

Enseigner la mécanique moléculaire dans les cursus universitaires

Partie 1 : Théorie

Alexandre Hocquet* professeur agrégé

Summary : *Teaching molecular mechanics in the chemistry curriculum. Part 1 : Concepts*

The concepts are simple, software and hardware more and more available, chemistry on the computer is attractive, practical applications are numerous, the field is growing in academic research as well as in the industry : these are the reasons why we should incorporate molecular modeling in our chemistry curriculum. This article presents molecular mechanics, the classical basis of molecular modeling, to the chemistry teacher. In this first part, the concepts are introduced. In a second part, the possibilities of including molecular mechanics in a chemistry curriculum are discussed.

Mots clés : *Modélisation moléculaire, enseignement, mécanique moléculaire.*

Key-words : *Molecular modeling, teaching, molecular mechanics.*

Qui dit modélisation moléculaire, dit modèle moléculaire. Il s'agit de construire une théorie (pour le modèle moléculaire de type Dreiding ou Cochranes, les atomes sont des boules, les liaisons des bâtons de différentes longueurs) pour décrire l'objet chimique qui est la molécule, afin d'en prédire certaines propriétés (pour le modèle moléculaire, les propriétés sont forcément grossières : approche qualitative de l'encombrement stérique, par exemple).

Modélisation moléculaire : une définition

Le développement de la modélisation moléculaire est récent. Or, les théories associées à cette modélisation ne le sont pas : que le modèle soit quantique, ou non, les concepts utilisés sont connus depuis plusieurs dizaines d'années. Le succès de la discipline provient plutôt de la puissance et de l'accessibilité de ceux qui font les calculs : les ordinateurs. Il est aujourd'hui possible de répondre à des questions inconcevables il y a seulement cinq ans.

La conséquence de ce phénomène est que la modélisation moléculaire était le domaine exclusif d'une petite quantité de chercheurs : chimistes théoriciens, thermodynamiciens... Si ces chercheurs sont, à juste titre, toujours ceux qui développent ces méthodes, ceux qui les utilisent représentent maintenant toute la communauté des chimistes (et des biophysiciens). On comprend dès lors la différence de domaine scientifique entre chimie théorique et modélisation moléculaire.

La modélisation moléculaire correspond à l'utilisation pratique des domaines de la chimie théorique. Il est à noter que les deux domaines qu'on oppose en anglais sont la « theoretical chemistry » et la « computational chemistry ». Or, ce dernier terme n'est pas l'exacte traduction de modélisation moléculaire (molecular modeling). Il regroupe en fait tous les domaines de l'utilisation de l'informatique en chimie, de la synthèse assistée par ordinateurs à la chimométrie en passant par la gestion de bases de données structurales ou le design de molécules [1]. Le terme « computational chemistry » n'a pas de réel équivalent en français.

L'attrait de la modélisation moléculaire est lié aussi à la représentation des molécules. Les logiciels de modélisation moléculaire proposés aujourd'hui par les distributeurs rivalisent de séduction. C'est qu'il existe un marché du logiciel de modélisation, provoqué par l'intérêt des puissantes industries pharmaceutiques pour ce domaine. La conséquence de ce système concurrentiel est le développement de produits à la fois simples d'accès et attirant l'œil. Cette facilité a un effet pervers : l'utilisation de cet outil dans des conditions pas forcément appropriées, par des utilisateurs pas forcément avertis.

* Facultad de Ciencias Físicas, Universidad de Chile, Blanco Encalada, 2008, Santiago Centro, Chili.
Tél. : +56 (2) 678 45 19, Fax : +56 (2) 696 73 59.
E-mail : ahocquet@famarugo.cec.uchile.cl

Définition de notre champ d'études : la mécanique moléculaire

Les méthodes de modélisation moléculaire sont nombreuses. Certaines reposent sur l'utilisation de méthodes quantiques de chimie théorique. C'est le cas de la méthode de Hückel, par exemple, qui est maintenant au programme de nos classes préparatoires. Une autre catégorie reste dans le cadre de la mécanique classique. C'est en particulier le cas de la méthode dite de la mécanique moléculaire. C'est à cette dernière que nous nous limiterons ici.

Parmi les méthodes quantiques les plus fréquemment employées, deux catégories sont distinguées : les méthodes « ab initio » correspondent à une résolution itérative de l'équation de Schrödinger par un calcul de type LCAO. Les méthodes « semi-empiriques » réalisent le même type de calcul avec diverses approximations, et en paramétrant de nombreuses intégrales pour réduire le temps de calcul. La méthode de Hückel correspond à une méthode semi-empirique aux approximations draconiennes [2]. Elle rend les calculs accessibles à une calculatrice mais son ambition prédictive reste qualitative.

En mécanique moléculaire, les atomes sont représentés par des boules, et les liaisons par des ressorts. Cette théorie représente donc la molécule comme une espèce de modèle moléculaire flexible. L'avantage de cette méthode par rapport à une méthode quantique est la simplicité des équations qui seront mises en jeu, ce qui permet donc de traiter, pour le même investissement en temps ou en moyens, des systèmes moléculaires plus complexes. L'inconvénient majeur est que cette théorie est complètement indépendante de la structure électronique de la molécule... et pourtant, elle permet de nombreuses prédictions.

Outre la mécanique moléculaire, on regroupe traditionnellement dans les méthodes classiques la dynamique moléculaire, qui correspond à la modélisation des mouvements d'une molécule, ainsi que les méthodes dites statistiques, qui correspondent à des applications de la thermodynamique statistique pour la prédiction de grandeurs macroscopiques [3].

Le choix pédagogique de la méthode de la mécanique moléculaire est à la fois restreint (pas de référence à la théorie quantique de la liaison chimique), et ambitieux (introduction pour les étudiants d'un nouveau concept).

Les avantages sont :

1 - Un modèle théorique à la portée des élèves : donc une compréhension de ce qui se passe à l'intérieur de la « boîte noire ». De plus, le modèle mathématique de l'énergie d'une molécule est directement lié à deux notions de mécanique du programme de physique des classes préparatoires ou des premiers cycles universitaires : l'énergie potentielle et l'oscillation harmonique, pour lesquelles il constitue une application.

2 - Une manipulation aisée de l'ordinateur : l'utilisation d'un logiciel de type « souris et menu » permet d'éviter des lourdeurs mathématiques ou informatiques.

3 - L'aura de « l'infographie » : la représentation de molécules à l'écran attire les élèves par son côté esthétique.

4 - De nombreuses applications dans le cadre du cours et des TP (voire des TIPE, en ce qui concerne les classes préparatoires) de chimie... et de biochimie :

- analyse conformationnelle,
- liaisons de faible énergie,
- équilibres, stabilité, réactivité en chimie organique (ou inorganique),
- spectroscopies RMN ou IR.

Une présentation du concept de modélisation moléculaire a déjà fait l'objet d'un article dans cette même revue, il y a maintenant dix ans [4]. L'objet de cet article était de présenter un outil aux chercheurs chimistes. Dix ans plus tard, toutes les conditions sont réunies pour que cet outil soit utilisé par les enseignants et leurs étudiants. Pour en savoir plus, on pourra se référer avec profit à cet article [4] ainsi qu'à celui de Lipkowitz et Boyd [5], paru dans le *Journal of Chemical Education*. Enfin, dans la collection de monographies bon marché Oxford Primers, le livre de Grant et Richards [3] propose une introduction très complète à la modélisation moléculaire.

