



République Algérienne Démocratique et Populaire



Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

UNIVERSITE ABOU BEKR BELKAID TLEMCCEN

Faculté des Sciences

Département De Chimie

Laboratoire de recherche

Substances Naturelles et Bioactives (LASNABIO)

THESE

Présentée par : **M^{me} BOUDGHENE STAMBOULI Née SARI MOHAMED SOUAD**

En vue de l'obtention du Diplôme de Doctorat

En Chimie Physique

Option : Chimie Théorique et Modélisation Moléculaire

Thème

Etude de l'Interaction Protéine-Protéine par Modélisation Moléculaire. Application aux Processus Biologiques

Devant le jury composé de :

Prof. Malika MOKHTARI	Université de Tlemcen	Présidente
Prof. Tewfik BOUCHAOUR	Université de Tlemcen	Examineur
Prof. Aissa BELOUATEK	Centre Universitaire-Relizane	Examineur
Prof. Mohamed KOTBI	Université de Tlemcen	Examineur
Prof. Driss AINAD-TABET	Université de Sidi Belabbes	Examineur
Prof. Said GHALEM	Université de Tlemcen	Directeur

- Janvier 2013 -

DEDICACE

A mon père, pour son amour et son encouragement

*A ma mère, symbole d'affection, d'amour et de sacrifices,
avec toute ma reconnaissance.*

*A mon mari Abderrahim, pour son soutien de tous les jours
et son aide dans les moments difficiles, pour sa gentillesse,
sa patience ... et tellement plus ...*

A mes deux enfants Choukri et Meriem.

*A ma belle-mère pour sa compréhension et son soutien
moral.*

A mes frères et à ma belle-sœur Meriem.

A mes beaux-frères et mes belles sœurs.

A ma très chère amie Imane.

A tous les membres du laboratoire LASNABIO.

A toute ma famille et ma belle-famille réunies

REMERCIEMENTS

*Le présent travail a été réalisé au sein du **Laboratoire de Substances Naturelles et Bioactives (LASNABIO)**, de l'université **Abou Bekr Belkaïd de Tlemcen**, sous la direction du **Professeur Saïd Ghalem**. Je tiens tout particulièrement à lui exprimer ma profonde reconnaissance d'avoir accepté de diriger cette thèse. Je tiens aussi à lui témoigner toute ma gratitude pour l'intérêt qu'il a constamment manifesté pour ce travail.*

*J'adresse mes vifs remerciements à **Mademoiselle Malika MOKHTARI**; Professeur à l'université de Tlemcen, qui m'a fait l'honneur de présider le jury de cette thèse.*

*J'adresse mes plus sincères remerciements à Messieurs les professeurs **Tewfik BOUCHAOUR** et **Mohamed KOTBI** de l'université de Tlemcen, à **Driss AINAD-TABET** de l'université de Sidi Belabbès et à **Aïssa BELOUATEK** du centre universitaire de Relizene ; qui ont accepté d'examiner ce travail et de participer aux membres de jury.*

Mes remerciements s'adressent également à tous mes collègues du laboratoire LASNABIO : I. ABDELLI, M. BENJAROU, N. MERAD, N. KHÉBICHAT, M. MESMOUDI, I. DAOUD, F. BOUKLI-HACENE, W. SOUFI, K. NEGAZ, pour leurs conseils et leur aide précieuse.

Enfin, j'exprime toute ma reconnaissance à ma famille et mes amis pour leur amour et leur soutien.

TABLE DE MATIERES

Remerciements

Liste des abréviations

Introduction Générale	1
------------------------------------	---

Chapitre I : Interactions Protéine-Protéine

1. Structure des protéines.....	8
a) La résolution.....	8
b) Le facteur.....	9
c) Choix d'une structure cristallographique.....	9
d) Facteur d'agitation thermique.....	10
e) Détermination du site actif.....	10
f) Les interactions entre le ligand et le récepteur.....	10
2. La flexibilité d'une protéine.....	10
3. Les complexes macromoléculaires.....	11
3.1. La stabilité des complexes.....	11
3.2. Les complexes protéine-protéine.....	13
3.3. Caractéristique des interfaces protéine-protéine.....	14
3.4. Propriétés physico-chimiques des interfaces des complexes protéiques.....	16
4. Importances des interactions protéiques.....	16
5. Interactions protéines-protéines.....	17
5.1. Mise en évidence des interactions protéines-protéines.....	20
5.2. Le rôle du solvant.....	21
5.3. Les changements de conformation.....	21
5.4. Identification des interactions protéines-protéines à grande échelle.....	22
5.5. Les bases de données des interactions protéines-protéines.....	24
5.6. Pourquoi porter autant d'intérêt aux interactions avec les protéines ?.....	25
Références	28

Chapitre II : Protéine Kinase et Peptides

1. La famille des protéines kinases comme cibles thérapeutiques et analyse de leurs principaux inhibiteurs.....	35
2. Classification des protéines kinases.....	37
3. Protéine kinase A.....	37
4. Mécanisme d'activation des protéines kinases.....	39
5. Les inhibiteurs de protéines kinases.....	40
a. Ligands compétitifs de l'ATP.....	41
b. Ligands compétitifs du substrat.....	42
6. Inhiber la protéine kinase ; faut-il cibler son état actif ou inactif ?.....	43
7. Peptides à activité biologique.....	48
8. Rôles attribués aux peptides biologiquement actifs.....	48
9. Les peptides antihypertenseurs.....	49
10. L'activité biologique.....	50
Références	51

