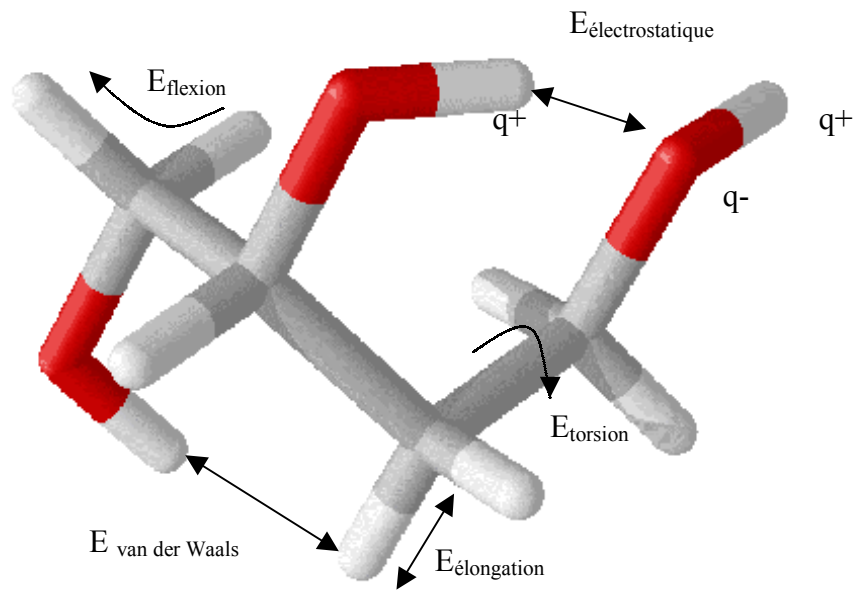
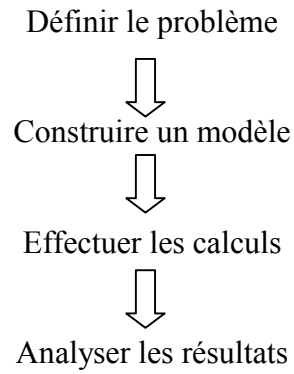


### 3 LA MODELISATION MOLECULAIRE.



Champ de forces :

$$E_{tot} = \sum_{bonds} K_r (r - r_{eq})^2 + \sum_{angles} K_\theta (\theta - \theta_{eq})^2 + \sum_{dihedrals} \frac{V_n}{2} [1 + \cos(n\phi - \gamma)] + \sum_{i < j} \left[ \frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^6} + \frac{q_i q_j}{\epsilon R_{ij}} \right]$$

### **3.1 Introduction.**

La recherche et la synthèse de nouveaux composés chimiques et biochimiques sont aujourd'hui souvent associées à une étude par modélisation moléculaire. La modélisation moléculaire est une technique permettant, non seulement de représenter les propriétés et les réactions chimiques mais aussi de manipuler les modèles des structures en deux ou trois dimensions.

Le développement important au cours de ces dernières années des moyens informatiques (mémoire plus importante, infographie plus conviviale, superordinateurs plus puissants, échange d'information grâce à l'Internet, etc.) a permis à la chimie de s'enrichir d'outils informatiques spécialement dédiés à la représentation des molécules. Ainsi, que ce soit avec des superordinateurs ou de simples ordinateurs de bureau (PC ou Mac), tout chimiste peut aujourd'hui réaliser une étude plus ou moins exhaustive et précise des systèmes moléculaires grâce à la modélisation moléculaire. Autrefois, principalement à cause de la faible puissance des ordinateurs, la modélisation moléculaire ne permettait que l'étude des petits systèmes moléculaires. Aujourd'hui, comme les ordinateurs sont de plus en plus puissants, nous travaillons d'une part sur des systèmes contenant plus d'atomes (molécules de solvant explicite pouvant atteindre des dizaines de milliers d'atomes, grandes surfaces moléculaires, protéines) et d'autre part sur des équations théoriques moins simplifiées. De ce fait, les temps de simulations se sont nettement allongés (de l'ordre de la nanoseconde) et un phénomène de compétition entre la taille des systèmes moléculaires et la puissance des superordinateurs<sup>48</sup> est apparu. La naissance de logiciels plus conviviaux ces dernières années, la parallélisation des ordinateurs, l'apparition de Linux (Linus B. Torvald, 1991 ; un système d'exploitation semblable à ceux des superordinateurs mais avec un coût d'achat nettement inférieur) ont apporté une aide significative au développement de la modélisation moléculaire. Contrairement à ses débuts, la modélisation moléculaire est aujourd'hui mieux reconnue.

---

48 - Van Gunsteren, W. ; Berendsen, H. J. C. *Angew. Chem. Int. Ed. Engl.* **1990**, 29, 992.

D'ailleurs, en 1998, le prix Nobel de chimie a été décerné à John. A. Pople et Walter Kohn pour leurs travaux dans les domaines de la chimie informatique et de la modélisation moléculaire (chimie quantique). De même en 1992, Rudolph A. Marcus reçut le prix Nobel pour ses travaux théoriques sur le transfert des électrons, et en 1981, Kenichi Fukui, Roald Hoffmann reçurent le même prix pour leurs travaux théoriques sur les réactions chimiques.

