

II-1. DESCRIPTION DE LA MÉTHODE

II-1-1. Introduction

Avec l'arrivée de méthodes théoriques de calcul de plus en plus sophistiquées et les ressources de computation plus accessibles, la *chimie informatique* est maintenant devenue un outil de plus en plus utile à la fois pour l'industrie et le milieu académique [1].

La modélisation par ordinateur d'une molécule implique généralement une présentation graphique de la géométrie ou de la configuration des atomes de la molécule, suivie de l'application d'une méthode théorique [2].

La modélisation moléculaire est un terme général qui englobe différentes techniques de graphisme moléculaire et de chimie computationnelle permettant d'afficher, simuler, analyser, calculer et stocker les propriétés des molécules [3].

La modélisation moléculaire (Mécanique et Dynamique Moléculaire) est une méthode qui permet la détermination de la structure et de l'énergie d'entités moléculaires.

II-1-2. Base de la chimie quantique

II-1-2-1. Equation de Schrödinger

La mécanique quantique postule la dualité onde-particule de la matière et définit une particule comme une onde c'est à dire distribuée à travers l'espace plutôt que possédant une trajectoire bien définie.

L'évolution au cours du temps t du système de N particules avec leurs positions symbolisées par r_1, r_2, \dots, r_N , décrit par sa fonction d'onde $\Psi(r_1, r_2, \dots, r_N, t)$, est déterminée par l'équation suivante:

$$\hat{H}\Psi(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N, t) = -i\hbar \frac{\partial \Psi(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N, t)}{\partial t} \quad \text{(II-1)}$$

qui est appelée équation de Schrödinger dépendante du temps, avec H l'Hamiltonien du système.

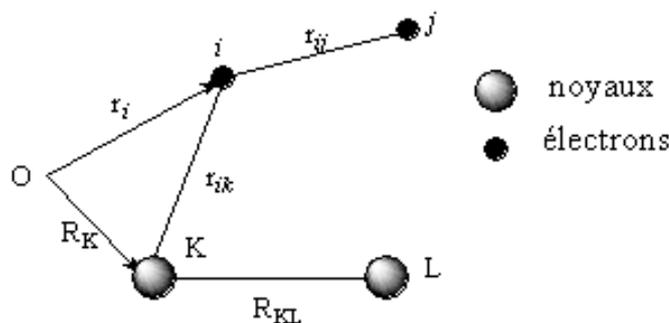
Donc :

$$\hat{H}\Psi = E\Psi \quad (\text{II-2})$$

où E est l'énergie associée à la fonction d'onde $\Psi(r_1, r_2, \dots, r_N)$, obtenue en résolvant l'équation de Schrödinger indépendante du temps. [4]

II-1-2-2. Hamiltonien moléculaire :

Soit un système constitué de N noyaux de numéro atomique Z_K , de masse M_K et de n électrons de charge e et de masse m . Leurs positions peuvent être repérées par des vecteurs R_K et r_i relatifs à une origine quelconque O .



L'Hamiltonien \hat{H} de ce système s'écrit sous la forme d'une somme d'opérateurs associés aux divers termes de l'énergie cinétique et potentielle.

$$\hat{H} = -\sum_{K=1}^N \frac{\hbar}{2M_K} \Delta(K) - \sum_{i=1}^n \frac{\hbar}{2m_i} \Delta(i) + \frac{1}{4\pi\epsilon_0} \left\{ \sum_{K<L}^N \frac{eZ_K Z_L}{R_{KL}} + \sum_{i,K=1}^{n,N} \frac{-e^2}{r_{iK}} + \sum_{j>i}^n \frac{e^2}{r_{ij}} \right\} \quad (\text{II-3})$$

On pose :

- énergie cinétique des noyaux :

$$\hat{T}_N = -\sum_{K=1}^N \frac{\hbar}{2M_K} \Delta(K) \quad (\text{II-4})$$

- énergie cinétique des électrons :

$$\hat{T}_e = -\sum_{i=1}^n \frac{\hbar^2}{2m_i} \Delta(i) \quad (\text{II-5})$$

Les énergies électrostatiques font intervenir les distances entre les particules, qui peuvent être exprimées en fonction des R_K et des r_i mais qui seront écrites selon des notations plus simplifiées :

- interaction noyaux-noyaux :

$$\hat{V}_{NN} = \frac{1}{4\pi\epsilon_0} \left\{ \sum_{K<L}^N \frac{eZ_K Z_L}{R_{KL}} \right\} \quad (\text{II-7})$$

- interaction noyaux-électrons :

$$\hat{V}_{Ne} = \frac{1}{4\pi\epsilon_0} \left\{ \sum_{i,K=1}^{n,N} \frac{-eZ_K}{r_{iK}} \right\} \quad (\text{II-8})$$

-interaction électrons-électrons :

$$\hat{V}_{ee} = \frac{1}{4\pi\epsilon_0} \left\{ \sum_{i<j}^n \frac{e}{r_{ij}} \right\} \quad (\text{II-9})$$

Comme la masse des électrons est très petite "environ 1800 fois moins pour l'hydrogène" devant celle des noyaux, On peut considérer en première approximation, que la répartition des électrons change pratiquement instantanément, lors des mouvements relatifs des noyaux, au cours des réactions chimiques ou des vibrations moléculaires. La fonction d'onde électronique peut donc être calculée, à un instant donné en considérant que les noyaux sont immobiles "approximation classique de Born-Oppenheimer". Il en résulte que :

$$\hat{T}_N = 0$$

et

$$\hat{V}_{NN} = \text{constante}$$

L'Hamiltonien devient donc :

$$\hat{H}' = \hat{T}_e + \hat{V}_{Ne} + \hat{V}_{ee} + \hat{V}_{NN} \quad (\text{II-10})$$

On pose :

$$\hat{H}_{el} = \hat{T}_e + \hat{V}_{Ne} + \hat{V}_{ee} \quad (\text{II-11})$$

où \hat{H}_{el} est l'Hamiltonien électronique. Il suffit donc de résoudre l'équation aux valeurs propres :

$$\hat{H}_{el}|\psi\rangle = E_{el}|\psi\rangle \quad (\text{II-12})$$

On aura alors :

$$\hat{H}_{el}|\psi\rangle = (E_{el} + V_{NN})|\psi\rangle \quad (\text{II-13})$$

L'énergie totale est obtenue en ajoutant à E_{el} le terme de répulsion nucléaire V_{NN} .

II-1-2-3. Fonction d'onde électronique

La résolution de l'équation aux valeurs propre, impose impérativement une fonction d'onde qui doit satisfaire le principe de Pauli ; pour cela, elle doit inclure le spin de l'électron, sous la forme de spin-orbitale. Sa détermination exigera en outre des approximations qui lui imposent une forme a priori : l'approximation orbitale, et le développement en une combinaison linéaire d'un nombre limité d'orbitales atomiques (LCAO) [5].

II-1-3.Comparaison sommaire des différentes méthodes.

Les principales méthodes qui permettent de déterminer la géométrie d'une molécule sont les méthodes de diffractions, les méthodes spectroscopiques et les techniques de calcul. Les informations que ces méthodes peuvent apporter concernant la structure moléculaire sont de deux sortes : *qualitatives* (forme générale et symétrie d'une molécule) et/ou *quantitatives* (positions relatives de tous les atomes dans une molécule, angles et longueurs des liaisons et angles de torsion).

Parmi les méthodes dont dispose le chimiste théoricien, les méthodes quantiques et les méthodes dites de champ de force ou "**MÉCANIQUE MOLÉCULAIRE**" qui conduisent à la connaissance de la structure **3D** des molécules.

La mécanique moléculaire est basée sur un champ de force empirique. C'est une méthode non quantique. En effet le système électronique n'est pas considéré explicitement mais les effets sont pris en compte indirectement. Les atomes sont assimilés à des boules reliées par des tiges ou ressorts. Une déformation de la structure entraîne un changement de l'énergie potentielle du système[1].

Les méthodes quantiques les plus simples sont : La méthode de **Hückel**, qui ne prend en considération que les électrons (π) et la méthode de Hückel étendue (E.H.T) laquelle prend en considération tous les électrons de valences, ces deux méthodes sont purement qualitatives.

Les méthodes semi-empiriques (ex: **CNDO, INDO, MNDO, PM3, AM1, ...**) ne considèrent que les électrons de valence et néglige certaines intégrales ou les remplacent par une paramétrisation adaptée.

Les méthodes quantiques les plus sophistiquées sont les méthodes dites ab-initio et la DFT qui tiennent compte de tous les électrons mais qui souffrent cependant de certaines limitations, en particulier les résultats dépendent de la base choisie (définition de la fonction d'onde) et nécessitent un temps de calcul très long, même pour des molécules de taille moyenne [6].

Quel que soit la méthode quantique utilisée la différence entre le temps d'un calcul par Mécanique Moléculaire et par une méthode quantique est de l'ordre de quelques puissances de dix. Cette différence énorme augmente en fonction de la taille de la molécule ; ceci explique pourquoi les industriels s'y intéressent [7], elle permet, donc de passer en revue des séries de grosses molécules, des produits pharmaceutiques ou des colorants. La Mécanique Moléculaire permet de faire un **tri** avant de passer au stade expérimental [8].

