

M/ 5047-90/01

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE  
MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE  
SCIENTIFIQUE



UNIVERSITE ABOU-BAKR BELKAID DE TLEMCCEN

FACULTE DES SCIENCES – DEPARTEMENT DE CHIMIE

LABORATOIRE DE CHIMIE ORGANIQUE, SUBSTANCES NATURELLES ET ANALYSES  
(COSNA)

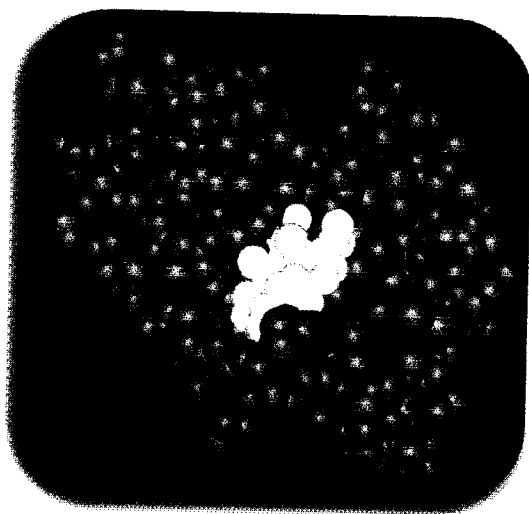
MEMOIRE

PRESENTE A L'UNIVERSITE ABOU-BAKR BELKAID DE TLEMCCEN POUR  
L'OBTENTION DU DIPLOME DE : MASTER

Spécialité : Chimie Bio-organique et thérapeutique

Présenté Par :

M<sup>r</sup> MISSOUM Noureddine



Thème :

# MODELISATION DE L'INTERACTION ENZYME-SUBSTRAT

Soutenue publiquement le : 10 juin 2009 devant le jury composé de :

- |  |            |
|--|------------|
| Mr H. ALLALI, MC, Université Abou Bekr BelKaid-Tlemcen,        | Président. |
| Mr B. TABTI, Pr, Université Abou Bekr BelKaid-Tlemcen,         | Examineur. |
| Mr J.KAJIMA MULENGI, Pr, Université Abou Bekr BelKaid-Tlemcen, | Examineur  |
| Mr O. BENSALD Pr, Université Abou Bekr BelKaid-Tlemcen,        | Examineur. |
| Mr A. ATMANI, MC, Université Abou Bekr BelKaid-Tlemcen,        | Examineur. |
| Mr Z. ARRAR, MC, Université Abou Bekr BelKaid-Tlemcen,         | Examineur. |
| Mr S. GHALEM, Pr, Université Abou Bekr BelKaid-Tlemcen,        | Encadreur. |



Inscrit le 24 SEP. 2009  
Date le  
Code: 36 67

## DEDICACES

*Je dédie ce travail à :*

Inscrit Sous le N°  
Date le 13 DEC. 2016  
Code: 1993

### Mes parents :

Mes très chers parents pour leurs dévouements, leurs amours, leurs sacrifices, leurs encouragements et leur bénédiction durant toute ma vie. Que ce travail soit, pour eux, un faible témoignage de ma profonde affection et tendresse.

Et aussi Mes très chers grands parents qui m'ont beaucoup soutenu et encourager par leurs présence en moi et leurs bénédiction pour accomplir ce modeste travail.

Mes frères : Amine, Sidahmed, Rachid, et Kheireddine.

Mes neveux : Aymene et Louaye.

La famille : MISSOUM, MAROUF, BENAI ET BENMOUSSAT.

Mes ami(e)s et Toutes les personnes que j'aime.

# Remerciements

*Le présent travail a été réalisé au sein de laboratoire COSNA (Chimie Organique, Substances Naturelles et Analyses) dans l'équipe de modélisation moléculaire de la Faculté des Sciences, Département de Chimie, de l'Université de Tlemcen sous la direction de Monsieur Pr. Said GHALEM.*

*Mes plus vifs remerciements vont à Monsieur le Professeur « GHALEM Said », de m'avoir accueilli dans son équipe, pour m'avoir fait confiance, pour son soutien et qui a su me laisser la liberté nécessaire à l'accomplissement de mes travaux, tout en y gardant un oeil critique et avisé. Son aide, lors de la rédaction de ce mémoire, a été particulièrement précieuse, ses propositions et corrections ont aidé à améliorer significativement la qualité de ce mémoire qui a permis cette délivrance formidable : la soutenance. J'espère que ce mémoire sera un remerciement suffisant au soutien et à la confiance dont il a fait preuve à mon égard.*

*J'exprime ma profonde reconnaissance à Monsieur H. ALLALI,, Maître de conférence à l'université de Tlemcen et COSNA, qui m'a fait l'honneur de présider le jury.*

*J'exprime également mes remerciements à Messieurs : Mr. B. TABTI, Pr. Mr. O. BENSALID, Pr. Mr. A. ATMANI, MC. Mr. Z. ARRAR, MC. et Mr Pr. J.KAJIMA MULENGI, Pr. à l'université Aboubekr Belkaid de Tlemcen qui ont accepté d'examiner ce travail et de participer au membre de jury.*

*J'exprime ma profonde reconnaissance à M<sup>elle</sup> I. ABDELLI pour sa disponibilité, ses conseils et ses nombreuses discussions qui m'ont aidé tout au long de ce stage.*

*Je tiens à remercier toute l'équipe de modélisation moléculaire et surtout I. DAOUD pour sa grande disponibilité, et son soutien.*

*Un grand merci à tous mes camarades de la promo C.B.O.T: KENICHE Assia, CHOUIKHI Dalila, BENSABEUR Khadra, YOUNES Kawter, MEZRAI Abdelmoumin, MESMOUDI Morad, CHIKHI Ilyes, MELAH Ilyes, et ZIRAR Mohamed qui ont contribué à créer une atmosphère conviviale et détendue, et qui m'ont toujours accordé tout au long de ce parcours un regard amical, un soutien moral, et la joie d'un sourire.*

*Je n'oublierai pas de remercier tous les étudiants en post-graduation à COSNA pour leur encouragement et leur amitié, et à tous ceux qui ont contribué de près ou de loin à la réalisation de ce mémoire.*

*Enfin j'adresse tous mes chaleureux remerciements aux enseignants qui ont contribué à ma formation, pour leur dévouement et abnégation.*

# TABLE DES MATIERES

<b>INTRODUCTION GENERALE</b> .....	1
Références .....	5
<b>CHAPITRE I : LES METHODES DE LA MODELISATION MOLECULAIRE</b>	
Introduction .....	6
I. Mécanique Moléculaire .....	7
I.1. Champ de force .....	8
I.2. Différents champs de force .....	10
II. Dynamique Moléculaire.....	12
II.1. Principe.....	12
II.2. Calcul de dynamique moléculaire .....	14
III. Docking moléculaire (Arrimage).....	16
Références.....	18
<b>CHAPITRE II : ANALYSE CONFORMATIONNELLE</b>	
I. Minimisation de l'énergie stérique .....	21
II. Surface d'énergie potentiel .....	22
III. Analyse conformationnelle .....	23
IV. Algorithme de minimisation.....	26
IV.1 La methode de steepest descent	26
IV.2. Méthode du gradient conjugué.....	26
IV.3. Méthode de Newton Raphson.....	26
IV.4. Méthode de simplexe.....	27
IV.5. Méthode de recuit simulé.....	27
Références.....	28
<b>CHAPITRE III : ACIDES- AMINES, PROTEINES, ENZYMES</b>	
I. Acides aminés .....	30
I.1. Propriétés physiques.....	33
I.2 Propriétés chimiques.....	33
II. Peptides.....	34
II.1. Structure de la liaison peptidique.....	34
II.2. Propriétés physiques.....	35
II.3. Propriétés chimiques.....	35
III. Protéines.....	36
III.1. Structure des protéines.....	37

III.2. Repliement des protéines.....	40
III.3. Fonctions des protéines.....	41
III.4. Pourquoi modéliser les protéines ?.....	41
IV. Enzymes.....	41
IV.1. Introduction.....	41
IV.2. Site actif.....	42
IV.4. Cinétique enzymatique.....	43
IV.5. Comment fonctionnent les enzymes ?.....	45
IV.6. Catalyse enzymatique.....	46
IV.7. Complexe enzyme- substrat (E-S).....	47
IV.8. Spécificité de l'association protéine- ligand.....	48
Références.....	49

## CHAPITRE IV : APPLICATION, RESULTATS ET DISCUSSIONS

Introduction.....	50
I. Préparation de l'enzyme.....	52
II. Preparation des ligands.....	54
III. Minimisation de l'énergie.....	55
III.1. Optimisation de la CTPR 3.....	55
III.2. Optimisation des Substrats .....	56
IV. Dynamique moléculaire.....	57
IV.1. Dynamique moléculaire de la CTPR 3.....	57
IV.2. Dynamique moléculaire des substrats.....	59
V. Docking des substrats et construction des complexe.....	61
V.I Dynamique moléculaire des complexe .....	61
V.II Energie d'interaction.....	64
Références.....	66
CONCLUSION GENERALE.....	67

# *Introduction Générale*

## INTRODUCTION GÉNÉRALE

---

Toute vie dépend du bon fonctionnement d'une grande variété de catalyseurs organiques ou enzymes. Les structures biologiques correspondent à des fonctions bien définies et la vie d'un organisme dépend du rôle des structures biologiques, des membres aux divers organes, et jusqu'aux agents chimiques du métabolisme, comme les enzymes.

Les enzymes ont fait l'objet de recherches intensives ayant permis de dévoiler les principes du fonctionnement cellulaire. Leur importance s'est encore accrue avec le lancement de projets de séquençage des génomes à grande échelle. De plus, leur intérêt va bien au delà de la collecte de connaissances fondamentales puisqu'elles trouvent des applications dans des domaines variés tant biomédical (réactifs, kits de diagnostic, méthodes thérapeutiques) qu'industriel (mise en œuvre de technologies à base d'enzymes dans l'industrie agro-alimentaire).

La caractéristique la plus évidente des organismes vivants est qu'ils sont complexes et extrêmement organisés. Par exemple, les organismes assez grands pour être visibles à l'œil nu sont formés de nombreuses cellules, le plus souvent de types différents. A leur tour, ces cellules contiennent des structures subcellulaires, qui sont des assemblages complexes de très grosses molécules. Ces macromolécules elles-mêmes ont un remarquable degré d'organisation avec une structure tridimensionnelle complexe, bien qu'elles soient constituées d'un petit nombre d'éléments chimiques comme les sucres et les acides aminés.

En fait, la structure tridimensionnelle complexe caractéristique des macromolécules résulte des interactions entre ses constituants, en fonction de leurs propriétés chimiques individuelles [1].

Le traitement informatique des processus enzymatiques s'est développé spectaculairement dans les deux dernières décennies [2]. Bien que nous sommes loin de pouvoir exécuter des calculs quantitatifs avec précision, le graphisme moléculaire, la dynamique moléculaire et la mécanique quantique fournissent de nouvelles informations essentielles non disponibles par les expériences les plus sophistiquées. Mais la meilleure manière d'obtenir de bons aspects structuraux des processus enzymatiques est de combiner les méthodes expérimentales et informatiques. Bien que les données primaires soient obtenues par la cristallographie des rayons X ou par spectroscopie RMN, l'importance de cette tendance a été déjà soulignée il y a plus de vingt ans [3]. Le développement rapide des ordinateurs et des méthodes informatiques permet d'étudier des systèmes entre 50-100 atomes par mécanique quantique et jusqu'à 50 000 atomes par mécanique moléculaire [4].



# INTRODUCTION GÉNÉRALE

---

## I. MODELISATION MOLECULAIRE

Les chercheurs en chimie et en biologie, font de plus en plus largement appel à des outils qui leur permettent de prévoir avec toujours plus de finesse les effets des molécules en fonction de leurs structures et, ce qui est encore plus intéressant pour les concepteurs, d'imaginer la structure de nouveaux édifices qui est susceptible d'induire un comportement donné et de produire l'effet recherché.

Dès 1950, l'outil informatique a été mis en œuvre dans le domaine chimique dans le but de modéliser les structures moléculaires. Depuis cette époque l'augmentation de la puissance de calcul des ordinateurs a permis à la modélisation moléculaire de s'intéresser à des structures de plus en plus complexes d'une manière plus rigoureuse et réaliste. Ainsi, la modélisation est devenue une nouvelle technique de compréhension des phénomènes chimiques et de calcul des surfaces d'énergie potentielle des molécules organiques [5].

Le développement important au cours de ces dernières années des moyens informatiques (mémoire plus importante, superordinateurs plus puissants, échange d'information grâce à Internet,...) a permis à la chimie de s'enrichir d'outils informatiques spécialement destinés à la représentation des molécules.

Aujourd'hui, la modélisation moléculaire nous permet de traiter des systèmes contenant plus d'atomes (molécules de solvant explicite pouvant atteindre des dizaines de milliers d'atomes, grandes surfaces moléculaires et protéines). De ce fait, un phénomène de compétition entre la taille des systèmes moléculaires et la puissance des superordinateurs est apparu [6].

Contrairement à ses débuts, la modélisation moléculaire est aujourd'hui mieux acceptée. D'ailleurs, en 1998, le prix Nobel de chimie a été décerné à John. A. Pople et Walter Kohn pour leurs travaux dans les domaines de la chimie informatique et de la modélisation moléculaire (chimie quantique).

La modélisation moléculaire se sert, pour simuler des systèmes moléculaires, des méthodes de calculs théoriques telles que la dynamique moléculaire (DM), la mécanique moléculaire (MM).

Les molécules peuvent adopter de nombreuses conformations dont certaines correspondent à l'interaction avec un site actif ou au comportement en solution. La stabilité relative de ces conformations peut être évaluée par le calcul de l'énergie potentielle qui leur

## INTRODUCTION GÉNÉRALE

---

est associée. Trois méthodes principales de calcul de l'énergie sont mises en œuvre, elles sont généralement couplées à des méthodes de recherche de conformations et des minimiseurs mathématiques permettant de calculer les minima de la fonction d'énergie.

### **I-1. LA MECANIQUE MOLECULAIRE :**

L'énergie conformationnelle est la sommation de termes d'interactions évaluées à partir de tables pré calculées (les champs de force) contenant les constantes d'élongation, de déformation, de torsion, d'attraction et répulsion électrostatique. C'est la méthode la plus fréquemment retenue car elle produit les temps de calcul les plus courts.

### **I-2. LA DYNAMIQUE MOLECULAIRE :**

Elle prend en compte l'effet de la température, de la pression et du temps, calcule les trajectoires des atomes et approche la flexibilité des molécules, leur comportement en solution ou les modes de vibration.

L'objet de ce mémoire est d'étudier les interactions de la tetratricopeptide (CTPR 3) avec un inhibiteur synthétisé, par des méthodes non-quantiques comme la mécanique moléculaire et la dynamique moléculaire. et un deuxième que nous avons conçu comme modèle pour pouvons comparer les stabilités et par suite l'efficacité vis-à-vis de la CTPR 3

Un bref aperçu sur la CTPR 3 sera exposé dans un premier temps.

## **II. TETRATRICOPEPTIDE 3 (CTPR 3) ET L'INHIBITEUR**

La séquence de tetratricopeptide (TPR) a été identifiée en 1990 [7] dans les gènes (CDC23). C'est une structure hélicoïdale composée de 34 acides aminés qui se produit dans plus de 300 protéines différentes et qui sont associés à une gamme diversifiée de fonctions biologiques. Trois protéines ont été construites avec un, deux, et trois copies ou plus du motif de TPR (CTPR1, CTPR2 et CTPR3).

La CTPR3 qui fait l'objet de notre travail elle est constituée de sept motif de TPR son code PDB est le « 1na0 » elle est cristallisée se forme d'un dimère [8].

C'est une enzyme dotée de différentes activités biologiques selon le changement de ces motifs, impliquée notamment dans les mécanismes de cancérisation et les infections virales.

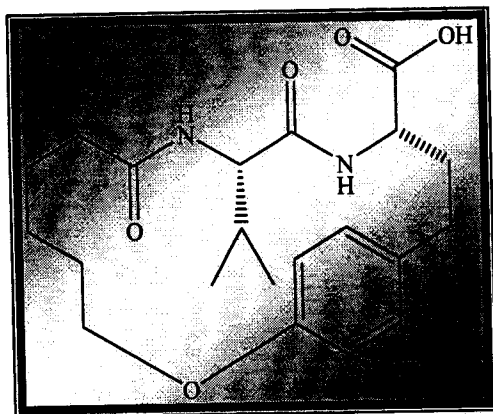
## INTRODUCTION GÉNÉRALE

Les inhibiteurs étudiés dans ce mémoire ce sont des peptides cycliques ils sont considérés comme des inhibiteurs de protéase [9].

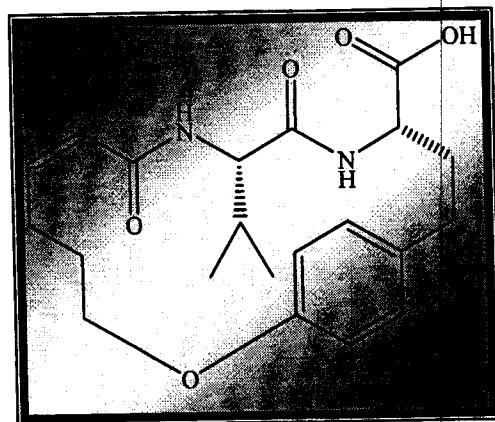
La protéase est une enzyme qui catalyse les liaisons amides dans les peptides et les protéines [10]. Elle joue un rôle très importants dans plusieurs maladies tel que les maladies cardiovasculaires [11], le désordre neurodégénératif [12], les infections virales [13] et cancer [14].

Par exemple, beaucoup de tumeurs ont des niveaux élevés des protéases à toutes les étapes de progression et de métastase de cancer, qui font à ces enzymes les cibles attrayantes pour des thérapies anticancéreuses [15].

D'où le but de notre travail qui consiste a étudiée l'interaction entre la CTPR 3 et les deux substrats représentés ci-dessous :



**Substrat 01 (Synthétisé)**



**Substrat 02 (Conçus)**

## INTRODUCTION GÉNÉRALE

---

### REFERENCES:

- [1] Reginald H. GARRETT, Charles M. GRISHAM, *Biochimie*, De Boeck University s.a. 2000, pp.3.
- [2] L.A. Eriksson (Eds.), *Theoretical biochemistry*, Processes and Properties of Biological Systems, Elsevier, Amsterdam, 2001.
- [3] G. Náray- Szabó (Eds.), Steric effects in biomolecules, Proceedings of International Symposium, Eger, Akadémiai Kiado/ Elsevier, Budapest/ Amsterdam, 1981, pp.1982.
- [4] Gábor Náray-Szabó, Imre Berente, *Journal of Molecular Structure (Theochem)*, Computer Modeling of Enzyme Reaction, 666-667, 2003, pp.637-644 .
- [5] S. D. Ittle, L.K. Johnson, M. Brookhart, *Chem. Rev*, 2000, 100, pp.1169.
- [6] W. Van Gunsteren, Berendsen, H. J. C. *Angew. Chem. Int. Ed. Engl.* 1990, 29, pp. 992.
- [7] T. Hirano N. Kinoshita Snap helix with knob and hole: essential repeats in S . pombe nuclear protein nuc2+. 1990. 60, 319
- [8] E. Main Y.Xiong, Design of stable  $\alpha$ -helical Arrays from an idealized TPR motif, *Elsevier science Ltd.* 2003. 11, 497
- [9] X. Fengtian, T. Christopher. Structure- activity studies of cyclique ketone inhibitors of the serine protease plasmin: Design, synthesis, and biological activity. *Bioorganic and Medicinal Chemistry.* 2006 , 14, 8467
- [10] C. Seife, *Science.* 1997, 277, 1602
- [11] M. Stubbs; W. Bode, *thrombi. Res.* 1993, 69, 1.
- [12] D. Selkoe ; J.Science. 1997, 275, 630.
- [13] W. Gibson; M. Hall. *Drug des. Disco.* 1997. 15, 39.
- [14] S. Yan, M. Sameni; B . Solane; *J. Biol. Chem* 1998. 379, 113
- [15] J. Rakic , C. Maillard; M. Jost; *Cell. Mol. Life Sci.* 2003, 60, 463.