Lorsque nous utilisons un modèle particulier pour prédire les propriétés d'un système moléculaire, nous devons être conscients des hypothèses (ex : validité de la reproduction des valeurs expérimentales), des approximations (ex : champ de forces utilisé) et des limitations (ex : temps de simulation) impliquées par le modèle.

La modélisation moléculaire se sert pour simuler des systèmes moléculaires, des méthodes de calculs théoriques telles que la mécanique quantique (QM) (*ab initio* ou semi-empirique), la dynamique moléculaire (DM), la mécanique moléculaire (MM) et, depuis quelques années, l'association QM/MM qui paraît très porteuse.

Dans cette thèse, la mécanique quantique, la mécanique moléculaire et la dynamique moléculaire nous serviront de support. De nombreux livres<sup>49,50,51</sup> décrivent ces méthodes ainsi que leurs différents champs de force. C'est pour cette raison que nous allons uniquement décrire les deux champs de force utilisés en MM et DM, c'est à dire MM3 et AMBER.

## **3.2 La mécanique moléculaire et le champ de force de MM3.**

### **3.2.1 Fondements théoriques.**

L'expression "Mécanique Moléculaire" désigne actuellement une méthode de calcul qui permet, *a priori*, d'obtenir des résultats de géométries et d'énergies moléculaires en se

---

49 - Leach, A. R. *Molecular Modelling*, Glaxo Wellcome Research and Development and the University of Southampton, 1996.

50 - Chapitres spécifiques apparaissant dans chaque volume de *Reviews in Computational Chemistry*, Boyd and Lipkowitz, 1990-2000.

51 - Burkert, U. ; Allinger, N. L. *Molecular Mechanics. American Chemical Society monograph*, 1982, 177, Washington, DC.

basant sur la mécanique classique. La MM est apparue en 1930<sup>52</sup>, mais s'est développée à partir des années soixante, quand les ordinateurs furent plus accessibles et plus performants. La MM est basée sur l'approximation de Born-Oppenheimer selon laquelle les électrons sont beaucoup plus rapides que les noyaux et ces derniers sont donc implicitement traités. La MM est une méthode empirique où les atomes (les noyaux) sont représentés par des masses ou des sphères, et les liaisons par des ressorts de différentes forces. Les valeurs de ces forces proviennent de données expérimentales de spectroscopie vibrationnelle, de données de diffraction des RX ou de valeurs obtenues par les méthodes *ab-initio*. Le calcul de l'ensemble de ces forces, selon les règles établies pour les oscillateurs harmoniques, permet d'obtenir l'énergie moléculaire et de déterminer la structure tridimensionnelle. Aujourd'hui, la MM est utilisée pour interpréter les données de RMN. L'avantage principal de la MM par rapport aux autres méthodes est la rapidité des calculs. Ceci est dû à une simplification des équations du champ de forces ; il est donc possible de traiter des systèmes moléculaires très grands. En revanche, cette méthode ne prend pas en compte la structure électronique moléculaire. Il est donc impossible d'étudier des systèmes dans lesquels les effets électroniques sont prédominants. Ainsi, la validité des résultats (résultats des énergies) ne dépend que de la paramétrisation originale des constantes internes au champ de force. De plus, la valeur de l'énergie stérique obtenue en MM n'a aucun sens physique dans l'absolu ; c'est uniquement la comparaison entre deux valeurs obtenues pour deux conformations différentes qui peut être interprétée.

Les champs de force les plus répandus (avec leur domaine d'application) sont :

AMBER (Kollman)<sup>53</sup> : protéines, acides nucléiques

CHARMM (Karplus)<sup>54</sup> : protéines

CVFF (Hagler)<sup>55</sup> : recherche de fréquences de vibration

ECEPP (Scherega)<sup>56</sup> : protéines

MM2<sup>57</sup>, MM3<sup>58</sup>, MM4<sup>59</sup> (Allinger) : molécules organiques de petite taille

---

52 - Andrews, D. H. *Phys. Rev.* **1930**, 36, 544.

53 - Weiner, P. K. ; Kollman, P. A. *J. Comput Chem.* **1981**, 2, 287.

54 - Karplus, M. *J. Comput. Chem.* **1983**, 4, 187.

55 - Ermer, O. *Struct. Bonding Berlin* **1976**, 27, 161.

56 - Momany, F. A. ; Carruthers, L. M. ; McGuire, R. F. ; Scheraga, H. A. *J. Phys. Chem.* **1974**, 78, 1595.

BIOSYM (MSI)<sup>60</sup> : protéines, molécules organiques

Sybyl (Tripos)<sup>61</sup> : molécules organiques.

### 3.2.2 Le champ de force de MM3 pour une molécule apolaire.

L'énergie stérique (SE) d'un système moléculaire, calculée à partir de la MM, est évaluée suivant la somme de contributions de plusieurs énergies dont chacune est caractérisée par une constante de force et une valeur à l'équilibre (repos). L'énergie stérique s'exprime par l'équation suivante :

$$SE = \sum E_{str}(l) + \sum E_{bnd}(\theta) + \sum E_{tors}(\omega) + \sum E_{elec} + \sum E_{nb}(r) + \text{termes croisés}$$

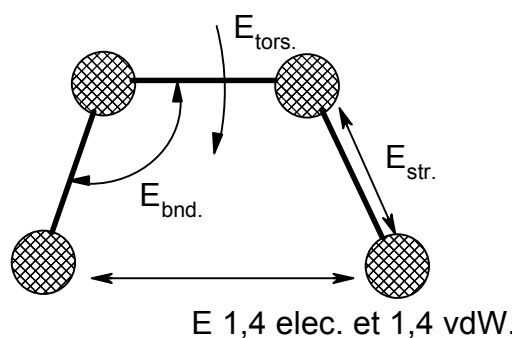


Figure 10 : Les déformations en mécanique moléculaire.