II-2. REPRÉSENTATION DES MÉTHODES DE CALCUL.

II-2-1. Mécanique moléculaire

II-2-1-1. Introduction

La mécanique moléculaire est une méthode appelée méthode de champ de force, ce dernier est un modèle mathématique basé sur les principes fondamentaux de la spectroscopie vibrationnelle ; l'énergie potentielle y est qualifiée d'énergie stérique ; pour une molécule donnée, sa valeur dépend du champ de force adopté.

Cette énergie est exprimée sous la forme d'une somme de contributions bidimensionnelles, associées aux écarts de la structure par rapport aux paramètres de références (longueur de liaison, angle de valence,..) et d'une somme de contributions 3D caractéristiques d'une énergie stérique (énergie de Van der Waals, énergie coulombienne,).

L'énergie stérique minimale calculée est la somme des contributions apportées par les énergies d'élongation, de flexion, de torsion, de Van der Waals, électrostatique,.... [9]

$$E(\text{stérique}) = E(\text{totale}) = E(L) + E(\theta) + E(\phi) + E(\text{VdW}) + E(e),\dots \quad (\text{II-14})$$

L'expression 'champ de force' recouvre l'ensemble des fonctions ainsi que les paramètres associés aux différents types d'atomes qu'elles contiennent. Les paramètres ne sont pas des constantes de force et sont caractéristiques d'un champ de force et non d'une molécule, ils vérifient donc le principe dit de "transférabilité" sur lequel repose l'intérêt de cette méthode. Le but est de construire un modèle mathématique qui décrive un ensemble de molécules [10].

Ce modèle mathématique est établi pour représenter aussi précisément que possible les variations de l'énergie potentielle avec la géométrie moléculaire.

II-2-1-2. Energie d'élongation

Les liaisons entre les atomes dans un édifice moléculaire ont souvent tendance à s'allonger ou à se contracter (figure II-1).

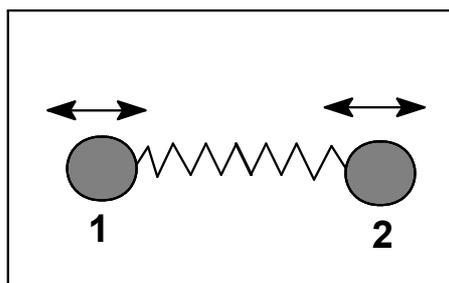


Figure II-1 : Elongation entre deux atomes.

Cette déformation est régie en première approximation par loi de "Hooke" d'élongation des ressorts. On peut ainsi lui associer une énergie d'élongation de la forme:

$$E(L) = 1/2[K_1(L-L_0)^2] \quad (\text{II-15})$$

Où K_1 : est la constante d'élongation ou constante de Hooke

L_0 : la longueur de la liaison de référence.

L : la longueur de la liaison dans le modèle.

Tous ces termes d'élongation sont sommés sur l'ensemble des liaisons de la molécule. Un terme cubique $(L-L_0)^3$ est généralement ajouté pour des déformations importantes. Le calcul de cette énergie impose donc de connaître au minimum les deux paramètres indissociables $\{K_1 \text{ et } L_0\}$ qui représentent un sous-ensemble du champ de force. En effet il ressort du développement en série de l'expression mathématique de la courbe de Morse traduisant l'interaction existante, entre deux atomes en fonction de leur distance respective[1].

II-2-1-3. Energie de flexion

La fluctuation des atomes autour de leur position d'équilibre engendre une déformation des angles de valence (figure II-2) [1].

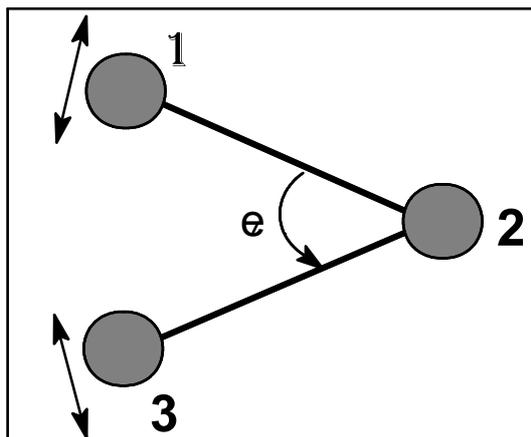


Figure II-2 : Déformation des angles de valence.

Ce phénomène est régi par une énergie de flexion qui peut s'exprimer sous les mêmes formes que précédemment à savoir, pour la plus simple :

$$E(\theta) = 1/2[K_f(\theta - \theta_0)^2] \quad (\text{II-16})$$

K_f : constante de flexion

θ_0 : angle de valence de référence

θ : angle de valence dans la molécule

Le couple $\{ K_f, \theta_0 \}$ représente ici encore un sous-ensemble du champ de force.

II-2-1-1-4. *Energie de torsion*

Elle concerne l'angle dièdre formé par les atomes 1-2-3-4. Elle rend notamment compte de la structure 3D de la molécule (figure II-3) [1].

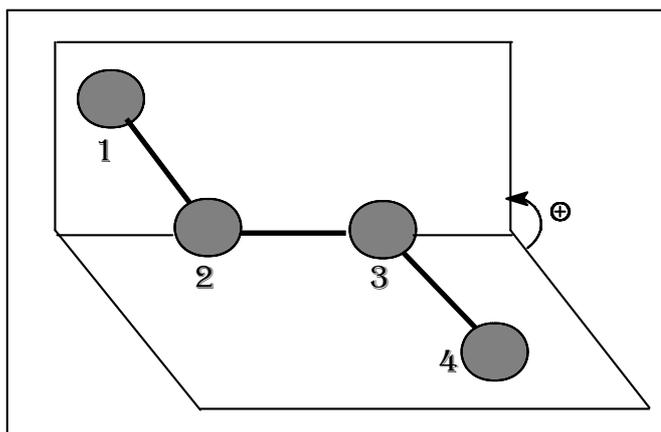


Figure II-3 : Angle dièdre formé par les atomes 1-2-3-4.

L'évaluation de cette énergie $E(\phi)$ se fait par une fonction développée en série de Fourier.

$$E(\phi) = 1/2 [V_1(1+\cos\phi)+V_2(1 - \cos2\phi)+V_3(1+\cos3\phi)] \quad (\text{II-17})$$

L'angle dièdre (ϕ) définit la torsion autour de la liaison 2-3.

V_1, V_2, V_3 sont les constantes du potentiel de l'énergie de torsion.

II-2-1-1-5. Energie de Van der Waals

Cette énergie concerne les atomes non liés entre eux et non liés à un atome commun [9]. Elle se compose de deux parties, l'une répulsive et l'autre attractive, et peut être exprimée par l'équation suivante[1] :

$$E(\text{vdw})= \epsilon^*[- C_1(r^*/r)^6 +C_2 \exp(-C_3(r/r^*))] \quad (\text{II-18})$$

ϵ^* : paramètre d'énergie qui caractérise la profondeur du puits de potentiel a la distance r^* , appelé aussi "dureté".

r^* : somme des rayons de VdW des atomes interagissants.

r : distance interatomique.

C_1, C_2, C_3 : constantes du champ de force.

On peut donc représenter cette énergie en fonction de la distance interatomique "r" de la manière suivante:

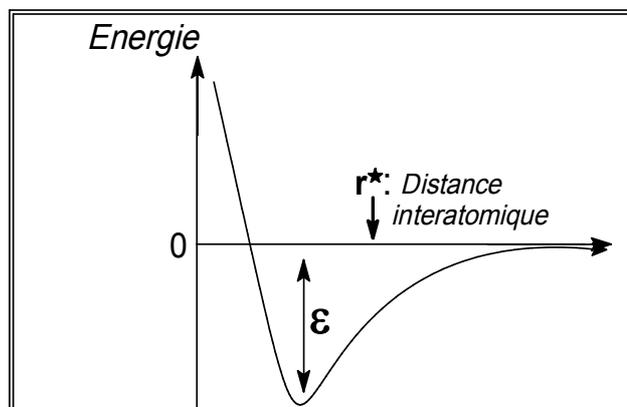


Figure II-4: Courbe d'énergie de Van der Waals

II-2-1-1-6. Energie électrostatique

Les interactions électrostatiques peuvent, dans certains cas, prendre une importance considérable, en particulier dans le cas des molécules comportant deux ou plusieurs hétéroatomes Meyer et al [11] ont proposé d'introduire un terme électrostatique même pour les hydrocarbures. IL peut s'exprimer à partir des charges atomiques ou des moments dipolaires de chaque liaison [1].

Dans le premier cas :

$$E(e) = \sum q_1 q_2 / D \cdot d_{12} \quad (\text{II-19})$$

D: constante diélectrique locale du milieu

q₁q₂: charges partielles atomiques des atomes 1,2

d₁₂: distance interatomique.