Chapitre III : Méthodes et Logiciels Utilisés

1. Introduction.....	57
2. Base de la mécanique moléculaire.....	57
3. Base de la dynamique moléculaire.....	58
3.1. Influence de la température.....	60
3.2. Problème du solvant.....	61
3.3. Intérêt de la dynamique moléculaire en amont du docking.....	62
4. Utilisation de la mécanique et de la dynamique pour la découverte et la conception de nouvelles molécules.....	62
4.1. Cas 1 : la structure 3D est connue.....	64
4.2. Cas 2 : la structure 3D n'est pas connue.....	66
5. Problèmes liés à l'approche par modélisation moléculaire.....	67
6. Modélisation d'une protéine.....	67
6.1. Modélisation par champ de force.....	69

6.2.	Simulation de biomolécules par dynamique moléculaire.....	69
6.3.	Le docking moléculaire.....	71
6.3.1.	Les programmes d'assemblage moléculaire.....	72
6.3.2.	Les divers programmes de docking moléculaire.....	74
7.	Logiciels utilisés.....	78
7.1.	Hyperchem8.....	78
7.2.	Hex 6.3.....	78
Références	79

Chapitre IV : Résultats et discussions

1.	Introduction.....	83
2.	Préparation du récepteur.....	83
3.	Construction des peptides.....	85
4.	Minimisation d'énergie.....	86
5.	Dynamique moléculaire.....	87
6.	Docking des peptides : construction des complexes.....	88
7.	Résultats et discussions.....	88
7.1.	Optimisation des peptides.....	88
7.2.	Optimisation de la protéine kinase A.....	89
7.3.	Dynamique moléculaire de la protéine kinase A.....	90
7.4.	Dynamique des peptides.....	92
7.5.	Formation des complexes.....	96
7.6.	Energie d'interaction.....	103
Références	110
Conclusion Générale	113

Liste des abréviations

3D : tridimensionnel

PDB : Protein Data Bank

RMN : Résonance Magnétique Nucléaire

Rx : Rayons x

D (Asp) : acide aspartique

E (Glu) : acide glutamique

F (Phe) : phénylalanine

G (Gly): glycine

Y (Tyr): tyrosine

W (Trp): tryptophane

WEG: L-Trp-L-Glu –L-Gly

feG: D-Phe-D-Glu-L-Gly

FEG: L-Phe-L-Glu-L-Gly

yeG: D-Tyr-D-Glu-L-Gly

wdG: D-Trp-D-Asp-L-Gly

weG: D-Trp-D-Glu-L-Gly

PK: protéines kinases

PKA: protéine kinase A

AMPC: adenosine monophosphatecyclique

ATP: adenosine triphosphate

ARN: Acide RiboNucléique

ADN: Acide DésoxyriboNucléique

IPP : interaction protéine-protéine

Ki : constante d'inhibition

IC₅₀ : constante inhibitrice à 50% de l'inhibition totale.

EC₅₀ : concentration effective à 50% de l'effet total.