# Chapitre I :

## *Méthodes de la Modélisation Moléculaire*

### Introduction

La recherche et la synthèse de nouveaux composés chimiques et biochimiques sont aujourd'hui souvent associées à une étude par modélisation moléculaire.

Grâce au développement informatique de ces dernières années et à l'essor du calcul parallèle intensif en particulier, la modélisation moléculaire est devenue un véritable enjeu. En effet les systèmes moléculaires qui sont amenés à être étudiés tendent à devenir de plus en plus complexes. Cette complexité non seulement liée à la taille des molécules envisagées (plusieurs centaines de milliers d'atomes pour les molécules biologiques par exemple) ainsi qu'à la structure intrinsèque des atomes eux-mêmes, mais aussi au degré de précision exigé pour le calcul de certaines grandeurs physiques [1].

La modélisation moléculaire est un terme général qui englobe différentes techniques de graphisme moléculaire et de chimie computationnelle permettant de dessiner, afficher, simuler, analyser, calculer et stocker les propriétés des molécules. En outre, elle permet d'investiguer les changements de conformations liées à des mutations à partir de structures 3D expérimentales. Le comportement des molécules peut être simulé de façon statique ou dynamique. Cette dernière inclut notamment le comportement et la flexibilité des molécules en solution, les mécanismes réactionnels et les interactions mises en jeu dans la catalyse enzymatique, la reconnaissance moléculaire et la complexation.

La modélisation moléculaire est une application des méthodes théoriques et des méthodes de calcul pour résoudre des problèmes impliquant la structure moléculaire et la réactivité chimique [2]. Ces méthodes peuvent être relativement simples et utilisables rapidement ou au contraire, elles peuvent être extrêmement complexes et demandaient des centaines d'heures de temps d'ordinateur, même sur un super-ordinateur. En plus, ces méthodes utilisent souvent des moyens infographiques très sophistiqués qui facilitent grandement la transformation de quantités impressionnantes de nombres en quelques représentations graphiques facilement interprétables [3].

### I. Mécanique moléculaire

L'expression « Mécanique moléculaire » (MM) désigne actuellement une méthode de calcul qui permet, d'obtenir des résultats de géométrie d'énergies moléculaires en se basant sur la mécanique classique.

La mécanique moléculaire est apparue en 1930 [4], mais s'est développée à partir des années soixante, quand les ordinateurs furent plus accessibles et plus performants.

Le concept est apparu, 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 ordinateur étaient réalisées. C'est que la MM 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 la fréquence d'absorption IR pouvait correspondre à la vibration d'une liaison particulière, d'une molécule à une autre [5].

La mécanique moléculaire est une méthode non quantique, mais elle a un intérêt pour les grands systèmes, comme dans le cas des systèmes biologiques qu'on ne peut aborder avec les méthodes quantiques [6]. La mécanique moléculaire, appelée parfois « calcul par champ de force empirique » [7-8], nous permet de minimiser l'énergie afin d'obtenir des conformations à basse énergie de notre système moléculaire et de réduire des forces initiales trop grandes qui aboutiront à une trajectoire aberrante [9].

La mécanique moléculaire ressemble aux modèles de type « tiges et boules ». Elle a pour but de calculer l'énergie potentielle d'une molécule (ou d'un système de molécules) en fonction de coordonnées d'atomes :

$$E_p = f(r_1, r_2, \dots, r_n)$$

Où  $r$  représente le vecteur position de l'atome  $i$ .

La mécanique moléculaire utilise les approximations suivantes :

- chaque atome constitue une particule ;
- l'atome est considéré comme une sphère rigide ayant un rayon et une charge déterminée ;

- les énergies sont calculées par des formules dérivées de la mécanique classique [2].

Un calcul de mécanique moléculaire aboutit à une disposition des noyaux telles que la somme de toutes les contributions énergétiques est minimisée ; ses résultats concernent surtout la géométrie et l'énergie du système. L'idée directrice de cette méthode est d'établir, par le choix des fonctions énergétiques et des paramètres qu'elles contiennent, un modèle mathématique, le « **CHAMP DE FORCE** », qui représente aussi bien que possible les variations de l'énergie potentielle avec la géométrie moléculaire [10].

### I. 1. Champ de force

Compte tenu de la taille des systèmes étudiés en biologie, l'utilisation de la mécanique quantique n'est pas possible (voir exemple d'un temps de calcul SCF d'un peptide).

Les macromolécules sont représentées comme un ensemble d'atomes ponctuels dont les interactions sont décrites par un potentiel semi-empirique ou champ de force.

On appelle champ de force le modèle mathématique représentant l'énergie potentielle d'une molécule en mécanique moléculaire.

Le champ de force exprime réellement la moyenne des interactions électroniques entre les atomes [11].

Le champ de force permet d'accéder à l'hypersurface énergétique d'une molécule en établissant un lien entre les déformations structurales du système et son énergie potentielle. Il désigne à la fois l'équation mathématique (fonction d'énergie potentielle) et les paramètres qui la composent [12]. La fonction d'énergie potentielle définit une énergie empirique, l'énergie totale étant décomposée en une somme de termes additifs représentant chacun des interactions inter atomiques. Elle est exprimée comme une somme de contributions de plusieurs types d'interaction [13-14]. Elle peut se décomposer en terme d'interaction intramoléculaire et terme d'interaction intermoléculaire.

Les interactions intramoléculaires ne dépendent que des coordonnées internes des molécules c'est-à-dire des liaisons, des angles de valence, et de torsions. En fait, pour affiner l'expression du terme potentiel est rendre plus fidèle la description du système, des termes de



## CHAPITRE I : Méthode de la modélisation moléculaire

couplages entre différents atomes ont été introduit. Le potentiel intramoléculaire peut s'écrire de la façon suivante :

$$V_{\text{intramoléculaire}} = \sum_{\text{liaison}} V_{\text{élongation}} + \sum_{\text{angles}} V_{\text{courbure}} + \sum_{\text{angles dièdres}} V_{\text{torsion}} + \sum V_{\text{croisé}}$$



Les interactions intermoléculaires prennent en compte les interactions qui n'interagissent pas par des termes de liaison, d'angle de courbure et d'angle de torsion. Le potentiel non liant s'exprime en deux termes : un terme de Van der Walls et un terme d'énergie électrostatique. On a donc ;

$$V_{\text{intermoléculaire}} = \sum_{\text{atomes non liés}} V_{\text{van der Walls}} + \sum_{\text{atomes non liés}} V_{\text{électrostatique}}$$

Donc l'énergie stérique s'exprime par l'équation suivante :

$$E = E_{\text{stretching}} + E_{\text{bending}} + E_{\text{torsion}} + E_{\text{VdW}} + E_{\text{elec}} + E_{\text{hydrogen}}$$

Le terme « Stretching » représente l'élongation des liaisons.

Le terme « Bending » représente la variation des angles.

Le terme « Torsion » se rapporte à l'énergie de torsion des angles dièdres.

Le terme « VdW » rend compte des énergies d'interaction non covalente entre atomes non liés.

Le terme « Elec » rend compte des énergies d'interaction électrostatiques entre atomes non liés.

Le terme « Hydrogen » rend compte des liaisons d'hydrogènes.

Il est important de noter que les champs de forces constituent une approche purement empirique. Ils sont paramétrés de manière à ce que l'ensemble des différentes contributions permette de simuler une série de résultats expérimentaux. Un champ doit donc toujours être considéré comme une entité indivisible et il est exclu de combiner les termes et/ou les paramètres de deux champs dans l'espoir d'obtenir un meilleur ensemble [15]. En outre, le paramétrage vise souvent à ce que le champ de forces puisse traiter de manière réaliste une catégorie particulière de composés. On peut ainsi trouver des champs destinés plus spécialement à la modélisation de petites molécules organiques, de macromolécules, de complexes organométallique [16], ou encore de nucléotides [17]. Il est donc dans un premier temps important de vérifier si le champ choisi est adapté au système à étudier.

### I. 2. Différents champs de force en mécanique moléculaire

Différents champs de force utilisent le même type de termes énergétiques mais paramétrés de manières différentes. Les champs de forces en MM peuvent être groupés en trois classes principales [18] :

- Champs de force contenant uniquement les termes harmoniques.
- Champs de force utilisant les termes d'ordre supérieur (cubique, quadratique,...).
- Champs de force suggérés par *Allinger et col* [19] ne considérant pas que les termes de la mécanique moléculaire classique mais aussi les effets chimiques comme l'électronégativité.

#### - MM2/MM3/MM4 :

MM2 est le premier champ de force développé par Allinger et col. [20-21]. Il a été conçu au début pour les molécules simples (alcane, alcènes, alcynes non conjugués, amines...), mais ses versions améliorées MM3 (1989) [22] et MM4 (1996) [23] lui permet de traiter des molécules organiques de plus en plus complexes.

- **OPLS :**

Le programme OPLS (Optimized Potentials for Liquid Simulations), comme l'indique son nom, est conçu pour optimiser le potentiel qui permet la description des propriétés de solvation. Il est développé par W. L. Jorgensen et J. Tirado Rives [24].

- **GROMOS :**

GROMOS (Groningen Molecular Simulation Program Package), est développé par Van Gunsteren [25] et conçu spécialement pour les biomolécules en milieu aqueux en vue de l'étude des interactions entre les molécules d'eau et les groupements polaires des protéines.

- **CHARM (Bio+) :**

Développé par Karplus et col [26-27], pour le calcul de biomolécules. Son concept est semblable à celui d'AMBER. Bien qu'au début, ce champ de force conçu pour les acides aminés et les protéines, maintenant il traite d'autres biomolécules.

- **EMO :**

Le programme Emo (Energy Of Molecule), est développé par B. Blaise [28-31], il est basé sur le champ de force MM2. Il est articulé autour de trois menus :

- Menu 1 : entrée de la molécule à l'aide d'un clavier.
- Menu 2 : manipulation géométrique sur la molécule.
- Menu 3 : minimisation de l'énergie par mécanique moléculaire.

- **AMBER :**

AMBER (Assisted Model Building with Energy Refinement), a été écrit par Kollman [32]. Le champ est paramétré pour les protéines et les acides nucléiques (UCSF, 1994). Il a été utilisé pour les polymères et pour d'autres petites molécules. Dans le cadre de ce travail, nous utiliserons ce champ de force, puisque on va traiter les protéines.

- **SPASIBA :**

(Spectroscopic Potential Algorithm for Simulating biomolecular conformational Adaptability), élaboré par Gérard Vergoten et col. (1995). Il combine le champ de force spectroscopique modifié de Urey-Bradley-Shimanouchi [33] et le champ de force AMBER. Il permet de trouver les structures, les énergies de conformation et les fréquences vibrationnelles au minimum énergétique d'une molécule [34].

## II. Dynamique moléculaire

La Dynamique Moléculaire simule, à partir des lois de la mécanique classique, les trajectoires des atomes en phase cristalline, en solution ou en phase gazeuse [35-36]. Elle donne l'évolution d'un système dans l'espace conformationnel au cours du temps. Une résolution numérique des équations classiques du mouvement est utilisée pour calculer l'énergie du système le long de son parcours sur la surface de potentiel du système.

### II. 1. Principe

Les simulations de dynamique moléculaire consistent à calculer les positions et les vitesses d'un système d'atomes [37].

En dynamique moléculaire classique [38-39], chaque atome de la molécule est considéré comme une masse ponctuelle dont le mouvement est déterminé par l'ensemble des forces exercées sur lui par les autres atomes en fonction du temps.

A une température donnée, une vitesse initiale est attribuée à chaque atome. A partir de ces vitesses, considérées constantes, la position des atomes change en  $v_i \delta t$ , ou  $v_i$  est la vitesse de l'atome  $i$  et  $\delta t$  est l'incrément de temps. Avec ces nouvelles coordonnées, on détermine les forces qui agissent sur chaque atome (équation 1) selon le champ de force et le potentiel utilisé, et selon la seconde loi de Newton, on attribue les nouvelles vitesses et positions.

$$\vec{F}_i = m_i \vec{a}_i = m_i \frac{d^2 \vec{r}_i(t)}{dt^2} \quad (1)$$

Ou  $\vec{F}_i$  : Vecteur force agissant sur l'atome  $i$ .

$m_i$  : Masse de l'atome  $i$ .

$\vec{a}_i$  : Vecteur accélération de l'atome  $i$ .

$\vec{r}_i$  : La position de l'atome  $i$ .

Comme la montre l'équation de Newton, la vitesse et le sens du mouvement atomique dépendent des forces qui s'exercent sur les atomes. En pratique, les atomes possèdent une vitesse initiale propre qui se conforme à l'énergie cinétique totale du système. Cette énergie va dépendre de la température de simulation étudiée.

Tout en simulant une montée en température ; on fournit une énergie cinétique aux atomes correspondant à une certaine température. Cette énergie cinétique autorise le système à franchir des barrières de potentiel que les autres minimiseurs ne peuvent franchir.

Des vitesses  $v$  sont assignées aux atomes par tirage aléatoire suivant une distribution de Maxwell :

$$E_c = \frac{1}{2} \sum m_i \langle v_i \rangle^2 = \frac{3}{2} NK_b T$$

Ou  $\langle v_i \rangle^2$  : valeur du moyenne carré de la vitesse.

$N$  : nombre d'atomes du système.

$K_b$  : Constante de Boltzmann.

$T$  : température absolue.

On peut alors démarrer une simulation, pour cela on calcule

\* les accélérations  $a_i$  de l'atome  $i$  suivant la loi de Newton, (1) avec  $F_i$  : force appliquée sur l'atome  $i$  est calculée selon la relation :  $\vec{F}_i = -dE_p/r_i(t)$

\* Les vitesses de chaque atome à partir de la connaissance des accélérations :

$$\vec{a}_i = d\vec{v}_i/dt$$

\* les positions atomiques à partir des vitesses atomiques :  $\vec{v}_i = d\vec{r}_i/dt$

Des intégrations successives de ces équations sont faites en subdivisant la trajectoire en une série d'état discret séparés par des intervalles de temps très courts, ce qui conduit à une trajectoire en fonction du temps.

Connaissant la vitesse et l'accélération de l'atome  $i$  à l'instant  $t$  on peut connaître sa position à l'instant  $t + \Delta t$  :

$\Delta t$  : Pas d'intégration ( $\approx 10^{-15} s$ )

$$r_i(t + \Delta t) = r_i(t) + v_i \cdot \Delta t + 1/2 a_i \cdot \Delta t^2 \quad (2)$$

### II. 2. Calcul de dynamique moléculaire

Une dynamique moléculaire se compose de quatre phases distinctes :

#### *Initialisation :*

La dynamique moléculaire requiert un jeu de coordonnées et vitesses initiales. La géométrie de départ utilisée provient d'une structure établie expérimentalement lorsque c'est possible, si non elle générée à partir du champ de force utilisé. L'énergie du système est minimisée afin d'éliminer toutes les interactions fortement défavorables, qui, si elles étaient conservées, perturberaient la simulation. Les vitesses initiales sont attribuées selon la statistique de Boltzmann.

#### *Thermalisation :*

Il s'agit de chauffer le système pour l'amener à la température souhaitée.

On utilise comme structure initiale la structure minimisée.

On tire aléatoirement le jeu de vitesse initiale, soit à zéro à des vitesses correspondant à quelques dizaines de Kelvin.

On intègre ainsi les équations du mouvement en multipliant régulièrement les vitesses afin d'augmenter la température jusqu'à la température de référence (généralement 300K).

Une montée en température demande en général 2 à 10ps de simulation (200 à 1000 pas d'intégration  $\Delta t$ ).

#### *Équilibration :*

C'est une phase importante au cours de laquelle on stabilise la température du système, il y a alors un échange important entre énergie potentielle et énergie cinétique.

C'est une phase où on contrôle régulièrement la température que l'on ramène dans la fenêtre souhaitée. La température de déviation se fait suivant un processus de relaxation où la température est recalculée après chaque étape et ramenée à la température de référence par la formule :

$$d/dtT = \tau^{-1} (T_0 - T)$$

Avec  $\tau$  : temps de relaxation.

$T_0$  : Température de référence.

$T$  : Température du système.

Durant la phase d'équilibration la valeur de  $\tau$  doit être assez petite pour permettre un couplage efficace (généralement  $\tau = 10$  fs)

Cette phase peut durer de 10 à 20 ps.

### ***Dynamique ou production :***

C'est la phase réellement exploitable qui dure en moyenne de 60 à 100 ps.

Au cours de cette simulation, on doit sauvegarder les conformations, les énergies... toutes les 0.05 ou 0.1 ps.

La dynamique la plus classique se fait à température constante avec un scaling des vitesses, mais un temps de relaxation plus grand (0.1 à 0.4 ps) pour permettre d'évaluer l'équilibre propre du système.

Il est conseillé d'enlever les contraintes ( $\Delta t = 2$  fs) durant cette phase de dynamique.

### III. Docking moléculaire (Arrimage moléculaire)

Connaitre la structure en trois dimensions des macromolécules et comprendre leurs mécanismes de compléxation est fondamental pour la compréhension des systèmes biologiques, et donc essentiel dans de nombreux domaines.

Dans ce but, la chimie apporte sa collaboration à la biologie structurale qui s'intéresse quant à elle au rapport qui existe entre la structure des molécules et leur fonction biologique.

Cette union débouche sur des avancées couvrant différentes applications, telles que :

- La recherche pharmacologique (conception de médicaments en trouvant un ligand, c'est-à-dire une molécule médicamenteuse complémentaire, à assembler avec un récepteur membranaire connu).
- La biologie cellulaire (interaction entre macromolécules, protéines ou acides nucléiques, intervenant dans les cycles cellulaires)
- La recherche médicale (élaboration de nouveaux traitements visant les cellules une à une ...)
- Les nanobiotechnologies (conception de puces à ADN, de biocapteurs ...)

Cependant les outils de modélisation moléculaire sont particulièrement performants pour l'étude au niveau atomique les interactions entre deux entités moléculaires et ce type d'étude fait appel le plus souvent aux techniques de « DOCKING ».

