

## **Cours 1 : Initiation à la modélisation moléculaire.**

### **(Les outils de la modélisation Moléculaire)**

#### **1. La modélisation moléculaire : Pour qui ? Pour quoi ?**

La modélisation moléculaire est un outil destiné aux chercheurs préoccupés par la structure et la réactivité des molécules

La connaissance de la structure des édifices moléculaires permet de comprendre ce qui est réalisé dans une transformation physique, chimique ou biologique. Elle peut permettre aussi de prévoir de telles transformations. La compréhension comme la prévision sont considérablement facilitées lorsque l'on peut **visualiser** les structures.

#### **2.Représenter une molécule sur l'écran**

Une molécule est correctement décrite par sa géométrie et ses propriétés thermodynamiques. La visualisation doit rendre compte de l'ensemble de ces caractéristiques. La question essentielle est **de représenter une molécule sur l'écran de la façon la plus proche possible de la "réalité"**. L'utilisation de l'informatique a permis de mettre au point un outil performant : **la modélisation moléculaire**.

**Modélisation** : Transformation d'un ensemble le plus grand possible d'observations expérimentales en un ensemble le plus petit possible de paramètres.

**Modèle** = Ensemble des paramètres et des fonctions mathématiques permettant une représentation simplifiée de la réalité

Le modèle (M) idéal serait :**Précis** ,**Général** et **Réel** .

***Modélisation moléculaire : Elaboration et application d'un modèle mathématique qui permet de représenter les molécules à l'échelle microscopique.***

La modélisation moléculaire permet au chimiste de mieux comprendre et prévoir les phénomènes, principalement à l'échelle moléculaire

### **3.La modélisation moléculaire : Comment ?**

#### **a) Principe**

Modéliser une molécule consiste à préciser, à partir de calculs, la position des atomes qui la constituent, dans l'espace et de calculer l'énergie de la structure ainsi engendrée. Une représentation "la plus proche possible de la réalité" correspondra à une structure de plus basse énergie (voir les détails dans les paragraphes suivants)

#### **b) Calculs**

Les utilisateurs de la modélisation moléculaire se divisent en deux groupes :

- ceux qui font des calculs relativement précis sur des petites molécules (environ 100 atomes)
- ceux qui cherchent par des méthodes plus approximatives à déterminer la structure des macromolécules.

Les méthodes de calculs utilisées répondent plus ou moins bien à ces deux types de préoccupation :

### **4.Méthodes de calculs de Modélisation Moléculaire**

Les méthodes de la modélisation moléculaire comprennent : les méthodes quantiques, la mécanique moléculaire et la dynamique moléculaire.

#### **4.1. La Mécanique Moléculaire**

Cette technique calcule l'énergie des atomes (et non plus des électrons) au moyen d'approximations semi-classiques. La simplification considérable des calculs qui en résulte permet de travailler sur des molécules de grande taille, telles que les macromolécules biologiques (protéines, acides nucléiques), ou sur des systèmes comportant un grand nombre de molécules.

Le calcul de l'énergie est basé sur la mécanique classique

- Champ de forces = fonctions et paramètres

#### **Avantage :**

1. Molécules de grande taille (macromolécules, protéines, ...)
2. Calcul des modes de vibration
3. Propriétés thermodynamiques
4. Dynamique moléculaire, Docking

#### **Inconvénients :** Impossible de casser des liaisons -> pas de réactivité

La mécanique Moléculaire voire les atomes et les liaisons comme :

- Atomes : sphères dures non interpénétrables
- Liaisons : ressorts (loi de Hooke)

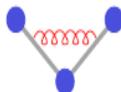
Donc l'énergie totale de système est calculée à partir des énergies mentionnées dans la figure suivante :

Elongation des liaisons



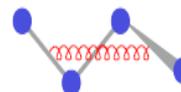
$$V_r = \sum_{i, \text{liaisons}} \frac{k_{r,i}}{2} (r_i - r_{i,0})^2$$