Cours de mécanique moléculaire

Présentation

Construisons le cyclohexa-1,4-diène avec un modèle moléculaire de type Dreiding : une structure plane de cette molécule paraît instable, et le modèle semble de lui-même vouloir se transformer en une forme bateau pour que les tensions entre les angles de liaisons soient moins fortes (figure 1). On a longtemps cru qu'une forme bateau était effectivement plus stable, résultat confirmé par la structure cristalline élucidée par diffraction des rayons X. Or, un calcul de mécanique moléculaire montre que la conformation plane est en fait plus stable, ce qui est en accord avec plusieurs résultats de RMN, qui ont longtemps été mal interprétés à cause de cette fausse impression [5].

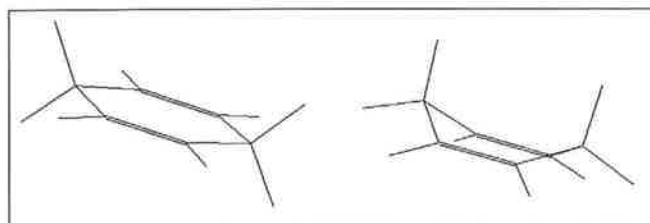


Figure 1 - Conformations plane et bateau du cyclohexa-1,4-diène.

La mécanique moléculaire est donc un modèle moléculaire plus sophistiqué. Le concept est apparu au début des années 70, à un moment où la théorie quantique de la liaison chimique était déjà reconnue depuis longtemps, et où les premiers calculs quantiques de molécules organiques par ordinateurs étaient réalisés. C'est que la mécanique moléculaire n'est pas une méthode de chimiste théoricien mais a été mise au point par la communauté des spectroscopistes, à un moment où il devenait admis que les fréquences d'absorption IR pouvaient correspondre à la vibration d'une liaison particulière, d'une molécule à l'autre [3]. De là l'idée d'un potentiel représentant l'énergie de la molécule qui serait une somme de termes correspondant à la vibration d'une liaison (bond), d'un angle, d'un angle dièdre (dihedral)... (figure 2). En première approximation, ces termes sont des potentiels

de type harmonique, caractérisés par une constante de force et une valeur de la variable à l'équilibre (K_b et r^0 pour le terme V_b) :

$$V = V_b + V_\theta + V_\tau + V_\chi + V_{nb} \quad (1)$$

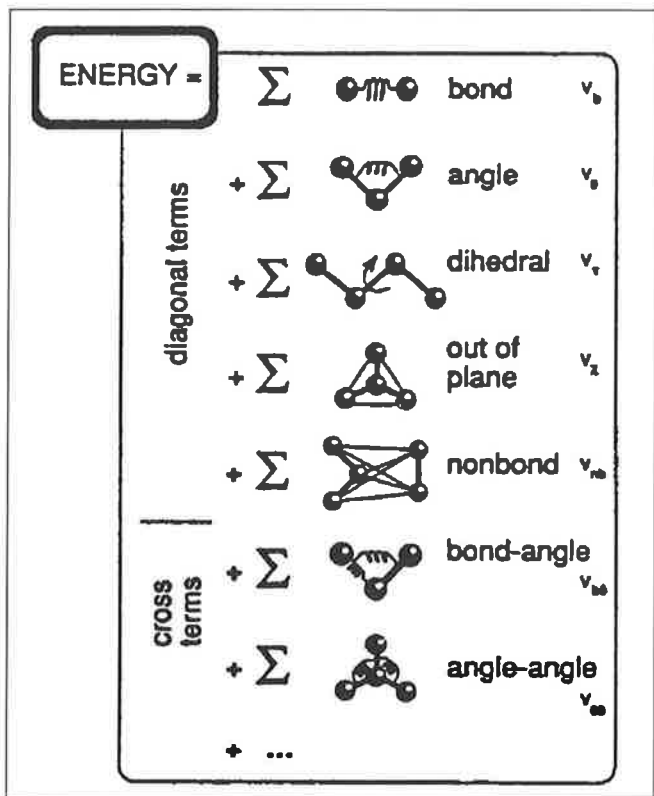


Figure 2 - Les différents termes de l'énergie de contrainte peuvent être vus comme des potentiels d'oscillateurs harmoniques. D'après réf. [17]. Ici, un terme d'étirement de liaison (bond), de déformation d'angle (angle), de torsion (dihedral), d'écartement du plan (out of plane) et d'interaction entre atomes non liés (nonbond) constituent les termes diagonaux (diagonal terms). Les termes correctifs (cross term) représentés ici sont le couplage étirement-déformation d'angle (bond angle) et entre deux déformations d'angle (angle-angle)¹.

Cette somme de termes vibratoires se nomme l'énergie de contrainte (strain energy), parfois appelée énergie stérique ; elle correspond à l'écart par rapport à l'énergie d'une molécule qui aurait pour paramètres tous les r^0 , θ^0 , ϕ^0 ... C'est elle qui va caractériser l'énergie de la molécule. On peut remarquer, dès à présent, que sa valeur absolue n'a aucune signification physique puisqu'elle dépend des paramètres qu'on choisit. Ce sont toujours des différences d'énergie (entre deux conformations, entre deux stéréoisomères...) que l'on comparera.

Ces paramètres sont choisis de manière empirique : à partir de données structurales, il s'agit que la structure calculée reproduise, au mieux et pour le plus grand nombre de molécules, les structures expérimentales : d'où le nom de champ de force empirique (empirical force field : EFF) appliqué à ce potentiel.

Puisqu'on a alors une fonction V de plusieurs variables, variables représentant les degrés de liberté d'une molécule, le calcul consiste donc à trouver le minimum de cette fonction. L'énergie de contrainte en fonction des coordonnées internes de la molécule peut donc être considérée comme une surface de potentiel (potential energy surface : PES) dont on recherche les puits.

Construction d'un champ de force

Un champ de force représente l'ensemble des fonctions et des paramètres utilisés pour définir analytiquement l'énergie de contrainte d'une molécule. Construire un champ de force, c'est rechercher un compromis : un champ de force trop simple et des paramètres trop généraux impliquent une faible précision du modèle. Un champ de force trop sophistiqué a deux défauts : sa complexité augmente le temps nécessaire pour les calculs. Sa trop grande précision fait diminuer sa transférabilité d'une molécule à l'autre. Ainsi, certains champs de force se veulent « généralistes », alors que d'autres sont spécialisés pour tel ou tel type de molécules (peptides, acides nucléiques...).

Terme d'étirement (stretch)

Chaque liaison de la molécule, considérée comme un ressort, a donc une énergie du type :

$$V_b = \sum_{\text{bond}} K_b (r - r^0)^2 \quad (2)$$

où K_b représente la constante de force de la liaison, ou en d'autres termes, la capacité de la liaison à résister à un étirement, et r^0 représente la distance de liaison à l'équilibre.

Par exemple, dans les champs de force Amber, la liaison C=O d'un carbonyle possède une distance d'équilibre de 1,229 Å et une constante de 570 kcal.mol⁻¹.Å⁻². Le minimum est à 1,526 Å pour une liaison C-C, et la constante de force est de 310 kcal.mol⁻¹.Å⁻². Cette dernière liaison est donc plus déformable, le puits de potentiel correspondant est plus large [6] (figure 3).