Le docking vise à prédire la structure d'un complexe formé par deux molécules, la formation de ces complexes est fondée sur la reconnaissance de la structure tridimensionnelle d'un ligand par un site récepteur et contrôle de l'activité de nombreuses molécules.

Le plus souvent, le récepteur est une protéine qui possède un ou plusieurs sites actifs spécifiques, plus ou moins accessibles selon les cas. Le ligand, est généralement une molécule flexible étrangère de petite taille. La liaison entre le ligand et le récepteur introduit un signal biologique qui peut avoir diverses formes [40].

Par exemple, certains récepteurs sont des protéines membranaires permettant le passage sélectif de certaines molécules (ions, nutriments) à l'intérieur de la cellule, d'autres sont des enzymes dont l'activité est fonction de l'affinité spécifique avec un substrat . . .



Le Docking moléculaire a pour but de déterminer le mode d'interaction d'un complexe formé de deux ou plusieurs molécules, en cherchant des orientations dans l'espace et des conformations favorables pour la fixation d'un ligand à un récepteur [41].

Toutes les méthodes de « Docking moléculaire » peuvent être décomposées en trois étapes :

- Caractérisation du récepteur et l'identification du site de fixation.
- L'orientation du ligand dans le site de récepteur.

L'évaluation de l'orientation du ligand et de son adéquation au récepteur [42].

**REFERENCES**

- [1] C.Audouze, Vers une parallélisation par bandes en chimie quantique, *Laboratoire de Mathématique, UMR CNRS 8628, Université Paris-Sud*, 2003, 1-8.
- [2] D.Liotta, *Advances in Molecular Modeling*, Ed. JAI Press, 1988, p.1.
- [3] J.Debord, *Introduction à la modélisation moléculaire*, 2004, p.37-41.
- [4] D.H.Andrews, *Phys.Rev*, 1930, p.36, 544.
- [5] G, Grant, W. Richards, *Computational Chemistry*, Oxford Primers, Oxford, 1995.
- [6] P.Chaquin, *Manuel de chimie théorique, application à la structure et la réactivité en chimie moléculaire*, 2000, p.190.
- [7] J.E.Williams, P.R.Schleyer, *Ann.Rev.Phys.Chem*, 1968, p.19,531.
- [8] U.Burkert, N.L.Allinger, *Molecular Mechanics*, ACS Monographe 177, American Chemical Society, Washington, D.C. 1982,
- [9] M. Baaden, Outils pour étudier la structure et la dynamique des peptides et des protéines, (Institut de Biologie Physico-Chimique, PARIS) *CNRS UPR9080*, 2003, 36.
- [10] J.S.Lomas, *l'actualité chimique, La Mécanique Moléculaire, une méthode non quantique pour le calcul de la structure et de l'énergie d'entité moléculaire*, 1986, p.7,
- [11] F.Jensen, *Introduction to Computational Chemistry*, John Wiley & Sons, Chichester, 1999,
- [12] G.Monard, *Introduction à la Modélisation Moléculaire*, Formation continue CNRS-Nancy, 2003,
- [13] J.P .Brown, N.L.Allinger, K.B.Boyd, *reviews in computational chemistry*, VCH, New York, 1991, p. 2, 81.
- [14] P.Comba, T.W.Hamblay, *Molecular Modelling of Inorganic Compounds*, VCH, New York, 1995.

- [15] K.B.Lipkowitz, Abuses of molecular mechanics : Pitfalls to avoid. *J. Chem.* 1995, p. 72, 1070-1075.
- [16] Biosym/MSI, Discover, user guide, ed. Biosym/MSI. 1995, San Diego.
- [17] S.J.Weiner, P.A.Kollman, D.T.Nguyen, D.A.Case, An all-atom force field for simulations of proteins and nucleic acids, *J. Comp. Chem*, 1986, p. 7, 230.
- [18] U. Dinur, A.Hagler, K.B.Lipkowitz, D.B.Boyd, Reviews in computational Chemistry, VCH, Weinheim. , 1991, p.2, 99.
- [19] N. L. Allinger, K. Chen, J. A. Katzenellenbogen, S. R. Willson, G. M. Anstead, *J. Comp. Chem.*, 1996, p.17, 747.
- [20] N. L. Allinger, *Adv. Phys. Org. Chem*, 1976, p.13, 1.
- [21] N. L. Allinger, *J. Am. Chem. Soc*, 1977, p.99, 8127.
- [22] N. L. Allinger, Y. H. Yuh, J-H. Li, *J. Am. Chem. Soc*, 1989, p.111, 8522.
- [23] N. L. Allinger, K. Chen, J-H. Lii, *J. Comp. Chem*, 1996, p. 17, 642.
- [24] W. L. Jorgensen, J. T. Rives, *J. Am. Chem. Soc.* 1988, p. 110, 1657.
- [25] W. F. Van Gunsteren, M. Karplus, *Macromolecules*. 1982, p. 15, 1528.
- [26] B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, M. Karplus, *J. Comp. Chem.*, 1983, p. 4, 187.
- [27] J. C. Smith, M. Karplus, *J. Am. Chem. Soc.* 1992, p. 114, 805.
- [28] A. Bouraoui, M. Fathallah, B. Blaive, R. Gallo, *J. Chem. Soc. Perkin Trans.* 1990, 2, pp. 1211-1214.
- [29] A. Bouraoui, M. Fathallah, F. M'Henni, B. Blaive, R. Gallo, "Modeling of Molecular Structures and Properties". *Proceeding of an International Meeting* Nancy, 11-15, 1989, ed, J.L. Rivail., *Studies in Physical and Theoretical chemistry*, p.71, 381-393.
- [30] A. Zinelabidine, A. Bouraoui, M. Fathallah, F. M'Henni, B. Blaive, R. Gallo, *J. Mol. Struct. (Theochem)*, 1993, 286, 267-274.

- [31] B. Blaive, G. Legsaï, R. Lai, *J. Mol. Struct.* **1995**, p.354, 245-250.
- [32] S. J. Weiner, P. A. Kollman, T. Nguyent, D. A. Case, *J. Comput. Chem.* **1986**, p.7, 230.
- [33] T. Shimanouchi, *Pure Appl. Chem.* **1963**, p.7, 131.
- [34] F. Tristram, V. Durier, G. Vergoton, *Journal of Molecular Structure.*, **1996**, 378, 249-259.
- [35] D. A. Case, D. A. Pearlman, J. W. Caldwell, T. E. Cheatham III, J. Wang, W. S. Ross, C. L. Simmerling, T. A. Darden, K. M. Mer, R. V. Stanton, A. L. Cheng, J. J. Vincent, M. Crowley, V. Tsui, H. Gohlke, R. J. Radmer, Y. Duan, J. Pitera, I. Massova, G. L. Seibel, U. C. Singh, P. K. Weimer et P. A. Kollman, *AMBER7*, **2002**.
- [36] P. A. Kollman, *Acc. Chem. Res.* **1985**, p.18, 105.
- [37] W. F. van Gunsteren, *Prot. Eng.* **1988**, p.2, 5.
- [38] J. A. McCammon; S. C. Harvey, *Dynamics of Proteins and Nucleic Acids Cambridge*, **1987**.
- [39] D. Frenkel, B. Smit, *Understanding molecular simulations: From algorithms to applications*, Academic press: San, Diego, **1996**.
- [40] W. F. Van Gunsteren, H. J. C. Berendsen, *Computer simulation of molecular dynamics: Methodology, applications, and perspectives in Chemistry*, *Angew: Chem.Int. Ed. Engl.* **1990**, p.29, 99-1023.
- [41] P. Bamborough et F. E. Cohn, *Modeling protein-ligand complexes*, *University of California*, San Fransisci, **1995**.
- [42] D. Duhovny, R. Nussinov, H. J. Wolfson, *Efficient unbound docking of rigid molecules*, **2002**.

# Chapitre II :

## *Analyse Conformationnelle*

L'énergie potentielle d'un système est une fonction complexe et multidimensionnelle des coordonnées des atomes qui le composent. La manière dont cette énergie varie en fonction de ces coordonnées est appelée surface d'énergie potentielle. Dans la majorité des cas, on s'intéresse plus particulièrement aux états les plus stables du système, c'est à dire à ceux qui possèdent une énergie potentielle minimale. Il est donc important de disposer d'algorithmes qui permettent d'optimiser rapidement la géométrie d'une molécule [1].

### I. Minimisation de l'énergie :

La mécanique moléculaire nous permet de minimiser l'énergie calculée [2]. Le rôle de la minimisation est essentiellement de relaxer la structure initiale et d'éliminer les mauvais contacts interatomiques. Les structures « minimisées » correspondent généralement au minimum local le plus proche de la structure de départ plutôt qu'au minimum global d'énergie.

La minimisation permet aussi de trouver le minimum d'énergie potentielle dans le champ de force étudié et sous les contraintes imposées. Ces méthodes autorisent généralement des changements de conformation plus importants grâce à la réduction du nombre et de la nature des variables ou la simplification du champ de force.

Les méthodes de minimisation sont basées sur le calcul des dérivées de la fonction d'énergie et utilisent des processus itératifs : à partir de l'énergie potentielle et du gradient, calculés pour un jeu de coordonnées, ces algorithmes génèrent un nouveau jeu de coordonnées correspondant à une énergie potentielle plus basse.

Le principe de base de la minimisation est qu'à partir d'une géométrie très approximative, on recherche le jeu de coordonnées cartésiennes qui réduit à son minimum la somme de toutes les contributions énergétiques dues aux déformations des coordonnées internes et aux interactions entre atomes non liés. 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 ou, pour chaque coordonnée  $r_i$  :  $\partial E / \partial r_i = 0$

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

Presque toutes les méthodes de minimisation ont au moins un point en commun : on commence à un endroit donnée 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 [3].

Les minima locaux sont inévitables car les longueurs de liaison et les angles de valence sont rapidement optimisés au stade initial de la minimisation ; mais quand ces derniers ont atteint leur valeur d'équilibre, d'autres mouvements torsionnels sont inhibés puisque pour réduire d'avantage l'énergie totale du système il faut pouvoir déplacer plusieurs atomes [4].

### II. Surface d'énergie potentielle :

Une surface d'énergie potentielle ou plus exactement l'hypersurface d'énergie potentielle, notée HSEP, est l'ensemble des points générés par la fonction  $E(R)$ ,  $R$  étant le vecteur dont les composantes  $(r_1, r_2, \dots, r_{3n})$  sont les variables géométriques du système moléculaire. L'hypersurface sera de dimension 1 si l'énergie du système ne dépend que d'une seule variable géométrique, dans ce cas  $E(R)$  sera représentée par une courbe de potentiel. Si  $E(R) = E(r_1, r_2)$  l'ensemble des points générés par cette fonction est un plan. En général, si un système moléculaire est repéré par  $n$  coordonnées géométriques, l'hypersurface sera de dimension  $n$ .

Chaque point sur l'hypersurface correspond à une géométrie précise du système moléculaire. Les points pour lesquels

$$\partial E / \partial r_i = 0$$

sont appelés points stationnaires, et correspondent aux minima et aux états de transition [5].

### III. Analyse conformationnelle :

La plupart des programmes modernes de mécanique moléculaire ont une interface graphique qui permet à l'utilisateur d'entrer la structure d'une molécule et de minimiser son énergie dans cette conformation. Cependant la conformation trouvée n'est pas nécessairement la structure de plus basse énergie ce qui lui confère une stabilité relative. En effet, il est tout à fait possible que la structure minimisée ait une énergie fortement élevée par rapport au minimum global ce qui lui donne peut d'intérêt pratique [6].

L'analyse conformationnelle est une méthode qui permet de calculer l'énergie relative associée à la conformation d'une molécule [4]. Son but est donc de trouver le plus de minima possible, et de calculer la population de Boltzmann [6]. Cette dernière nous renseigne sur la population des niveaux occupés à une température donnée.

$$\left. \begin{array}{l} N_i = N_0 \exp(-E_i / KT) \\ N_j = N_0 \exp(-E_j / KT) \end{array} \right\} \Rightarrow N_i / N_j = \exp(-\Delta E_{ij} / KT)$$

$E_i$  : Energie de la molécule.

$K$  : Constante de Boltzmann ( $K=10.38 \times 10^{-23}$  J/K).

$T$  : Température (Kelvin). [5]

Les molécules flexibles présentent un nombre considérable de conformations différentes, mais les plus importantes sont celles ayant une énergie libre suffisamment basse pour qu'elles puissent être peuplées à des températures raisonnables [7].

Les propriétés physiques observées d'une molécule fortement flexible ne résultent pas d'un seul état de conformation global optimal, mais plutôt d'une distribution thermique d'états représentant diverses régions de l'espace conformationnel. Cette réalisation a profondément influencé les croyances concernant le comportement des macromolécules biologiques [8].

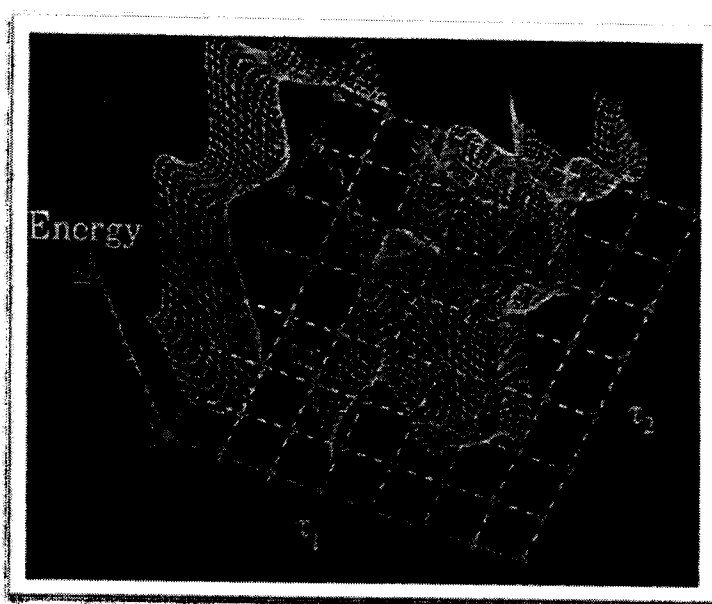


## CHAPITRE II : Analyse conformationnelle

---

Pour cette raison, il est important de trouver tous les conformères de basse énergie de la molécule étudiée [9].

Un des modèles les plus réussis utilisés en chimie computationnelle est la surface d'énergie potentielle (Potential Energy Surface « PES »). C'est une représentation remarquable des molécules individuelles, des complexes moléculaires et des interconversions conformationnelles. Elle permet aussi, de faire le lien entre les concepts chimiques clés et les caractéristiques géométriques de base des surfaces. Dans le cas de l'analyse conformationnelle, les minima et les points de selle (saddle points) sur la PES sont les caractéristiques les plus importantes. Les minima correspondent à des conformations d'énergie minimale et les points de selle à des états de transition associés à des interconversions conformationnelles [7].

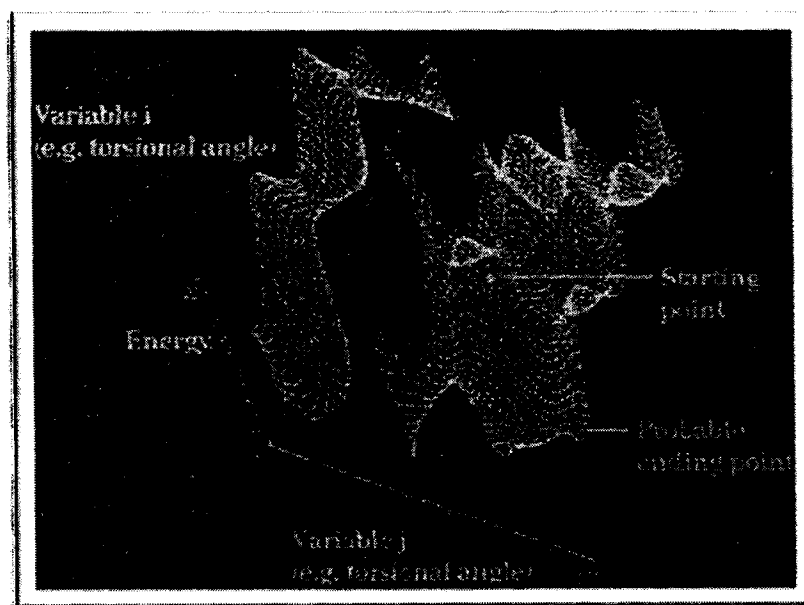


En analyse conformationnelle, un grand nombre de géométries de départ de haute énergie sont générées, minimisées et stockées. Pour que la recherche soit fiable, les géométries de départ brutes doivent couvrir toute la surface d'énergie potentielle : dans le cas contraire, on n'est pas sûr de trouver tous les minima [6].

De nombreuses méthodes de recherche conformationnelle sont décrites dans la littérature. Parmi elles, on trouve celles qui sont destinées au traitement des petits cycles, des peptides acycliques et des molécules organiques [9-11], d'autre pour les petites protéines [12], des polymères [13] et les hydrates de carbone [14].

## CHAPITRE II : Analyse conformationnelle

L'analyse conformationnelle permet de trouver les minima et les points de selle, et d'identifier les interconversions impliquant des petites barrières énergétiques. Pour cela il faut générer un ensemble de structures d'essai, et passer à la minimisation ensuite. Dans un système moléculaire sans contrainte et avec un nombre d'atomes  $n > 2$ , le nombre de degrés de liberté est égal au nombre de coordonnées internes c'est à dire  $3n-6$ . Lors de la minimisation de l'énergie, l'ensemble des structures d'essai est généré en faisant varier les degrés de liberté les plus flexibles à savoir les angles dièdres [7]. De façon générale, pour les liaisons simples, la modification d'un angle dièdre coûte moins cher en Kcal/mol que la déformation d'un angle de valence qui, à son tour, est plus facile que l'élongation d'une liaison. On dit que les angles dièdres sont plus mous que les autres coordonnées internes [3]. Pour pouvoir faire varier tous les angles dièdres d'une molécule, on dresse une grille pour obtenir toutes les possibilités combinatoires. Ainsi, une molécule avec trois angles dièdres va générer 216 conformations si l'incrément d'angle est de  $60^\circ$  c'est à dire  $6^3$ . Pour une molécule ayant six angles dièdres on atteindra déjà 46656 conformères, tandis qu'avec neuf angles on arrive à plus de 10 millions de conformations possibles [4].



La recherche et la localisation des structures des états d'équilibre et de transition sur l'hypersurface d'énergie potentielle d'un système moléculaire sont nécessaires pour la détermination des propriétés thermodynamiques et cinétiques, de même que pour l'interprétation des spectres de vibration et de rotation. Des techniques spéciales d'optimisation de géométrie ont été élaborées pour cela [16].

#### **IV. Algorithmes de minimisation :**

L'optimisation de la géométrie d'un système moléculaire consiste à trouver la géométrie de ce système qui correspond à l'énergie la plus basse du système. C'est pourquoi ces méthodes sont dites de minimisation. Il est important de préciser que ces dernières s'appliquent au voisinage des points stationnaires, donc pour des petites variations de la géométrie. Ceci est dû au fait que ces méthodes partent toutes du développement de l'énergie en série de Taylor qui n'est valable que pour un petit écart de la position d'équilibre recherchée. Elles peuvent être en défaut et déceler un minimum global englobant ainsi plusieurs minimums locaux, si la géométrie de départ est trop éloignée de celle de l'état d'équilibre [5]. Plusieurs méthodes de minimisation existent, elles sont citées succinctement ci-après.