Déformation des angles de valence



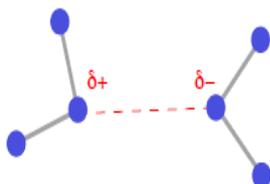
$$V_\theta = \sum_{i, \text{angles}} \frac{k_{\theta,i}}{2} (\theta_i - \theta_{i,0})^2$$

Torsion



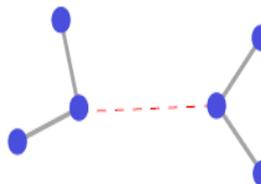
$$V_\tau = \sum_{n, \text{torsions}} \frac{V_n}{2} [1 + \cos(n\omega - \gamma)]$$

Electrostatique



$$V_{el} = \sum_{i=1}^N \sum_{j>i}^N \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}}$$

Interaction de van der Waals



$$V_{vdW} = \sum_{i=1}^N \sum_{j>i}^N \left( 4\epsilon_{ij} \left[ \left( \frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left( \frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] \right)$$

Energie totale d'une molécule :  $\mathcal{V}(\mathbf{r}) = V_r + V_\theta + V_\tau + V_{el} + V_{vdW}$

## 4.2. La dynamique moléculaire

Cette technique a pour but de calculer les mouvements des molécules au cours du temps, le plus souvent à partir des énergies de la mécanique moléculaire (partir de l'énergie potentielle calculée par la mécanique moléculaire), en appliquant les lois de la mécanique classique. Elle permet donc de calculer l'énergie cinétique.

### Principe du calcul

Le problème de la dynamique moléculaire consiste à calculer la position d'un atome  $i$  à l'instant  $t + \Delta t$  connaissant sa position à l'instant  $t$  :

$$\vec{r}_i(t) \rightarrow \vec{r}_i(t + \Delta t)$$

La méthode consiste à partir de l'énergie  $E_p$  calculée par la mécanique moléculaire pour déterminer successivement la force, l'accélération, la vitesse et la position de l'atome  $i$  :

$$E_p \rightarrow \vec{F}_i \rightarrow \vec{a}_i \rightarrow \vec{v}_i \rightarrow \vec{r}_i$$

### 4.3. Les méthodes quantiques (La chimie Quantique) (en utilise dans la suites de tous les chapitres )

Ces méthodes sont basées sur la résolution de l'équation de Schrödinger et le calcul des orbitales moléculaires (OM). Leur complexité augmente rapidement avec le nombre d'électrons.

$$\hat{H}\psi(r_1, r_2, \dots, r_n, R_1, R_2, \dots, R_N) = E\psi(r_1, r_2, \dots, r_n, R_1, R_2, \dots, R_N)$$

**H** : Hamiltonien totale d'une molécule comportant **N** noyaux et **n** électrons. Il est défini par la somme de cinq termes :

(cinétique des électrons, cinétique des noyaux, répulsion entre électrons, répulsion entre noyaux et le terme d'attraction électrons-noyaux).

$$H = -\frac{\hbar^2}{2m_e} \sum_i^n \Delta_i - \frac{\hbar^2}{2M_K} \sum_K^N \Delta_K + \sum_{i>j}^n \frac{e^2}{r_{ij}} + \sum_{K>L}^N \frac{Z_K Z_L e^2}{r_{KL}} - \sum_{K=1}^N \sum_{i=1}^n \frac{Z_K e^2}{R_{Ki}}$$

$$\hat{H}\psi(r_1, r_2, \dots, r_n, R_1, R_2, \dots, R_N) = E\psi(r_1, r_2, \dots, r_n, R_1, R_2, \dots, R_N)$$

$$r_i = (x_i, y_i, z_i) \quad \text{Coordonnées de l'électron } i$$

$$R_A = (X_A, Y_A, Z_A) \quad \text{Coordonnées du noyau } A$$

$$\hat{H} = \hat{T}_N + \hat{T}_e + \hat{V}_{eN} + \hat{V}_{ee} + \hat{V}_{NN} \quad \text{Opérateur hamiltonien}$$

$$\psi(r, R) \quad \text{Fonction d'onde}$$

La résolution **exacte** de cette équation n'est possible que pour l'atome d'hydrogène et les systèmes mono-électroniques. Pour les systèmes poly-électroniques, on fait appel aux méthodes d'approximation, pour la résolution approchée de l'équation de SCHRÖDINGER