Dans le deuxième :

$$E(e) = \mu_1 \mu_2 (\cos X - 3 \cos \alpha_1 \cdot \cos \alpha_2) / D \cdot r_{12}^3 \quad (\text{II-20})$$

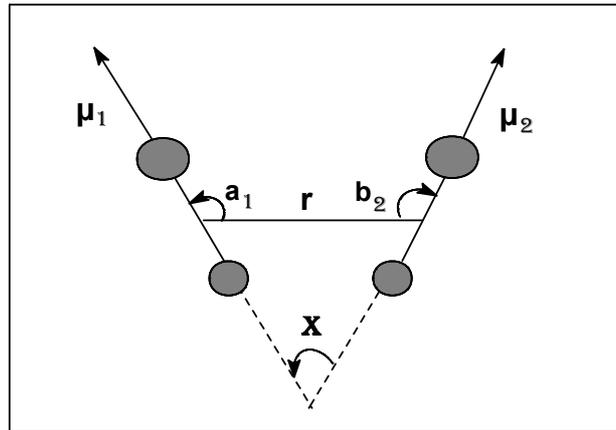


Figure II-5 : Interactions électrostatiques entre deux atomes

r : distance entre les milieux des deux liaisons

μ_1, μ_2 : représentent respectivement les moments dipolaires des deux liaisons.

X : l'angle entre les deux vecteurs moments.

a_1, b_2 : angles formés respectivement entre μ_1 et r et μ_2 et r

II-2-1-1-7. Adjonction de termes croisés

L'ensemble des termes énergétiques répertoriés jusqu'ici représente des mouvements simples des atomes dans la molécule. Pour tenir compte des effets de couplages entre deux mouvements non indépendants, il est d'usage de rajouter un ou plusieurs termes supplémentaires d'interaction [12]. Les plus couramment formulés sont[1] :

Le terme d'élongation-flexion $E(L, \theta)$ (figure II-6) qui s'exprime par

$$E(L, \theta) = K_{lf}(L - L_0)(\theta - \theta_0) \quad (\text{II-21})$$

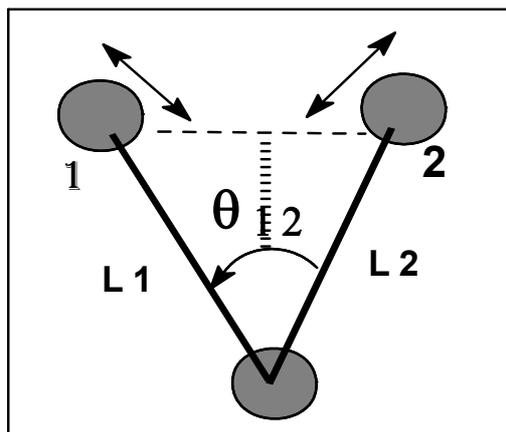


Figure II-6 : Couplage élongation-flexion

et le terme de flexion-torsion $E(\phi, \theta)$ qui s'exprime par:

$$E(\phi, \theta) = K_{ft} (\theta - \theta_0)(\phi - \phi_0). \quad (\text{II-22})$$

II-2-2. Dynamique moléculaire

II-2-2-1. Introduction

Nous avons utilisé la dynamique moléculaire et les simulations de Monte Carlo, pour la recherche conformationnelle et pour raffiner les structures obtenues par RX. La dynamique moléculaire est une méthode d'étude des mouvements et de l'évolution de la configuration spatiale des systèmes moléculaires. En pratique, on opère par résolution des équations classiques du mouvement de Newton, étant donné une fonction énergie potentielle et son champ de force associé.

Le mouvement des atomes est inhérent à tout système chimique par suite des vibrations atomiques dépendant de la température. Au niveau moléculaire, cet effet se traduit par une compression/extension des liaisons ou par la variation des angles de liaison ou de torsion.

II-2-2-2 Principes de la dynamique moléculaire.

La dynamique moléculaire est une technique de modélisation sur ordinateur par laquelle l'évolution en fonction du temps ou trajectoire, d'une molécule est décrite par les

principes de la mécanique classique Newtonienne [13].

En dynamique moléculaire, on essaie de simuler les mouvements intramoléculaires que l'on peut visualiser ensuite en temps réel. Ces mouvements correspondent à des vibrations autour d'un minimum ou au passage d'un minimum à un autre minimum d'énergie. Si l'énergie fournie virtuellement au système est suffisamment élevée, des barrières énergétiques importantes peuvent être franchies.

L'avantage majeur de la dynamique moléculaire est de conduire à des conformations qui dépendent moins de la structure initiale que celles obtenues par mécanique moléculaire. Depuis 1990, l'accès à des moyens informatiques de plus en plus puissants a permis de développer des méthodes de calcul de dynamique moléculaire qui ont un double but :

a) Simuler les mouvements intramoléculaires et si possible aboutir à une description thermodynamique du système étudié (entropie, énergie libre, ...).

b) Optimiser les structures d'une façon plus efficace en évitant plus facilement les minima multiples en particulier lorsqu'on introduit dans le calcul un recuit simulé. L'échelle des temps actuelle est environ de 10^{-14} à 10^{-10} s par simulation. D'autre part, la dynamique moléculaire simule les mouvements instantanés et présents d'un système moléculaire. Pour cela, chaque atome est traité comme une particule obéissant à la loi d'action de masse de Newton ;

$$\vec{\Sigma F} = m\vec{a} \quad (\text{II-23})$$

Des intégrations successives de cette équation en fonction du temps conduit à une **trajectoire** de l'atome en fonction du temps sous la forme d'une série de positions et de vitesses dans l'espace. L'analyse est effectuée par périodes de temps de l'ordre de 1-100 ps. Plusieurs mouvements sont alors effectués durant ces périodes.

Les simulations de dynamique moléculaire introduisent donc la dimension du temps dans l'étude des molécules. Cette évolution temporelle (trajectoire) d'une molécule est décrite par la résolution des équations de mouvement de Newton [14].

La molécule est donc perçue comme une entité dynamique dont les atomes modifient leurs positions spatiales au cours du temps. Ces changements sont influencés par leur énergie cinétique ($1/2k_B T_o$ par degré de liberté, si la simulation a été réalisée à la température T_o) et la résultante des forces qui s'exercent sur eux par les atomes environnants. Au temps t , chaque atome i , de vecteur position \vec{x}_i et de masse m_i subira une accélération \vec{a}_i telle que selon la loi de Newton :

$$\vec{\Sigma F} = \vec{m} \vec{a} = \frac{m d\vec{V}}{dt} = \frac{m d^2 \vec{r}}{dt^2} \quad (\text{II-24})$$

\vec{m} = masse des atomes

\vec{a} = accélération (2ième dérivée)

t = temps

\vec{r} = coordonnées cartésiennes de l'atome

Ainsi une molécule se comporte comme une structure dynamique ayant des coordonnées atomiques qui vont changer en fonction du temps à cause de leur énergie cinétique et des forces exercées par les atomes environnants :

$$\frac{d^2 \vec{x}_i}{dt^2} = \vec{a}_i = \frac{\vec{F}_i}{m_i} ; \vec{F}_i = \frac{d\vec{V}}{d\vec{x}_i} \quad (\text{II-25})$$

V = fonction de l'énergie potentielle totale d'interaction auquel est soumise la molécule

On peut alors intégrer les équations de mouvement et obtenir une **trajectoire** pour chaque atome en fonction du temps. Pour cela on utilise l'algorithme de précision de 2e ordre de Verlet (en troncature de la série de Taylor). Sa validité est aisément testée à l'aide de la 3^{ième} loi fondamentale de la thermodynamique qui implique que l'énergie totale d'un système doit être conservée au cours d'une simulation.

En dynamique moléculaire on explore une surface conformationnelle, c'est-à-dire des points dans tout l'hyperespace. Cependant, le résultat n'est pas nécessairement réaliste car les

structures obtenues ne sont pas minimisées. On a donc une image de l'ensemble en fonction du temps « time dépendent Picture ».

II-2-2-1-3. Applications de la dynamique moléculaire

Une application importante de la dynamique moléculaire est l'analyse des modes normaux de vibration le long de la trajectoire. Une autre application est l'optimisation et le raffinement des structures 3D d'après les données de la cristallographie et/ou de la RMN. La mise en œuvre de cette méthode requiert néanmoins des moyens de calcul particulièrement puissant (CRAY par exemple) et elle est coûteuse en temps et en argent. Elle se généralise cependant pour les études de peptides et de petites protéines [15].

En effet, un des problèmes de fond de la technique de dynamique moléculaire est l'inadéquation entre l'échelle de temps des phénomènes physiques observés et celle des simulations accessibles par la technologie actuelle. L'une des plus longues simulations (en absence de solvant est à 298 K) a durée 300 ps et a été réalisée sur la myoglobine (1423 atomes). Tandis que les échelles de temps des processus dynamiques observés varient de la femtoseconde (e.g. l'élongation des liaisons covalentes) à la seconde, voire plus (e.g. la vibration des cycles aromatiques des résidus d'acides aminés).