L'énergie stérique peut être divisée en deux groupes :

- Interaction entre atomes liés par deux ou trois liaisons chimiques (énergie de *stretching*, *bending*, torsion et torsions impropres, termes croisés).
- Interaction entre atomes non liés (ou séparés de plus de trois liaisons chimiques) (énergie électrostatique, de liaisons hydrogènes, de van der Waals).

---

57 - Allinger, N. L. *J. Am. Chem. Soc.* **1977**, *99*, 8127.

58 - Allinger, N. L. ; Yuh, Y. H. ; Lii, J-H. *J. Am. Chem. Soc.* **1989**, *111*, 8522.

59 - Allinger, N. L. ; Chen, K. ; Lii, J-H. *J. Comp. Chem.* **1996**, *17*, 642.

60 - Molecular Simulation Inc., San Diego.

61 - Tripos Inc., St. Louis, MO 63144, USA.

### 3.2.2.1 L'énergie de *stretching* ou d'élongation.

L'énergie d'élongation de 2 atomes liés d'une distance  $l$ , est basée sur la simplification du développement en série limitée de la loi de Hooke :

$$E_{str.} = \sum k_l(l - l_0)^2$$

$k_l$  représente la force de la liaison (ou raideur du ressort), et  $l_0$ , la valeur de la distance de la liaison à l'équilibre ou au repos. Chaque valeur dépend de la nature des atomes.

Dans le champ de force de MM3, l'équation appliquée à tous les atomes que nous utiliserons, est :

$$E_{str.} = 71.94k_l(l - l_0)^2 \left[ 1 - 2.55(l - l_0) + (7/12)2.55(l - l_0)^2 \right]$$

Celle-ci est d'ordre 4 de façon à être plus précise lorsque  $l$  s'éloigne de l'équilibre, et ainsi cette équation se rapproche de la fonction plus exacte de Morse (forme exponentielle).

### 3.2.2.2 L'énergie de *bending* ou de flexion.

L'énergie de flexion est fonction de l'angle  $\theta$  formé entre trois atomes par rapport à la valeur à l'équilibre (ou au repos)  $\theta_0$  et à la constante de force de flexion selon la loi générale:  $E_{bnd.} = \sum k_\theta(\theta - \theta_0)^2$ . Le champ de force de MM3 contrairement à ses prédécesseurs, possède des termes d'ordre 6 de façon à améliorer les résultats géométriques et de vibrations.

$$E_{bnd.} = 0.021914k_\theta(\theta - \theta_0)^2 \left[ 1 - 0.014(\theta - \theta_0) + 5.6(10^{-5})(\theta - \theta_0)^2 - 7.0(10^{-7})(\theta - \theta_0)^3 + 9.0(10^{-10})(\theta - \theta_0)^4 \right]$$

### 3.2.2.3 L'énergie de torsion.

Cette énergie correspond à la rotation d'une liaison selon l'angle dièdre  $\omega$  formé par quatre atomes. Pour reproduire l'énergie de torsion, le champ de force de MM3 utilise 3 termes d'une série de Fourier :

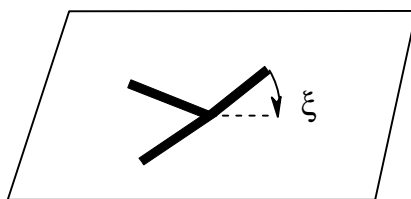
$$E_{tors.} = \left(\frac{V_1}{2}\right)(1 + \cos \omega) + \left(\frac{V_2}{2}\right)(1 - \cos 2\omega) + \left(\frac{V_3}{2}\right)(1 + \cos 3\omega)$$

Les termes  $V_1$ ,  $V_2$  et  $V_3$  sont des constantes dépendant du type d'atomes. Le terme  $V_2$  représente la constante pour des liaisons de type  $sp^2-sp^2$  et le terme  $V_3$  pour des liaisons de type  $sp^3-sp^3$ .

### 3.2.2.4 L'énergie des angles dièdres impropres.

Il peut exister une déformation pour trois atomes liés à un quatrième atome, par exemple pour les atomes  $sp^2$ . L'équation qui la régit est :

$E_\xi = 1/2 k_\xi (\xi - \xi_0)^2$  où  $\xi$  représente l'angle formé entre la liaison d'un atome et l'atome central ( $sp^2$ ) et le plan formé par les deux autres atomes et l'atome  $sp^2$  (Figure 11).