### Approximation de Born-Oppenheimer

Cette approximation est basée sur le fait que les électrons se déplacent beaucoup plus rapidement que les noyaux, ceci étant dû à la masse beaucoup plus faible des électrons. Par conséquent, les électrons réagissent quasi instantanément à une modification de la position des noyaux. Les électrons sont ainsi considérés comme se déplaçant dans un champ moyen créé par des noyaux immobiles, et sont donc sujets à un potentiel nucléaire statique. L'équation de Schrödinger peut ainsi être séparée en une partie nucléaire et une partie électronique.

- Born-Oppenheimer

$$\psi(r, R) = \psi_R^{el}(r) \Phi_N(R)$$

$$\hat{H}_{el} \psi_R^{el}(r) = E_{el}(R) \psi_R^{el}(r)$$

$$E = E_{el} + V_{NN}$$

### Approximation orbitale Approximation orbitale (Mono-électronique)

L'approximation orbitale consiste à exprimer la fonction d'onde, solution de l'équation de Schrödinger limitée à l'hamiltonien électronique H décrit plus haut, par un produit de fonctions dépendantes chacune des coordonnées d'un seul électron:

$$\Psi(1, 2, \dots, n) = \psi_1(1) \psi_2(2) \dots \psi_n(n)$$

Les fonctions mono électroniques  $\psi_i$  sont appelées orbitales

Malheureusement, cette fonction d'onde ne satisfait pas au principe de Pauli qui stipule que la fonction d'onde décrivant un système multiélectronique doit changer de signe lors de la Permutation des coordonnées de deux électrons quelconques. Dans L'approximation orbitale, la fonction d'onde qui satisfait ce principe se met sous la forme d'un déterminant de Slater:

$$\Psi = \frac{1}{\sqrt{n!}} \begin{vmatrix} \chi_1(1) & \chi_2(1) & \dots & \chi_n(1) \\ \chi_1(2) & \chi_2(2) & \dots & \chi_n(2) \\ \dots & \dots & \dots & \dots \\ \chi_1(i) & \chi_2(i) & \dots & \chi_n(i) \\ \dots & \dots & \dots & \dots \\ \chi_1(n) & \chi_2(n) & \dots & \chi_n(n) \end{vmatrix}$$

$$\psi_R^{el}(r) = \frac{1}{\sqrt{n!}} \det |\phi_a(r_1)\phi_b(r_2) \dots \phi_z(r_n)|$$

$\phi(r)$  Orbitale moléculaire

Généralement, les orbitales moléculaires sont développées en une combinaison linéaire d'orbitales atomiques. Cette approche est appelée LCAO: (Linear Combinaison of Atomic Orbitals)