##### **IV. 1. La méthode de Steepest Descent :**

Cette méthode appelée aussi méthode de la plus grande pente, se base sur le calcul de la dérivée première de la fonction énergétique  $\partial E/\partial r$  par rapport aux coordonnées géométriques afin de déterminer la distance au minimum énergétique le plus proche et dans quelle direction modifier les coordonnées de chaque atome de la structure.

Cette méthode est souvent utilisée en première approximation, dans le but de dégrossir les calculs. Elle marche bien lorsque le gradient  $r$  est grand et possède une faible convergence car ce dernier devient faible rapidement à l'approche du minimum [16].

##### **IV. 2. La méthode du gradient conjugué :**

- Elle trouve un minimum en moins d'étapes que la méthode de « Steepest Descent ».
- Elle nécessite des temps de calcul importants lorsque la structure est éloignée du minimum.

##### **IV. 3. La méthode de Newton-Raphson :**

- Cette méthode se base sur le calcul de la dérivée seconde de la fonction énergétique par rapport aux coordonnées géométriques.
- Elle peut trouver le minimum plus vite que les méthodes du gradient seul.
- On peut rencontrer des problèmes lorsque la conformation initiale est éloignée du minimum [13].

### **IV. 4. La méthode du Simplex :**

- Elle se limite à un calcul direct de l'énergie et observer si elle augmente ou diminue.
- Fonctionne bien si on a 20 variables ou moins.
- Elle consiste en essais successifs représentant une triangulation de l'hypersurface ; c'est à dire qu'on examine le sens et l'amplitude des variations de la fonction dans différentes directions pour suivre celles qui mènent vers un minimum.

Néanmoins, il devient nécessaire de disposer d'une procédure qui part d'une structure approchée et déplace les atomes vers leurs coordonnées optimales d'une façon moins aléatoire [4].

### **IV. 5. La méthode du recuit simulé :**

Les méthodes précédentes ont la particularité de faire décroître à chaque pas la fonction F ; elles 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 du recuit simulé, développée par Kirkpatrick (1983), 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 [4].

En somme, un minimiseur est une boîte noire mathématique qui demeure un outil important pour la recherche d'un minimum énergétique. En général, on utilise plusieurs minimiseurs en passant de l'un à l'autre si ça ne converge pas assez vite. Tout dépend aussi du nombre de variables ou du nombre de variations d'angle de liaison introduit.

En principe on calcul l'énergie du système avec des fonctions énergétiques empiriques. Chaque logiciel a ses valeurs à lui qui sont plus au moins meilleures qu'un autre, et en fin, on va chercher un minimum qui converge [4].

A ce titre, l'analyse conformationnelle reste un moyen efficace de prédiction des conformations stables d'une structure moléculaire considérée et celle d'une structure moléculaire correspondant aux minima de son énergie intramoléculaire. L'énergie calculée mesure la différence entre l'énergie de la structure considérée et celle d'une structure hypothétique dont toutes les

## CHAPITRE II : Analyse conformationnelle

---

coordonnées internes prendraient leurs valeurs de référence, sa valeur absolue dépend du champ de force considéré, c'est-à-dire des fonctions d'énergie potentielle et des paramètres utilisés. Les valeurs de cette énergie sont toutefois très utiles lorsqu'il s'agit de comparer la stabilité relative des conformations ou des stéréo-isoméries d'une même molécule [4].

### REFERENCES:

- [1] A. R. Leach, Energy minimisation and related methods for exploring the energy surface, in *Molecular modeling: Principles and applications*, Addison Wesley Longman Ltd: Harlow. **1996**, 211-260.
- [2] M. Baaden, Outils pour étudier la structure et la dynamique des peptides et des protéines, (Institut de Biologie Physico-Chimique, PARIS) CNRS UPR9080, **2003**, 36.
- [3] J. S. Lomas, l'actualité chimique- France, **Mai 1986**
- [4] H. Dugas, Principes de base en modélisation moléculaire, *Aspects théoriques et pratiques*, quatrième édition, Librairie de l'Université de Montréal, **1996**.
- [5] S. Djennane-bousmaha, Thèse de doctorat, USTHB Alger, **1999**.
- [6] J. Bartol, P. Comba, M. Melter, M. Zimmer, *Journal of computational Chemistry*, **1999**, 20, 1549-1558.
- [7] G. Keserü et I. Kolossvary, *Molecular Mechanics and Conformational Analysis in Drug Design*, Department of chemical Information Technology, Technical University of Budapest, Blackwell Science Ltd, **1999**.
- [8] C. L. Brooks, III, M. Karplus, Pettitt, B. M. *Proteins. Advances in Chemical Physics*, Willey: New York, LXXI, **1998**.
- [9] J. Thomas Ngo and M. Karplus, *J. Am. Chem. Soc.* **1997**, 119, 5657-5667.
- [10] I. Kolossvary and W. C. Guida, *J. Am. Chem. Soc.* **1996**, 118, 5011-5019.
- [11] H. Senderowitz, F. Guarieri, W. C. Still., *J. Am. Chem. Soc.*, **1995**, 117, 8211.
- [12] M. J. Dudeck, H. A. Scheraga, *J. Comput. Chem*, **1990**, 11, 121.

## CHAPITRE II : Analyse conformationnelle

---

- [13] E. Leontidis, J. J. de Pablo, M. Laso, U. W. Suter, *Adv. Polym. Sci.*, **1994**, 116, 283.
- [14] C. W. Von Der Lieth, T. Kozar, W. E. Hull, *Theochem. J. Mol. Struct.*, **1997**, 395, 225.
- [15] A. Hinchliffe, "Modelling Molecular structures", John Wiley and Sons, New York, **1994**.
- [16] « Modélisation moléculaire » Cours du Professeur G. Vergoten à l'Université des Sciences et Technologies de Lille.

# Chapitre III :

*Acides aminés, Peptides, Protéines, Enzymes*

Dans notre nourriture, les éléments nutritifs sont l'eau, les lipides, les protéines, les hydrates de carbone, les vitamines et les minéraux. Il s'agit ici de la façon habituelle de parler et plusieurs s'y retrouvent. Pourtant, il serait plus précis de faire référence aux acides aminés plutôt qu'aux protéines. En effet, les protéines contenues dans notre nourriture (viande, poisson et légumineuses surtout) sont d'abord défaits en filaments dans l'estomac, puis en unités minuscules dans les intestins. Ces particules sont les acides aminés. Ils traversent la paroi intestinale pour se déverser dans la veine cave qui les amène au foie. Là, ils sont soit assemblés en de nouveaux peptides ou protéines soit modifiés en d'autres acides aminés; sinon, ils sont déversés dans le sang pour parvenir aux autres cellules du corps.

Notre organisme s'approvisionne donc en acides aminés de la même façon qu'il retire les gras, les vitamines et les autres éléments de la nourriture. Les protéines alimentaires seraient la principale source des acides aminés nécessaires à notre corps.

### I-Acides aminés

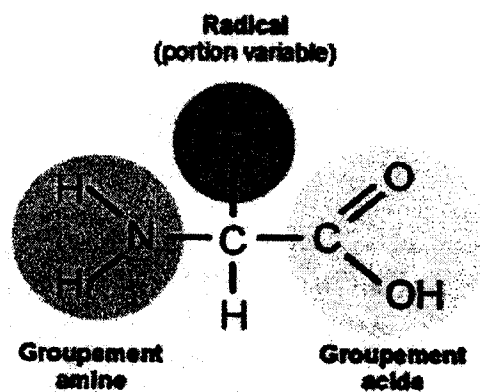
Les acides aminés sont des acides carboxyliques qui portent une fonction aminée. Les protéines de toutes les espèces sont constituées de seulement vingt acides aminés différents.

Ils sont tous constitués d'un atome de carbone lié à :

- Un groupement carboxylique  $\text{-COOH}$ .
- Une fonction aminée  $\text{-NH}_2$ .
- Un atome d'hydrogène.
- Une partie variable R.

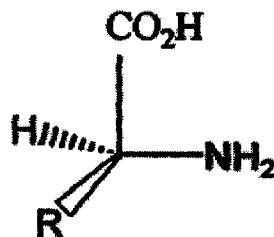
- Dans la nature, les acides les plus courants sont les 2-aminoacides, appelés les  $\alpha$ -aminoacides.





**Figure 1** : structure générale des acides aminés

- Tous les acides aminés naturels adoptent la configuration S (sauf la glycine) et appartiennent à la série L.

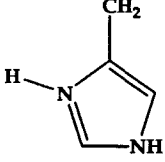
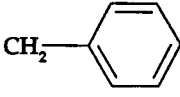
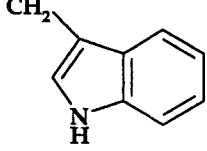
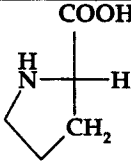
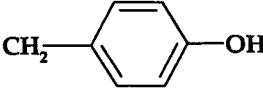


**Figure 2** : représentation des L- $\alpha$ -aminoacide

- Il y a huit aminoacides essentiels que les êtres humains ne peuvent pas Biosynthétiser (Valine, Leucine, Isoleucine, Thréonine, Méthionine, Lysine, Phénylalanine et le tryptophane) et qui doivent être apportés par l'alimentation.

**CHAPITRE III : Acides aminés, Peptides, Protéines, Enzymes**

**Tableau:** Principaux acides aminés.

$\begin{array}{c} \text{NH}_2 \\   \\ \text{H} - \text{C} - \text{COOH} \\   \\ \text{R} \end{array}$							
R	Nom	Code à trois lettres	Code à une lettre	R	Nom	Code à trois lettres	Code à une lettre
H	Glycine	Gly	G	CH <sub>2</sub> SH	Cystéine	Cys	C
CH <sub>3</sub>	Alanine	Ala	A	CH <sub>2</sub> CH <sub>2</sub> SCH <sub>3</sub>	Méthionine	Mét	M
CH(CH <sub>3</sub> ) <sub>2</sub>	Valine	Val	V	(CH <sub>2</sub> ) <sub>4</sub> NH <sub>2</sub>	Lysine	Lys	K
CH <sub>2</sub> CH(CH <sub>3</sub> ) <sub>2</sub>	Leucine	Leu	L	$\begin{array}{c} \text{NH} \\    \\ (\text{CH}_2)_3\text{NHCNH}_2 \end{array}$	Arginine	Arg	R
$\begin{array}{c} \text{CHCH}_2\text{CH}_3 \\   \\ \text{CH}_3 \end{array}$	Isoleucine	Ile	I		Histidine	His	H
	Phénylalanine	Phe	F		Tryptophane	Trp	W
	Proline	Pro	P	CH <sub>2</sub> COOH	Acide aspartique	Asp	D
CH <sub>2</sub> OH	Serine	Ser	S	$\begin{array}{c} \text{O} \\    \\ \text{CH}_2\text{CNH}_2 \end{array}$	Asparagine	Asn	N
$\begin{array}{c} \text{CHOH} \\   \\ \text{CH}_3 \end{array}$	Tréonine	Thr	T	CH <sub>2</sub> CH <sub>2</sub> COOH	Acide glutamique	Glu	E
	Tyrosine	Tyr	Y	$\begin{array}{c} \text{O} \\    \\ \text{CH}_2\text{CH}_2\text{CNH}_2 \end{array}$	Glutamine	Gln	Q

La différence entre ces vingt acides aminés se situe au niveau de leur radical, en effet on dénombre vingt radicaux différents. Ces groupements spécifiques sont dits « chaîne latérale » et selon lesquels, on peut classer les acides aminés comme suit :

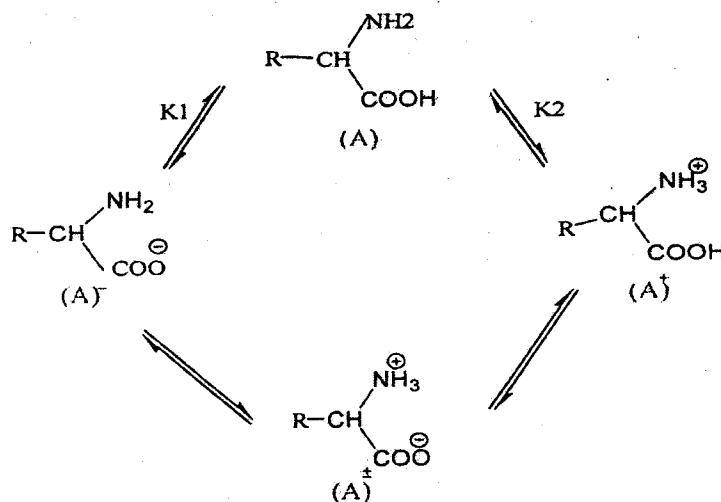
- Ceux dont les radicaux sont basiques (arginine, histidine, lysine).
- Ceux dont les radicaux sont acides (acide aspartique, acide glutamique).
- Ceux qui possèdent des radicaux non chargés (glutamine, serine, thréonine).

### I-1. Propriétés physiques

- les acides aminés (AA) sont des solides blancs cristallisés qui se décomposent avant de fondre (ils perdent l'acide carboxylique).
- La solubilité des AA est en rapport avec leurs structures polaires et dépend de plusieurs facteurs :
  - o La nature de la chaîne R.
  - o Le pH de la solution.
  - o La nature des ions présents dans la solution.
- Pouvoir rotatoire : les solutions d'acides aminés sont optiquement actives. Les aminoacides naturels appartiennent à la série L (les formes D sont exceptionnelles, on les retrouve surtout chez les bactéries).

### I-2. Propriétés chimiques

- Tous les aminoacides possèdent au moins deux groupes ionisables.

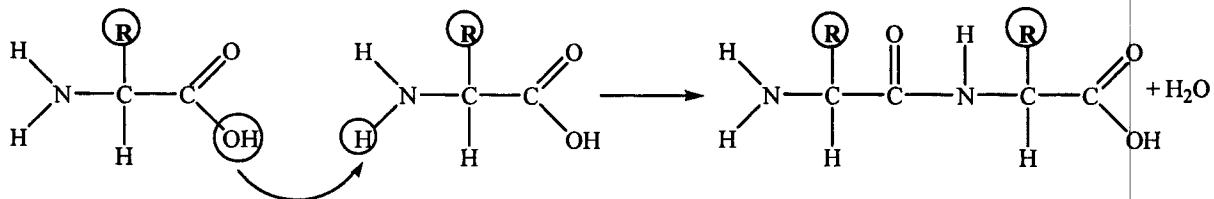


**Figure 3:** les formes ioniques des acides aminés

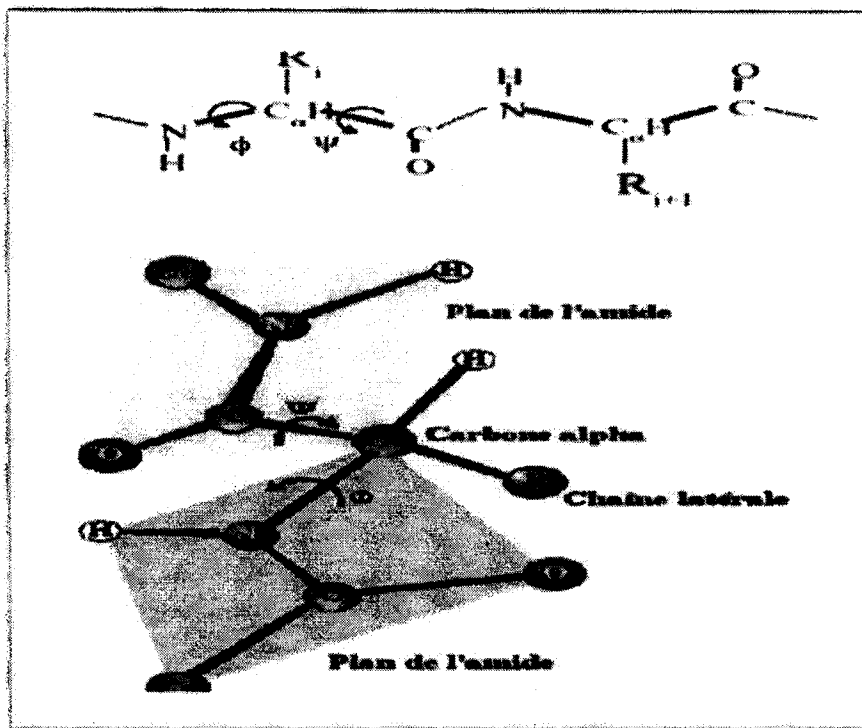
Le point isoélectrique ou isoionique : est un pH intermédiaire où l'acide aminé est neutre et n'a aucune charge. Ce pH est calculé par :  $pH = (pK_i + pK_a)/2$ .

## II-Les peptides :

Les acides aminés peuvent se lier les uns aux autres par une liaison peptidique qui se fait entre le groupement acide d'un acide aminé et le groupement aminé de l'autre. Au cours de la réaction, une molécule d'eau est éliminée ; il s'agit donc d'une réaction de condensation.



### II-1. Structures de la liaison peptidique :



**Figure 4** : les angles dièdres de la chaîne principale.

- Le plan rigide de la structure peptidique interdit la rotation des atomes des fonctions acide et amine. Les seules liaisons dont l'orientation reste libre sont celles qui entourent chacun des carbones asymétriques porteurs des radicaux des acides aminés.
- La liaison ( $C_{\alpha}$ -NH) située à l'extrémité  $NH_{\alpha}$ -terminale peut tourner librement : l'angle de rotation de cette liaison est appelée l'angle  $\Phi$  (phi).

La liaison (C $\alpha$ -CO) situé à l'extrémité COOH-terminale peut tourner librement : l'angle de rotation de cette liaison est appelé angle  $\Psi$  (psi).

- La liaison peptidique (NH-CO) est de caractère partiellement double dont la rotation est restreinte, l'angle de rotation de cette liaison est appelé angle  $\Omega$  (oméga) et adoptera généralement la valeur de 180° (trans) ou parfois de 0° (cis). Les angles  $\Phi$  et  $\Psi$  demeurent eux aussi habituellement près de 180° (trans) mais avec un peu plus de variation que l'angle  $\Omega$ , grâce à leur caractère de liaison simple. La structure secondaire d'une protéine dépend de la valeur de ces deux angles, pour chaque acide aminé, en fonction des liaisons hydrogènes qui unissent les atomes (carbonyle et aminé) des liaisons peptidiques.

Les peptides sont des composés naturels ou synthétiques qui résultent de l'enchaînement d'un nombre limité d'acides aminés par des liaisons peptidiques.

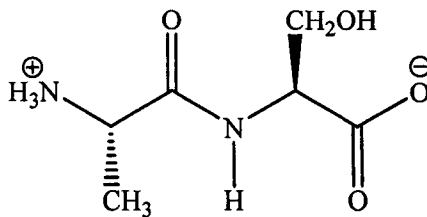
### II-2. Propriétés physiques

Les peptides sont des solides blancs, généralement cristallisables et plus ou moins soluble dans l'eau, ils présentent un caractère amphotère. Ils présentent dans l'ultra violet (U.V) une forte bande d'absorption entre 180 nm et 230 nm caractéristique de la liaison peptidique.