**Figure 11 : Angle dièdre impropre  $\xi$**

### 3.2.2.5 L'énergie d'interaction de van der Waals.

L'énergie d'interaction de van der Waals correspond à l'interaction à une distance déterminée entre deux atomes non liés. Elle est répulsive à très courte distance et attractive jusqu'à l'infini. Le champ de force de MM3 décrit cette interaction selon une fonction de type Buckingham :

Sachant que  $P=r_v/r$ , où  $r_v$  représente la somme des rayons de van der Waals des deux atomes et  $r$  représente la distance interatomique

si  $P \leq 3.020$  alors

$$E_{vdw} = \varepsilon \left[ 1.84 (10^5) e^{-12.00 \left( \frac{r}{r_v} \right)} - 2.25 \left( \frac{r_v}{r} \right)^6 \right]$$

si  $P > 3.020$  alors

$$E_{vdw} = \varepsilon 192.270 P^2$$

où  $\varepsilon$  correspond à la moyenne de la dureté des deux atomes, il représente aussi la profondeur du puits de potentiel à la distance  $r_v$ . Le terme exponentiel correspond à l'interaction répulsive, et le terme en puissance de 6 correspond à l'interaction attractive.

N.B. : Dans d'autres champs de force, l'expression de l'énergie de van der Waals est souvent représentée sous la forme d'un potentiel de Lennard-Jones :

$$E_{vdw} = \sum \left[ A \left( \frac{r_v}{r} \right)^{12} - B \left( \frac{r_v}{r} \right)^6 \right].$$

A l'intérieur de cette interaction de van der Waals, il existe l'interaction hydrogène ayant une fonction du type : si  $P \leq 3.02$  alors

$$E_{hb} = (\varepsilon_{hb} / D) \left[ 1.84 (10^5) e^{-12.00 \left( \frac{r}{r_v} \right)} - 2.25 \left( \frac{r_v}{r} \right)^6 \cos \theta (l / l_0) \right]$$

si  $P > 3.02$  alors

$$E_{hb} = (\varepsilon_{hb} / D) [192.270 P^2 + 1706.96 - 1706.96 \cos \theta (l / l_0)]$$

où  $\varepsilon_{HB}$  représente le paramètre énergétique de la liaison hydrogène,  $D$  est la constante diélectrique,  $l_0$  est la distance idéale de l'hydrogène à l'atome auquel il est lié, et  $l$  la distance réelle,  $\theta$  est l'angle formé par les trois atomes (O-H...O) qui participent à la liaison hydrogène.



### 3.2.2.6 L'énergie d'interaction électrostatique.

Suivant les champs de force, l'énergie d'interaction électrostatique dépend des charges atomiques  $q$  ou des moments dipolaires  $\mu$ . Pour une molécule neutre, le programme MM3 associe un moment dipolaire aux liaisons et les interactions électrostatiques sont donc des interactions dipôle-dipôle. L'énergie d'interaction électrostatique augmente avec la polarité des atomes et dépend de l'interaction du milieu :

$$E_{elec.} = 14.39418\mu_a\mu_b(\cos\chi - 3\cos\alpha_a\cos\alpha_b) / R^3\epsilon$$

où  $\mu_a$  et  $\mu_b$  sont les moments dipolaires,  $\chi$  est l'angle formé par les deux moments dipolaires,  $R$  la distance entre les deux atomes,  $\alpha_a$  et  $\alpha_b$  sont les angles formés entre le moment dipolaire et la distance  $R$ ,  $\epsilon$  correspond à  $4\pi\epsilon_0$ ,  $\epsilon_0$  étant la permittivité dans le vide.

### 3.2.2.7 L'énergie des termes croisés.

Il existe d'autres termes, appelés termes croisés qui décrivent les interactions entre les trois premiers termes, c'est pourquoi on les désigne aussi sous le nom de *stretching-bending*, *torsion-stretching*, *bending-bending*. Dans le champ de force de MM3, ces interactions sont exprimées par les équations suivantes :

$$E_{sb} = 2.5111k_{sb}[(l - l_0) + (l' - l'_0)](\theta - \theta_0)$$

$$E_{ts} = 11.995(k_{ts} / 2)(l - l_0)(1 + \cos 3\omega)$$

$$E_{bb} = -0.021914k_{\theta\theta'}(\theta - \theta_0)(\theta' - \theta'_0)$$

où  $k_{sb}$ ,  $k_{ts}$  et  $k_{\theta\theta'}$  représentent les constantes des termes croisés.

Il n'existe pas de termes croisés *torsion-bending*.

Contrairement au champ de force MM2, les paramètres des paires d'électron non liantes ne sont pas pris en compte dans le champ de force MM3.

## 3.3 La dynamique moléculaire et le programme AMBER.

“All the bugs are not removed from amber” (Peter Albert Kollman)

### 3.3.1 La dynamique moléculaire<sup>62,63</sup>.

Les premiers pas de la dynamique moléculaire n'ont pu se faire que grâce à l'arrivée des premiers ordinateurs (1957)<sup>64</sup>. Mais les premières réelles simulations ont été faites par Rahman<sup>65</sup>, grâce à ses travaux sur la simulation de l'argon liquide en 1964 avec un temps de simulation de  $10^{-11}$ s pour un temps de CPU de 0.05 h, puis de l'eau liquide<sup>66</sup> en 1971.