II-2-3. Simulations de Monte Carlo

II-2-3-1. Introduction

Nous avons utilisé les simulations de Monte Carlo pour la recherche multiconformationnelle. En effet, les méthodes de Monte Carlo consistent à des simulations de problèmes mathématiques dans lesquelles des nombres aléatoires sont utilisés pour trouver une solution qui peut ne pas être aléatoire [16].

Historiquement, le développement de cette méthode coïncide avec l'élaboration de la bombe atomique. En effet, la diffusion aléatoire de neutrons dans un matériau fossile était un problème majeur qui allait être soumis avec succès à l'expérimentation numérique.

Vers 1948, Métropolis applique cette méthode à la chimie théorique et obtenaient des estimations des valeurs propres de l'équation de Schrödinger par une méthode de Monte Carlo [17].

II-2-3-2. Principes de base des simulations de Monte Carlo

Cette procédure de Monte Carlo est une méthode **stochastique** qui consiste à générer un enchaînement de conformations où les propriétés thermodynamiques et structurales sont obtenues par moyennage.

Il faut générer un très grand nombre de conformations ou de configurations. Une distribution de **Boltzmann** du type $\exp(-\Delta E_i/kT)$ permet ce moyennage. Contrairement à une minimisation d'énergie, l'approche Monte Carlo ne trouve pas un minimum d'énergie mais échantillonne un ensemble d'états moléculaires avec des états énergétiques croissants avec la température.

Mais les états trop hauts en énergie auront peu de contributions sur la stabilité du système car si l'énergie du système augmente, le terme de Boltzmann devient petit. L'algorithme qui effectue cet échantillonnage est celui de **Métropolis**.

Par opposition à la dynamique moléculaire, en Monte Carlo on peut prendre des sauts plus grands pour optimiser. De plus, on peut effectuer des pas différents pour les rotations, les translations, les coordonnées internes, et optimiser chacun indépendamment. La dynamique moléculaire est plutôt un processus **déterministique**. Il n'y a pas d'aspect aléatoire dans les calculs. Tout est déterminé au point initial de départ et on laisse évoluer le système.

L'algorithme Métropolis [17] sert à déterminer si la nouvelle configuration était retenue ou non. La probabilité que la nouvelle structure soit gardée est son facteur de Boltzmann, qui vaut :

$$B = \exp(-\Delta E_i/kT) \quad (\text{II-26})$$

Autrement dit, un nombre i de l'intervalle $[0,1]$, tiré au hasard, est comparé au facteur de Boltzmann B . Si i est inférieur à B , la nouvelle configuration sera acceptée. Sinon, elle sera rejetée et l'ancienne configuration sera de nouveau soumise à une autre perturbation élémentaire, en résumé l'algorithme **Métropolis** :

1. sélection au hasard d'un atome et déplacement au hasard de Δx , Δy , Δz
2. calcul du changement ΔE_{pot} . après le déplacement de l'atome
3. si $\Delta E < 0 \Rightarrow$ accepte nouvelle conformation
4. si $\Delta E > 0 \Rightarrow$ choisie un nombre i au hasard entre $[0,1]$
 - a) si $\exp(\Delta E/kT) < 1 \Rightarrow$ accepte la nouvelle conformation
 - b) si $\exp(\Delta E/kT) > 1 \Rightarrow$ garde la conformation originale et retourne à (1.)

L'ensemble des configurations retenues est ainsi progressivement construit et forme une **chaîne de Markov**. C'est une façon naturelle de faire une recherche de l'espace conformationnel. Les structures obtenues peuvent être le point de départ pour une minimisation future. Si sur un essai, le déplacement d'un atome est dans la même direction que le lien covalent, la surface énergétique sera "accentuée" et il y aura une probabilité élevée que ce déplacement ne sera pas accepté.

II-2-3-3. Analyse conformationnelle par la Monte Carlo

Dans une première approximation, on tient compte uniquement des forces intramoléculaires pour calculer les propriétés conformationnelles d'une molécule. Une nouvelle approche maintenant consiste à développer des fonctions générant des conformations multiples ("multiconformer search"). Plutôt que de se baser sur une technique de recherche systématique ("systematic search"), la méthode d'analyse utilise l'approche **Monte Carlo (MC)**.

Elle est basée sur une évaluation progressive des angles de torsion ("torsion search method"), développée par Still et Lipton [18] . Cette approche est plus rapide car elle n'optimise que les angles dièdres plutôt que chaque angle de liaison. Ce sont les angles dièdres qui fluctuent le plus lors d'une minimisation.

Rappelons que contrairement à une minimisation d'énergie, la méthode Monte Carlo ne cherche pas une énergie minimum mais fait plutôt un échantillonnage d'un **ensemble** d'état conformationnel, sondant des états énergétiques de plus haute énergie en augmentant la température.

II-2-3-4. Quelques Applications de la Monte Carlo

a) Molécules cycliques

Cette technique ne semble pas appropriée à l'étude des molécules cycliques. En effet, pour une molécule linéaire, les combinaisons d'angles dièdres possibles peuvent varier à l'infini. A l'inverse, pour les systèmes cycliques, seules certaines combinaisons fermeront le cycle [19].

b) Evaluation de l'énergie de solvation

La meilleure méthode est d'entourer le soluté de plusieurs centaines de molécules de solvant et d'utiliser soit la dynamique moléculaire, soit la méthode Monte Carlo. Ainsi on offre une fraction des configurations du solvant disponible afin d'obtenir une moyenne des énergies de solvation. Pour estimer la polarisation, il faut une méthode de perturbation d'énergie libre. On utilise le programme **OPLS** (efficace en solvation mais pas en analyse conformationnelle).

II-2-4. Méthode quantique : Ab initio

Les méthodes Ab initio sont caractérisées par l'introduction d'une base arbitraire pour étendre les orbitales moléculaires et alors le calcul explicite toutes les intégrales exigées qui impliquent cette base.

Les calculs Ab initio peuvent être exécutés au niveau d'approximation de Hartree-Fock, qui est équivalent à un calcul du champ auto – cohérent SCF (*Self Consistent Field*). L'option et Les niveaux Hartree-Fock inclut les effets de corrélation qui n'est pas incluse au niveau d'approximation d'Hartree-Fock d'une solution non - relativiste pour l'équation de Schrödinger.

II-2-4-1. Méthode Hartree-Fock

L'approximation orbitale décrit la fonction d'onde à n électrons comme un produit de n fonctions d'espace et de spin associées à chaque électron (ou spin-orbitales). Ceci est valable pour un modèle de particules indépendantes où le terme de répulsion inter-électronique de l'Hamiltonien est omis. L'équation de Schrödinger à n électrons peut être séparée en n équations à un électron. Cependant la fonction d'onde à n électrons décrite de cette manière ne satisfait pas au principe de Pauli (les électrons sont des fermions), car elle n'est pas asymétrique sous une permutation de deux électrons. Ce problème est levé par la description de la fonction d'onde comme un déterminant de Slater construit à partir des n spin-orbitales (les n spin-orbitales Ψ proviennent des $n/2$ orbitales spatiales combinées avec deux fonctions de spin α et β).

$$\Psi = \frac{1}{\sqrt{(2n)!}} \begin{vmatrix} \psi_1(1) & \psi_1(2) & \dots & \psi_1(n) \\ \dots & \dots & \dots & \dots \\ \psi_n(1) & \dots & \dots & \psi_n(n) \end{vmatrix} \quad (\text{II-27})$$

De par la méthode variationnelle, le choix des meilleures spin-orbitales correspond au minimum de l'énergie électronique (E^e). Cette fonction d'onde électronique optimale, atteinte pour l'énergie la plus basse, sera toujours au-dessus de l'énergie exacte. La procédure de minimisation de l'énergie est une procédure auto-cohérente avec des contraintes d'orthonormalité pour les orbitales. Les valeurs propres de cette équation sont les énergies mono-électroniques associées aux fonctions d'ondes qui sont, dans ce cas, les orbitales.

$$H^{\text{HF}} \phi_i = \varepsilon_i \phi_i \quad (\text{II-28})$$

Une solution exacte des équations HF, avec l'obtention des orbitales atomiques, est possible pour les atomes comprenant un noyau et un électron. Des développements supplémentaires sont nécessaires pour résoudre l'équation de Schrödinger pour des systèmes poly-électroniques. La méthode de Roothaan est basée sur la méthode LCAO ("linear combination of atomic orbitals") qui décrit les orbitales moléculaires en terme de combinaison linéaire

d'orbitales atomiques. Les orbitales atomiques utilisées pour représenter les orbitales moléculaires constituent ce qui est appelé les "bases " du système; plus ces bases sont étendues plus les orbitales moléculaires seront proches de celles obtenues en résolvant directement les équations HF, avec la limite HF pour des bases de dimensions infinies. La carence majeure de la méthode HF est le manque de corrélation électronique.