### II-3. Propriétés chimiques

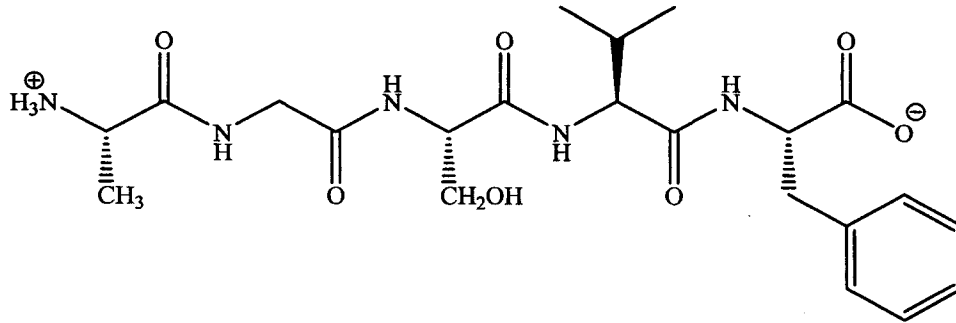
Elles sont liées à la composition des peptides, la seule propriété qui soit particulièrement intéressante de la liaison peptidique : la mobilité de l'hydrogène et la présence des électrons libres rendent possible la formation des complexes. En milieu alcalin, les peptides forment avec les ions cuivriques un complexe violet dont la réaction est dite de biuret.

- Un dipeptide : une molécule obtenue par la formation d'une liaison peptidique entre deux acides aminés.



**Figure 5** : alanylsérine

- Polypeptide : une molécule formée par l'enchaînement de plusieurs acides aminés. Par convention le groupement aminé du premier acide aminé (appelé aussi N-terminal) est placé à gauche et le groupement carboxylique du dernier acide aminé (C-terminal) est placé à droite [1].



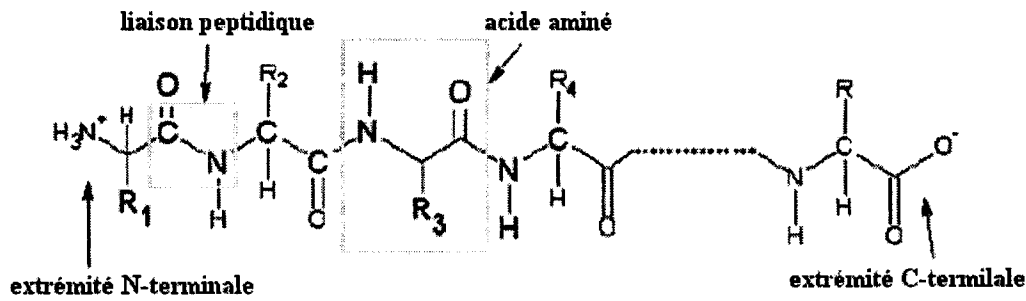
**Figure 6** : Alanyl-glycyl-seryl-valyl-phenylalanine

### III- Protéines

Les protéines représentent plus de la moitié de la masse d'une cellule, à l'exclusion de l'eau, et sont responsables de la plupart, des fonctions intra et inter cellulaires. Elles gèrent par exemple la réplication de l'ADN et le maintien des équilibres chimiques. Ces fonctions sont fortement liées aux formes spatiales que prennent les protéines dans le milieu cellulaire [1].

Les protéines forment la classe la plus nombreuse et la plus diversifiée des molécules biologiques, elles constituent plus de 50% du poids sec des cellules. Cette abondance reflète le rôle ubiquitaire des protéines dans pratiquement tous les aspects structuraux et fonctionnels de la cellule. Les protéines sont les agents des fonctions biologiques ; elles expriment aussi l'information génétique.

Les protéines sont de grosses molécules ou macromolécules qu'interviennent dans de nombreuses fonctions organiques. Elles sont constituées d'acides aminés (ou résidus) liés entre eux par des liaisons peptidiques entre le groupe  $\alpha$ - carboxylique d'un acide aminé et le groupe  $\alpha$ - aminé de l'acide aminé suivant. Une protéine est donc une chaîne d'acides aminés, appelée chaîne polypeptidique, pouvant contenir jusqu'à des dizaines de milliers de résidus. Ces chaînes de résidus ont une structure flexible ce qui leur permet de se replier en une forme appelée structure native [2].



**Figure 7** : Formule développée d'une protéine de n acides aminés. Les  $R_i$  désignent les différentes chaînes latérales des résidus.

Plusieurs techniques expérimentales permettent d'obtenir la structure 3D des protéines : la diffraction des rayons X sur des cristaux de protéines, la résonance magnétique nucléaire (RMN), la cryomicroscopie. Ces méthodes nous fournissent les coordonnées spatiales des atomes constituant les protéines. Il existe environ de 37000 structures qui ont été résolues à ce jour et sont disponibles via la Protein Data Bank [3].

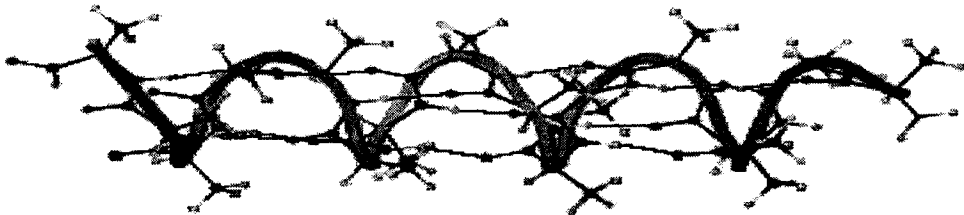
### III-1. Structure des protéines

**a/ La structure primaire** : La structure primaire est l'ordre d'enchaînement des acides aminés de la chaîne protéique. On nomme la liste des résidus en commençant par la terminaison amine (ou ammonium) et en terminant par le résidu portant la fonction acide carboxylique (ou carboxylate). Le premier résidu est alors nommé N-terminal et le dernier C-terminal. Le code à une lettre des acides aminés est alors très pratique pour décrire les protéines (voir l'exemple figure 10).

**b/ La structure secondaire** : est un premier niveau de repliement adopté par des portions de la protéine, résultant d'interaction entre des acides aminés voisins de la chaîne par des liaisons hydrogène qui sont formées entre les groupements C=O et H-N des acides aminés. Les deux structures secondaires les plus fréquentes sont les hélices  $\alpha$  et les feuillets  $\beta$  qui sont les deux formes répétitives les plus importantes des protéines représentant chacune respectivement 30 % et 20 % des résidus. Cette importance vient de leur stabilité énergétique particulière [4].

On distingue deux structures principales. L'hélice  $\alpha$  et Le feuillet  $\beta$ .

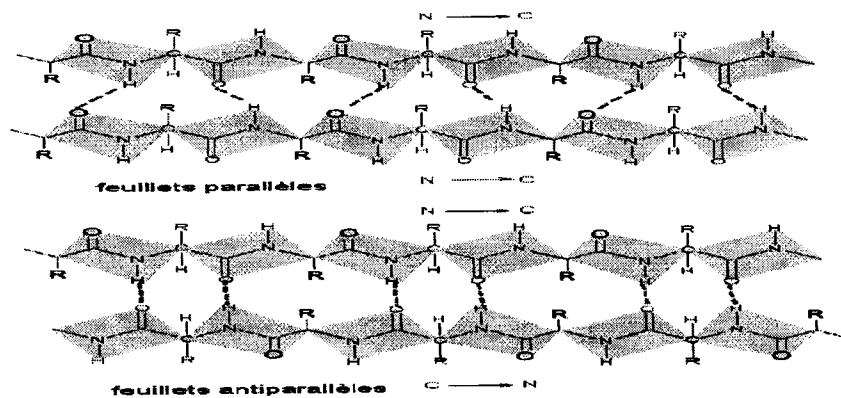
- **L'hélice  $\alpha$**  : résulte de l'enroulement de la chaîne protéique de façon à ce que les liaisons peptidiques qui en constituent l'épine dorsale aient la possibilité de former des liaisons hydrogènes entre elles. L'hélice  $\alpha$  est stabilisée par des liaisons hydrogènes entre les résidus. Ces liaisons sont formées entre l'atome d'hydrogène d'un azote peptidique et l'atome d'oxygène d'un groupement carbonyle distant de quatre résidus dans la structure primaire.



**Figure 8** : Liaisons hydrogène dans une hélice  $\alpha$ . Les liaisons H sont représentées en bleu.

Dans une hélice  $\alpha$ , chaque résidu est relié au précédent par une translation de  $1.5\text{\AA}$  le long de l'axe par rotation de  $100^\circ$  ; ce qui donne environ 3.6 résidus par tour d'hélice et un pas de  $5.4\text{\AA}$ .

- **Le feuillet  $\beta$**  : Dans le feuillet  $\beta$ , les liaisons hydrogène intermoléculaires stabilisent l'alignement ordonné des chaînes peptidiques. Les chaînes polypeptidiques (ou brins) voisines sont alors dites parallèles si leurs bouts N-terminaux sont tous du même côté et antiparallèles dans le cas contraire (figure 9).



**Figure 9** : Feuillet  $\beta$ , parallèles et antiparallèles. Les liaisons hydrogène sont dessinées en rouge.



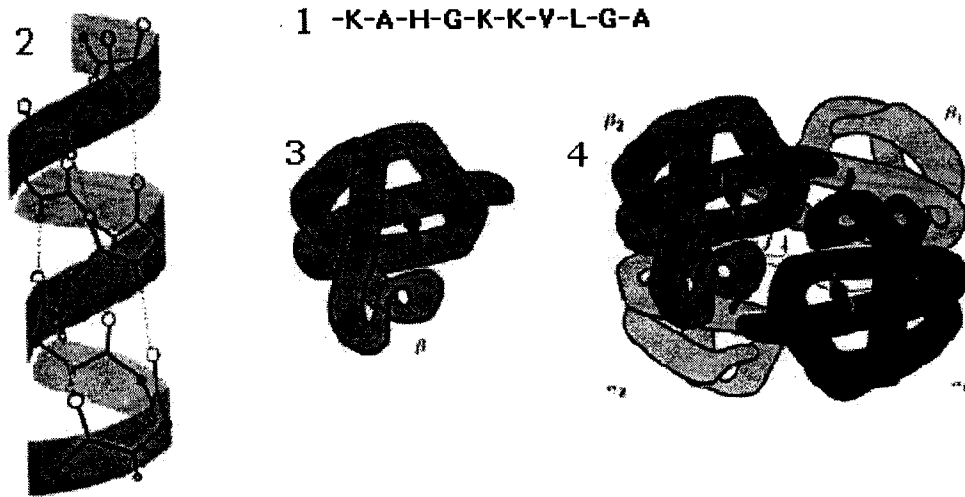
**c/ La structure tertiaire :** La structure tertiaire d'une protéine est sa disposition tridimensionnelle : c'est la donnée de l'agencement des structures secondaires et de l'organisation spatiale des chaînes latérales. Un certain nombre d'interactions stabilisent les structures tertiaires :

- ✓ Les liaisons disulfures entre deux cystéines forment un lien covalent entre deux parties d'une chaîne qui sont espacées dans la structure primaire.
- ✓ Les liaisons hydrogène, qui stabilisent aussi les structures secondaires, contribuent à la stabilisation des structures tertiaires. Ces liaisons impliquent les squelettes polypeptidiques (chaînes principales) et les chaînes latérales.
- ✓ Les ponts salins se forment entre deux acides aminés ionisés. L'attraction coulombienne entre les ions positifs et les ions négatifs stabilisent leur rapprochement.
- ✓ Les interactions hydrophobes sont formées entre groupements non polaires. En solution aqueuse, les groupements polaires sont tournés en général vers l'extérieur des protéines globulaires alors que les groupements non polaires sont confinés à l'intérieur pour interagir préférentiellement entre eux plutôt qu'avec les molécules d'eau. Ces interactions sont plus faibles que les liaisons hydrogène et les ponts salins mais elles sont en général suffisamment nombreuses dans les régions au coeur des protéines pour permettre la stabilisation de la structure.

**d/ La structure quaternaire :** La structure quaternaire est le niveau le plus élevé d'organisation des protéines. Elle concerne les protéines constituées de plusieurs chaînes polypeptidiques et détermine l'arrangement spatial des différentes sous-unités entre elles.

Les zones de contact entre sous-unités sont très semblables à celles à l'intérieur d'une protéine à une seule sous-unité. Elles contiennent des chaînes latérales non polaires regroupées, des liaisons hydrogène et dans certain cas des ponts disulfure intercaténaires.

Les différents niveaux de description d'une protéine sont résumés dans la figure 10. Il faut noter qu'en général, la donnée de la structure primaire détermine la structure tertiaire des protéines [4].



**Figure 10** : Différents niveaux de description d'une protéine : primaire (1), secondaire (2) (exemple d'une hélice  $\alpha$ ), tertiaire (3) et quaternaire (4).

Le problème principal de la prédiction de structures est de prédire la structure d'une protéine à partir de sa séquence d'acides aminés. Ce problème est extrêmement difficile à résoudre et est reconnu comme l'un des grands défis de la bio-informatique actuelle [5- 9].

### III-2. Le repliement des protéines

Les radicaux des acides aminés ont des propriétés chimiques différentes. Certains sont hydrophobes, d'autres hydrophiles, certains s'ionisent négativement et d'autres positivement. Certains radicaux peuvent former des liaisons chimiques plus ou moins fortes avec d'autres radicaux. Il peut donc y avoir dans une chaîne d'acides aminés des interactions entre les radicaux. Certains se repoussent et d'autres se rapprochent et forment des liens chimiques. La chaîne d'acides aminés aura donc tendance à se replier sur elle-même pour adopter une structure tridimensionnelle précise.

Les forces qui stabilisent la forme pliée :

- Liaisons hydrogènes intramoléculaire
- Entropie de déshydratation.
- Interactions hydrophobes.
- Ponts salins.
- Interactions dipolaires.

Les forces qui déstabilisent la forme pliée :

- Liaisons hydrogènes avec l'eau.
- Perte d'entropie configurationnelle.

### **III-3. Fonctions des protéines**

Les fonctions des protéines résultent de la forme précise prise par celles-ci dans le milieu cellulaire. Ces fonctions peuvent être liées à l'activité de catalyse, de reconnaissance, de changement de conformation ou même d'une combinaison des trois. Ceci explique que certaines protéines possèdent plusieurs fonctions. Dans ce cas comment définir la fonction d'une protéine ? P.Bork et collaborateurs proposent exemple trois niveaux de description des fonctions : phénotypique, cellulaire et iculaire [10] ; mais il n'existe pas, à l'heure actuelle, de définition unique des ions des protéines [11].

### **III-4. Pourquoi modéliser les protéines ?**

En un mot, parce qu'elles sont le moteur derrière les processus biochimiques. En effet, qu'il s'agisse d'un processus normal tel que la croissance d'une plante ou d'un processus anormal à l'origine d'une maladie, il y a fort à parier que celui-ci est dû à la présence d'une protéine. En conséquence, les protéines représentent des cibles tout à fait intéressantes dans le cadre de traitements médicamenteux [1].

Les structures cristallines des protéines sont disponibles pour seulement une fraction de protéines connues. Avec ces structures, nous pouvons visualiser la protéine et comprendre son fonctionnement.

## **IV- Enzymes**

### **IV-1. Introduction**

Dans les systèmes biologiques ; c'est à dire à des températures en général inférieur ou égale à 37°C, les réactions chimiques du métabolisme se produisent rarement spontanément, mais sont le plus souvent catalysées par des protéines particulières que l'on appelle des enzymes.

Le mot enzyme vient des racines grecques *en* (dans) *zume* (levain). Les premières enzymes identifiées ont été extraites des levures, ces microorganismes responsables de la transformation du glucose en alcool.

Dans les cellules, les enzymes permettent aux réactions de se produire avec peu d'énergie. Souvent, la seule énergie cinétique des molécules à la température ambiante est suffisante pour que la réaction se produise. La plupart des réactions qui se déroulent dans les cellules ne pourraient pas s'y produire sans les enzymes appropriées.

Les enzymes sont des protéines qui jouent le rôle de catalyseur cellulaire. Elles sont constituées des milliers d'acides aminés liés en une chaîne linéaire. Ces acides aminés diffèrent dans leur nature chimique, sont liés entre eux par des liens covalents appelés liens PEPTIDIQUES. Un catalyseur est une substance qui modifie la vitesse d'une réaction chimique mais qui se retrouve inchangée à la fin de la réaction. En termes plus techniques, un catalyseur est une substance qui abaisse l'énergie d'activation d'une réaction chimique.

Actuellement plusieurs milliers d'enzymes sont connues, la séquence des acides aminés de près de trois milles d'entre elles a été élucidée. Nombre d'entre elles sont utilisées industriellement (biotechnologies, lessives, jus de fruits, production de sucre à partir d'amidon, traitement des peaux, des fibres textiles etc.) et d'autres sont de merveilleux outils dans la recherche biomédicale.

### **IV-2. Site actif**

Le site actif d'une enzyme est la région privilégiée de l'enzyme qui interagit avec le substrat. Les acides aminés du site actif peuvent être divisés en deux groupes :

- Ceux qui interviennent à la reconnaissance spatiale du substrat, en formant avec lui des liaisons non-covalentes (site de fixation).
- Ceux qui participent à la transformation chimique du substrat en produit (site catalytique).

Les autres acides aminés sont nécessaires soit au maintien de la conformation tridimensionnelle active de l'enzyme, soit au maintien de l'enzyme vis-à-vis d'autres protéines.

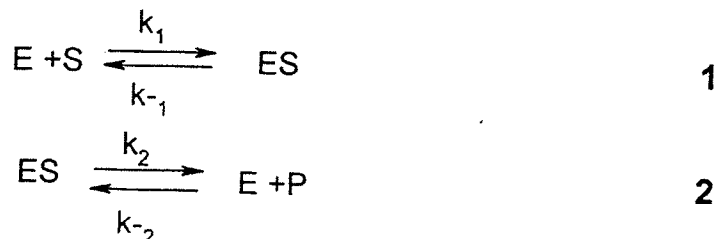
L'insertion des fragments (atomes) dans le site actif est basé sur le principe de complémentarité stérique et électronique entre ligand et récepteur.

### **IV-3. Cinétique Enzymatique**

La cinétique est la science qui a pour objet l'étude de la vitesse des réactions chimiques. La cinétique enzymatique englobe le rôle biologique des catalyseurs enzymatiques et les mécanismes qui permettent leur activité catalytique. Avec la cinétique enzymatique, on cherche à déterminer la vitesse maximale de la réaction que l'enzyme catalyse et à mesurer son affinité pour les substrats et les inhibiteurs. La connaissance précise de la cinétique

enzymatique est indispensable pour la conception rationnelle d'un nouveau médicament et le succès de son action pharmacologique [12].