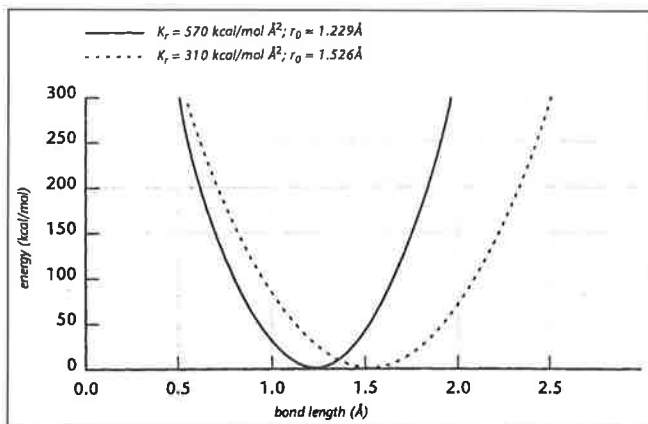


Figure 3 - Énergies d'étirement de liaison comparées pour une liaison C=O carbonyle (en trait plein) et pour une simple liaison C-C (en pointillés) en fonction de la longueur de liaison (bond length). D'après référence [6]¹.

L'énergie d'une liaison en fonction de la distance entre les deux atomes peut être représentée par une courbe de Morse (figure 4). Le terme harmonique V_b peut donc être vu comme un développement limité de cette fonction. L'inconvénient évident est que pour des distances loin de l'équilibre, le modèle perd en précision. Dans certains champs de force, un terme correctif cubique, voire de puissance supérieure est rajouté. D'autres expriment analytiquement la courbe de Morse [1].

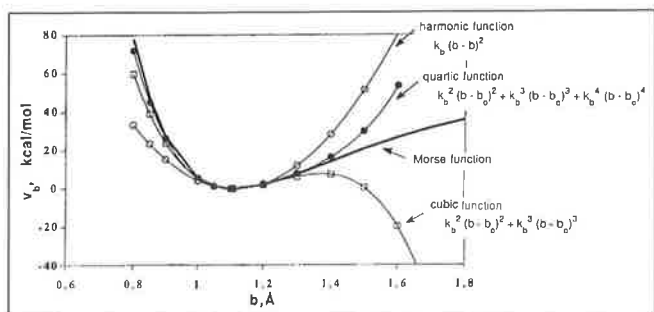


Figure 4 - Énergies de liaison représentées par : une courbe de Morse (en trait gras), une fonction harmonique (cercles), une fonction cubique (carrés), une fonction quadratique (cercles pleins). D'après référence [1].

Terme de flexion (bend)

De la même façon, le terme énergétique propre aux variations d'un angle entre deux liaisons est défini de manière harmonique :

$$V_{\theta} = \sum_{\text{angle}} K_{\theta} (\theta - \theta^{\circ})^2 \quad (3)$$

Les significations des paramètres sont équivalentes. Si on compare les ordres de grandeur de K_b et K_{θ} , on s'aperçoit qu'il est plus facile énergétiquement de déformer un angle qu'une longueur de liaison : par exemple, pour $K_b = 634 \text{ kcal.mol}^{-1} \cdot \text{Å}^{-2}$ (liaison C-C, champ de force publié dans la référence [10]), il coûte $3,2 \text{ kcal.mol}^{-1}$ pour allonger (ou réduire) la liaison de 10 pm, alors que pour un angle de liaisons C-C-C, avec $K_{\theta} = 82 \text{ kcal.mol}^{-1} \cdot \text{rad}^{-2}$, il faut seulement $0,3 \text{ kcal.mol}^{-1}$ pour ouvrir l'angle de 10° [7]. Comparativement, le modèle moléculaire en tiges métalliques de type Dreiding considère ces deux opérations comme aussi difficiles l'une que l'autre.

Terme de déformation d'angle dièdre (torsion)

Ce terme énergétique concerne cette fois trois liaisons donc quatre atomes A, B, C, D. L'angle dièdre ϕ correspond à l'angle entre les liaisons AB et CD, donc à la rotation autour de la liaison BC. Ce terme énergétique correspond à la barrière d'énergie rencontrée lors de la rotation d'une liaison. Il s'écrit :

$$V_{\tau} = \sum_{\text{dihedral}} V_n / 2 (1 + \cos(n\phi - \phi_n)) \quad (4)$$

Chaque terme correspondant à un angle dièdre ϕ contient plusieurs termes en V_n . le terme en V_1 correspond à une rotation complète autour de la liaison, le terme en V_2 à un demi tour, etc. Ainsi, pour une liaison simple, le terme V_3 sera prépondérant, pour une liaison double, le terme V_2 .

La torsion autour de la liaison HN-CO d'une amide, dans le champ de forces Amber possède un terme en V_1 et un terme en V_2 beaucoup plus important, caractérisant la barrière de rotation pour une liaison partiellement conjuguée (figure 5). On remarque aussi que la liaison sera plus stable en position trans [6].

D'une manière générale, en ce qui concerne les liaisons simples, il faut beaucoup moins d'énergie pour déformer un angle dièdre qu'un angle ou une distance de liaison. Les valeurs respectives des K_b , K_{θ} et V_n sont paramétrées en conséquences. Quand on décompose l'énergie de contrainte en ses différents termes, il est donc courant de trouver des

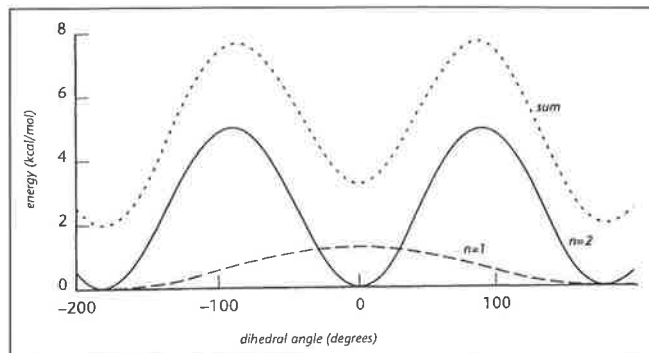


Figure 5 - Énergie de torsion autour d'une liaison N-C d'amide en fonction de l'angle dièdre (dihedral angle) caractéristique de la rotation autour de cette liaison. Terme en V_1 (pointillés larges), terme en V_2 (trait plein) et somme des deux termes (pointillés courts). Pour la lisibilité de la figure, la somme des deux termes a été décalée vers le haut. D'après référence [6].

valeurs plus faibles (c'est-à-dire de faibles écarts par rapport aux distances et angles d'équilibre) pour les deux premiers termes que pour le troisième. Certains champs de force, particulièrement en dynamique moléculaire, qui gèrent des macromolécules suppriment même ces deux premiers termes, fixant ainsi les longueurs et angles de liaisons et ne faisant varier que les angles dièdres.

Termes croisés ; termes correctifs

Pour certaines molécules particulièrement contraintes comme les petits cycles organiques, la déformation de l'angle de liaison s'accompagne d'un allongement de la liaison. Pour en rendre compte, certains champs de force incluent un terme dit « croisé » stretch bend où l'énergie dépend en même temps des deux déformations.

$$V_{b\theta} = \sum_{\text{bond angle}} K_{b\theta} (r - r^{\circ}) (\theta - \theta^{\circ}) \quad (5)$$

Des termes croisés entre angles et angles dièdres, entre longueurs et angles dièdres peuvent aussi exister. Ces termes deviennent essentiels si le but du champ de force est de reproduire correctement les fréquences de vibrations des molécules, afin d'en simuler le spectre IR [5].

Dans le même souci de reproductibilité des comportements de certaines molécules, un terme dit « out of plane bending » ou « improper dihedral » peut assurer à des carbones trigonaux de rester plans, ou à des centres chiraux de ne pas s'inverser accidentellement. On le définit par un terme où la variable est la distance de l'atome à un plan formé par ses trois voisins [3].