La mécanique moléculaire permet de calculer la position des atomes et des molécules dans leurs états énergétiques minima, à 0 K c'est à dire sans la moindre vibration ni le moindre mouvement. La dynamique moléculaire tente de simuler le mouvement des atomes et des molécules au cours du temps pour une température supérieure à 0 K. Ce mouvement est basé sur les lois de Newton de la mécanique classique. A une température donnée, une vitesse initiale est attribuée à chaque atome. A partir de ces vitesses, considérées constantes, la position des atomes change en  $v_i \delta t$ , où  $v_i$  est la vitesse de l'atome  $i$  et  $\delta t$  est l'incrément de temps. Avec ces nouvelles coordonnées, on détermine les forces qui agissent sur chaque atome selon le champ de force et le potentiel utilisé, et avec la seconde loi de Newton, on attribue les nouvelles vitesses et positions :

$$\vec{F}_i = m_i \vec{a}_i = m_i \frac{d^2 \vec{r}_i(t)}{dt^2}$$

où  $\vec{F}_i$  est le vecteur force agissant sur l'atome  $i$ ,  $m_i$  est la masse de l'atome  $i$ ,  $\vec{a}_i$  est le vecteur accélération de l'atome  $i$  et,  $\vec{r}_i$  la position de l'atome  $i$ .

Grâce aux vitesses et aux positions de chaque atome dans le temps, il est possible d'évaluer les données macroscopiques, comme l'énergie cinétique et la température.

L'énergie cinétique se calcule selon la formule :

---

62 -van Gunsteren, W. F. ; Berendsen, H. J. C. *Angew. Chem. Int. Ed. Engl.* **1990**, 29, 992.

63 - Dugas, H, "Modélisation Moléculaire. Stéréochimie et analyse conformationnelle". 2<sup>ème</sup> Ed., La Librairie de L'Université de Montréal, **1992**.

64 -Alder, B. J. ; Wainwright, T. E. *J. Chem. Phys.* **1957**, 27, 1208.

65 - Rahman, A. *Phys. Rev.* **1964**, 136, A 405.

66 - Rahman, A. ; Stillinger, F. H. *J. Chem. Phys.* **1971**, 55, 3336.

$$E_c = \sum_{i=1}^N \frac{|p_i|^2}{2m_i} \text{ où } p_i \text{ est la quantité de mouvement de l'atome } i.$$

La température s'obtient à partir de l'énergie cinétique :

$$E_c = \frac{k_b T}{2} (3N - N_c)$$

$k_b$  est la constante de Boltzmann,  $N_c$  le nombre de contrainte,  $3N - N_c$  le nombre total de degré de liberté.

Habituellement, la température est lentement augmentée par étapes successives de 0 K à 298 K, amenant ainsi le système moléculaire à un état d'équilibre.

La force  $\vec{F}_i$  qui s'exerce sur un atome  $i$  se trouvant en position  $r_i(t)$  est déterminée par différentiation (ou calcul du gradient) de la fonction d'énergie potentielle au voisinage de la position  $r_i(t)$  :

$$\vec{F}_i = - \frac{d\vec{E}(r_1, \dots, r_n)}{dr_i(t)}$$

L'énergie potentielle  $\vec{E}$  se calcule à partir des paramètres du champ de force de la mécanique classique. Dans un premier temps, les accélérations atomiques sont calculées à partir des forces s'exerçant sur chaque atome. Ensuite, les vitesses de chaque atome sont évaluées à partir des accélérations atomiques précédemment calculées. Enfin, les positions des atomes sont calculées grâce aux vitesses atomiques selon :

$$\vec{a}_i = \frac{d\vec{v}_i}{dt} \text{ et } \vec{v}_i = \frac{d\vec{r}_i}{dt}$$

Pour générer les trajectoires du système moléculaire, il suffit d'intégrer les équations de mouvement de manière itérative jusqu'à l'intervalle de temps choisi. L'intégration de

l'équation de Newton ( $\vec{F}_i = m_i \vec{a}_i = m_i \frac{d^2 \vec{r}_i(t)}{dt^2}$ ) se fait dans de faibles intervalles de temps

$\Delta t$ , en général 1-10 fs. (1 fs. =  $10^{-15}$  s.) pour des systèmes moléculaires. Cet intervalle de temps doit être plus petit que la période du mouvement de la plus haute fréquence. Une simulation de dynamique moléculaire permet de générer des trajectoires de l'ordre du millier de ps (1 ps. =  $10^{-12}$  s). Si le complexe moléculaire comporte de nombreux atomes, le temps de

calcul s'en trouvera allongé. Il est donc fréquent d'utiliser la méthodologie SHAKE<sup>67, 68</sup> qui consiste à restreindre des degrés de liberté internes d'un système. La restriction habituelle consiste à appliquer des forces additionnelles pour maintenir des liaisons fixes dans leur état d'équilibre. L'utilisation de ces restrictions aux liaisons C-H, celles possédant la plus petite vibration, permet d'augmenter l'intervalle de temps d'intégration.

Les calculs de dynamique moléculaire effectués dans cette thèse ont été réalisés grâce aux champs de force AMBER et AMBER\* de MacroModel v5.0. AMBER représente un ensemble de programmes permettant d'effectuer des simulations de dynamique moléculaire, particulièrement sur des biomolécules. Il représente aussi le champ de force empirique inclus dans les programmes précédents et dans d'autres programmes (tel que MacroModel). Le champ de force AMBER est en libre distribution. Néanmoins la suite de programmes est distribuée sous licence personnelle.