La cinétique enzymatique a fait l'objet d'études au début des années 1900. Leonor Michaelis et Maud L. Menten ont proposé en 1913 [13] une théorie générale de l'action enzymatique compatible avec les résultats expérimentaux. Leur modèle cinétique est basé sur la formation du complexe enzyme- substrat selon les étapes suivantes :



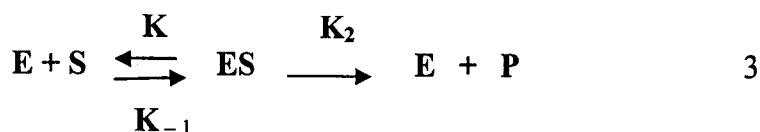
Avec  $k_1$ ,  $k_{-1}$ ,  $k_2$  et  $k_{-2}$  sont des constantes de vitesse.

Michaelis et Menten résolurent le système réactionnel en posant les hypothèses simplificatrices suivantes :

- les mesures de cinétique seront toujours faites pour des concentrations de produit très faibles : c'est la mesure de la vitesse initiale ( $v_i$ ) pour  $[P] = 0$  et  $[S] = [S]_0$ , où  $[S]_0$  est la concentration du substrat à l'instant initial.
- La concentration totale du substrat  $[S]_0$  est grande devant celle de l'enzyme  $[E]_0$ .

Dès l'addition de l'enzyme dans la solution de substrat, il s'établit un équilibre rapide entre les formes libres de l'enzyme, du substrat et du complexe, on parle d'hypothèse du pré-équilibre.

Le schéma réactionnel s'écrit ainsi :



De ce modèle Michaelis et Menten ont déduit une équation pour déterminer la vitesse initiale de formation du produit P en fonction de la concentration du substrat S.

$$[E]_0 = [E] + [ES] \quad [E] = [E]_0 - [ES] \quad \quad \quad 4$$

Si on considère l'équation de vitesse pour ES :

$$\frac{d[ES]}{dt} = K_1 [E] [S] - K_{-1} [ES] - K_2 [ES] \quad \quad \quad 5$$

A l'état stationnaire:

$$\frac{d[ES]}{dt} = 0 \Rightarrow K_1 [E] [S] - K_{-1} [ES] - K_2 [ES] = 0 \quad \quad \quad 6$$

$$\Rightarrow K_1 ([E]_0 - [ES])[S] - K_{-1} [ES] - K_2 [ES] = 0$$

$$\Rightarrow K_1 [E]_0 [S] - K_1 [ES][S] - K_{-1} [ES] - K_2 [ES] = 0$$

$$\Rightarrow K_1 [E]_0 [S] - [ES](K_1 [S] + K_{-1} + K_2) = 0$$

Qui donne:

$$[ES] = \frac{K_1 [E] [S]}{K_1 [S] + K_{-1} + K_2} \quad 7$$

$$V = K_2 [ES]$$

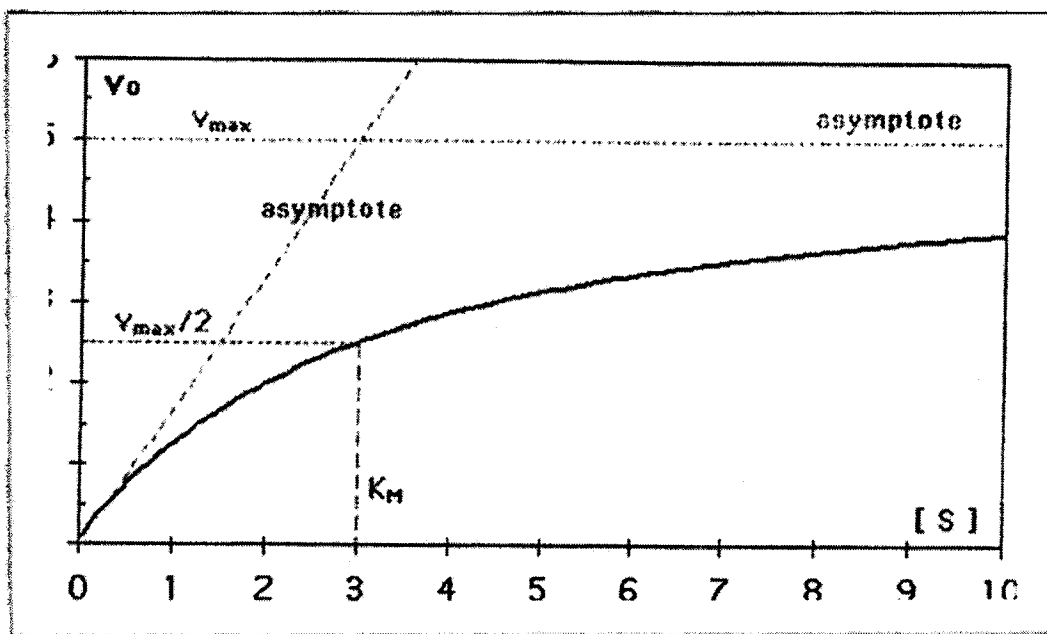
$$\Rightarrow V = \frac{K_2 [E]_0 [S]}{[S] + \frac{K_{-1} + K_2}{K_1}} \quad 8$$

$$K_m = \frac{K_{-1} + K_2}{K_1} \quad \text{Et } K_2 [E]_0 = V_{\max}$$

$$V = \frac{V_{\max} [S]}{[S] + K_m} \quad 9$$

Appelée équation de **Michaelis-Menten**

Le graphe de Michaelis-Menten est donné par la courbe  $V_0 = f[S]$  :



**Figure 11** : Graphe de Michaelis-Menten.

Le graphe de Michaelis- Menten indique dans quel rapport la formation du produit dépend de [S] quand on maintient [E] constant. A concentration basse de S, la vitesse de réaction est proportionnelle à [S] ; quand on augmente [S], la vitesse n'augmente pas indéfiniment en proportion de [S], mais va atteindre V<sub>Max</sub>, vitesse indépendante de [S]. V<sub>Max</sub> est proportionnelle à [E] et à une constante catalytique  $k_w$  qui est une propriété caractéristique de chaque enzyme ; en réduisant [E] de moitié, on réduit de moitié la vitesse de la réaction, pour toute valeur de [S].

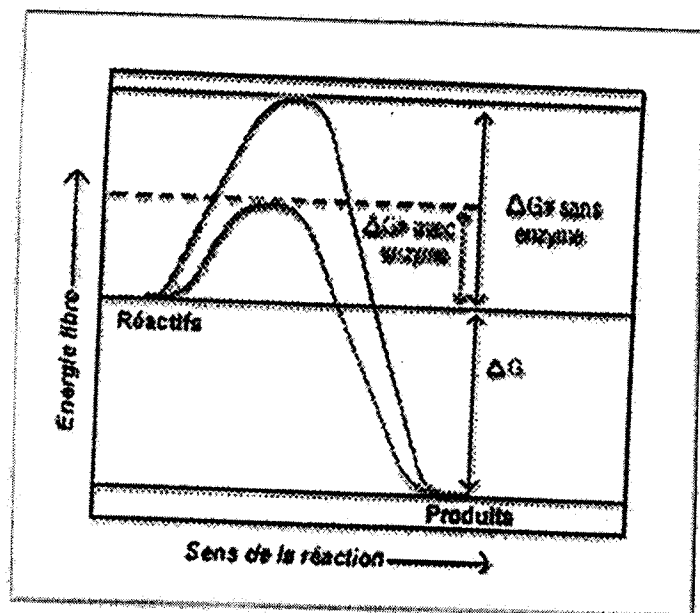
#### **IV-4. Comment fonctionnent les enzymes ?**

Les enzymes accélèrent des réactions chimiques qui peuvent se produire naturellement, mais à des vitesses beaucoup plus lentes. Une des caractéristiques des enzymes est qu'elles ne sont pas utilisées ou transformées durant une réaction chimique. Elles peuvent donc servir plusieurs fois. Le principe d'action est basé sur le fait que la plupart des réactions endothermiques nécessitent un apport d'énergie afin de démarrer. Cette énergie initialement requise est appelée l'ENERGIE D'ACTIVATION. Alors qu'on peut amorcer une réaction en apportant de l'énergie, une enzyme agit plutôt en se liant aux réactifs de façon à faciliter leur réaction et ainsi abaisser l'énergie d'activation requise. La température de l'environnement cellulaire est alors suffisante pour faire franchir le seuil de l'énergie d'activation aux réactifs et les faire réagir.

#### **IV-5. Catalyse enzymatique**

Les enzymes sont de bons catalyseurs car elles peuvent augmenter la vitesse de réaction d'un facteur allant de  $10^4$  à  $10^6$ . Elles ne modifient pas la position de l'équilibre de la réaction mais seulement la vitesse d'inter- conversion entre substrat et enzyme.

Un catalyseur a pour effet d'abaisser l'énergie d'activation en favorisant la stabilité de l'état de transition. Les énergies des réactifs et des produits restent inchangées, ce qui explique pourquoi le rapport à l'équilibre de ces diverses molécules ne varie pas.



**Figure 12 :** Représentation de la catalyse enzymatique.

- **Catalyse acide-base:** dans les sites actifs des enzymes, il y a des groupements acides carboxyliques ou aminés qui sont donneurs ou accepteurs de protons. Ceci permet donc de réaliser des réactions catalysées par les bases ou les acides.
- **Catalyse covalente:** des groupements tel les thiols ou les aminés sont des nucléophiles susceptibles d'attaquer les groupements électrophiles de substrat pour former des intermédiaires covalents hautement réactifs, ceux ci peuvent être attaqués à nouveau par de l'eau de façon à libérer le produit.
- **Catalyse par la chymotrypsine:** cette enzyme coupe la liaison peptidique des chaînes latérales aromatiques. Cette spécificité est due à la présence de groupements hydrophobes sur la chaîne. Le substrat est maintenu en place par l'interaction entre la poche hydrophobe et la chaîne latérale aromatique.

#### **IV-6. Le complexe enzyme – substrat**

Une fois que le substrat s'est rapproché de l'E, on constate une restriction des possibilités de rotation et de translation du substrat. Ceci est dû à des interactions faibles entre enzyme et substrat. Ces interactions sont nécessaires à la réaction. En effet, les réactions intramoléculaires sont accélérées lorsque les mouvements internes des molécules sont bloqués. De plus, les enzymes favorisent l'orientation des molécules de substrat: en effet, lorsque le substrat se fixe sur l'enzyme, ceci provoque un changement conformationnel qui amène les groupements fonctionnels du site actif dans une position favorable à la catalyse. Il y



a donc création d'interactions faibles entre E et S. Cette adaptation induite permet également la formation d'interactions faibles supplémentaires qui vont stabiliser l'état de transition.

La formation initiale d'un complexe enzyme - substrat E-S (NON covalent) fût suggérée d'après les observations suivantes :

a. Le haut degré de spécificité de la reconnaissance d'un substrat par une enzyme. Pour l'expliquer, Emil Fisher suggéra en 1894 que cette reconnaissance résulte d'une très forte complémentarité des structures (mais aussi de la nature chimique des "groupements réactionnels) du substrat et de l'enzyme qui le fixe, comme le sont une clé et la serrure ; dans laquelle elle entre.

b. La formation de la courbe dite de saturation : vitesse initiale de la réaction enzymatique en fonction de la concentration en substrat ( $v_i = f([S])$ ) [14, 15].

c. Le fait que les substrats protègent souvent les enzymes de l'inactivation. L'hypothèse "clé -serrure", bien qu'extrêmement satisfaisante, ne peut rendre compte de certaines observations :

a. par exemple, certains composés qui ressemblent chimiquement à un substrat mais qui ont des groupements moins volumineux ne sont pas catalysés, bien qu'ils doivent encore mieux s'insérer dans le site actif

b. Il existe un mécanisme enzymatique à deux substrats appelé "fixation ordonnée" pour lequel un substrat B ne peut se fixer que si le substrat A l'est déjà. Or, selon l'hypothèse "clé - serrure", le substrat B devrait se fixer d'emblée.

#### **IV-7. Spécificité de l'association protéine - ligand, site actif :**

*Quels sont les facteurs qui expliquent l'extrême spécificité de la reconnaissance entre une protéine et un ligand ?*

a. Toutes les protéines se replient dans une conformation dite native et c'est dans cette conformation qu'elles acquièrent leur activité biologique (leur pouvoir de catalyseur dans le cas des enzymes). Ce repliement aboutit à une structure tridimensionnelle unique de la protéine.

b. Cette structure globale de la macromolécule permet à une région particulière, (souvent enfouie au sein de la protéine), d'adopter elle aussi une structure spatiale qui est reconnue par le ligand spécifique de la protéine (et, le cas échéant, par un petit nombre de molécules dont la structure est proche de celle du ligand).

c. Le site actif est constitué d'un petit nombre d'acides aminés, qui le plus souvent ne sont pas proches dans l'enchaînement de la chaîne polypeptidique. Ces acides aminés sont

### **CHAPITRE III : Acides aminés, Peptides, Protéines, Enzymes**

---

caractérisés par une chaîne latérale dont à la fois la nature chimique (groupement ionisable ou polarisable) et la structure (encombrement stérique) sont particulières.

d. La stéréochimie qui résulte de cet agencement unique des acides aminés qui constituent le site actif est la cause de la stéréospécificité de reconnaissance entre ces acides aminés et le (ou les) ligand(s).

Les enzymes ne fixent pas seulement un ligand (un substrat) ; elles le transforment en un produit lors d'une réaction chimique. Certains acides aminés du site actif ont pour fonction, non pas de fixer le substrat, mais de fournir les groupements chimiques nécessaires à la réaction catalysée par l'enzyme. Dans le cas des enzymes, on distingue donc au sein du site actif, les acides aminés qui constituent le site de fixation (ces acides aminés n'ont pas de fonction chimique impliquées dans la réaction) et les acides aminés qui constituent le site catalytique.

On peut citer divers types d'association entre une protéine et un ligand : les complexes enzyme - substrat , enzyme - régulateur (inhibiteur, activateur ...) , antigène - anticorps , histones - ADN et protéines - hormones.

**REFERENCES :**

- [1] L. Stryer. La biochimie. Edition Flammarion 4eme édition, chapitre 2 et 16.
- [2] Graham L. PATRICK, Chimie pharmaceutique (traduction de la 2<sup>ème</sup> édition), 2003, pp.20-31.
- [3] H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, and P. E. Bourne. The protein data bank. *Nucleic Acids Research*, 2000, 28 , Pp.235-242.
- [4] R. Srinivasan and G. Rose. A physical basis for protein secondary structure. *PNAS*, 1999, 96, 14258-14263.
- [5] C. Branden and J. Tooze. Introduction to protein structure. Garland Publishing, 1999.
- [6] H. J. Greenberg, W. E. Hart, and G. Lancia. Opportunités for combinatorial optimization in computational biology. *INFORMS Journal on Computing*, 16(3), 2004.
- [7] T. Head-Gordon and J. C. Wooley. Computational challenges in structural and functional genomics. *IBM Systems Journal*, 2001, 40 , pp.265-296.
- [8] T. Lengauer. Computational biology at the beginning of the post-genomic era. In R. Wilhelm, éditeur, *Informatics : 10 Years Back -10 Years Ahead*, 2001, volume 2000 of Lecture Notes in Computer Science, pages 341\_355. Springer- Verlag,.
- [9] J. C. Setubal and J. Meidanis. Introduction to Computational Molecular Biology. PWS Publishing Company, 1997.
- [10] P. Bork, T. Dandekar, Y. Diaz-Lazcoz, F. Eisenhaber, M. Huynen and Y. Yuan. Predicting Function : from genes to genomes and back. *Journal of Molecular Biology*, Nov. 1998 : 283(4), pp.707-725.
- [11] Guillaume Collet, la reconnaissance de repliements de protéines, Rapport bibliographique de stage de Master, Université de Rennes1, 2005.
- [12] Reginald H. Garrett, Charles M. Grisham, Biochimie (traduction de la 2<sup>ème</sup> édition américaine), 2000, pp.427- 435.
- [13] L. MICHAELIS, M. L. MENTEN, „ Die Kinetik der Invertinwirkung *Biochemi.* 1913, 9, pp.333-369.
- [14] V. C. Henri, „ *r. hebd. Acad. Sci. Paris* , 1902, 135, pp.916-919.
- [15] V. Henri, Lois Générales de l'Action des Diastases, Hermann, Paris, 1903.

# Chapitre IV :

*Application, Résultats et Discussions*

### **Introduction :**

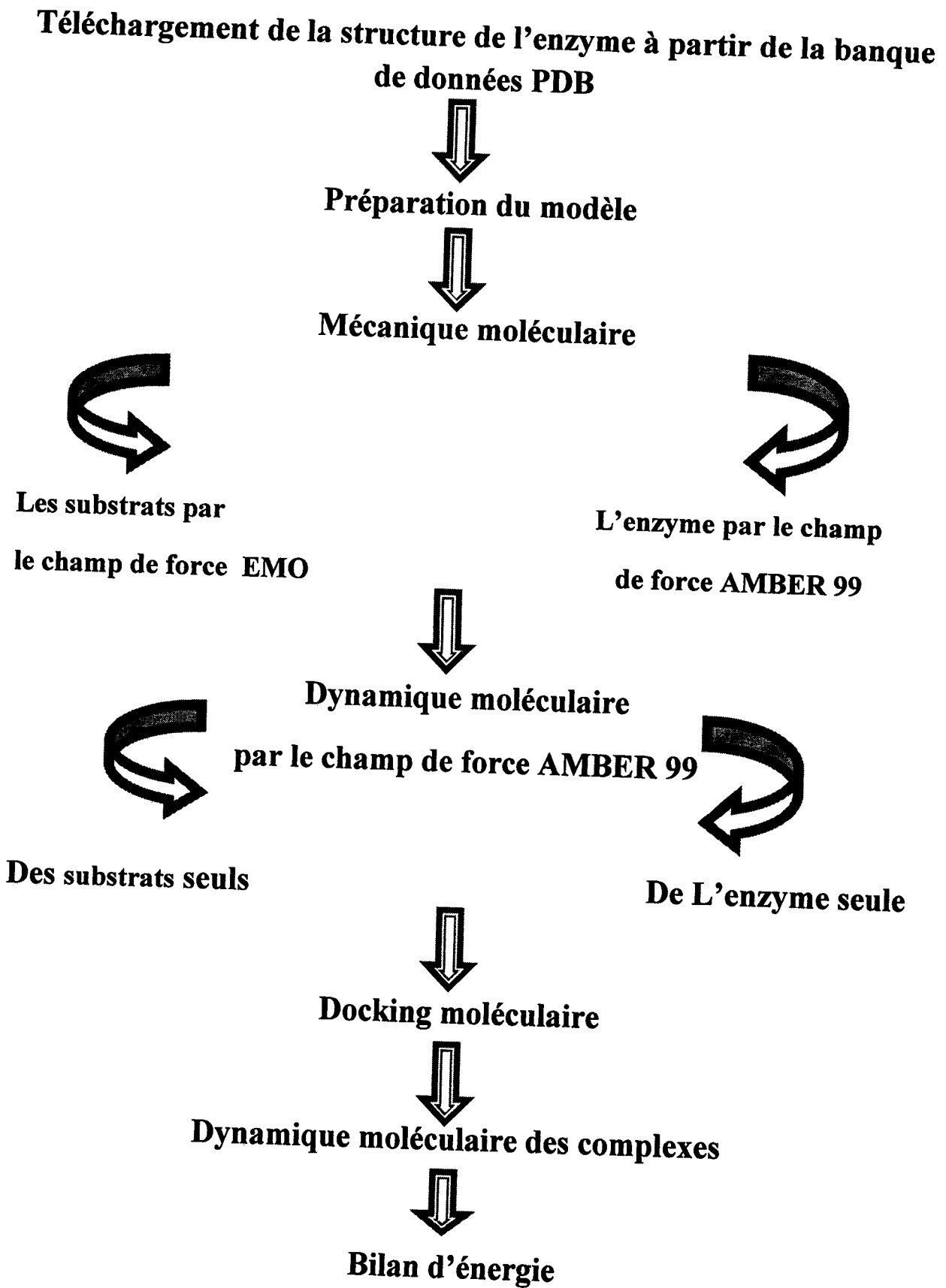
La discussion que nous allons engager est basée sur les énergies d'interactions et les distances entre certains groupements de la chaîne latérale de l'enzyme et ceux des ligands.