Jusqu'ici, les différents termes énergétiques ne représentaient des interactions qu'entre atomes liés. Par opposition, les termes suivants sont appelés « non-bonded interactions ». Leur point commun est qu'ils dépendent maintenant de la distance entre deux atomes, non liés, la somme doit donc être réalisée sur l'ensemble des couples d'atomes de la molécule. Certains champs de force, pour de grosses molécules, prévoient une distance au delà de laquelle (cutoff) ces interactions ne sont plus prises en compte.

Terme d'énergie de van der Waals

L'énergie de van der Waals, en mécanique moléculaire, est une extrapolation aux atomes de l'énergie potentielle du

même nom, qui, elle, concerne les molécules. Selon les champs de force, on a un potentiel du type Lennard-Jones ou Buckingham. Dans les deux cas, on définit un rayon de van der Waals pour chaque type d'atome.

$$V_{vdW} = \sum_{i,j} A_{ij}/r_{ij}^{12} - B_{ij}/r_{ij}^6 \quad (6)$$

Dans le potentiel de Lennard-Jones, le premier terme est répulsif et est prépondérant à faible distance, le deuxième est attractif. Il existe donc un minimum pour le terme V_{vdW} dont la distance correspond à la somme des rayons de van der Waals des deux atomes interagissant (figure 6).

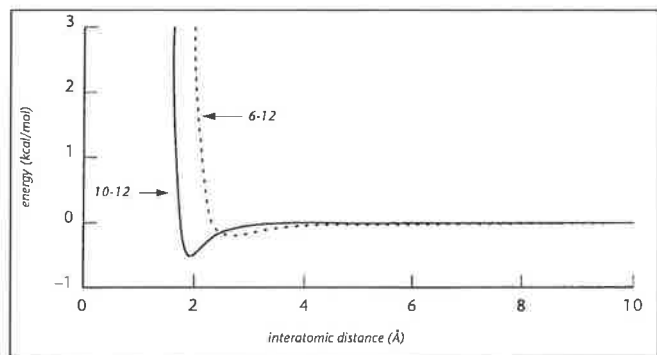


Figure 6 - Énergie de van der Waals (en pointillés) et énergie de liaison hydrogène (en trait plein) en fonction de la distance entre les deux atomes concernés (interatomic distance). L'expression de l'énergie pour ces deux termes diffère par l'exposant du terme attractif. D'après référence [6]¹.

Terme d'énergie électrostatique

La deuxième interaction de ce type est l'énergie électrostatique entre deux atomes non liés. Selon l'électronégativité des atomes, des charges partielles sont attribuées à chacun et l'énergie électrostatique entre deux atomes i et j correspond à :

$$V_{elec} = \sum_{i,j} q_i q_j / \epsilon r_{ij} \quad (7)$$

Dans certains champs de force, pour rendre compte de phénomènes de solvation, une permittivité diélectrique relative ϵ_r peut être incluse, éventuellement dépendante de la distance. On rend ainsi compte de la polarisabilité de certains groupements [6]. Dans d'autres champs de force, l'énergie électrostatique est considérée comme une interaction de dipôles, dipôles paramétrés pour chaque type de liaison [7].

Dans les champs de force spécialisés dans les biomolécules, le terme électrostatique devient d'une importance cruciale. D'une part, les molécules possèdent beaucoup d'hétéroatomes. Il s'agit donc de modéliser les différences d'électronégativité entre chaque élément par ces charges dites partielles, localisées sur chaque atome. Ce type de grandeur est issu de calculs quantiques car aucune technique expérimentale ne permet d'y accéder. Or, définir une charge partielle localisée sur un atome est difficile dans le cadre d'une théorie qui a pour base la délocalisation des électrons. Plusieurs définitions existent, ce qui rend la paramétrisation d'autant plus difficile [8]. D'autre part, le milieu étudié (par exemple une protéine en solution aqueuse) n'a certainement pas la même permittivité relative à l'intérieur de la protéine et à l'extérieur [9]. D'où la nécessité de faire varier la permittivité diélectrique relative.

Certains champs de force prévoient un terme spécifique à la liaison hydrogène. la plupart se contentent de considérer cette dernière comme étant décrite par la combinaison des termes électrostatiques et de van der Waals. Le champ de force AMBER possède un potentiel spécifique dont l'expression est similaire à un potentiel de Lennard-Jones (figure 6). L'exposant du terme attractif n'est plus -6 mais -10 [6].

Enfin, les interactions d'électrons π sont difficiles à modéliser en mécanique moléculaire. Pour certains champs de force, la possibilité de conjugaison dans une molécule insaturée est vérifiée par un calcul quantique. Si tel est le cas, les paramètres des liaisons simples et doubles impliquées sont corrigés selon le degré de liaison partielle calculé [3].

En analysant les différents termes de l'énergie de contrainte, il est nécessaire de rester conscient de l'interdépendance de ceux-ci. Par exemple, la barrière d'énergie pour la rotation de la liaison C-C de l'éthane est représentée par le terme en V_3 de l'énergie de torsion de l'éthane. En ce qui concerne le butane, on sait que la rotation est influencée par un facteur similaire à l'éthane plus une interaction stérique entre les groupements $-CH_3$. Ainsi, des interactions de type van der Waals sont vraisemblablement contenues dans le terme V_r . D'un champ de force à l'autre, la reproduction de certains résultats peut se traduire par des choix de paramétrisation différents qui impliquent une répartition différente de l'énergie de contrainte entre ses différents termes. L'interprétation de cette répartition doit donc se faire avec précaution, par exemple en comparant plusieurs champs de force [1] [7].

Paramétrisation

Les paramètres d'un champ de force, ce sont toutes les constantes de force et les grandeurs d'équilibre intervenant dans les expressions de l'énergie. En toute rigueur, chaque molécule possède ses propres paramètres, pour chaque atome. En pratique, le premier type de paramètres à définir est le type d'atome. En effet, un atome de carbone trigonal et un atome de carbone tétraédrique n'auront pas les mêmes valeurs d'angle à l'équilibre. Il faut donc définir un type d'atome pour le Csp2 et un autre pour le Csp3. Selon la précision du champ de force, les types seront plus ou moins nombreux. Par exemple, un champ de force spécialisé dans les sucres sera obligé de définir un type spécial d'atome de carbone anomérique pour pouvoir rendre compte de l'effet anomère.

Il existe à l'heure actuelle un grand nombre de champs de force. Le choix des fonctions analytiques et des paramètres répond à deux buts :

- obtenir une fonction analytique à la fois précise et simple pour permettre les calculs à la fois les plus exacts et les plus rapides possibles ;
- définir une gamme de paramètres la plus petite possible mais assez exhaustive pour que les résultats soient transférables d'une molécule à une autre.

Pas assez de paramètres nuit à la précision, trop de paramètres nuit à la simplicité. Lors de l'élaboration d'un champ de force, on fait des choix selon le type de molécules auquel on s'intéresse, et le type de propriétés qu'on veut obtenir du calcul.

Puisque le champ de force est empirique, l'obtention de ces paramètres se fait par essai et erreur. Plus grand sera le nombre de résultats expérimentaux sur lesquels se baser, meilleurs seront les paramètres. Typiquement, les structures de molécules obtenues par diffraction de rayons X, par RMN, par spectres de microondes, diffraction de neutrons en phase gaz, spectres IR permettent la définition des paramètres utilisés en mécanique moléculaire. Les champs de force de nouvelle génération se basent de plus en plus sur une paramétrisation provenant de calculs quantiques *ab initio*. L'avantage est que la paramétrisation devient alors une technique plus systématique.