QM : mécanique quantique

MM : mécanique moléculaire

MC : Monte-Carlo

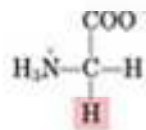
DM : Dynamique moléculaire

FFT : Fast fourier transform

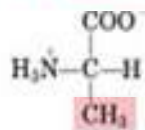
PFT : polar fourier transform

les vingt acides aminés

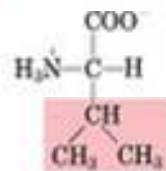
non polaires



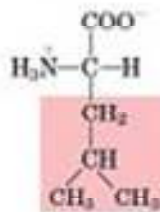
glycine



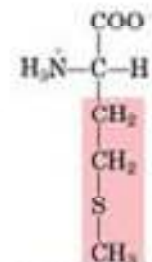
alanine



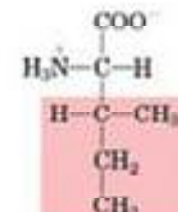
valine



leucine

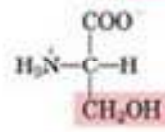


méthionine

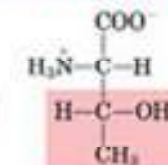


isoleucine

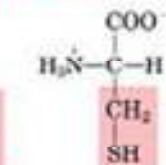
polaires mais non-chargé



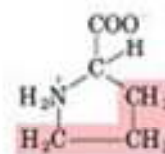
sérine



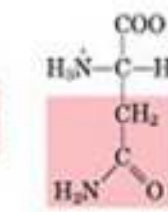
thréonine



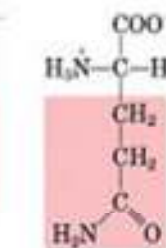
cystéine



proline

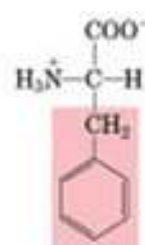


asparagine



glutamine

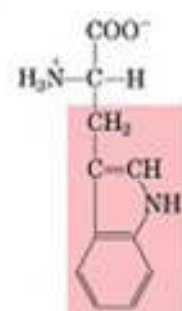
aromatiques



phénylalanine

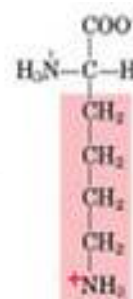


tyrosine

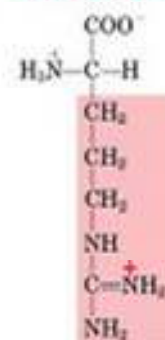


tryptophane

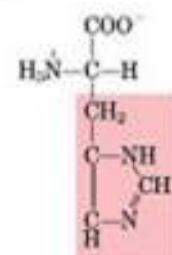
chargés positif



lysine

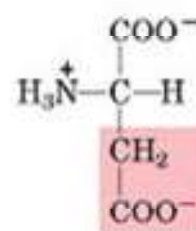


arginine

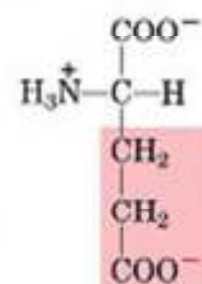


histidine

chargés négatif



aspartate



glutamate



INTRODUCTION GENERALE

Les interactions protéine-protéine sont un aspect essentiel des processus biologiques. Elles sont fortement impliquées dans la formation de structures macromoléculaires, dans la signalisation, dans la régulation et dans les différentes voies métaboliques. Leur étude est donc cruciale pour la compréhension des réseaux d'interaction protéiniques, but majeur dans l'étude des systèmes biologiques. Les interactions protéine-protéine ont un rôle conséquent dans l'induction de beaucoup d'états pathologiques et dans les processus importants pour la pathogenèse des infections bactériennes et virales [1].

La haute spécificité de ces interactions fait d'elles des cibles idéales pour les agents thérapeutiques [2,3]. Dans certains cas, notamment chez les plantes, l'identification des interactions protéine-protéine est une étape critique dans la détermination des fonctions de protéines inconnues. En effet, l'absence d'outils technologiques comme une vaste collection d'anticorps ne permet pas dans certains cas d'étudier directement la fonction d'une protéine donnée.

Toutefois, c'est l'augmentation des technologies d'études des interactions ainsi que leur perfectionnement qui reste primordial, puisque toute étude bioinformatique doit être validée par une étude expérimentale. Le nombre d'interaction entre protéines est considérable. L'exploration de ces réseaux protéiques complexes nécessite des méthodologies adaptées ainsi que des outils bioinformatiques puissants pour les analyser.

Techniques d'études des interactions [4,5] :

1. Méthodes génétiques.

L'intérêt pour les interactions protéine-protéine a précédé les outils pour les étudier. La bonne connaissance de certains génomes comme celui de *Saccharomyces cerevisiae* ou de *Drosophila melanogaster* a permis d'étudier indirectement les interactions entre protéines via la génomique. Ces techniques sont basées sur la bonne connaissance des gènes codant les protéines d'intérêt et leurs mutations. Ces techniques sont devenues anciennes avec l'apparition des méthodes biochimiques.

2. Méthodes biochimiques

Les techniques biochimiques sont les plus utilisées parce qu'elles sont les plus nombreuses, les moins chères et les plus accessibles ; généralement elles ne nécessitent pas de matériel complexe. Elles sont toutes basées sur la mise en évidence d'une interaction entre une protéine d'intérêt et une protéine cible. La protéine d'intérêt sélectionne la protéine cible qui va être analysée pour être identifiée. Chacune de ces méthodes a sa spécificité et sa sensibilité. Une haute spécificité signifie que la plupart des interactions détectées existent dans la réalité. Une forte sensibilité signifie que la plupart des interactions existante vont être détectées. L'utilisation de l'une de ces méthodes dépend donc du type de résultat recherché.

3. Méthodes biophysiques et théoriques

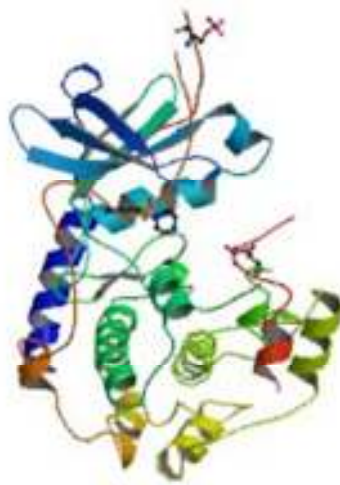
Depuis 10 ans, les progrès des technologies basées sur des propriétés physiques et notamment l'application de la spectrométrie de masse à la biologie a permis à de nouvelles techniques de voir le jour pour étudier les interactions protéine-protéine. Elles permettent d'obtenir des informations supplémentaires sur la structure des complexes mais aussi de réaliser des cribles des protéomes cellulaires. Toutefois, ces techniques restent encore difficilement applicables, le matériel nécessaire étant coûteux et nécessitant des formations particulières.

La bioinformatique apporte elle aussi sa contribution à l'étude des interactions protéine-protéine par des méthodes et des logiciels comme « molecular dynamics » [6] et « protein-protein docking » [7].

Les protéines kinases, plus particulièrement, représentent une des plus grandes familles de protéines, ce sont des enzymes comptant plus de 500 membres dans le génome humain. Quatre de ces protéines kinases (PK) sont impliquées dans la plupart des voies de transduction du signal régulant le fonctionnement cellulaire sous tous ses aspects. La plupart des cancers sont associés au dérèglement de protéines kinases, tels que des mutations génétiques qui conduisent à une sur-expression ou à une activation constitutive de ces protéines. Elles sont également impliquées dans autres pathologies.