Par conséquent l'optimisation des interactions de Van der Waals de type attractif a lieu lorsque les géométries du site actif de liaison de l'enzyme et celle du substrat deviennent davantage complémentaires (modèle clé-serrure). Il est à noter qu'une parfaite complémentarité de forme entre l'enzyme et le substrat est très rare.

Il faut signaler que tant que les atomes du substrat et de l'enzyme ne sont pas trop proches, les interactions de type de Van der Waals restent de nature attractive et certainement vont stabiliser le complexe formé. Il existe des situations défavorables dans lequel nous avons un très grand espace « vide » entre l'enzyme et le substrat.

Cette situation peut engendrer un changement conformationnel qui se produit afin de pouvoir accommoder l'établissement d'une nouvelle interaction entre enzyme et le substrat. Une interprétation possible suggère la création d'une grande différence entre ces deux entités du point de vue énergétique. Ceci ne peut être compensé par la formation d'une nouvelle interaction.

La demarche suivie :



## I- Préparation de l'enzyme :

Le téléchargement de la CTPR3 a été exécuté à partir de la base de données **Bookhaven Protein Data Bank** (code d'accès 1na0).

### Visualisation tridimensionnelle des molécules :

Depuis longtemps la modélisation moléculaire occupe une place primordiale dans la recherche en chimie, l'évolution des ordinateurs ces dernières décennies est devenu une part intégrante de la vie du scientifique avant 1984, peu de scientifiques utilisaient un ordinateur personnel ; il était réservé au chimiste très expérimenté qui utilisait des programmes complexes difficile à maîtriser.

La possibilité de visualiser de grosses molécules biologiques par ordinateur est très intéressante parce qu'il est très difficile de parler par exemple de la structure d'une hélice  $\alpha$  et des interactions qui en assurent la stabilité sans observer véritablement donc la visualisation tridimensionnelle de la structure moléculaire est nécessaire à la compréhension des processus biologiques.

L'hypothèse concernant l'importance de la forme des molécules avait été émise dès 1946. Pauling et Astbury en 1951 proposent des modèles théoriques de structures secondaires hélice  $\alpha$  et le feuillet  $\beta$ , les logiciels de visualisation et de manipulation des molécules sont actuellement disponibles avec des fonctionnalités conviviales (par exemple Hex4e [1], Rasmol version 2.6 [2] et Hyperchem 7.5 version Professionnelle [3]).

- **Rasmol**

Le logiciel Rasmol :

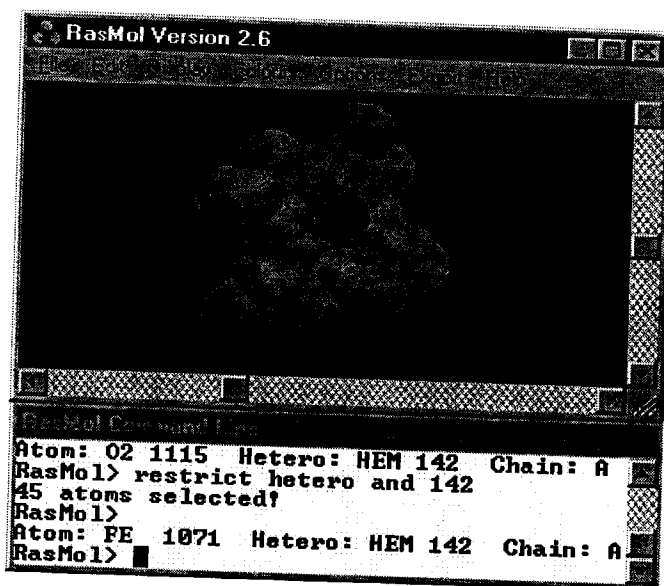
- Fonctionne sous Windows
- Il est exploitable à partir d'un ensemble de deux fenêtres :  
Une fenêtre donnant l'accès à des macro commandes (menu) et à l'affichage de l'image et une deuxième dans laquelle, à la fois on récupère des informations (séquence, position, distance...) et on signifie les instructions en utilisant un langage simple de commande.

Il ouvre des fichiers de molécules assez volumineuses (plus de 1000 atomes).

- Un comporte des fonctions d'affichage des informations (nom de la

molécule, sa séquence par chaîne, ses structures) qui sont lues dans le fichier PDB.

- Il permet d'avoir différents modes de visualisation du modèle moléculaire : sphères, bâtonnets, sphères et bâtonnets, ruban,...



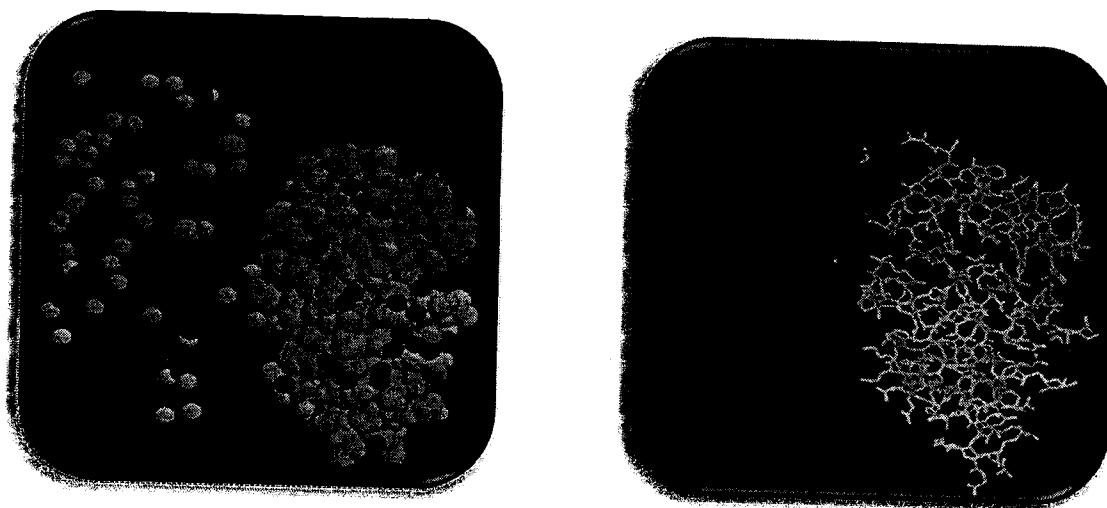
**Figure 1** : Aperçu sur les deux fenêtres de Rasmol.

- Tous les éléments composant une molécule sont accessibles en cliquant directement sur l'écran. On peut ainsi les identifier, obtenir des indications sur les distances ou les angles entre atomes.

La structure tridimensionnelle de la CTPR3 a été obtenue par diffraction en rayon X avec une résolution (1.60 Å) [4]. On remarque que la CTPR3 cristallise sous la forme d'un dimère (figure 1), chaque monomère est constitué de 983 atomes.

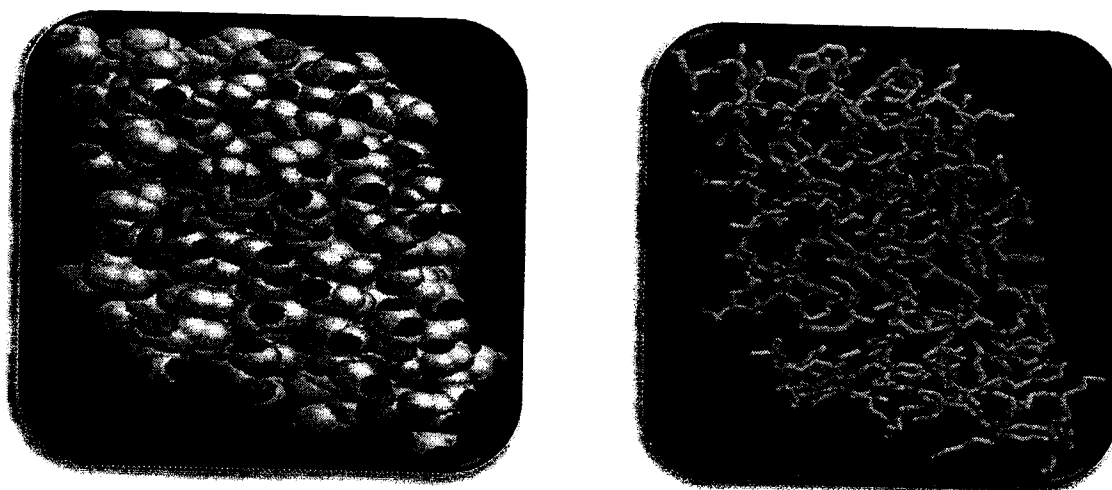
Avec l'effet réducteur de la modélisation moléculaire, nous avons simplifié le modèle de l'enzyme et nous avons retenu un seul monomère (figure 2).





**Figure 2** : la forme dimère de la CTPR3

Avec l'effet réducteur de la modélisation moléculaire, nous avons simplifié le modèle de l'enzyme et nous avons retenu un seul monomère (figure 2).



**Figure 3** : le model simplifié de la CTPR 3

## **II- Préparations des ligands :**

Il est nécessaire d'attribuer à chaque atome un potentiel, c'est à dire une charge partielle électronique. Ce potentiel va dépendre de l'environnement chimique dans lequel chaque atome est placé et du champ de force qui sera utilisé. L'objectif d'un champ de force est de décrire une classe entière de molécules avec une précision raisonnable. Il représente une des approximations dans la modélisation moléculaire. La qualité d'un champ de force, son application au système d'une part, et sa bonne qualité prédictive des propriétés

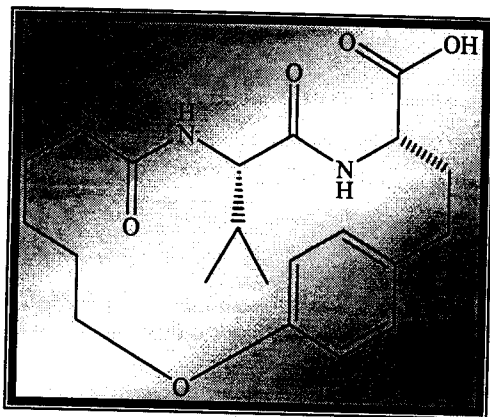
mesurées lors des simulations, d'autre part, influencent directement sur la validité des résultats.

De là, vient la nécessité d'utiliser le champ de force le plus performant possible et qui soit adapté au type de molécules à modéliser.

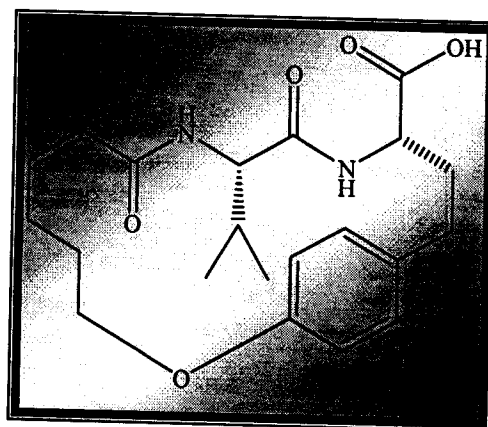
La valeur de l'énergie d'optimisation est :

$$E_{\text{pot}} = -981.107 \text{ Kcal/mol}$$

**III- 2- Optimisation des substrats:**



Substrat 01



Substrat 02

La construction et l'optimisation de ligand étaient faite par le programme EMO (énergie of molécule)

- La minimisation de l'énergie stérique de ce ligand a été effectuée en utilisant l'optimisation SCAN, qui permet :

- Un balayage de la surface d'énergie.
- D'éliminer les conformères les moins probablement stables.
- D'éviter les minima locaux.

- Les résultats obtenus sont reportés dans le tableau suivant :

**Tableau 01:** Résultats obtenus à l'aide du programme EMO

<b>Énergie stérique (KJ/mole)</b>	<i>Élongation</i>	<i>Valence</i>	<i>Torsion</i>	<i>VDW.</i>	<i>Électro.</i>	<i>Solvant</i>	<i>Totale</i>
<i>Substrat 01</i>	10.39	63.61	-01.61	96.35	-52.87	00.00	115.86
<i>Substrat 02</i>	12.56	169.68	103.16	130.01	-50.87	00.00	364.53

### IV- Dynamique Moléculaire :

La connaissance à un niveau atomique des aspects structuraux et dynamiques des systèmes organisés est particulièrement importante pour mieux comprendre les fonctions de ces édifices moléculaires complexes. Dans de nombreux cas, l'obtention du détail microscopique par les techniques expérimentales conventionnelles s'avère impossible. Cependant, la véritable évolution des moyens informatiques amorcée depuis une dizaine d'années, et le développement d'algorithmes performants, rendent possible l'étude d'assemblages supramoléculaires de complexité croissante par les méthodes de chimie théorique [5].

Le principe de la dynamique moléculaire est de fournir de l'énergie cinétique au système pour franchir des barrières énergétiques.

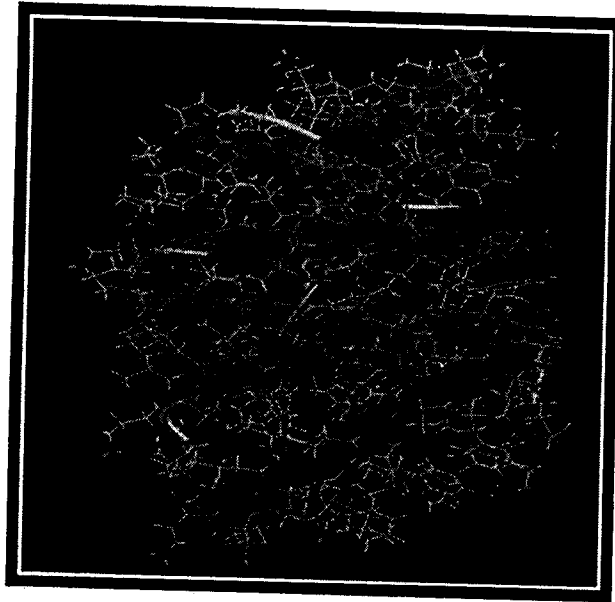
En générale, l'algorithme de la plus grande pente ou «Steepest descent» est utilisé au début, pendant 100 à 200 pas. Après l'algorithme de gradient conjugué ou «Conjugate Gradient» peut être utilisé pour compléter la minimisation jusqu'à la convergence. La convergence dans l'algorithme de «Steepest Descent» est lente, mais cette méthode est extrêmement robuste. Cet algorithme est surtout utilisé lorsque les conformations sont éloignées de leur minimum d'énergie. L'algorithme du gradient conjugué est la méthode de choix pour les grands systèmes moléculaires, puisqu'il présente l'avantage de la convergence rapide.

La dynamique moléculaire a été effectuée sur la CTPR3 seule, les ligands seules et les complexes.

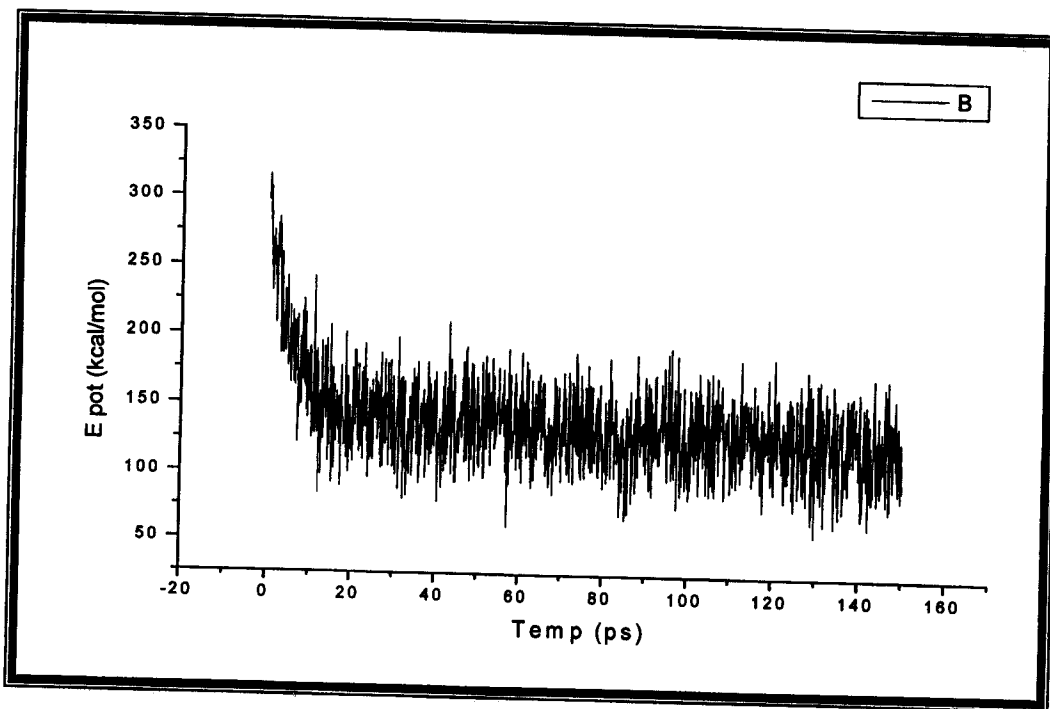
#### IV- 1. Dynamique moléculaire de la CTPR 3 :

Nous avons commencé la dynamique par une initialisation du système : à  $t=0$  nous avons  $r(t)=0$ , c'est-à-dire la structure initiale, précédemment minimisée. Puis nous avons chauffé le système jusqu'à 300 K pendant 1000 pas avec un pas d'intégrations de 1 fs.

A 300 K, il y a une équilibration : les vitesses sont réajustées pour conserver la température constante (il y a échange entre l'énergie cinétique et l'énergie potentielle). Ensuite, il y a production de conformations. Le temps de simulation de la dynamique moléculaire est de 100 picosecondes.