### **3.3.2 Le programme AMBER et son champ de force.**

AMBER, *Assisted Model Building with Energy Refinement*<sup>69</sup>, possède plusieurs champs de force comme OPLS<sup>70</sup>, all\_nuc94.in, all\_amino94.in, parm94.dat<sup>71</sup>, parm96.dat<sup>72</sup>, parm98.dat<sup>73</sup>. C'est le champ de force parm94.dat qui sera utilisé lors du travail de recherche avec AMBER v5<sup>74</sup> et v6<sup>75</sup>. AMBER a été développé par Peter Albert Kollman *et al.* pour

---

67 - Ryckaert, J. P. ; Ciccoti, G. ; Berendsen, H.J.C. *J. Comput. Phys.* **1977**, 23, 327.

68 - Ryckaert, J.P. *Mol. Phys.* **1985**, 55, 549.

69 - Weiner, P.K. ; Kollman, P. A. *J. Comp. Chem.* **1981**, 2, 287.

70 - Jorgensen, W. L. ; Tirado-Rives, J. *J. Am. Chem. Soc.* **1988**, 110, 1657.

71 - W.D. Cornell, W.D. ; Cieplak, P. ; Bayly, C.I. ; Gould, I. R. ; Merz, K.M. ; Jr. ; Ferguson, D.M. ; Spellmeyer, D.C. ; Fox, T. ; Caldwell, J.W. ; Kollman, P.A. *J. Am. Chem. Soc.* **1995**, 117, 5179-5197.

72 - Kollman, P. A. ; Dixon, R. ; Cornell, W. ; Fox, T. ; Chipot, C. ; Pohorille, A. *Computer Simulation of Biomolecular Systems*, Vol. 3 A.

73 - Cheatham, T.E. ; Cieplak, P. ; Kollman, P.A. *J. Biomol. Struct. Dyn.* **1999**, 16, 845-861.

74 - Case, D.A. ; Pearlman, D.A. ; Caldwell, 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. ;

étudier les protéines, les acides aminés et bien d'autres petites molécules<sup>76</sup>. L'équation utilisée dans le programme AMBER peut-être simplifiée par l'équation suivante :

$$E_{tot} = \sum_{bonds} K_r (r - r_{eq})^2 + \sum_{angles} K_\theta (\theta - \theta_{eq})^2 + \sum_{dihedrals} \frac{V_n}{2} [1 + \cos(n\phi - \gamma)] + \sum_{i < j} \left[ \frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^6} + \frac{q_i q_j}{\epsilon R_{ij}} \right]$$

L'énergie de liaison, de flexion et de torsion sont représentées par une simple expression harmonique diagonale. L'énergie d'interaction de van der Waals est représentée par un potentiel 6-12, et l'énergie électrostatique est modélisée par des interactions Coulombiennes de charges centrées sur les atomes (*atom-centered point charges*). Dans les champs de force précédents (parm91.dat<sup>77</sup>), l'interaction des liaisons hydrogènes incluse dans l'énergie de van der Waals était exprimée avec un potentiel 10-12.

$$E_{vdw} = \sum \left[ \frac{C}{r^{12}} - \frac{D}{r^{10}} \right]$$

Paramètres de van der Waals	R* (Å)	ε (kcal.mol <sup>-1</sup> )
OW	1.768	0.152
HW	1.000	0.020

Dans les champs de force suivants, et d'après les suggestions d'Hagler *et al.*<sup>78</sup>, R\* et ε sont égaux à zéro pour les hydrogènes de solvant participant à des liaisons hydrogènes. Les

Seibel, G.L. ; Singh, U.C. ; Weiner, P.K. ; Kollman, P.A. **1997**, AMBER 5, University of California, San Francisco.

75 - Case, D.A. ; Pearlman, D.A. ; Caldwell, 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. ; Tsui, V. ; Radmer, R.J. ; Duan, Y. ; Pitera, J. ; Massova, I. ; Seibel, G.L. ; Singh, U.C. ; Weiner, P.K. ; Kollman, P.A. **1999**, AMBER 6, University of California, San Francisco.

76 - Wipff, G. ; Weiner, P. ; Kollman, P.A. *J. Am. Chem. Soc.* **1982**, 104, 3249.

77 - (a) Weiner, S.J. ; Kollman, P.A. ; Case, D.A. ; Singh, U.C. ; Ghio, C. ; Alagona, G. ; Profeta, S. ; Weiner, P. *J. Am. Chem. Soc.* **1984**, 106, 765 (b) Weiner, S.J. ; Kollman, P.A. ; Nguyen, D.T. ; Case, D.A. *J. Comput. Chem.* **1986**, 7, 230.

78 - Hagler, A. ; Euler, E. ; Lifson, S. *J. Am. Chem. Soc.* **1974**, 96, 5319.

valeurs des charges et des paramètres de van der Waals ont été modifiées et il n'existe plus d'équations de van der Waals exprimée avec un potentiel 10-12. Ainsi le modèle explicite d'eau TIP3P<sup>79</sup> possède pour les hydrogènes (HW) des valeurs  $R^*$  et  $\epsilon$  égales à zéro.

Paramètres de van der Waals	$R^*$ (Å)	$\epsilon$ (kcal.mol <sup>-1</sup> )
OW	1.768	0.152
HW	0.000	0.000

Contrairement au champ de force MM3 qui évalue les énergies électrostatiques à partir des moments dipolaires, AMBER calcule l'énergie électrostatique grâce aux interactions de charges atomiques. L'obtention des charges atomiques est donc primordiale pour les résultats des simulations.