Selon l'objet du champ de force, l'accent sera mis sur telle ou telle provenance des paramètres. Ces données expérimentales dépendent des conditions expérimentales : l'exemple le plus typique est que les structures déterminées par diffraction des rayons X, correspondent à des molécules en phase cristalline, donc soumises à l'influence de leurs voisines, alors que le calcul est réalisé pour une molécule isolée. D'où la possibilité de mauvaises interprétations de résultats comme l'exemple du cyclohexa-1,4-diène.

Les qualités et les défauts des champs de force dépendent entièrement de ces choix. Par exemple, un champ de force comme AMBER dont le terme d'étirement est harmonique sera peu fiable loin de l'équilibre, contrairement à MM2 qui contient une correction [1]. Avant d'appliquer un champ de force à une molécule, il faut aussi, dans le même ordre d'idées, s'assurer que les paramètres du champ de force sont optimisés pour cette famille de molécules. Ainsi, de nombreux champs de forces, spécialisés dans les protéines n'auront pas les paramètres adéquats pour étudier des molécules organiques.

Un autre précaution importante est la justification physique de la valeur de l'énergie de contrainte. Puisque celle-ci dépend de la paramétrisation, on ne peut comparer que deux molécules qui ont exactement les mêmes paramètres : aussi bien constantes de force que valeurs d'équilibres. Ne seront comparables donc que deux molécules qui auront exactement le même nombre de chaque type d'atomes, de liaisons, d'angles et d'angles dièdres. En pratique, sauf cas particulier, on ne pourra comparer que des conformères ou des stéréoisomères entre eux. Un moyen de remédier à ce défaut est le calcul de grandeurs thermodynamiques ($\Delta_r H^\circ$) à partir de l'énergie de contrainte [10]. Certains champs de force en sont spécialistes. Il faut toutefois ne pas oublier qu'il s'agit d'enthalpies et non d'enthalpies libres. Les énergies sont calculées pour des molécules gelées, et aucun effet entropique n'est pris en compte. La comparaison de telles grandeurs

avec des constantes d'équilibre doit donc se faire avec précaution.

Choisir un champ de force, c'est donc se préoccuper de son domaine de validité. Dans le cas de molécules organiques, les champs de force créés par Allinger (MM1, puis MM2 et MM3) et leurs variantes (MM+, MMX...) font autorité [7]. Dans le domaine de la biophysique, de nombreux champs de forces existent pour un spectre de molécules restreint (protéines, acides nucléiques, sucres) : le plus répandu est AMBER. OPLS en est une version destinée à prendre en compte des effets de solvation. CHARMM est un champ de force généraliste dont les différentes variantes correspondent chacune à un domaine de molécules.

Minimisation

Une fois l'énergie de contrainte précisément définie, on peut donc la calculer pour une molécule donnée. Le but étant de trouver la structure la plus stable possible, il faut donc trouver un minimum de V en fonction de toutes les coordonnées de chaque atome, c'est-à-dire se déplacer sur la surface de potentiel jusqu'à en trouver un puits. Les méthodes mathématiques utilisées dans ce but calculent les dérivées premières et parfois secondes de V et travaillent avec les coordonnées cartésiennes de chaque atome. L'avantage immédiat de la mécanique moléculaire est d'avoir une expression analytique de V , donc facilement dérivable.

Plusieurs méthodes de minimisations existent. Les plus précises requièrent le calcul de matrices correspondant aux dérivées premières et secondes de V selon toutes les coordonnées. C'est la technique de Newton Raphson. L'avantage est de tenir compte à la fois de la pente et du rayon de courbure de la surface de potentiel, l'inconvénient est le besoin en mémoire pour calculer la matrice des dérivées secondes (Hessian matrix). Une approximation (méthode block diagonal Newton-Raphson) consiste à ne prendre en compte que les termes diagonaux de cette matrice, c'est-à-dire à négliger les dérivées croisées.

Une approximation beaucoup plus sévère consiste à la considérer constante. C'est la méthode de la plus grande pente (steepest descent) qui ne tient compte que de la dérivée première, c'est-à-dire que, à chaque itération, les coordonnées sont changées selon la direction et la taille du gradient d'énergie. La conséquence en est que le processus est de plus en plus lent près du minimum, et que la minimisation peut passer à côté d'un puits de potentiel sans le voir. En contrepartie, c'est la méthode la plus efficace quand on est loin d'un minimum (figure 7).

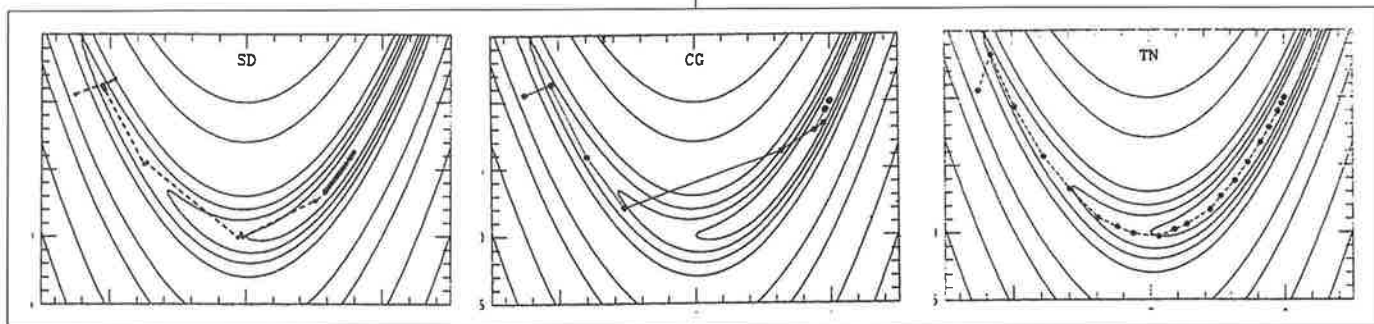


Figure 7 - Sur ces trois figures sont représentées les itérations parcourues lors d'une minimisation (de la gauche vers la droite) par méthode steepest descent (SD), conjugate gradient (CO) et Newton-Raphson (TN) sur une surface de potentiel représentée par des courbes isopotentielle. On remarque pour la méthode steepest descent, la vitesse au début de la minimisation et la lenteur proche de la fin. D'après référence [18].

Une méthode plus fine consiste à prendre en compte la direction du gradient de l'itération précédente dans le calcul du gradient suivant : c'est la méthode des gradients conjugués (conjugate gradients). Les minimums sont alors atteints de manière plus rapide [3].

Dans tous les cas, la paramétrisation des champs de force, réalisée à partir de structures à l'équilibre, implique que les surfaces d'énergie potentielle sont moins réalistes quand on s'écarte d'un minimum. Quand on cherche à reproduire des états de transition, il faut donc prendre des précautions. Une fois encore, la comparaison de champs de force différents permet de mieux appréhender cet inconvénient [1].

Enfin et surtout, la minimisation mathématique ne peut trouver qu'un minimum local (un puits de potentiel) sans pour autant qu'on puisse affirmer qu'il s'agit de la structure de plus faible énergie (minimum global). Il faut alors partir de structures de départ (initial guesses) différentes pour balayer le mieux possible l'espace conformationnel. Ainsi des résultats complètement différents peuvent être trouvés avec les mêmes champs de forces et les mêmes méthodes de minimisation selon les coordonnées initiales de la molécule.

Construire une structure initiale peut se faire de deux manières : partir d'une structure déjà connue (souvent un résultat de diffraction de rayons X) ou construire une molécule à partir du logiciel qu'on utilise. La plupart des logiciels possèdent en effet un module « build » qui correspond, en fait, à un champ de force très simplifié et à des paramètres les plus généraux possibles. On obtient ainsi une structure approchée de la molécule qui permettra au calcul effectif de démarrer par une structure la moins fantaisiste possible.