$$\psi_i = \sum_j c_{ij} \phi_j$$

### Méthodes Ab initio (Hartree Fock)

#### **Méthode Hartree-Fock**

$$\hat{f}_i \phi_a(r_i) = \epsilon_a \phi_a(r_i)$$

$\hat{f}$  Opérateur de Fock

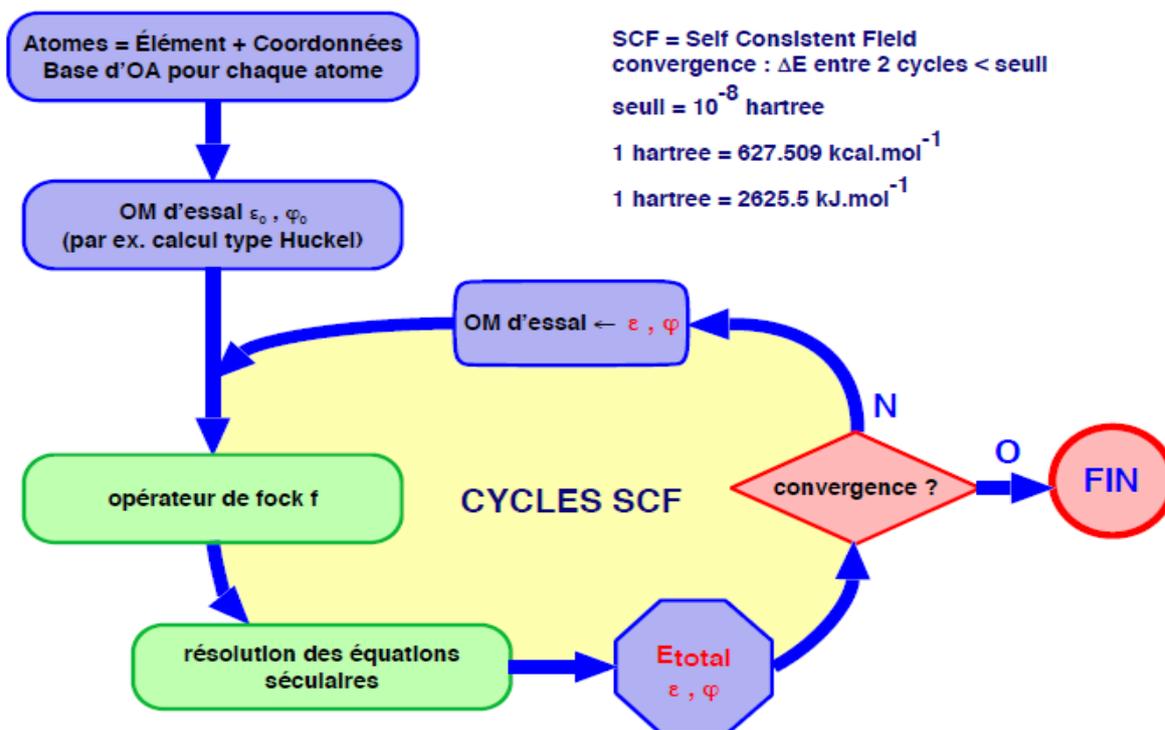
$\phi_a(r_i)$  OM occupée par l'électron de coordonnées  $r_i$

$\epsilon_a$  Énergie de cette OM

#### **LE problème**

L'expression de  $\hat{f}$  dépend des  $\phi$  -> il faut connaître la solution pour développer  $\hat{f}$

## Champ auto-cohérent



### Bases d'orbitales atomiques

#### Quelques définitions

Dans les méthodes ab initio (HF), on utilise un ensemble de fonctions de bases ("basis set") à partir duquel on construit l'ensemble des orbitales moléculaires du système chimique étudié (méthode LCAO). Cet ensemble de fonctions de bases est fini et, dans le cas général, non orthonormé.

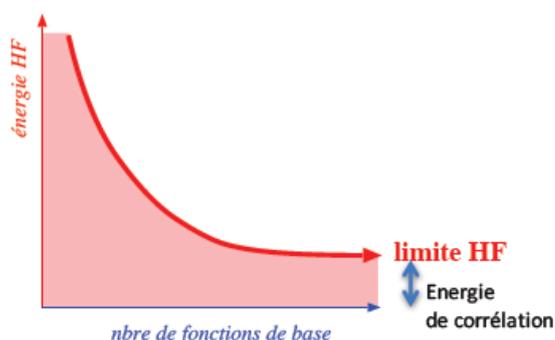
- Approximation LCAO (Linear Combination of Atomic Orbitals)

$$\phi_a(r_i) = \sum_{p=1}^M c_{pa} \chi_p(r_i)$$

$\{\chi_1, \chi_2, \dots\}$  = base d'OA

- Limite Hartree-Fock

- L'énergie d'une molécule baisse si le nombre de fonctions augmente
- Dans la limite où le nombre de fonctions de base est infini, on obtiendrait l'énergie HF limite (ce n'est pas l'énergie exacte!)
- On définit une base d'OA tronquée



Il existe deux grand types de fonctions de bases (appelées orbitales atomiques, bien qu'elles ne soient pas solutions de l'équation de Schrödinger pour les atomes) : les orbitales de type Slater (STOs) et les orbitales de type gaussienne (GTOs).