Les kinases sont des enzymes de groupes des transférases, catalysant les réactions de phosphorylation par l'ajout d'un ion phosphate, provenant du nucléotide adénosine triphosphate (ATP), sur une molécule cible appelée substrat, qui peut être une protéine, un acide nucléique, un lipide, un sucre ou même une autre kinase.

Notre objectif de travail est d'étudier par les méthodes de modélisation moléculaire l'inhibition de la protéine kinase A par des fragments de protéine (tripeptides) présentant une activité anti inflammatoire.



Structure de la protéine kinase A

Les peptides, au même titre que les protéines, sont considérés comme des molécules thérapeutiques. Ils sont utilisés depuis de nombreuses années dans le domaine de la biotechnologie, par exemple pour la production d'anticorps à l'usage des laboratoires de biologie. Ces molécules sont utilisées pour d'autres applications et plus particulièrement dans le domaine de l'industrie pharmaceutique.

On peut classer les peptides thérapeutiques en trois catégories [8] :

- Les peptides naturels, aussi appelés peptides bioactifs qui sont soit produits par l'organisme, soit dérivés de protéines [9].
- Les peptides issus de bibliothèques génétiques [10].

- Les peptides issus de bibliothèques chimiques produites par synthèse en phase solide [11].

Les peptides concernés pour notre étude sont les dérivés du tripeptide feG (D-Phe-D-Glu-Gly). Ce tripeptide est un anti-inflammatoire qui réduit la gravité des réactions d'hypersensibilité, et inhibe la chimiotaxie des neutrophiles et l'adhésion aux tissus [12].

La modélisation moléculaire de systèmes biologiques, vise à étudier la structure et la dynamique des molécules biologiques. Les méthodes retenues pour mener à bien cette étude sont la mécanique moléculaire, la dynamique moléculaire et le docking moléculaire [13-15]. Les calculs ont été effectués par les logiciels Hyperchem8 et Hex6.3. L'interprétation des résultats est basée sur l'énergie d'interaction des complexes kinase-tripeptides formés. Le complexe de plus basse énergie d'interaction présentera probablement une meilleure inhibition de la protéine kinase A.

Références

- [1] Karimova, G., Ladant, D. and Ullmann, A. *International Journal of Medical Microbiology* (2002) 292, 17-25.
- [2] Park, S.H. and Raines, R.T. *Nature Biotechnology* (2000) 18, 847-851.
- [3] Vidal, M. and Endoh, H. *Trends in Biotechnology* (1999) 17, 374-381.
- [4] Phizicky, E.M. and Fields, S. *Microbiological Reviews* (1995) 59, 94-123.
- [5] Suter, B., Kittanakom, S. and Stagljar, I. *BioTechniques* (2008) 44, 681-691.
- [6] Amisaki, T., Toyoda, S., Miyagawa, H. and Kitamura, K. *Journal of Computational Chemistry* (2003) 24, 582-592.
- [7] Lawrence, H.R. et al. *Bioorganic & Medicinal Chemistry Letters* (2009) 19, 3756-3759.
- [8] Sato A.K., Viswanathan M., Kent R.B. & Wood C.R., Therapeutic peptides: technological advances driving peptides into development. *Curr. Opin. Biotechnol.*, (2006) 17(6), 638-642.
- [9] Watt P.M., Screening for peptide drugs from the natural repertoire of biodiverse protein folds. *Nature Biotech.*, (2006) 24, 177-183.
- [10] Sergeeva A. et al.,. Display technologies: Application for the discovery of drug and gene delivery agents. *Adv. Drug Delivery Rev.*, (2006) 15, 1622-1654.
- [11] Shin D.S., Kim D.H., Chung W.J. & Lee Y.S., Combinatorial solid phase peptide synthesis and bioassays. *J. Biochem. Mol. Biol.*, (2005) 38(5), 517-525.
- [12] R.D. Mathison et al. Identification of a binding site for the anti-inflammatory tripeptide feG, / *Peptides* (2003) 24 1221–1230.
- [13] I. Abdelli , S. Sari. & S. Ghalem, Molecular Modelling of Green Tea Catechins Inhibitors of Human Brain BACE1, *IBScientific Journal of Science*, (2008) 3, 27-29.

[14] S. SARI, W. MEDJAHED, J. KAJIMA MULENGI and S. GHALEM, Approach of the Interaction Enzyme-Substrate by Molecular Modelling, Asian Journal of Chemistry, (2007) 19, 2117-2124.

[15] I. ABDELLI, S. SARI, C. ZIANI-CHÉRIF and S. GHALEM, Theoretical Study of the Inhibition of β -Secretase. Asian Journal of Chemistry. (2007) 19, 5407-5416.

CHAPITRE I :

INTERACTIONS PROTEINE-PROTEINE

1. La structure des protéines

Trois méthodes expérimentales permettent aujourd'hui de déterminer la structure des protéines : la résonance magnétique nucléaire (RMN), la microscopie électronique et la cristallographie par rayons X. Cette dernière technique est responsable de la majorité des structures issues d'une base de données de structures accessibles gratuitement appelée la «Protein Data Bank» (PDB), c'est-à-dire plus de 40 000 structures protéiques (juin 2007). La RMN fournit, quant à elle, 6 200 structures contre seulement 150 dans le cas de la microscopie électronique.

a) La résolution

La résolution en angström de la protéine est une des données reflétant la qualité des structures ayant permis de construire le modèle cristallographique. Généralement, la résolution est limitée par la manière dont les cristaux diffractent, le temps nécessaire pour collecter des données de meilleures résolutions, la dynamique de la protéine, la qualité de l'appareillage et la température du système. Une résolution proche de 1 Å permet de distinguer tous les atomes y compris les hydrogènes. Une résolution de l'ordre de 6 Å permet seulement de distinguer que des structures de types «hélice α » ou «feuille β » par exemple (Figure 1).