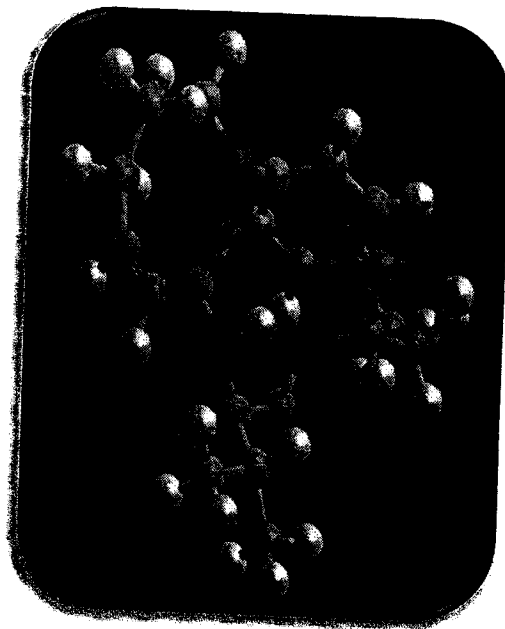
1na0



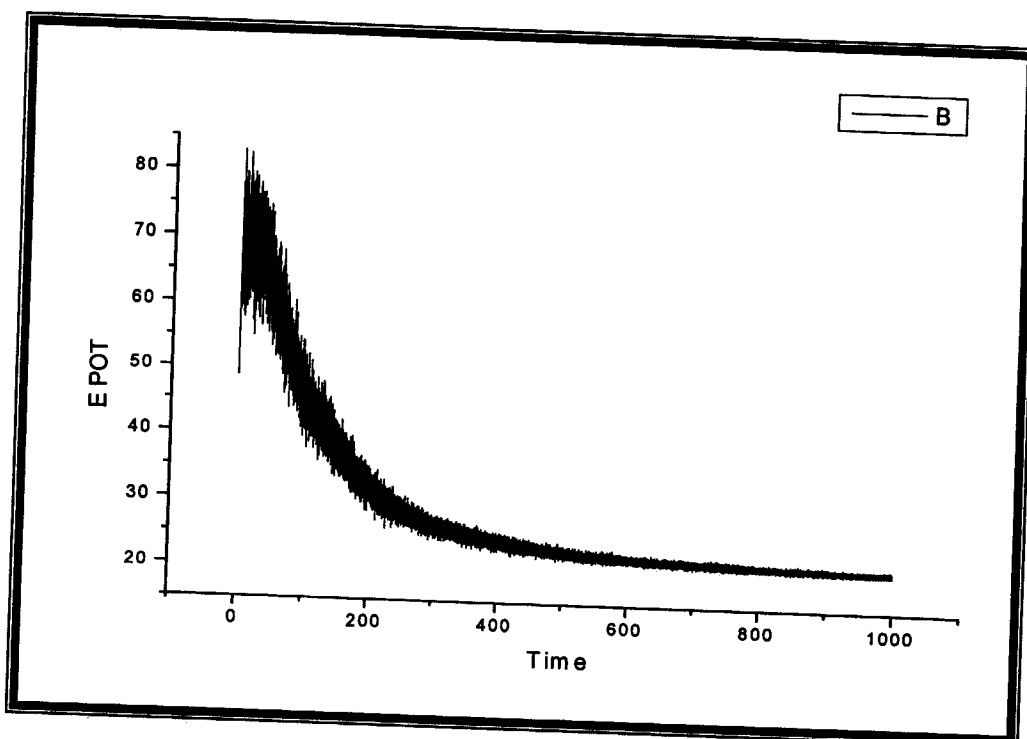
**Figure 4** : variation de l'énergie potentiel de la 1na0 seul en fonction du temps

IV- 2. Dynamique moléculaire des substrats :

Substrat 01 :



Substrat 01

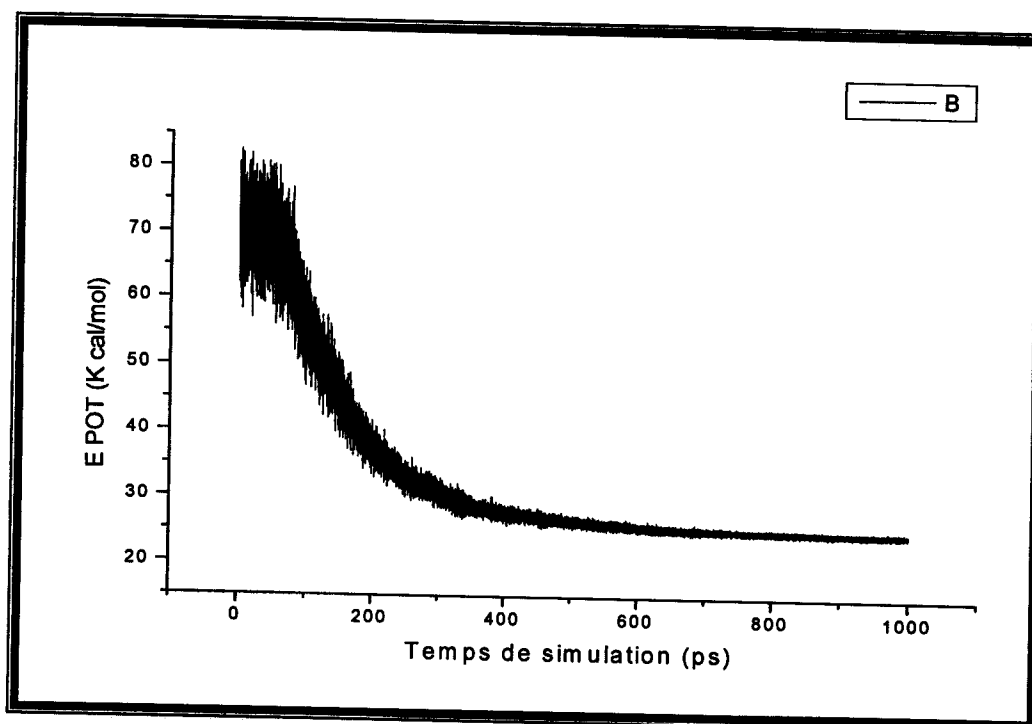


**Figure 5 :** variation de l'énergie potentiel du Substrat en fonction du temps

**Substrat 02**



**Substrat 02**



**Figure 6** : variation de l'énergie potentielle de substrat 02 en fonction du temps

**V- Docking des substrats et construction des complexes :**

L'étape suivante, après la construction des ligands, est le positionnement de ces molécules dans le site actif de la CTPR 3. Pour ce faire, nous avons utilisé le module Docking Moléculaire à l'aide logiciel Hex4e

Une fois que le complexe ligand- récepteur est formé, celui-ci va adapter la conformation la plus stable, c'est à dire avec le niveau énergétique le plus faible.

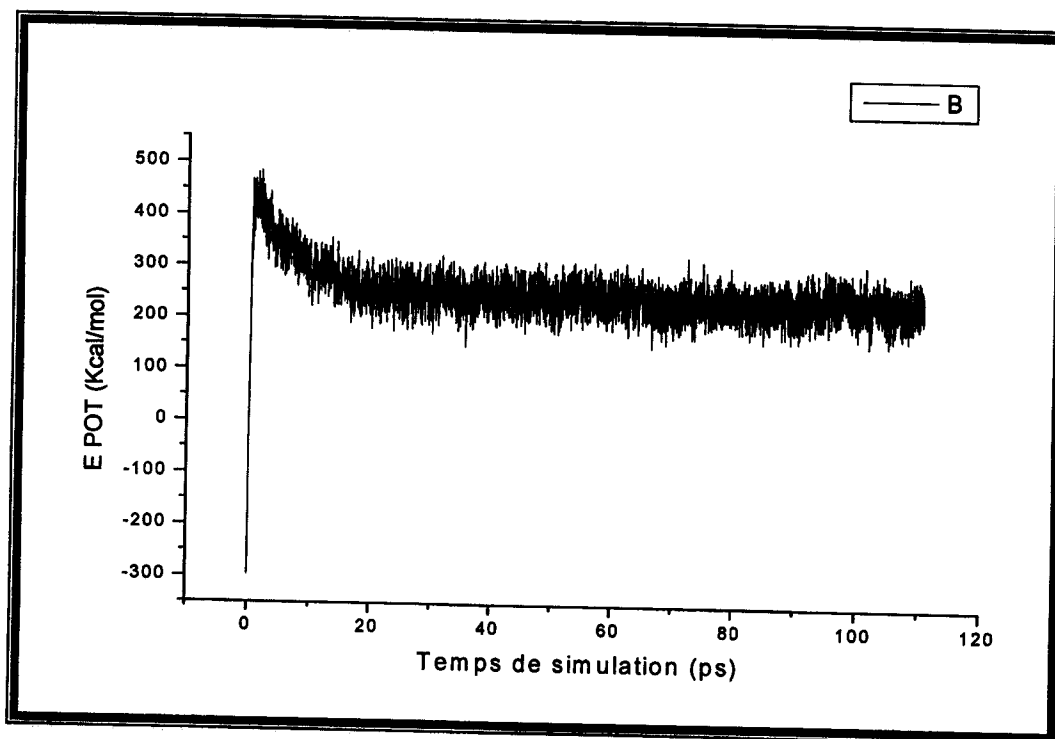
**VI. Dynamique moléculaire des complexes :**

Une fois tous les complexes formés, nous allons effectuer une optimisation de la géométrie et un calcul de dynamique moléculaire pour chercher la conformation la plus stable.

Complexe 1na0-substrat 01



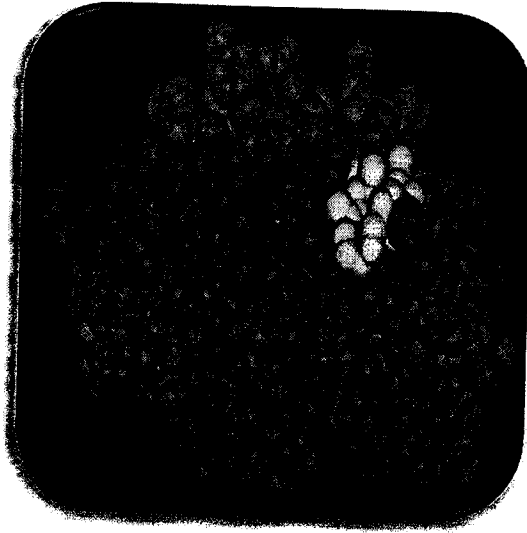
Complexe 1na0-substrat 01



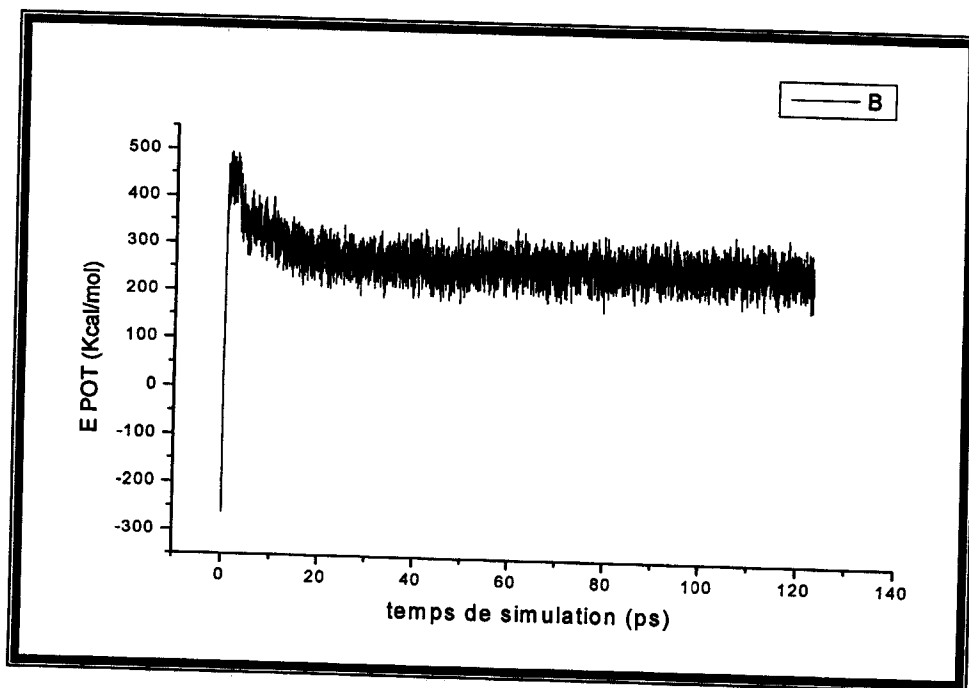
**Figure 7** : variation de l'énergie potentielle du complexe 1na0-substrat 01 en fonction du temps



Complexe 1na0-substrat 02



Complexe 1na0-substrat 02



**Figure 8:** variation de l'énergie potentielle du complexe 1na0-substrat 02 en fonction du temps

**VII. Énergie d'interaction :**

Dans le but d'établir des corrélations entre les énergies d'interaction et les activités, et pour envisager par quel type de forces la nature des interactions CTPR3-substrat est gouvernée, on fait souvent référence à une décomposition de l'énergie potentielle totale de la molécule en plusieurs termes (Énergie électrostatique, Energie de VDW,...).

Les énergies d'interaction entre les différents substrats étudiés et la CTPR3 sont obtenues à l'aide de la relation suivante :

$$E_{\text{interactions}} = (E_{\text{potentielle totale complexe CTPR 3-substrat}}) - (E_{\text{potentielle totale CTPR3 seule}} + E_{\text{potentielle totale Substrat}})$$

**Tableau 02 :** Bilan des énergies en (Kcal/mol) :

	$E_{\text{tot du complexe}}$	$E_{\text{tot du substrat}}$	$E_{\text{Vdw du complexe}}$	$E_{\text{Vdw du substrat}}$	$E_{\text{inter. VdW}}$	$E_{\text{inter. totale}}$
<i>Substrat01</i>	236.05	21.01	-520.31	96.35	-227.36	122.13
<i>Substrat02</i>	223.80	24.91	-525.25	130.01	-373.02	105.97

**Tableau 03 :** Bilan des volumes molaires en ( $\text{Å}^3$ ) :

	<i>Volume molaire <math>\text{Å}^3</math></i>
<i>Substrat 01</i>	1786.60
<i>Substrat 02</i>	1420.55

A la lumière des résultats obtenus au cours de notre travail de mémoire, il ressort que les valeurs obtenues relatives aux énergies d'interaction totale et aux énergies de Van der Waals sont du même ordre de grandeur pour les deux substrats.

Il faut tenir compte de l'interaction de Van der Waals puisque c'est l'interaction entre atomes non liés qui stabilisent le complexe enzyme substrat.

Dans le souci de faire une comparaison entre ces deux substrats (substrat 01 : tiré de la littérature, substrat 02 : conçus) nous avons effectué la même démarche de calcul pour le substrat 02.

Nous allons axer notre discussion, sur les énergies de Van der Waals responsable de l'encombrement stérique, pour le substrat 1 (-227.37 kcal/mol) et pour le substrat 02 (-273,02 kcal/mol).

Notons une particularité entre le substrat 01 et le substrat 02, au niveau de ce dernier la présence d'une double liaison insérée dans le cycle à 18 chaînons.

Ce qui lui confère une conformation relativement stable d'énergie plus basse que le substrat 1; donc une interaction avec CTPR 3 plus importante.

Les volumes molaires montrent que le substrat 02 ( $1420.56 \text{ \AA}^3$ ), présente une géométrie inférieure à celle du substrat 01 ( $1796 \text{ \AA}^3$ ) ce qui laisse suggérer que la cavité du site actif est de même ordre de grandeur et que la complémentarité est relativement importante pour le substrat 02 que le substrat 01.

**REFERENCES:**

- [1] D. Ritchie. computing science university, copyright © 1996-2005.
- [2] H. Bertnetenie. Molecular graphics, copyright © 1998-1999.
- [3] Molecular Modeling system copyright © 2002 hypercube. inc.
- [4] R.G Ewan , Y. xiong, J. Melanie, L. D'andrea, and Lynne Regan. Design of stable  $\alpha$ -Helical arrays from an Idealized TPR Motif.2003 Elsevier science Ltd. Vol.11, 497-508.
- [4] D. A. Case, D.A. Pearlman, J.W. Caldwell, T.E. Cheatham III, W.S. Ross, C.L. Simmerling, T.A. Darden, K.M. Merz, R.V. Stanton, A.L. Cheng, J.J. Vincent, M. Crowley, V. Tsui, R.J. Radmer, Y. Duan, J. Pitera, I. Massova, G.L. Seibel, U.C. Singh, P.K. Weiner, P.A. Kollman, 1999, programme Amber 6, University of California, San Francisco.
- [5] Christophe Chipot, *Centre National de la Recherche Scientifique*, les méthodes numériques de la dynamique moléculaire, **septembre 2003**.
- [6] J. Gibart, J. Garnier, Cours de modélisation moléculaire, seconde édition, DRED- DS2, ATF « Chimie et Physico- Chimie des Grandes Molécules »,1993.

# *Conclusion générale*

## CONCLUSION GÉNÉRALE

L'objectif de notre travail consiste en l'étude des interactions enzyme-substrat par modélisation moléculaire et plus précisément, l'étude des interactions de tetratricopeptide (CTPR 3) avec des structures de type peptide cyclique synthétisées dans notre laboratoire (COSNA) par l'équipe de synthèse organique. Les outils informatiques retenus pour mener à bien cette étude sont la mécanique moléculaire, la dynamique moléculaire, et le Docking Moléculaire.

Cette étude comprend quatre chapitres :

1/ **Le chapitre I** est réservé à la description des différentes méthodes de la modélisation moléculaire (mécanique moléculaire et dynamique moléculaire), les techniques de «Docking Moléculaire» qui a pour but la détermination des modes d'interactions entre deux entités moléculaires. La définition du champ de force est aussi apportée, le champ de force représente l'une des approximations en modélisation moléculaire, et qui a pour objectif de décrire des classes entières de molécules avec une précision relativement probable. Il faut noter que la qualité d'un champ de force et son applicabilité au système d'une part, et sa bonne qualité prédictive des propriétés calculées lors des simulations, d'autre part, influencent directement la validité des résultats. D'où la nécessité du choix du champ de force adapté au type de molécules à étudier.

2/ **Dans le chapitre II**, nous avons donné un aperçu sur l'analyse conformationnelle qui demeure une méthode qui nous permet de calculer l'énergie relative associée à la conformation d'une molécule, dans le but de trouver le plus de minima possible

3/ **Au niveau de chapitre III**, des notions sur les acides aminés, peptides et protéines sont apportées en insistant sur les différents niveaux d'énergies structurales (primaire, secondaire, tertiaire et quaternaire). L'importance des différents types d'interactions est signalée. Ces dernières maintiennent les différentes conformations engendrant des activités conséquentes.

4/ **Le chapitre IV** est consacré aux résultats obtenus ainsi que la discussion et l'interprétation de ces derniers. La discussion est basée sur les énergies d'interactions, et les volumes molaires des substrats.

Ce travail nous a permis de faire un classement et d'opter pour le substrat 2. Il est à noter que la reconnaissance moléculaire des complexes Enzyme -substrat est hautement difficile et que notre contribution à la compréhension des processus d'interaction

## CONCLUSION GÉNÉRALE

---

reste certainement à compléter. Il est souhaitable d'avoir des tests biologiques pour une éventuelle validation du modèle utilisé.