II-2-4-2. Méthode post-Hartree-Fock

Les méthodes les plus importantes, dans l'utilisation courante, pour introduire la corrélation électronique sont d'une part les méthodes appelées post-HF avec l'interaction de configuration (CI), les méthodes de "coupled cluster" (CC) et les "many-body perturbation theory" (MP2, MP4, ...) et, d'autre part, les méthodes qui dérivent de la théorie de la fonctionnelle de la densité (DFT).

La résolution des équations de HF donne une fonction d'onde de référence sous la forme d'un déterminant de Slater.

La fonction d'onde de la méthode CI est une combinaison linéaire de déterminants de Slater représentant l'état fondamental et des configurations excitées. Ces dernières correspondent à l'excitation d'un, de deux ou plus électrons d'une orbitale occupée à une virtuelle. Les coefficients de cette combinaison linéaire sont déterminés selon le principe variationnel; l'énergie qui en découle est donc une limite supérieure à l'énergie exacte. L'IC totale ("Full CI") est la limite que l'on peut atteindre dans une base donnée, c'est à dire qu'elle comprend toutes les excitations possibles des n électrons. Cependant l'IC totale accroît énormément les calculs avec le nombre d'électrons et la dimension des bases utilisées. Pour ces raisons de tels calculs servent habituellement de référence pour des systèmes comprenant un petit nombre d'électrons ($n \leq 20$). La méthode CI limitée à un nombre d'excitation inférieure à la totalité des possibilités ("truncated CI") n'est pas "size consistent" (c'est à dire, par exemple, que l'énergie d'un dimère placé à distance infinie n'est pas égale à deux fois l'énergie du monomère). Cette propriété importante est garantie par des méthodes non variationnelles comme "many-body perturbation theory" ou "coupled cluster methods".

Dans le formalisme de Moller-Plesset, l'hamiltonien est représenté comme la somme de l'hamiltonien HF de l'état fondamental et d'une différence, entre l'hamiltonien exact et HF, traitée comme une perturbation. Dans la théorie de perturbation, la fonction d'onde et l'énergie, pour un état donné, donnent l'état appelé "zéro" ou fondamental auquel est ajouté des corrections successives résultant des différents ordres de perturbation pour le traitement du système. L'approximation MP2 comprend des substitutions simples et doubles, la théorie des perturbations d'ordre 4 (MP4) ajoute des substitutions triples et quadruples.

Les méthodes "coupled cluster" (CC) sont actuellement les plus puissantes des méthodes ab initio. La théorie CC commence par un postulat sur la fonction d'onde à n électrons à savoir que cette fonction d'onde est multipliée par une exponentielle naturelle d'une somme d'opérateurs d'excitation d'électrons (T); T2 implique une double substitution. CCSDT comprend jusqu'à un opérateur de triple excitation.

II-2-4-3. Théorie de la fonctionnelle de la densité (DFT)

La théorie de la fonctionnelle de la densité (DFT) s'est beaucoup développée ces dernières années. Dans cette approche l'énergie de l'état fondamental d'un système est une fonctionnelle d'une densité électronique tridimensionnelle. L'application du principe variationnel donne les équations appelées équations de Kohn-Sham qui sont similaires aux équations de Hartree-Fock. En principe, il suffit de remplacer la contribution d'échange de l'opérateur de Fock par un potentiel d'échange et de corrélation qui correspond à la dérivation de la fonctionnelle d'énergie d'échange et de corrélation par rapport à la densité. Le point crucial en DFT est que l'énergie d'échange et de corrélation n'est pas connue de façon exacte. Néanmoins les formules approchées pour cette énergie donnent des résultats qui sont comparables ou meilleurs que ceux donnés par MP2 à un moindre coût de ressource informatique.

Les premières approximations de la DFT sont similaires à celles appliquées aux méthodes HF. L'équation de Schrödinger est non-dépendante du temps et non-relativiste. A partir de l'approximation de Born-Oppenheimer le formalisme et les approximations divergent [4].

HyperChem (7.0) généralement exécute les calculs d'ab initio (SCF). Il peut aussi calculer l'énergie de la corrélation (peut être ajouté à l'énergie totale par la méthode SCF) par la procédure d'option Hartree-Fock, appelé MP2 qui font le calcul de Møller-Plesset au second ordre de perturbation. La procédure MP2 est disponible pour les calculs de 'single point'. L'énergie de corrélation MP2, peut être ajouté à l'énergie totale (SCF) à cette configuration du 'single point'.

II-2-5. Méthodes semi-empiriques

II-2-5-1. Introduction.

La mécanique quantique est une technique mathématique rigoureuse basée sur l'équation de Schrödinger. La solution de cette équation permet d'obtenir des informations précises sur les propriétés géométriques et électroniques de la molécule. Les calculs peuvent être de type *ab initio* ou semi-empirique (ex : *CNDO*, *PM3*). En *ab initio* on tient compte de tous les électrons de la molécule et on vise une solution rigoureuse de l'Hamiltonien [20].

Les calculs semi-empiriques traitent seulement les électrons de valence et utilisent un Hamiltonien plus simple ayant des facteurs de correction basés sur des données expérimentales. L'équation de Schrödinger d'un système moléculaire peut être résolue sans approximation (*ab initio*) ou en introduisant des approximations (*semi-empirique*).

II-2-5-2. Méthode semi-empirique (CNDO)

En mécanique quantique on se préoccupe de la distribution des électrons (orbitales) dans l'espace. Les meilleurs programmes comportent des processus d'optimisation de la géométrie [21].

L'objectif de la mécanique quantique est principalement de déterminer l'énergie et la distribution électronique. Ainsi les énergies moléculaires sont calculées en utilisant l'équation de Schrödinger avec le formalisme des orbitales moléculaire (MO). Par les méthodes MO, les énergies peuvent être extraites via le formalisme SCF ("*Self Consistent Field*") ou par une méthode de perturbation comme PCILO.

On utilise un potentiel Hartree-Fock qui consiste à manipuler en jeu d'orbitales sans paramétrisation. Le calcul s'arrête lorsque la convergence est jugée satisfaisante. On dit que le processus en cours est auto-cohérent ("Self-Consistent Field").

Pour les systèmes polyélectroniques, l'équation de Schrödinger n'est pas résolue car on traite ici d'un problème à N-corps. Il faut donc aussi faire des approximations orbitales en tenant compte de chaque électron de façon indépendante. On parle alors d'un développement linéaire de combinaisons d'orbitales atomiques ("LCAO") pour chaque électron. Notons qu'en mécanique moléculaire, les calculs sont plus rapides (~10X) car l'énergie obtenue provient d'une sommation simple de chacune des coordonnées atomiques de la molécule sans passer par une diagonalisation matricielle.

En mécanique quantique, on étudie la distribution électronique (orbitale) dans l'espace. Les propriétés moléculaires qui peuvent être calculées sont la géométrie, les spectres de vibration, les spectres d'absorption UV, les moments dipolaires, les spectres électroniques et les fonctions responsables de la réactivité chimique (Orbital HOMO, LUMO).

La méthode CNDO : Complete Neglecting of Differential Overlapping (Négligence complète du recouvrement différentiel) a été développée par Pople en vue d'étudier la structure électronique et spatiale des molécules.

Contrairement à la mécanique moléculaire, ici on est moins dépendant des paramètres empiriques et spécifiques. Les programmes commerciaux de mécanique quantique sont excellents et donnent d'excellents résultats, mais les faiblesses sont:

- a) la méthode est très exigeante en temps de calcul et demande de gros ordinateurs
- b) actuellement, la mécanique quantique est valable que pour des systèmes de moins de 40 atomes par *ab initio* et environ 100 atomes par des calculs semi-empiriques.

II-2-5-3. Méthode semi-empirique (PM3)

PM3 est une méthode semi-empirique (SCF) pour les calculs chimiques ; il est une paramétrisation de la méthode AM1.

PM3 et AM1 généralement sont les méthodes les plus rigoureuses dans le logiciel HyperChem (7.0), PM3 a été paramétré pour beaucoup des éléments principaux des groupes et quelques métaux de transition. PM3 est différent d'AM1 seulement dans les valeurs des paramètres.

II-2-6. Méthodes de corrélation quantitative (QSAR)

II-2-6-1. Introduction.

Le modèle QSAR (*Quantitative Structure-Activity Relationships*) est une corrélation basée sur un calcul statistique, de paramètres physico-chimiques avec l'activité biologique [22].

La comparaison de l'activité biologique de certaines molécules et de leur structure a permis d'établir dans de nombreux cas des corrélations entre les paramètres structuraux et les propriétés d'une molécule. L'association des variations de l'activité aux paramètres structuraux permet d'obtenir un système d'équations qui donne, *pour une série chimique donnée et pour une activité définie, une équation de corrélation.*

L'intérêt essentiel de cette équation est qu'elle doit permettre de déterminer les valeurs des paramètres qui correspondent à une activité maximale et ainsi de prédire l'activité des molécules qui n'ont pas encore été synthétisées [23].

La validité d'un modèle QSAR dépendra donc du choix que l'on aura fait sur les paramètres. Il est ainsi judicieux d'identifier et d'évaluer quantitativement les paramètres les plus pertinents à utiliser en fonction de l'activité et des molécules sélectionnées.