Utilisations de la mécanique moléculaire

La mécanique moléculaire trouve des applications, logiquement, dans tous les domaines concernés par la structure de molécules et, corrélativement, par les interactions entre molécules, donc par la réactivité. Dans les paragraphes qui suivent, quelques exemples significatifs de son utilisation sont présentés.

Analyse conformationnelle

Si la mécanique moléculaire calcule l'énergie d'une molécule et en trouve le minimum, l'application la plus évidente est l'analyse conformationnelle. De fait, la rapidité de calcul de cette technique permet l'analyse conformationnelle de macromolécules, dont le nombre de degrés de liberté rend impossible un calcul quantique. Le domaine le plus étudié est donc naturellement celui des protéines [8].

Les diagrammes de Ramachandran sont un moyen commode de visualiser les conformations adoptées par les différents acides aminés constituant la protéine. A l'origine utilisés en cristallographie de rayons X, ces diagrammes consistent à reproduire chaque valeur d'angle dièdre ϕ et ψ , caractéristiques de chaque liaison peptidique (figure 8), sur un graphe. Les domaines conformationnels ainsi représentés définissent des zones dans l'espace du graphe qui caractérisent les conformations traditionnelles des protéines (struc-

ture en hélice, en feuillet...) (figure 9). En mécanique moléculaire, ce type de diagramme peut se réaliser en calculant l'énergie d'un dipeptide (par exemple) en fonction de ϕ et de ψ . On peut alors tracer une surface d'énergie potentielle, sous forme de lignes de niveaux, qui montre ainsi les énergies relatives de chaque conformation et l'énergie nécessaire pour passer d'un puits de potentiel à l'autre (figure 10).

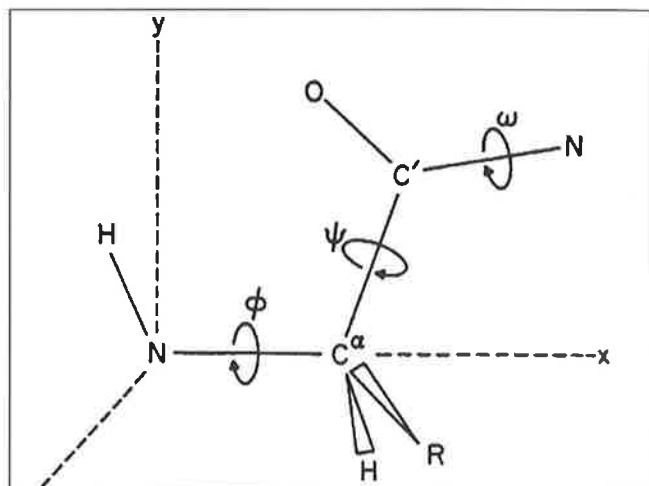


Figure 8 - Définition des angles dièdres ϕ et ψ d'une liaison peptidique. D'après référence [19].

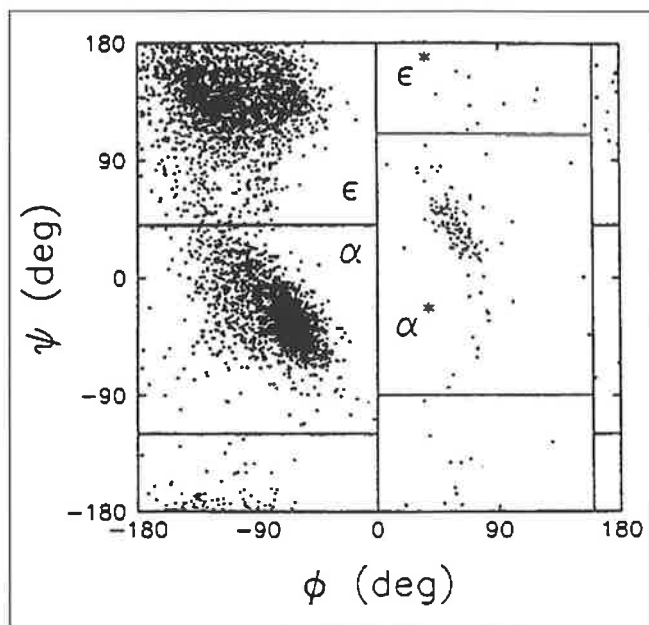


Figure 9 - Diagramme de Ramachandran de 5986 résidus autres que glycine et proline obtenus par structures cristallographiques de plusieurs protéines. Les régions ϵ , α et α^* correspondent respectivement aux familles de structures feuillet β , hélice α et hélice gauche. D'après référence [19].

Étude des liaisons de faible énergie

Puisque l'énergie de contrainte contient des termes d'interactions entre termes non liés, les interactions de faible énergie inter- ou intramoléculaires y sont donc modélisées. Les interactions de molécules avec un solvant, ou entre un substrat et le site d'un enzyme peuvent être caractérisées [3].

Le programme GRID a été conçu pour déterminer les sites d'interaction énergétiquement favorables entre un substrat et une macromolécule biologique [9]. GRID utilise son

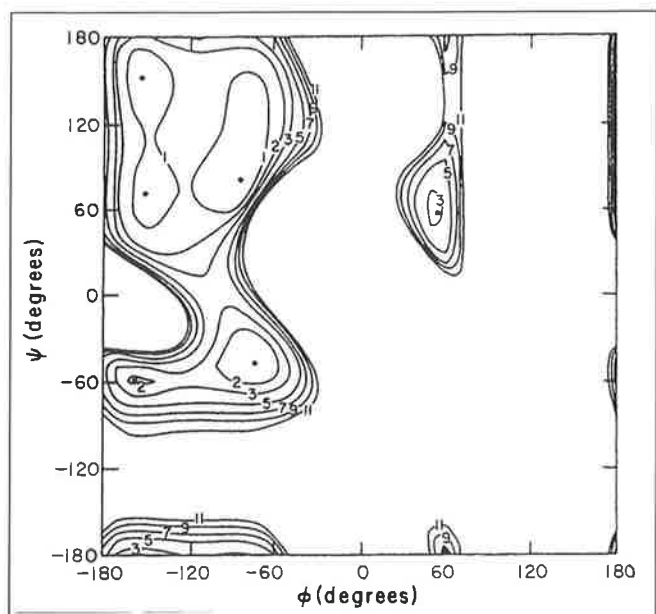


Figure 10 - Diagramme énergétique obtenu par mécanique moléculaire de la N-acétyl-N'-méthylamide. Les minimums trouvés correspondent à des régions de l'espace conformationnel qu'on retrouve dans les diagrammes de Ramachandran cristallographiques. D'après référence [19]¹.

propre champ de forces, dans lequel ne sont représentés que des termes d'interaction entre atomes non liés (terme de van der Waals, d'interaction électrostatique, de liaison hydrogène). Autour de la macromolécule à laquelle on s'intéresse, un cube est matérialisé, puis gradué dans les trois dimensions (figure 11). Sur chaque position, ainsi définie, est calculée l'interaction entre un groupement test, à choisir parmi une liste de groupements représentant toute une gamme d'interactions (de la molécule d'eau jusqu'au groupement méthyle, censé caractériser une interaction de type hydrophobe). Le but est de définir les régions de l'espace où les