### 3.3.2.1 Obtention des charges atomiques.

Les charges atomiques utilisées par AMBER sont calculées à partir du potentiel électrostatique restreint (RESP)<sup>80</sup>. Il faut donc dans un premier temps obtenir un champ électrostatique correct. RESP (*Restrained ElectroStatic Potential*) fixe, à la surface moléculaire, le potentiel électrostatique (ESP) calculé à partir de méthodes quantiques en utilisant un modèle de charges atomiques centrées. Ce processus permet de considérer plusieurs conformations de la molécule. Ainsi, les charges RESP dépendent de l'ensemble des conformations de la molécule.

Avant l'apparition de ces méthodes, les charges utilisées étaient des charges ESP obtenues à partir des méthodes semi-empiriques. Mais elles ne représentaient pas correctement les moments dipolaires, donnant une représentation erronée des modèles solvatés dans une cage de solvant de type TIP3P<sup>81</sup>. De plus, les charges ESP dépendaient des

---

79 - Jorgensen, W.L. ; Chandreskhar, J. ; Madura, J. D. ; Impey, R.W. ; Klein, M.L. *J. Chem. Phys.* **1982**, 79, 926.

80 - Bayly, C. ; Cieplak, P. ; Cornell, W. ; Kollman, P.A. *J. Phys. Chem.* **1993**, 97, 10269.

81 - Jorgensen W.L. ; Chandrasekhar, J. ; Madura, J.D. ; Impey, R.W. ; Klein, M.L. *J. Chem. Phys.* **1983**, 79, 926.

conformations des molécules. Enfin, par rapport à des résultats de calculs *ab-initio*, elles étaient surévaluées pour des molécules non polaires. Les charges RESP sont mieux estimées, puisqu'elles sont extraites à partir de plusieurs conformations.

Le potentiel électrostatique est déterminé à partir d'une base 6-31G\* à un niveau Hartree-Fock (HF). La base STO-3G donne un moment dipolaire égal ou légèrement inférieur par rapport au moment dipolaire obtenu en phase gazeuse, mais elle sous-évalue les moments quadripolaires. Dans le but d'utiliser des modèles de solvation explicite, la base 6-31G\* est indispensable.

Dans un premier temps, la molécule ou les fragments de la molécule sont optimisés avec la base STO-3G à un niveau de théorie HF en utilisant le programme Gaussian 94<sup>82</sup>, puis dans un deuxième temps, le champ électrostatique est calculé avec la base 6-31G\*. Pendant cette dernière étape, les mots clés suivants sont ajoutés : Pop=mk (méthode de Merz-Kollman<sup>83</sup>), iop(6/33)=2.

Une fois obtenu le potentiel électrostatique de chaque conformère ou résidu, on procède à la détermination des charges atomiques RESP. Le premier pas est de convertir le fichier initial au format Gaussian en format compréhensible pour le programme RESP. Ce convertisseur est disponible à la page Internet d'AMBER (cf. Annexe). Le programme RESP s'effectue ensuite en deux étapes. A la première étape on évalue toutes les charges atomiques, et à la deuxième on affine les charges des atomes externes, surtout les hydrogènes, ceci en fixant la charge des atomes internes. Par exemple, on effectue une restriction de symétrie pour les hydrogènes des groupes méthyles de façon à obtenir la même charge atomique pour chaque hydrogène d'un groupe méthyle. On peut aussi effectuer une restriction de conformation, où l'on considère que toutes les conformations introduites doivent avoir les

---

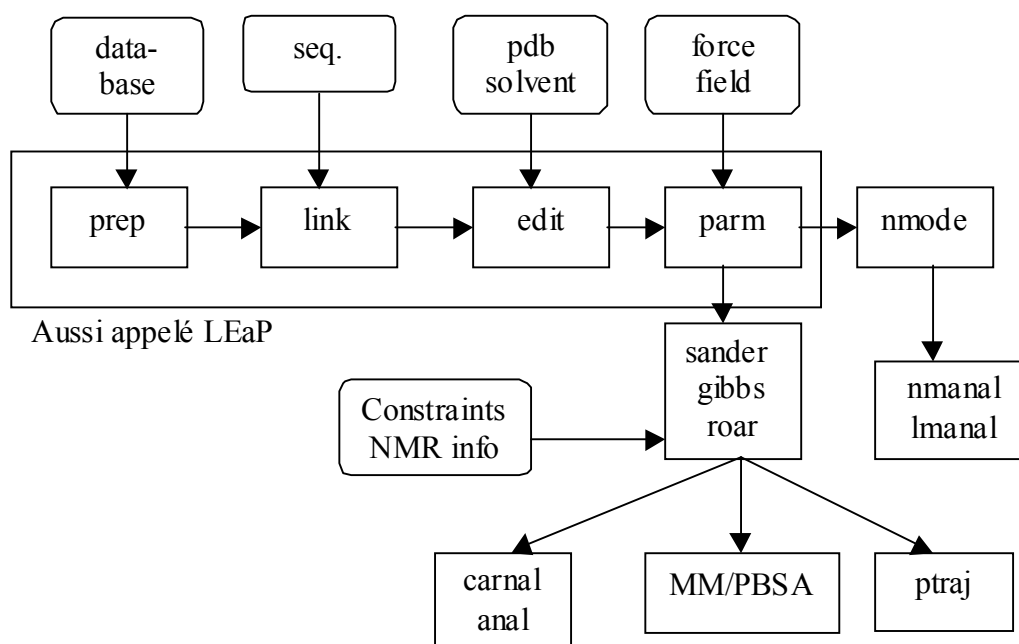
82 - Frisch, M.J. ; Trucks, G.W. ; Schlegel, H.B. ; Gill, P.M.W. ; Johnson, B.G. ; Robb, M.A. ; Cheeseman, J.R. ; Keith, T.A. ; Petersson, G.A. ; Montgomery, J.A. ; Raghavachari, K. ; Al-laham M.A. ; Zakrzewski, V.G. ; Ortiz, J.V. ; Foresman, J.B. ; Ciosowski, J. ; Stefanov, B.B. ; Nanayakkara, A. ; Challacombe, M. ; Peng, C.Y. ; Ayala, P.Y. ; Chen, W. ; Wong, M.W. ; Andres, J.L. ; Replogle, E.S. ; Gomperts, R. ; Martin, R.L. ; Fox, D.J. ; Binkley, J.S. ; Defrees, D.J. ; Baker, J. ; Stewart, J.P. ; Head-Gordon, M. ; Gonzales, C. ;Pople, J.A. Gaussian 94, 1995, Gaussian, Inc., Pittsburgh PA.