Les principaux paramètres entrant en jeu dans l'activité d'une molécule sont :

II-2-6-2. Volume et surface moléculaire :

Le volume moléculaire et la surface moléculaire sont défini par détermination du volume (ou la surface) occupé par l'enveloppe de Van der Waals de la molécule considérée.

le volume est défini par la relation

$$V = \frac{MW}{d} \quad (II-29)$$

où:

MW: est la masse moléculaire.

d: est la densité.

II-2-6-3. L'énergie d'hydratation :

L'énergie d'hydratation est le facteur clé, déterminant la stabilité des différentes conformations moléculaires.

Les liaisons hydrogènes constituent un cas particulier d'une interaction dipôle – dipôle. Les groupements donneurs sont : F-H, O-H, N-H, Cl-H, P-H, et parfois C-H . Les groupement accepteurs doivent obligatoirement porter un doublet disponible : F, O, N, mais aussi Cl, S et P et les entités chargées telles que l'anion carboxylate , par exemple .

L'énergie de liaison hydrogène est généralement de 2 à 10 Kcal/mol et les longueurs des liaisons hydrogène sont intermédiaires, elles sont plus longues que les liaisons covalentes mais plus courtes que les liaisons de Van der Waals.

L'union de deux moléculaires par une liaison hydrogène constitue un cas particulier, de ce que l'on nomme les complexes donneurs-accepteurs qui sont établis par attraction électrostatique entre une molécule riche en électrons et une molécule accepteuse d'électrons [24].

En effet dans les milieux biologiques les molécules polaires, s'entourent de molécules d'eau. Il s'établit des liaisons hydrogène entre une molécule d'eau et ces molécules. Les sites donneurs de proton interagissent avec l'atome d'oxygène de l'eau et les sites accepteurs de proton avec l'atome d'hydrogène.

Les premiers correspondent au complexe avec la plus forte liaison hydrogène. Ces molécules hydratées se déshydratent au moins partiellement avant et lors de leur interaction.

Ces interactions de faible énergie, que l'on observe notamment entre messagers et récepteurs, sont généralement réversibles [25].

II-2-6-4. Paramètre de lipophilie.

Le caractère lipophile d'une molécule traduit souvent des propriétés biologiques comme le métabolisme, la distribution dans les tissus, la liaison avec le site récepteur... . Une variation de la lipophilie entraîne ainsi une modification de la concentration de la molécule sur le site récepteur.

La lipophilie est par conséquent une *variable importante* dans les équations de corrélation utilisées pour prédire l'activité biologique (QSAR).

Pour évaluer les effets imputables à la lipophilie, on détermine le **coefficient de partage** appelé également coefficient de partition P.

Le partage d'une molécule entre une phase aqueuse et une phase lipidique conditionne en partie ses propriétés biologiques telles que le transport, le passage à travers les membranes, la biodisponibilité (distribution et accumulation), l'affinité pour un récepteur et la fixation par une protéine, l'activité pharmacologique ou encore la toxicité. S'agissant de contaminants, ce même partage conditionne leur devenir dans notre environnement en particulier leur accumulation dans les organismes aquatiques ; le partage est donc une propriété physico-chimique importante qui peut être utilisée pour représenter la nature lipophile ou hydrophile d'une molécule.

Le coefficient de partage P d'une molécule dans un système biphasique constitué de deux solvants non-miscibles (le plus souvent le système *n*-octanol/eau), est reconnu pour sa faculté à mimer le passage de cette molécule à travers les membranes biologiques. Pour des solutions diluées, ce coefficient de partage *n*-octanol/eau est le rapport de la concentration d'une molécule de soluté dans le *n*-octanol sur sa concentration dans l'eau lorsque le système biphasique est en équilibre. [26]

Plus communément le logarithme du coefficient de partage, Log P, dans un mélange 1-octanol/eau, en utilisant l'une des méthodes suivantes :

Méthode de Hansch (méthode semi-empirique):

$$\log P_{RX} = \log P_{RH} + \Pi_X + \Pi_{CORR} \quad (II-30)$$

où:

P_{RX} et P_{RH} : représentent respectivement les coefficients de partage des molécules R_H et R_X dans un système n-octanol/eau.

Π_X : est le paramètre lipophile du substituant X .

Π_{CORR} : est un terme correctif tenant compte de l'effet entraîné par les ramifications, les doubles liaisons, les fermetures de cycle, les liaisons intramoléculaire, les repliements de molécule

Méthode de Rekker (méthode fragmentale)

$$\log P = \sum_n a_n F_n + F_{CORR} \quad (II-31)$$

où:

a_n : nombre de fragments identiques dans la molécule.

f_n : constante fragmentale hydrophobe.

f_{CORR} : terme correctif multiple entier d'une *constante magique* (0,289) et décrivant des caractéristiques structurales spécifiques (atomes d'hydrogène attachés à des groupes polaires, conjugaison aryl-aryl, proximité de groupes polaires, ...).

Il peut être intéressant de remarquer qu'une propriété moléculaire aussi complexe que la lipophilie est, encore aujourd'hui, calculée principalement sur la base de valeurs de fragments ou d'atomes déterminées de façon empirique et en faisant abstraction des processus impliqués dans le phénomène de solvatation et des propriétés moléculaires qui les affectent. [26]

II-2-6-5. Polarisabilité :

On appelle polarisabilité, la facilité avec laquelle un nuage électronique se déforme sous l'influence d'un champ électrique. La molécule subit une certaine déformation et acquiert un moment électrique dipolaire induit proportionnel au champ E [27].

$$P_{(e)} = \epsilon_0 \alpha E \quad (II-32)$$

Où :

$P_{(e)}$: Coefficient de polarisabilité.

ϵ_0 : Constante diélectrique .

α : Moment électrique dipolaire induit .

II-2-6-6. Paramètres associés à la topologie 1D

Paramètre de corrélation des forces de dispersion des substituants dans les réactions biochimiques: la réfractivité moléculaire qui est définie par la relation suivante [10] :

$$MR = \frac{MW}{d} \frac{n^2 - 1}{n^2 + 1} \quad (II-33)$$

où:

n: est l'indice de réfraction

MW: est la masse moléculaire

d: représente la densité

L'indice de réfraction ne variant pas énormément pour la plupart des composés organiques, la réfraction molaire dépend essentiellement du volume.

II-3. MÉTHODE DE RECHERCHE DES MINIMA

II-3-1. Introduction

La minimisation donne parfois des hydrogènes et des doublets disposés autour de l'atome dans des positions structurales impossibles. Ceci est souvent dû à un mouvement initial inapproprié de ces atomes légers quand une structure fortement distordue est introduite. Cela peut également se produire si on démarre avec une structure planaire. C'est pourquoi le programme effectue une "seconde passe" de minimisation après avoir repositionné les atomes légers.

Une autre difficulté de la minimisation concerne le problème du minimum local. Les routines d'optimisation sous contrainte, ont en effet la fâcheuse tendance de trouver un minimum d'énergie le plus proche de la structure d'entrée [28].

En général, les distances et les angles de liaison sont correctement minimisés, si bien que le problème du minimum local peut se résumer à l'optimisation des angles dièdres (il faut beaucoup plus d'énergie pour déformer une liaison ou un angle de valence par rapport à un angle dièdre). Les approches suivantes ont pour but d'extraire la molécule de son puits de potentiel.

Presque toutes les méthodes de minimisation ont au moins un point en commun : on commence en un endroit donné de l'hypersurface et on descend vers le minimum le plus proche, sans savoir si ce minimum est local ou absolu. On doit donc présenter à l'ordinateur plusieurs conformations de départ, sous forme de coordonnées internes, en s'inspirant de modèles moléculaires[1].

II-3-2. Algorithmes de minimisation

Pour une molécule comprenant N atomes, la fonction à minimiser comprend donc 3N variables. Une telle fonction comprend en général un minimum global et des minima locaux. A partir de la géométrie initiale, on recherche le jeu de coordonnées cartésiennes qui réduit à son minimum la somme de toutes les contributions énergétiques.

En principe, il suffit de prendre la dérivée première de l'énergie stérique par rapport à chacun des degrés de liberté de la molécule et de trouver l'endroit sur l'hypersurface énergétique où, pour chaque coordonnée r_i , $(dE/dr_i) = 0$.

Les procédures pour atteindre ce but sont de deux types: Les unes utilisent uniquement la pente de la surface (dérivée première), les autres, à la fois cette pente et la courbure de la surface (les dérivées première et seconde).

II-3-2-1. La méthode de "steepest descent"

Le premier programme de minimisation pouvant effectuer une optimisation de géométrie est due à Wiberg [29] et utilise la méthode de la plus grande pente (steepest-descent). Après avoir calculé l'énergie correspondant à une géométrie initiale, on déplace chaque atome individuellement selon ses trois coordonnées cartésiennes et l'on recalcule l'énergie après chaque déplacement.