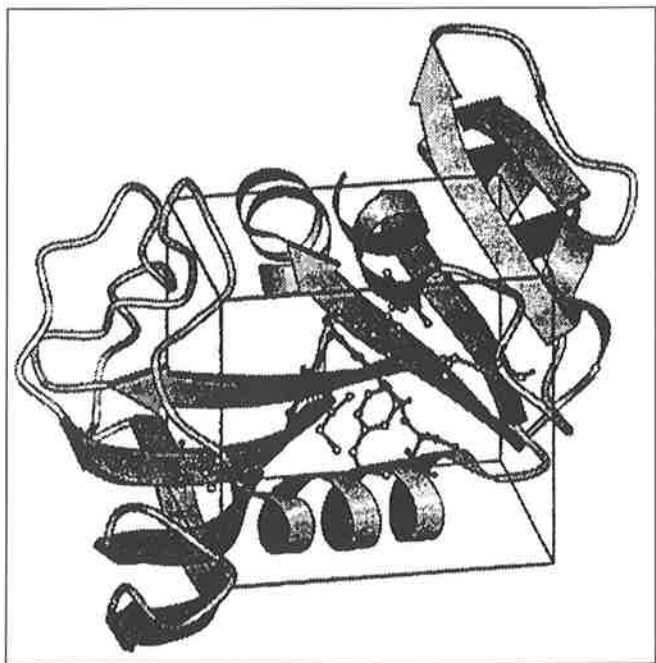


Figure 11 - Cage de 18 x 18 x 20 Å entourant le site actif de la dihydrofolate reductase de E. Coli. Cette cage est ensuite divisée en graduations de 1 Å, et sur chacun des points ainsi définis, l'interaction entre molécule test et l'ensemble du site actif est calculée. D'après référence [20].

interactions entre molécule cible et groupement test sont les plus favorables. En visualisant ces régions, les sites où une interaction avec un substrat est possible sont ainsi caractérisés (figure 12).

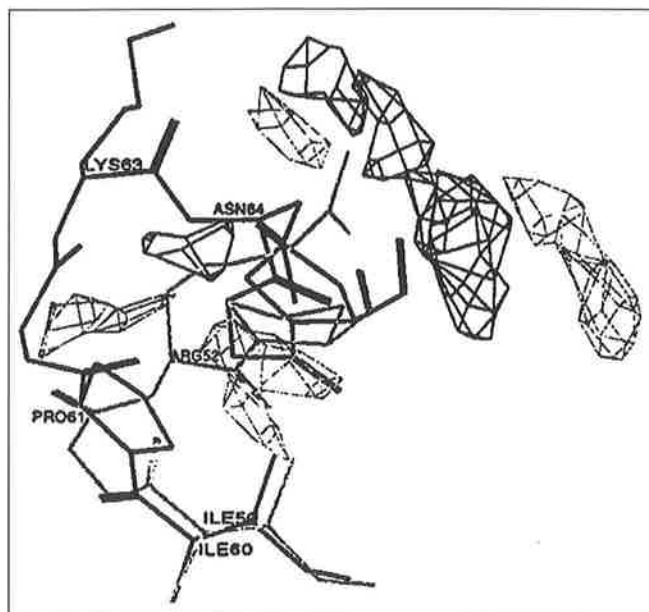


Figure 12 - Vue partielle du site actif de la dihydrofolate reductase de E. Coli. Le squelette de la protéine est inscrit en trait plein. Les volumes dessinés autour du site actif correspondent aux régions où l'interaction entre la molécule test et la protéine est particulièrement favorable. D'après référence [20].

Spectroscopies

La mécanique moléculaire trouve son origine dans le domaine de la spectroscopie de vibration. C'est pourtant avec la RMN qu'aujourd'hui les interactions sont les plus fortes. Puisque le champ d'application le plus important de ces deux techniques est celui de la macrobiomolécule, l'analyse conformationnelle de ces composés est souvent effectuée par la conjonction des renseignements fournis par ces deux techniques. Certains champs de force utilisent même des termes d'énergie correspondant aux renseignements spatiaux fournis par les constantes de couplage [11]. L'information peut fonctionner dans l'autre sens : la structure d'un sucre protégé a récemment été élucidée par RMN : les constantes de couplage des protons attachés au cycle correspondent à une conformation twist du sucre (figure 13). Cette conformation twist est plus stable que la conformation chaise, comme le confirme (pour une molécule isolée, et non

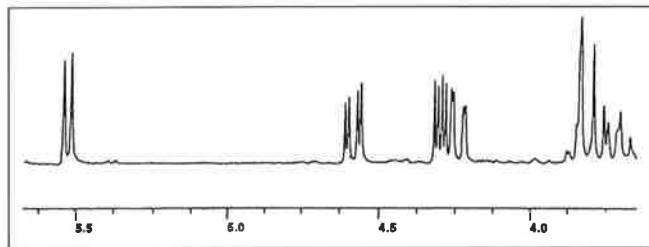


Figure 13 - Spectre RMN du proton du 1,2:3,4-di-o-isopropylidène-α-D-galactopyranose. Les constantes de couplage de chaque proton permettent non seulement d'attribuer chaque signal à chaque proton, mais en plus d'estimer la conformation du sucre grâce aux courbes de Karplus. D'après référence [12].

en solution) la mécanique moléculaire (figure 14), à partir de laquelle on a pu calculer des constantes de couplage théoriques d'après un calcul de Karplus liant constante de couplage entre deux protons vicinaux et angle dièdre correspondant à la liaison HCCH [12].

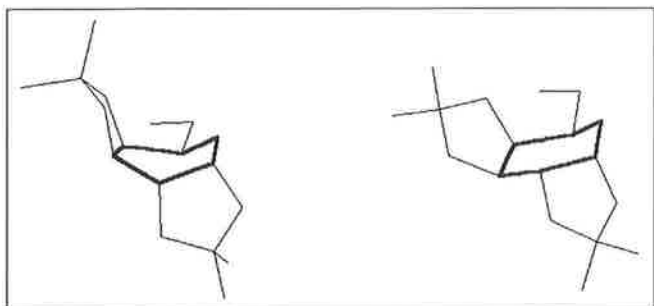


Figure 14 - Conformations twist (à droite) et chaise (à gauche) du 1,2:3,4-di-o-isopropylidène- α -D-galactopyranose (le cycle pyranose est en gras). La conformation chaise implique des tensions pour les deux cycles à cinq atomes. Elle est donc moins stable que la forme twist. D'après référence [12].

Réactivité, mécanismes

La mécanique moléculaire peut être utilisée pour calculer l'énergie de molécules instables, intermédiaires réactionnels, ou états de transition [13]. Localiser un état de transition sur une surface d'énergie potentielle est mathématiquement simple. Alors qu'un intermédiaire réactionnel (ou toute autre molécule stable) doit correspondre à un minimum d'énergie, et ce dans toutes les directions de l'espace, un état de transition, par définition, puisqu'il représente un col d'énergie, doit être un minimum vis-à-vis de toutes les coordonnées sauf une, pour laquelle il est un maximum. Il est donc analytiquement réalisable de localiser les états de transition sur une surface d'énergie potentielle. Le problème posé par la mécanique moléculaire est que la paramétrisation du champ de force est réalisée pour des molécules stables, donc la surface n'est correctement définie qu'autour des minimums. Une paramétrisation spécifique est souvent nécessaire [13].

