in or on a volume of a storage or distribution medium, is called an "aggregate" if the copyright resulting from the compilation is not used to limit the legal rights of the compilation's users beyond what the individual works permit. When the Document is included in an aggregate, this License does not apply to the other works in the aggregate which are not themselves derivative works of the Document.

If the Cover Text requirement of section 3 is applicable to these copies of the Document, then if the Document is less than one half of the entire aggregate, the Document's Cover Texts may be placed on covers that bracket the Document within the aggregate, or the electronic equivalent of covers if the Document is in electronic form. Otherwise they must appear on printed covers that bracket the whole aggregate.

8. TRANSLATION

Translation is considered a kind of modification, so you may distribute translations of the Document under the terms of section 4. Replacing Invariant Sections with translations requires special permission from their copyright holders, but you may include translations of some or all Invariant Sections in addition to the original versions of these Invariant Sections. You may include a translation of this License, and all the license notices in the Document, and any Warranty Disclaimers, provided that you also include the original English version of this License and the original versions of those notices and disclaimers. In case of a disagreement between the translation and the original version of this License or a notice or disclaimer, the original version will prevail.

If a section in the Document is Entitled "Acknowledgements", "Dedications", or "History", the requirement (section 4) to Preserve its Title (section 1) will typically require changing the actual title.

9. TERMINATION

You may not copy, modify, sublicense, or distribute the Document except as expressly provided for under this License. Any other attempt to copy, modify, sublicense or distribute the Document is void, and will automatically terminate your rights under this License. However, parties who have received copies, or rights, from you under this License will not have their licenses terminated so long as such parties remain in full compliance.

10. FUTURE REVISIONS OF THIS LICENSE

The Free Software Foundation may publish new, revised versions

of the GNU Free Documentation License from time to time. Such new versions will be similar in spirit to the present version, but may differ in detail to address new problems or concerns. See <http://www.gnu.org/copyleft/>.

Each version of the License is given a distinguishing version number. If the Document specifies that a particular numbered version of this License "or any later version" applies to it, you have the option of following the terms and conditions either of that specified version or of any later version that has been published (not as a draft) by the Free Software Foundation. If the Document does not specify a version number of this License, you may choose any version ever published (not as a draft) by the Free Software Foundation.

ADDENDUM: How to use this License for your documents

To use this License in a document you have written, include a copy of the License in the document and put the following copyright and license notices just after the title page:

```
Copyright (c) YEAR YOUR NAME.
Permission is granted to copy, distribute and/or modify this document
under the terms of the GNU Free Documentation License, Version 1.2
or any later version published by the Free Software Foundation;
with no Invariant Sections, no Front-Cover Texts, and no Back-Cover Texts.
A copy of the license is included in the section entitled "GNU
Free Documentation License".
```

If you have Invariant Sections, Front-Cover Texts and Back-Cover Texts, replace the "with...Texts." line with this:

```
with the Invariant Sections being LIST THEIR TITLES, with the
Front-Cover Texts being LIST, and with the Back-Cover Texts being LIST.
```

If you have Invariant Sections without Cover Texts, or some other combination of the three, merge those two alternatives to suit the situation.

If your document contains nontrivial examples of program code, we recommend releasing these examples in parallel under your choice of free software license, such as the GNU General Public License, to permit their use in free software.