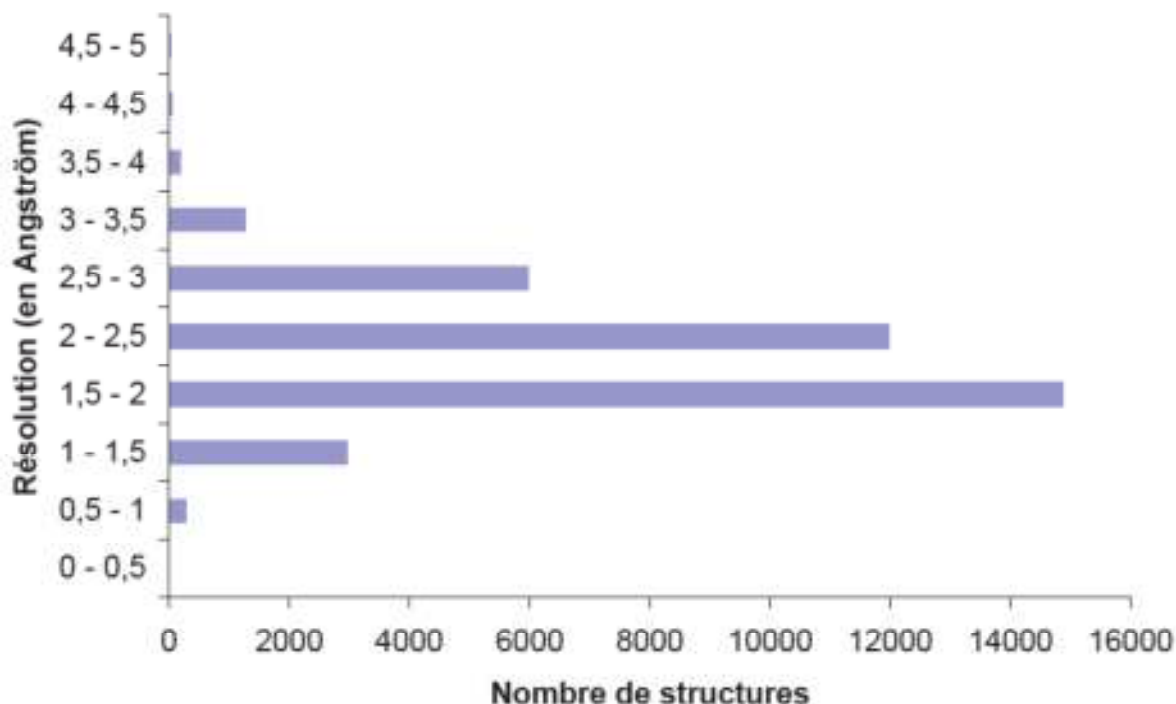


Figure 1. Nombre de structures protéiques par intervalle de résolution donné

L'histogramme précédent illustre parfaitement la répartition de la qualité des structures au sein de la PDB. La majorité des structures ont une résolution comprise entre 1,5 et 2,5 Å. Cet intervalle de résolution prouve que les structures sont de bonne qualité.

b) Le facteur R

Un autre indicateur est le facteur R qui est une grandeur indicatrice de l'écart entre les facteurs de structures observés et calculés. Le facteur R est compris entre 0 et 1 (plus le facteur R est proche de 0 et plus la prédiction est juste).

c) Choix d'une structure cristallographique

La combinaison des deux facteurs évoqués précédemment permet de réaliser un choix pertinent de structure cristallographique. Par exemple, une structure de protéine dont le facteur R est proche de 0 mais possédant une mauvaise résolution n'a aucune signification. Typiquement, une structure dont le facteur R avoisine la

valeur de 0,2 et possédant une résolution correcte (inférieure ou égale à 2 Å) correspond la plupart du temps à une structure précise et bien définie. Il existe à ce jour des outils capables d'évaluer la qualité des structures cristallographiques [1].

d) Facteur d'agitation thermique

Les atomes du cristal bougent autour d'une position atomique moyenne. Les rayonnements ne voient pas les atomes exactement à la même position dans les différentes mailles du cristal. Ce facteur permet d'observer la flexibilité des résidus.

e) Détermination du site actif

Une étape fondamentale dans la stratégie de criblage virtuel est l'identification des résidus d'acides aminés pouvant intervenir dans le processus de reconnaissance du ligand. Un ligand cristallisé avec le récepteur est susceptible de renseigner sur l'emplacement du site actif. Malgré tout, cette information doit être interprétée avec prudence car, pour un même récepteur, le site actif peut être différent selon le profil pharmacologique de la molécule (agoniste, antagoniste, agoniste inverse)

f) Les interactions entre le ligand et le récepteur

Les interactions entre une protéine et un ligand sont en général de nature non covalente. L'énergie libre de Gibbs se calcule selon l'Équation :

$$\Delta G = -RT \ln K_i = \Delta H - T\Delta S$$

R: constante des gaz (8,314 J. K⁻¹ mol⁻¹)

T: Température (K)

K_i: constante d'inhibition

La grandeur du ΔG en solution aqueuse est comprise entre -10 et -70 KJ/mol. Les composantes de l'énergie libre sont:

- la complémentarité géométrique, stérique et de surface entre le ligand et le récepteur.
- les contacts présents entre deux régions lipophiles des deux entités.
- les liaisons hydrogène

Les molécules d'eau dans les cavités des protéines peuvent parfois être un élément fondamental [2]. Elles sont capables d'assurer le relais entre le récepteur et le ligand et ainsi créer des réseaux de liaisons hydrogène.