L'attaque d'une fonction carbonyle par un nucléophile a été l'objet de nombreux modèles pour comprendre la stéréosélectivité de ces réactions et, très tôt, des études de mécanique moléculaire ont été conduites sur ce thème [4]. En utilisant une version modifiée du champ de forces MM2, on a pu mettre en évidence les stéréosélectivités comparées de cyclohexanones et de benzocycloheptanones [13] (figure 15). Les tensions relatives calculées dans les états de transition permettent de comprendre les préférences opposées de ces deux types de molécules pour l'approche axiale

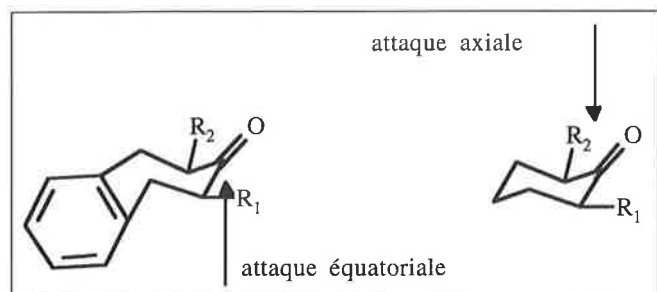


Figure 15 - En dépit de leur similitudes, les réactivités vis-à-vis de LiH des cyclohexanones et des benzocycloheptanones impliquent des stéréosélectivités opposées. D'après référence [14] et [15].

ou équatoriale. Intuitivement, les deux sites réactifs se ressemblent pourtant beaucoup [14, 15].

Dynamique moléculaire

La dynamique moléculaire consiste à donner une vitesse initiale à chaque atome de la molécule, puis à calculer le mouvement de cette molécule d'après le principe de la dynamique. On obtient ainsi une simulation newtonienne des mouvements de la molécule. Cette utilisation de la mécanique moléculaire constitue en fait un domaine de la modélisation à part entière [16] et a plusieurs avantages. Elle permet de simuler des grandeurs physiques dépendant du temps. Elle permet de balayer l'espace conformationnel de la molécule, si l'énergie cinétique de la molécule est suffisante pour sortir des puits de potentiel. En revanche, cette technique demande un temps de calcul beaucoup plus long.

C'est par dynamique moléculaire qu'on a réussi à caractériser le processus dynamique de l'oxygène dans l'hémoglobine [16]. D'après la structure cristallographique de cette protéine, l'hème censé interagir avec la molécule d'oxygène est profondément enterré à l'intérieur de la protéine, et apparemment inaccessible. Pour comprendre comment une molécule d'oxygène peut se frayer un chemin, une étude de dynamique fut entreprise. Pour supprimer l'inconvénient de l'improbabilité du succès de l'événement (une molécule d'oxygène au centre de la molécule a une probabilité faible de sortir de la protéine), plusieurs centaines de molécules d'oxygène furent placées autour de l'hème, toutes invisibles les unes par rapport aux autres, mais chacune pouvant interagir avec la protéine. Ainsi le labyrinthe énergétique a pu être résolu.

Conclusion

Les exemples d'utilisation de la mécanique moléculaire présentés ici sont loin d'être exhaustifs. Le domaine d'applications de cette technique peut intéresser aussi bien le physico-chimiste ou le spectroscopiste que le chimiste de synthèse, le chimiste organicien comme le chimiste inorganicien, le biophysicien comme le pharmacien, le polymériste comme le chimiste du solide. Concept simple d'application générale, la mécanique moléculaire rend des services dans quasiment tous les domaines de la recherche. Pour ces mêmes raisons, nous verrons dans la deuxième partie de cette présentation comment il peut s'insérer dans le cursus universitaire de chimie.

Remerciements

La société ChemCAD, le Laboratoire de chimie analytique de l'École Supérieure de Physique et Chimie Industrielles, le département de chimie de l'École Normale Supérieure de Cachan, Marie Terrien (université de Pau) Josette Fournier (université d'Angers), Céline Couot (lycée Janson de Sailly) et Jean-Michel Lefour (École Polytechnique) ont tous contribué, de manière directe ou indirecte à la réalisation de cet article. D'autre part, cet article a pour origine un stage de la Mission Académique de

Formation Professionnelle de l'Éducation Nationale (Académie de Paris) organisé par l'auteur et destiné aux enseignants de classes préparatoires. Tous les participants à ce stage ont collaboré par leurs remarques à l'amélioration de cet article.

Note

¹NDR : Les termes anglais des figures ne sont pas traduits car ils sont extraits de figures autorisées à être reproduites sans modification. Ils sont cependant explicités dans le texte ou dans les légendes.

Références

- [1] Lipkowitz K., Abuses of molecular mechanics, *J. Chem. Ed.*, **1995**, 72, p. 1070.
- [2] Atkins P., *Quanta : a handbook of concepts*, Oxford University Press, Oxford, **1991**.
- [3] Grant G., Richards W., *Computational chemistry*, Oxford Chemistry Primers, Oxford, **1995**.
- [4] Lomas J., La mécanique moléculaire, une méthode non quantique pour le calcul de la structure et de l'énergie d'entités moléculaires, *L'Actualité Chimique*, mai **1986**, p. 7.
- [5] Boyd D., Lipkowitz K., Molecular mechanics : the method and its underlying philosophy, *J. Chem. Ed.*, **1982**, 59, p. 269.
- [6] Hyperchem, Computational chemistry, Hypercube Inc., Waterloo, **1994**.
- [7] Eliel E., Wilen S., *Stereochemistry of organic compounds*, Wiley, New York, **1994**.
- [8] Bachrach S., Population analysis and electron densities from quantum mechanics, dans *Reviews in Computational Chemistry*, vol. 5, Boyd, Lipkowitz Ed., VCH, New York, **1994**.
- [9] Goodford P., A computational procedure for determining energetically favorable binding sites on biologically important macromolecules, *J. Med. Chem.*, **1985**, 28, p. 849.
- [10] Engler E., Adose J., von Ragué Schleyer P., Critical evaluation of molecular mechanics, *J. Am. Chem. Soc.*, **1973**, 95, 24, p. 8005.
- [11] Torda A., van Gunsteren W., Molecular modeling using nuclear magnetic resonance data, dans *Reviews in Computational Chemistry*, vol. 5, Boyd, Lipkowitz Ed., VCH, New York, **1994**.
- [12] Midland M., Beck J., Peters J., Rennels R., Asirwathan G., Synthesis and conformational analysis of 1,2:3,4-di-o-isopropylidene- α -D-galactopyranose, *J. Chem. Ed.*, **1994**, 71, p. 897.
- [13] Eksterowicz J., Houk K., Transition state modeling with empirical force fields, *Chem. Rev.*, **1993**, 93, p. 2439.
- [14] Wu Y., Houk K., *J. Am. Chem. Soc.*, **1987**, 109, p. 908.
- [15] Mukherjee D., Wu Y., Fronczek F., Houk K., *J. Am. Chem. Soc.*, **1988**, 110, p. 3328.
- [16] Karplus M., Petsko G., Molecular dynamics simulations in biology, *Nature*, **1990**, 347, p. 631.
- [17] DeKock R., Madura J., Rioux F., Casanova J., Computational chemistry in the undergraduate curriculum, dans *Reviews in Computational Chemistry*, vol. 4, Boyd, Lipkowitz Ed., VCH, New York, **1993**.
- [18] Schlick T., Optimisation methods in computational chemistry, dans *Reviews in Computational Chemistry*, vol. 3, Boyd, Lipkowitz Ed., VCH, New York, **1992**.
- [19] Scheraga H., Predicting three dimensional structures of proteins, dans *Reviews in Computational Chemistry*, vol. 3, Boyd, Lipkowitz Ed., VCH, New York, **1992**.
- [20] Pastor M., Cruciani G., A novel strategy for improving ligand selectivity in receptor based drug design, *J. Med. Chem.*, **1995**, 38, p. 4637.

Publication du Recueil des olympiades de la chimie

Du fait de l'abondance des articles, la publication du Recueil des Olympiades nationales de la chimie n'est pas assurée dans ce numéro.

Nous prions nos lecteurs de nous en excuser.

Cette publication sera reprise dans le numéro 8-9 d'août-septembre 1997.