Ceci revient à calculer la dérivée première uniquement. Ensuite on déplace tous les atomes sur une distance qui dépend de (dE/dri) , et ainsi de suite. Cet algorithme suivra donc la direction imposée par les forces interatomiques dominantes.

C'est pourquoi il se révèle très efficace pour supprimer les mauvais contacts ou les principaux problèmes stéréochimiques qui existent dans les coordonnées brutes d'une structure cristalline ou modélisée, tout en perturbant très peu cette dernière.

Cependant cette méthode aléatoire est généralement longue vers la fin de chaque cycle de minimisation et la convergence devient très lente au-delà des premiers cycles (phénomènes oscillants, remontée d'énergie).

En effet, la méthode consiste à rechercher la direction de plus grande pente au cours de laquelle la fonction objective $F(x)$ décroît le plus rapidement. La direction suivie sera celle indiquée par l'opposé du gradient d'énergie, c'est-à-dire la direction de la plus grande pente de la fonction d'énergie, qui est la direction dans laquelle l'énergie diminue le plus vite, du moins localement.

II-3-2-2. La méthode du gradient conjugué.

Cette méthode, fondée sur le même principe que la précédente (direction opposée au gradient d'énergie), prend également en compte les étapes précédentes, afin de déterminer plus finement la direction et le pas. Pour une surface d'énergie quadratique, fonction de $3N$ variables, cette méthode converge en $3N$ pas [30].

Elle conserve une bonne efficacité, mais est plus coûteuse en temps de calcul (un facteur 2 par rapport à "steepest descent"). Le pas est ajusté à chaque cycle pour obtenir la meilleure diminution d'énergie. L'intérêt de cet algorithme est d'éviter un comportement oscillatoire autour du minimum et d'accélérer la convergence. Il se révèle cependant moins efficace ou même inutilisable (pas de convergence) pour des structures qui présentent beaucoup de mauvais contacts, telles que les structures moyennées sur la trajectoire d'une dynamique moléculaire.

II-3-2-3. La méthode Newton-Raphson.

Une amélioration supplémentaire de la convergence peut encore être obtenue en ayant recours à une approximation quadratique Q de la fonction F, obtenue par développement en série de Taylor [31].

La méthode consiste à chercher à chaque pas le minimum du développement à l'ordre 2 de la fonction F. Cette méthode dite de "Newton-Raphson", a recours aux dérivées secondes. Maintenant on fait plutôt appel à cette technique d'optimisation. Elle évalue les dérivées secondes de l'énergie moléculaire par rapport aux paramètres géométriques et converge donc plus rapidement.

En résumé, l'optimisation complète selon la méthode Newton-Raphson demande de calculer la matrice complète des dérivées secondes (matrice des constantes de force). Cette matrice contient $3N \times 3N$ éléments pour une molécule de N atomes. Les dimensions $3N$ correspondent aux trois degrés de mouvement pour chaque atome.

L'optimisation de la géométrie requière cependant seulement $3N-6$ degrés de liberté puisque les trois translations et les trois rotations ne sont pas accompagnées par des changements d'énergie. Les constantes de forces sont alors manipulées sous la forme d'une matrice de dérivées secondes.

II-3-2-4. La méthode du recuit simulé.

Les méthodes que nous venons de décrire ont la particularité de faire décroître à chaque pas la fonction F ; ces méthodes ne peuvent donc pas échapper au minimum local proche de la structure de départ, et ont par conséquent, un rayon de convergence toujours

restreint. La méthode de *recuit simulé* “*anneal*”, développée par Kirkpatrick [32], autorise la fonction F à augmenter momentanément afin de franchir des barrières d'énergie pour retomber dans un minimum plus profond.

Le franchissement de ces barrières permet d'aller au-delà des minima locaux au voisinage de la structure initiale pour explorer de façon plus extensive l'espace conformationnel accessible, afin de découvrir des minima plus profonds et plus éloignés de la structure initiale que les minima locaux.

En conclusion, un minimiseur est une boîte noire mathématique qui demeure un outil important pour la recherche d'un minimum d'énergie. En général, on utilise plusieurs minimiseurs. On passe à un deuxième ou un troisième si ça ne converge pas assez vite. Tout dépend aussi du nombre de variables ou du nombre de variations d'angle de liaison introduit.

II-4.DOMAINE D'APPLICATION DE LA MODÉLISATION MOLÉCULAIRE

II-4-1. Introduction

Les études théoriques s'orientent actuellement vers la conceptualisation rationnelle “Rational design” qui signifie que la connaissance des relations entre les propriétés et la structure moléculaire des molécules connues permet au scientifique de développer des nouvelles molécules, avec une assez bonne anticipation tel que le produit final aura un ensemble sur mesure de propriétés prédéfinies [33] .

La mécanique moléculaire est la méthode de base pour les calculs de modélisation moléculaire destinée aux grosses moléculaires, elle s'est largement développée pour permettre de déterminer la conformation des molécules, leurs propriétés thermodynamiques et leurs spectres vibrationnels. C'est une méthode empirique résultant de l'ajustement (filtrage) de résultats expérimentaux sur des fonctions mathématiques simples. D'un certain point de vue, on peut considérer la mécanique moléculaire comme un système expert constitué d'un modèle mathématique et d'un jeu de paramètres permettant de représenter différentes molécules. Cet ensemble (potentiel/fonction mathématique) est appelé champ de force.

Il faut remarquer que la mécanique moléculaire ne saurait être un outil isolé ; par exemple on peut faire quelques calculs ab-initio sur des systèmes très simplifiés et intégrer les résultats dans un champ de force de la mécanique moléculaire ; la MM peut aussi constituer une base de départ pour un calcul quantique ultérieur, certains chercheurs ont en effet réalisé le couplage de la Mécanique Moléculaire et la Mécanique Quantique [34] . L'association des deux méthodes a permis alors d'étendre le domaine d'application à des systèmes conjugués et à ceux comportant plusieurs hétéroatomes.

On peut diviser les applications de la MM en plusieurs catégories selon les différentes informations obtenues [1] :

II-4-2. Etude de structures

Très souvent la Mécanique Moléculaire est utilisée pour obtenir une géométrie à laquelle on attache de l'intérêt. Elle peut en particulier guider l'interprétation des résultats provenant des études de structure par rayons X et diffraction électronique ou bien de les raffiner. Elle est prête aussi à servir et à modéliser une molécule pour les besoins de l'infographie[1].

II-4-3. Interprétation de la réactivité

La MM peut permettre l'interprétation des effets stériques sur la réactivité. On peut distinguer dans ce cas deux méthodes d'exploitation :

- 1) Les comparaisons dans un ensemble de molécules similaires pour caractériser des relations entre les changements de structure ou des substituants et la réactivité.
- 2) Les comparaisons intramoléculaires pour prévoir une voie de réarrangement, une régiosélectivité ou une stéréosélectivité.

II-4-4. Analyse conformationnelle

Quand aucune liaison n'est rompue, ni formée et qu'aucun intermédiaire chargé n'intervient, l'interconversion conformationnelle se prête particulièrement bien à une description par la MM. On peut obtenir grâce à cette analyse des informations structurales

sous forme d'un profil énergétique (en fonction d'un angle dièdre par exemple) ou des cartes énergétiques 3D.

La recherche multiconformère consiste à trouver un ensemble de conformères. Deux types de méthodes sont possibles: stochastique ou déterministe. Parmi les méthodes stochastiques citons par exemple la dynamique Brownienne, ou Monte Carlo Métropolis. Parmi les méthodes déterministes, il existe en particulier des méthodes de déformations spatiales séquentielles et des méthodes de déformations spatiales temporelles sous le nom générique de D.M. (Dynamique moléculaire).

De toutes ces méthodes il faut retenir qu'aucune d'elles ne permet à coup sûr d'obtenir toutes les conformations réellement existantes dans une fenêtre énergétique donnée quels que soit les durées offertes pour le calcul. Généralement on utilise plusieurs méthodes en même temps pour être plus sûr d'avoir toutes les conformations les plus probables.

Un ensemble de conformères issu de ces calculs peut être scindé en familles conformationnelles caractérisées par une spécificité géométrique donnée et une énergie moyenne [1].

On utilise, la distribution de Boltzmann [35] , pour l'examen du taux de population de chaque type conformationnel.

$$P_i = \frac{\exp(-E_i / kT)}{\sum_i^n g_i \exp(-E_i / kT)} \quad (\text{II-34})$$

avec $P_i = \frac{n_i}{n_t}$

P_i : la population du niveau i d'énergie E_i parmi n conformations.

n_i : nombre de conformations d'énergie E_i .

n_t : nombre totale de conformations.

k : constante de Boltzmann.

T : température en degré K.

II-5. GRANDEURS DERIVES DE L'ENERGIE STERIQUE

Nous exposons dans cette partie la détermination, à partir de l'énergie stérique totale, de deux autres grandeurs: l'enthalpie de formation qui constitue un point de repère expérimental très important et l'énergie de tension indispensable pour la comparaison des molécules dissemblables [36].