2. La flexibilité d'une protéine

Le concept ancien de la clé et de la serrure, figurant respectivement la molécule organique et la protéine, est aujourd'hui abandonné. Une image plus adaptée d'un tel complexe est celui d'une main dans un gant (respectivement la molécule organique dans la protéine). Chacun peut modifier sa forme, favoriser la complémentarité afin de limiter les interactions gênantes et augmenter le confort procuré par le contact entre les deux entités. La difficulté dans la majorité des protéines est qu'il existe toujours des zones très flexibles à proximité de zones rigides. Les zones flexibles correspondent généralement à des boucles verrouillant le ligand. Quant aux zones rigides, elles correspondent majoritairement à des sites catalytiques.

3. Les complexes macromoléculaires

3.1. La stabilité des complexes

Les molécules des êtres vivants sont capables de s'auto-associer les unes aux autres grâce à plusieurs types de liaisons faibles, dont les forces électrostatiques, les forces de Van der Waals, les liaisons hydrogène et les liaisons hydrophobes. Etant faibles, ces liaisons nécessitent un rapprochement des molécules les unes des autres. Les associations sont dues aux propriétés chimiques, électriques et géométriques des molécules. A l'inverse des liaisons covalentes, les liaisons faibles permettent une

grande flexibilité et plasticité des macromolécules et des structures cellulaires. Les protéines s'associent ainsi entre elles par des centaines de liaisons.

Les interactions qui stabilisent les structures tridimensionnelles des protéines sont principalement les mêmes que celles qui stabilisent le repliement des protéines, mis à part les ponts dissulfures :

– Les interactions électrostatiques concernent les interactions dipôle - dipôle et les ponts salins. Les ponts salins, correspondent aux interactions entre régions chargées déterminées par les lois de l'électrostatique. A pH physiologique, les résidus chargés positivement sont les lysines, les arginines, certaines histidines et l'extrémité N-terminale de la chaîne polypeptidique. Les résidus chargés négativement sont les aspartates, glutamates et l'extrémité C-terminale. Des études par RMN montrent que la formation des ponts salins n'a pas toujours un effet stabilisateur. La désolvatation des chaînes latérales chargées est en effet très défavorable d'un point de vue énergétique [3].

- Les forces de Van der Waals correspondent aux interactions entre molécules indépendantes à couches électroniques complètes. Ces forces comprennent une composante attractive à longue portée (force de dispersion de London) et une composante répulsive à très courte portée. Les forces de dispersion correspondent à un effet inductif entre dipôles instantanés qui se produisent lors de fluctuations du nuage électronique. Pour des atomes distants de 3-4 Å, l'énergie de liaison est seulement de 1 kcal/mol. Cependant, de par leur grand nombre, ces interactions jouent un rôle important dans la stabilisation des complexes et favorisent leur compactage.

– Les liaisons hydrogènes : une liaison hydrogène résulte de l'interaction électrostatique entre un atome d'hydrogène (H), lié par covalence à un atome électronégatif (O, N, S) (donneur) et un deuxième atome électronégatif possédant une paire d'électrons non partagée (accepteur). La liaison hydrogène est par exemple : -O-H ... O = C. Les liaisons hydrogène sont plus fortes que les liaisons de Van der Waals ; leurs énergies sont estimées entre 3 et 9 kcal/mol. Les acides aminés polaires peuvent ainsi former des liaisons hydrogène entre eux ou avec des

molécules d'eau. Les acides aminés hydrophiles et non chargés sont la sérine, la thréonine, la cystéine, la tyrosine, l'asparagine et la glutamine. Leur chaîne latérale possède un groupement hydroxyle, phénol, amide ou thiol.

– Les effets hydrophobes et de solvatation : ne pouvant interagir avec l'eau, les molécules non polaires tendent à se rapprocher afin de limiter les contacts avec l'eau, et à créer entre elles des liaisons de type interactions hydrophobes. Pour quantifier cet effet, on le définit comme l'énergie associée au transfert d'une surface hydrophobe d'un milieu aqueux vers l'intérieur d'une protéine. Cela induit des variations d'énergie dues aux changements des interactions de van der Waals subis par les groupements hydrophobes, mais aussi aux variations d'énergie libre associées à la restructuration de l'eau. L'énergie de transfert du milieu aqueux vers l'intérieur de la protéine est de l'ordre de $-25 \text{ kcal.mol}^{-1}$ par Å^2 pour une surface apolaire.

Les acides aminés hydrophobes ont des chaînes latérales non chargées et non polaires ; il s'agit de l'alanine, la valine, la leucine, l'isoleucine, la proline, la phénylalanine, le tryptophane et la méthionine. De même que les interactions de Van der Waals, les effets hydrophobes sont responsables du compactage très dense à l'interface dans les assemblages macromoléculaires.