- Expression d'une OA? -> Fonctions de Slater
  - Solutions exactes de l'équation de Schrödinger pour l'atome H
  - Extension aux atomes polyélectroniques

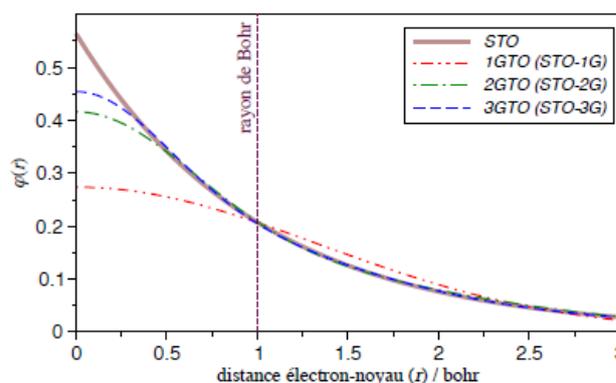
$$\chi(r, \theta, \varphi) = N r^{n-1} \exp(-\zeta r) Y_{lm}(\theta, \phi)$$

- Fonctions gaussiennes

$$\chi = \sum d_j g_j$$

$$g(r) = N \exp(-\alpha r^2)$$

Intérêt ? -> numérique (efficacité)



Les orbitales de type gaussienne sont employées de manière préférentielle par rapport aux orbitales de type Slater car elles facilitent grandement le calcul des intégrales électroniques dans les méthodes ab initio.

- Base STO-3G = 1 fonction (« simple-zeta ») de 3 gaussiennes par OA

- Base « Split-Valence »

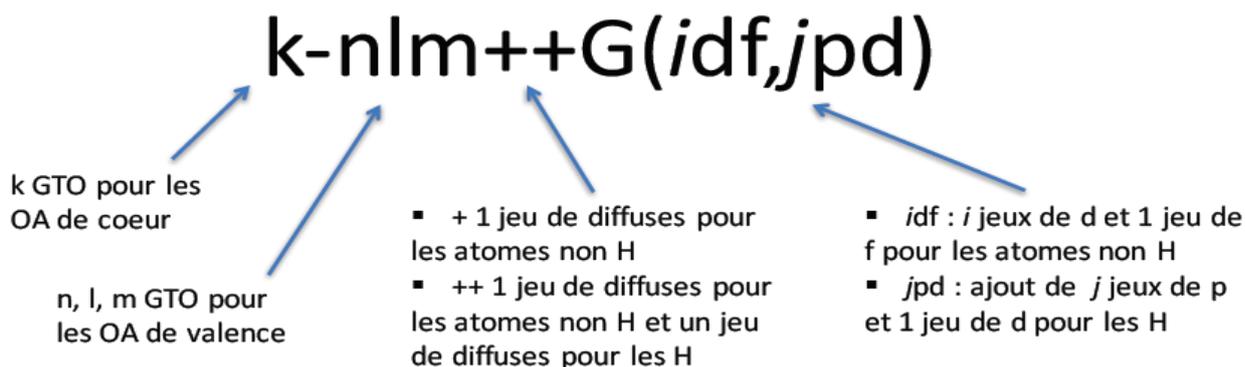
Exemple : 3-21G signifie une fonction de 3 gaussiennes pour chaque OA de « coeur » et 2 fonctions (composées de 2 gaussiennes et 1 gaussienne) pour chaque OA de valence (« double-zeta »)