83 - Besler, B.H. ; Merz, K.M. ; Kollman, P.A. *J. Comp. Chem.* **1990**, *11*, 431.

mêmes charges atomiques en fin de processus. On introduit aussi la valeur de la charge totale (zéro pour des molécules neutres).

### 3.3.2.2 Structure du programme.

Le programme AMBER est composé de plusieurs modules et chacun d'eux a une fonction bien déterminée. Aucun module ne contient d'interface graphique. Sous AMBER **tout** est à écrire en ligne de code. Chaque module comporte de nombreux paramètres. La structure du programme est résumée sur la figure suivante :



Dans la version 6 (fin 1999), les sous-programmes prep, link, edit et parm sont regroupés dans le sous-programme Leap.

PREP : crée ou ajoute un résidu de la base de données selon l'information de la topologie. Les types d'atomes sont ajoutés à ce stade. Il existe sous AMBER de nombreux choix de types d'atomes: 13 types d'atomes de carbone différents, 12 types d'atomes d'hydrogène, 7 type d'atomes d'azote. Les charges atomiques sont ajoutées dans ce sous-programme.

LINK : lie tous les résidus entre eux selon l'ordre spécifié et selon la topologie précédente.



EDIT : construit la molécule selon le fichier au format pdb (*protein data bank*) et à partir du fichier de topologie précédemment créé. Ici les coordonnées moléculaires sont entrées. Les molécules de solvant et les contre-ions sont ajoutés.

PARM : liste à partir du fichier de paramètres du champ de force, les données des liaisons, angles etc. Les paramètres non inclus dans le champ de force AMBER sont ajoutés à ce niveau.

SANDER : programme principal de minimisation et de dynamique moléculaire.

GIBBS : programme de perturbation d'énergie libre.

ROAR : programme permettant de calculer une partie du système (le cœur) en mécanique quantique et le reste en mécanique moléculaire.

ANAL, CARNAL, PTRAJ : programmes d'analyse des coordonnées internes et des trajectoires générales pour SANDER, GIBBS et ROAR.

MMPBSA (version 6) : analyse les énergies des instantanées (*snapshots*) obtenus en MD en utilisant un modèle de solvation implicite permettant d'estimer les énergies libres d'interactions des changements conformationnels d'un ligand avec le récepteur.

### 3.3.2.3 AMBER\* sous MacroModel.

Pendant nos travaux de recherche, nous avons utilisé le programme MacroModel v5.0<sup>84</sup> avec le champ de force AMBER\* (parm94.dat). Ce champ de force est légèrement différent de l'original. Un des principaux changements est l'absence de recherche de charges atomiques à effectuer puisque, sous MacroModel, les charges sont déjà préassignées. Le calcul de l'énergie électrostatique du champ de force de MacroModel est basé sur l'utilisation des dipôles.

De plus, AMBER original utilise deux équations différentes pour le calcul de l'énergie de torsion ( $V_n$ ) lorsque  $n=2$  et  $n=3$ , alors que MacroModel utilise la même équation.

$$E = \sum_{angles} \sum_{\eta} \frac{V_{\eta}}{2} [1 + \cos(\eta \Phi - \gamma)]$$

---

84 - Mohmadi, F. ; Richards, N.G.J. ; Guida, W.C. ; Liskamp, R. ; Lipton, M. ; Caufield, C. ; Chang, G. ; Hendrickson, T. ; Still, W.C. *J. Comput. Chem.* **1990**, *11*, 440.

AMBER résout l'équation du calcul de l'énergie de torsion uniquement avec des valeurs de potentiels de torsion positives. Cependant il ajuste la position du maxima et du minima avec un *offset* angulaire (gamma). AMBER\* utilise des valeurs négatives et positives pour calculer cette équation.

Par exemple : pour  $V_3/2$ , AMBER utilise  $n=3$  et un gamma de 0, Macromodel utilise  $V_n/2$  ; AMBER utilise  $n=3$  et un gamma de 180, Macromodel utilise  $-V_n/2$ .

Et aussi, pour  $V_2/2$ , AMBER utilise  $n=2$  et un gamma de 0, Macromodel utilise  $-V_n/2$  ; AMBER utilise  $n=2$  et un gamma de 180, Macromodel utilise  $V_n/2$ .