3.2. Les complexes protéine-protéine

Il est possible de distinguer deux grandes classes de complexes selon leur durée de vie :

– les complexes transitoires : les protéines se replient séparément et assurent indépendamment leurs fonctions avant de s'associer. Il s'agit de protéines impliquées dans l'adhésion cellulaire, dans les mécanismes d'oxydoréduction, dans les catalyses enzyme-substrat et dans la transduction du signal. Ces protéines s'associent et se dissocient très rapidement. Le temps de vie des complexes formés est compris entre 10^{-3} s et 1 s. Cette classe comprend aussi des complexes anticorps / antigènes, qui se montrent plus stables (10^3 s) et des complexes enzymes / inhibiteurs ($\sim 10^3$ s).

– les complexes permanents, comme les complexes oligomériques, les capsides de virus ou les fibres musculaires. Leur durée de vie est de l'ordre de 10^6

secondes. Les complexes de courte durée de vie sont impliqués dans des processus biologiques très importants. Peu d'entre eux sont présents dans la PDB. Pour ce type de complexe, les interactions protéine-protéine ressemblent davantage aux interactions non-spécifiques (durée de vie de 10^{-6} s) présentes dans les cristaux de protéines qu'à celles des assemblages permanents [4].

Les assemblages macromoléculaires comprennent les complexes binaires, constitués de deux composants, et les complexes multimériques. Les complexes binaires sont le plus souvent des complexes transitoires et se rencontrent souvent dans les mécanismes de signalisation et de régulation cellulaire. Les complexes multimériques englobent les machineries enzymatiques, qui s'associent et se dissocient très rapidement pour exercer leur fonction enzymatique. Les machineries enzymatiques englobent le ribosome, responsable de la biosynthèse des protéines, l'ARN polymérase qui catalyse la formation de l'ARN ou l'ATP synthase qui catalyse la formation de l'ATP. Les complexes multimériques comprennent aussi les assemblages stables à topologie restreinte. Les monomères s'assemblent selon un nombre défini de sous-unités (hexamère, octamère...) pour former une géométrie particulière. Parmi eux se trouvent des complexes présentant une activité, comme le protéasome, mais également les protéines structurales (que l'on retrouve dans les membranes cellulaires), comme les hélicases ou les pores. Les particules de virus sont également des assemblages stables à topologie restreinte. Enfin, certaines protéines s'associent selon un nombre illimité de monomères, pour former des filaments qui supportent des processus biologiques fondamentaux, tels que la recombinaison homologue (filament de Rad51 chez l'homme par exemple), la formation du cytosquelette (la tubuline) ou la motricité musculaire (filament d'actine). Ces filaments créent une haute force de tension, ce qui permet par exemple d'étirer un polymère tel que l'ADN d'un facteur 1.5 lors de la recombinaison homologue par la protéine RecA [5], protéine orthologue de Rad51 chez les procaryotes. La dynamique et la régulation de la longueur de ces filaments dépendent de l'hydrolyse de l'ATP.

3.3. Caractéristiques des interfaces protéine-protéine

Les différentes caractéristiques des interfaces protéine-protéine peuvent grossièrement être divisées en deux catégories :

- (i) les propriétés structurales, taille, forme et complémentarité géométrique ;
- (ii) les propriétés chimiques, potentiel de solvation, hydrophobicité, potentiel électrostatique, liaisons hydrogène et ponts salins.

La taille des interfaces des complexes protéine-protéine est généralement mesurée par la différence entre l'aire de la surface accessible (ASA) du complexe et celle des composés séparés:

$$B = ASA_{\text{prot.1}} + ASA_{\text{prot.2}} - ASA_{\text{complexe}}$$

La mesure B donne une indication de la force de liaison ; l'aire de la surface enfouie est en relation avec l'énergie hydrophobe de désolvation [6,7].

Lo Conte et collaborateurs ont analysé la structure atomique des sites de reconnaissance dans 75 complexes protéine-protéine : 24 complexes protéase / inhibiteur, 19 anticorps / antigène et 32 autres complexes, incluant 9 enzyme / inhibiteur et 11 complexes impliqués dans la transduction du signal [8]. Sur les 75 complexes, 52 ont une interface de 'taille standard' égale à 1600 (± 400) Å². Il s'agit principalement de complexes anticorps / antigène et protéase / inhibiteur. La plus petite des interfaces est de 1150 Å² (le cytochrome peroxydase et son substrat, le cytochrome C). Les petites interfaces sont sans doute caractéristiques des complexes de basse stabilité et de courte durée de vie. Par exemple, les complexes enzyme / substrat, dont les réactions catalytiques sont très rapides, ne nécessitent pas une forte affinité de liaison. Vingt complexes ont des interfaces qui s'étendent de 2000 à 4660 Å². Il s'agit principalement de protéases complexées avec une classe particulière d'inhibiteurs, de complexes comprenant les protéines G, et d'autres composants impliqués dans le système de la transduction du signal. Par leur taille, les interfaces de ces complexes ressemblent à celles des sous-unités protéiques des assemblages oligomériques [9].

Toutefois, elles sont moins hydrophobes et contiennent une proportion plus importante de groupes chargés. Dans ces complexes, l'association implique de

grands changements conformationnels, principalement un remodelage de la chaîne principale à l'interface.