- Nécessité d'introduire des fonctions de polarisation

Exemple : 6-31G(d) : ajout d'un jeu d'orbitales d pour les atomes non H

- Traitement d'effets à grande distance (liaisons H) -> fonctions diffuses

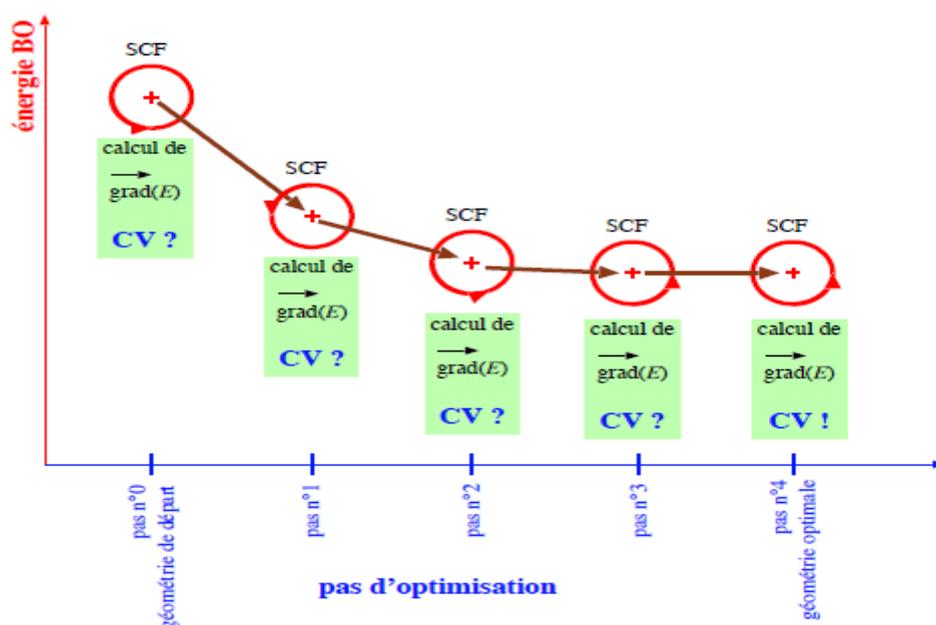
Exemple : 6-31+G(d) : ajout d'un jeu de fonctions diffuses pour les atomes non H



**6-311++G(3df,3pd)**

**Procédure d'optimisation d'un géométrie**

**Optimisation de géométrie et SCF**



**Les logiciels de Modélisation Moléculaire**

Un logiciel de modélisation moléculaire comprend de manière générale les modules suivants :

- 1 - Construction, visualisation et manipulation des molécules.
- 2 - Calculs
- 3 - Sauvegarde des structures et gestion des fichiers
- 4 - Etude des propriétés moléculaires.

Il existe plusieurs logiciels de calcul dans le domaine de la chimie quantique : par exemple ADF, Abinit, ,hyperchem, Gaussian. Nous avons essentiellement utilisé le programme Gaussian et son interface graphique GaussView

### Logiciel GAUSSIAN

La série de programmes Gaussian a été créée au début des années 70 par John Pople et ses collègues à l'université de Carnegie Mellon. La compagnie Gaussian, Inc. a été installée dans les années 80 pour distribuer le programme.

Gaussian70, Gaussian76, Gaussian77, Gaussian78, Gaussian80, Gaussian82, Gaussian83, Gaussian85, Gaussian86, Gaussian88, Gaussian90, Gaussian 92, Gaussian93, Gaussian 94, Gaussian95, Gaussian96, Gaussian 98, Gaussian 03, Gaussian 09. Gaussian 16.

### Construction, visualisation

La construction des molécules peut se faire selon plusieurs méthodes :

- **À partir de fragments préexistants** (Chem 3D). Ceux-ci sont associés par substitution, la structure est alors élaborée par des modifications successives du squelette initial (type de liaison, type d'atome, stéréochimie). Les paramètres sont ajustés au fur et à mesure des substitutions et modifications.
- **à partir d'une structure dessinée en deux dimensions** (Hyperchem). La transformation en structure en trois dimensions est ensuite obtenue par ajustement global des paramètres. La stéréochimie aléatoire peut alors être modifiée.
- **à partir de données géométriques** (coordonnées des atomes provenant d'une structure de RX) ou à partir de bases de données courantes (ex : Cambridge Structural Data Base, Brookhaven Protein Bank).

La visualisation des molécules peut être faite dans une grande variété de modes graphiques permettant dans chaque cas de mettre en valeur les informations demandées. Ci-dessous Deux représentations de la 2-méthylcyclohexanone : "Boules et bâtons" "Bâtons"