II-5-1. Enthalpie de formation

Elle s'écrit :

$$\Delta H^{\circ f} = E_{\text{tot}} + E(\mathbf{l}) \quad (\text{II-35})$$

$E(\mathbf{l})$: énergie de liaison donnée par la somme d'incrément a_x associés à chaque type de liaison, ce qui est équivalent à chaque type de groupe ($\text{CH}_3, \text{CH}_2, \text{CH}, \dots$);

Les incréments sont déterminés par comparaison entre les $\Delta H^{\circ f}$ expérimentales et les $E(\text{tot})$ calculées des molécules ayant servi à élaborer le champ de force [36] .

$$E(\mathbf{l}) = a_{\text{CH}_3} \cdot n_{\text{CH}_3} + a_{\text{CH}_2} \cdot n_{\text{CH}_2} + a_{\text{CH}} \cdot n_{\text{CH}} + \dots \quad (\text{II-36})$$

n_{CH_3} : nombre de groupes CH_3

II-5-2. Energie de tension

L'énergie stérique peut servir à comparer les stabilités thermodynamiques d'isomères de conformation ou d'isomères géométriques ou encore celles d'autres molécules isologues comportant le même nombre d'unités $\text{CH}_3, \text{CH}_2, \text{CH}$ etc....).

En revanche pour toute autre comparaison il faut avoir recours à une autre grandeur, "l'énergie de tension" définie par l'équation [36] :

$$E(\text{Ten}) = \Delta H^{\circ f} - E_{\text{LST}} \quad (\text{II-37})$$

ELST : est l'énergie de liaison "sans tension", qui est exprimée de la même façon que $E(l)$.

$$\mathbf{ELST} = \mathbf{b}_{\text{CH}_3} \cdot \mathbf{n}_{\text{CH}_3} + \mathbf{b}_{\text{CH}_2} \cdot \mathbf{n}_{\text{CH}_2} + \mathbf{b}_{\text{CH}} \cdot \mathbf{n}_{\text{CH}} + \dots \quad (\text{II-38})$$

Les incréments \mathbf{b} sont calculés pour que $\mathbf{E}(\mathbf{Ten})$ soit égale à zéro pour un certain nombre de molécules de référence.

Ce qui nous intéresse n'est pas la valeur absolue de l'énergie de tension mais la différence entre les valeurs relatives à deux ou plusieurs composés qui se ressemblent.

II-6. LIMITATION DE LA MODÉLISATION MOLÉCULAIRE

La Modélisation Moléculaire s'adresse surtout à des organiciens intéressés par des problèmes de réactivité et de structure de molécules comportant déjà un nombre significatif d'atomes, elle s'adresse aussi au biochimistes et pharmaciens préoccupés par la relation structure-activité [1].

Si l'on veut exploiter intelligemment les programmes disponibles pour le calcul et la visualisation, certains principes de base doivent être retenus, il est nécessaire de connaître les origines de la méthode, ses potentialités et ses limites.

Cette méthode empirique, ne s'applique bien que lorsqu'on étudie des molécules voisines de celles qui ont servi à établir le champ de force. Plus on sophistique le champ de force de la mécanique moléculaire ($MM2 \rightarrow MM3$) plus on a besoin de paramètres; il est difficile d'avoir un champ de force général et on s'oriente plutôt vers des champs de force spécifiques, sur les hydrocarbures conjugués [37], les protéines, les peptides [38] et les polymères,...etc. Enfin il faut toujours valider une étude en Modélisation Moléculaire par confrontation avec l'expérience (RX, RMN....) sur des molécules types.

II-7. PROGRAMMES UTILISES ET MATERIELS

Ce travail de ce mémoire a été effectué au sein du laboratoire de chimie informatique de Biskra.

L'investigation des différentes conformations préférentielles des macrocycles et le calcul structure-activité, a été réalisé par la modélisation moléculaire (mécanique moléculaire, Abinitio et QSAR), en utilisant le logiciel HyperChem (7.0) [39] dans un PC (MICRO PROCESSUS INTEL P IV, CPU 2.80 GHz, DISQUE DUR 40 G, 512 Mo de RAM).

Un calcul type de la pénicilline d'une première minimisation a duré quelques minutes par la mécanique moléculaire, alors que pour la recherche conformationnelle sur P.C environ deux semaines par Abinitio.

Pour faire la comparaison électronique du noyau de la pénicilline, le calcul par la méthode MM a duré 5 min, par la méthode PM3 a duré 20 min, par les méthodes ab initio 'minimum et Small' a duré 34 heures.

II-8. RÉFÉRENCES

1. S. Belaidi, *Thèse de doctorat, Université de Batna*, 2002.
 2. T. Clark, 'Handbook of Computational Chemistry', *Edition, Wiley*, London, 1985.
 3. P. Kollman, *Acc. Chem. Res.*, 1996, 29, 462.
 4. N. Vulliermet, *Thèse de doctorat, Université de Genève (Suisse)*, 2000.
 5. S. Meskaldji, *Thèse de doctorat, Université de Constantine*, 2008.
 6. W.C. Ripka & J.M. Blaney, 'Topics In Stereochemistry', *Eds .Elieil & Wilen*, 1991.
 7. J. S.Lomas, *act.chimique*, 1986, 5, 14.
 8. K. B. Lipkowitz et D. B. Boyd, "The Application of Quantitative Design Strategies in Pesticide Discovery". *Reviews in Computational Chemistry*, p. 119, 1990.
 9. N.L. Allinger, X. Zhou, J. Bergsma, *J. Mol. Structr. (Theochem)*, 1994, 312, 69.
 10. W.C. Still, 'Molecular Mechanics', *Columbia Univ. Press*, 1986.
 11. A.Y.Meyer and R.F.Forrest, *J.Am.Chem.Soc.* 1981, 103, 4664.
-

12. K.B. Lipkowitz, *J. Chem. Educ.* , 1995,72, 1070.
 13. J. Tirado-Rives & W.L. Jorgensen, *J. Am. Chem. Soc.* , 1990,112, 2773.
 14. M. Karplus & G.A. Petsko, *Nature*, 1990,347, 631
 15. J.A. McCammon & S.C. Harvey, "Dynamics of Proteins and Nucleic Acids",
Cambridge, Univ. Press, 1987.
 16. D.R. Ripoll & U.A. Scheraga, *J. Prot. Chem.*, 1989, 8, 263.
 17. N. Métropolis, A. Rosenbluth, M. Rosenbluth, A. Teller et E. Teller, *J. Chem. Phys.*,1953,21, 1087.
 18. Still et Lipton, *J. Comput. Chem.*, 1990 , 11, 440
 19. F. Chaplin, *Pour la Science* , Nov. 1985.
 20. J.L.Rivail "Elément de chimie quantique à l'usage des chimistes".
Inter Editions. Paris. 1994.
 21. J. A. Sordo, M. Probst, G. Corongiu, S. Chin et E. Clementi, *J.Am.Chem.Soc.*,
1987, 109, 1702 .
 22. V. N. Viswanadhan, M. R. Reddy, R.J. Bacquet et M.D. Erion, *Journal of computational chemistry*, 1993,14, 1019
 23. V. N. Viswanadhan , A. K. Ghose et J. N. Weinstein, *Biochimica et Biophysica Acta* , 1990, 1039, 356.
 24. Y. Cohen, "Pharmacologie moléculaire", *Masson*, paris, 1978.
 25. S. Belaidi, M. Omari, T. Lanez et A . Dibi, *J. Soc.Alger. Chim.*, 2004, 14, 27.
 26. A. Carpy, *ANALYSIS © EDP Sciences, Wiley-VCH*, 1999, 1 3, 27.
 27. B. Yavorski et A. Detlaf, "Aide-mémoire de physique",p 376, *Editions Mir, Moscou*,
1980.
 28. V.Brenner, *thèse de doctorat*, Université de Paris-Sud ,octobre, 1993
 29. K.B.Wiberg, *J.Am. chem. Soc.* 1965, 87,1070.
 30. W.H.Press, B.P.Flanning,S.A. Teukolsky et W.T. Vetterling," Numerical Recipes", *Cambridge U.P., New-York*,pp.274, 1986.
 31. A.Warshel and S.Lifson, *J.Chem.Phys*,1970, 53, 582.
 32. S. Kirkpatrick, C.D. Gellet, M.P. Vecchi, *Science*,1983, 220,671.
 33. H.R. Allcock, *Science*, 1992,225, 1106
 34. P.Kollmann, *Acc.Chem.Res.* 1985,18,105.
 35. J.Koca, *J.Mol.Structr.*1993, 291, 255.
 36. W. F.van Gunsteren et H. J. C. Berendsen, "Computer Simulation of
-

- Molecular Dynamics'' *Angew. Chem. International Edition England* , 1990.
37. N.L.Allinger, F.Li and L.Yan , *J.Am. chem. Soc.*1990,117, 868.
38. P.W. Schiller, et M. Goodman , *Biopolymers* , 1990, 29, 179.
39. HyperChem 7.0 (Molecular Modelling System) *Hypercube, Inc.*, 1115 NW 4th Street, Gainesville, FL 32601 ; USA, (2002).