3.4. Propriétés physico-chimiques des interfaces des complexes protéiques

Les propriétés physico-chimiques des interfaces de complexes protéiques ont été étudiées afin de déterminer les contributions majeures à la stabilité des interfaces. Glaser et collaborateurs ont démontré, en analysant une base de données de 621 interfaces de complexes protéiques, regroupant entre autres des oligomères ou des complexes enzyme inhibiteur, que la plupart des contacts entre paires de protéines aux interfaces de complexes impliquaient de larges résidus hydrophobes [10]. En outre, les auteurs ont remarqué une différence significative de la composition des interfaces entre de petites et de grosses interfaces. En effet, il apparaît que les résidus hydrophobes sont souvent retrouvés au sein de grosses interfaces, alors que les résidus polaires prévalent dans les petites interfaces. Compte tenu de ces résultats, il apparaît que l'effet hydrophobe, jouant un rôle primordial dans le repliement des protéines, a également un rôle majeur dans la stabilisation des complexes protéiques [10].

Les interactions électrostatiques et hydrophobes seraient les propriétés physico-chimiques qui contribuent le plus à la stabilité des complexes protéiques [11,12]. Un des aspects non appréhendé par l'étude de Glaser et collaborateurs concerne la distinction entre les complexes permanents et les complexes transitoires. Or en fonction du type de complexes, les forces qui régissent leur interaction sont différentes. En effet, il a été montré que les interfaces de complexes permanents étaient généralement plus hydrophobes que les interfaces de complexes transitoires [13,14].

4. Importance des interactions protéiques

Plusieurs facteurs thermodynamiques favorisent les interactions entre protéines.

- La cellule constitue un environnement restreint caractérisé par la présence d'une forte concentration de macromolécules, ce qui influence la cinétique et les réactions d'associations impliquées à la fois dans les repliements de protéines et l'assemblage des complexes [15,16]. Il est nécessaire que la cellule puisse réguler les associations entre protéines, *via* les complexes protéiques, pour une exécution efficace des programmes cellulaires dans lesquels elles sont impliquées.
- De plus, l'association de plus de deux protéines entraîne un niveau de complexité supérieur à celui d'une interaction binaire avec la manipulation du nombre de surfaces d'interactions augmentant la constance d'affinité [17]. Les liaisons entre les protéines dans un complexe sont donc plus fortes qu'entre deux protéines.
- Enfin, les changements allostériques et conformationnels associés aux interactions protéiques peuvent influencer les associations avec d'autres facteurs [18].

5. Interactions protéines-protéines

Si l'ADN représente le plan d'architecte de la vie, alors les protéines sont définies comme les machines moléculaires accomplissant ce plan. Les protéines réalisent souvent leurs fonctions en agissant avec d'autres protéines pour former des complexes protéiques. Ces complexes peuvent exister en tant qu'associations transitoires de courte durée (comme, par exemple, dans le cas de la catalyse enzymatique), de systèmes multimériques (tels que le ribosome), les facteurs de transcription, les protéines à la surface des cellules, ou encore les protéines formant les canaux ioniques. Cependant, même si nous connaissons le plan de montage – c.-à-d. la séquence ADN – nous ne savons que très peu de choses sur le fonctionnement des protéines au niveau moléculaire.

En effet, les études portant sur l'ensemble du génome [19-22], fournissent une liste croissante d'interactions protéine-protéine (IPP) potentielles; mais comprendre le fonctionnement de ces interactions exige davantage d'analyses

biochimiques et structurales. Par exemple, la levure est l'un des organismes les plus étudiés et est connue pour avoir à peu près 6.000 protéines, impliquant environ entre 38.000 et 75.000 IPP. Environ 50% de ces IPP ont été observées expérimentalement. D'autre part, le génome humain code approximativement 30.000 protéines, donnant de 154.000 à 370.000 IPP, dont seulement ~10% sont connus [23,24]. Comprendre comment les protéines interagissent est essentiel à la compréhension des mécanismes moléculaires provoquant des maladies. Par exemple, les médicaments fonctionnent souvent en modulant ou en bloquant les IPP, et donc ces IPP représentent une classe importante de cibles médicamenteuses [25,26].

En 1975, Chotia et Janin [27], furent les premiers à proposer un modèle d'interaction entre protéines : « L'hydrophobicité est le principal facteur stabilisant les associations protéines-protéines, alors que la complémentarité joue un rôle sélectif en décidant quelles protéines peuvent s'associer ». Et, bien que de nombreux progrès ont été réalisés dans le domaine, ce modèle de base est encore approuvé par de nombreux auteurs. Par la suite, les travaux effectués permirent une amélioration des connaissances mais souffrirent longtemps du faible nombre de structures 3D connues (Figure 2) et du peu de types d'interactions représentées : principalement des complexes enzyme (protéase)-inhibiteur et antigène-anticorps. Pour exemple, l'article de Chotia et Janin [27], se base sur seulement 3 dimères alors que les travaux actuels utilisent fréquemment plus de 1000 structures [28] et même jusque 1981 complexes pour Aytuna et al [29]. Actuellement, des efforts supplémentaires sont effectués en vue de collecter des structures de complexes protéiques par le biais de projets de 'structural genomics' [30]. Ces projets permettent de récolter à haut débit de telles structures mais avec peu ou pas d'informations sur la fonction de celles-ci.

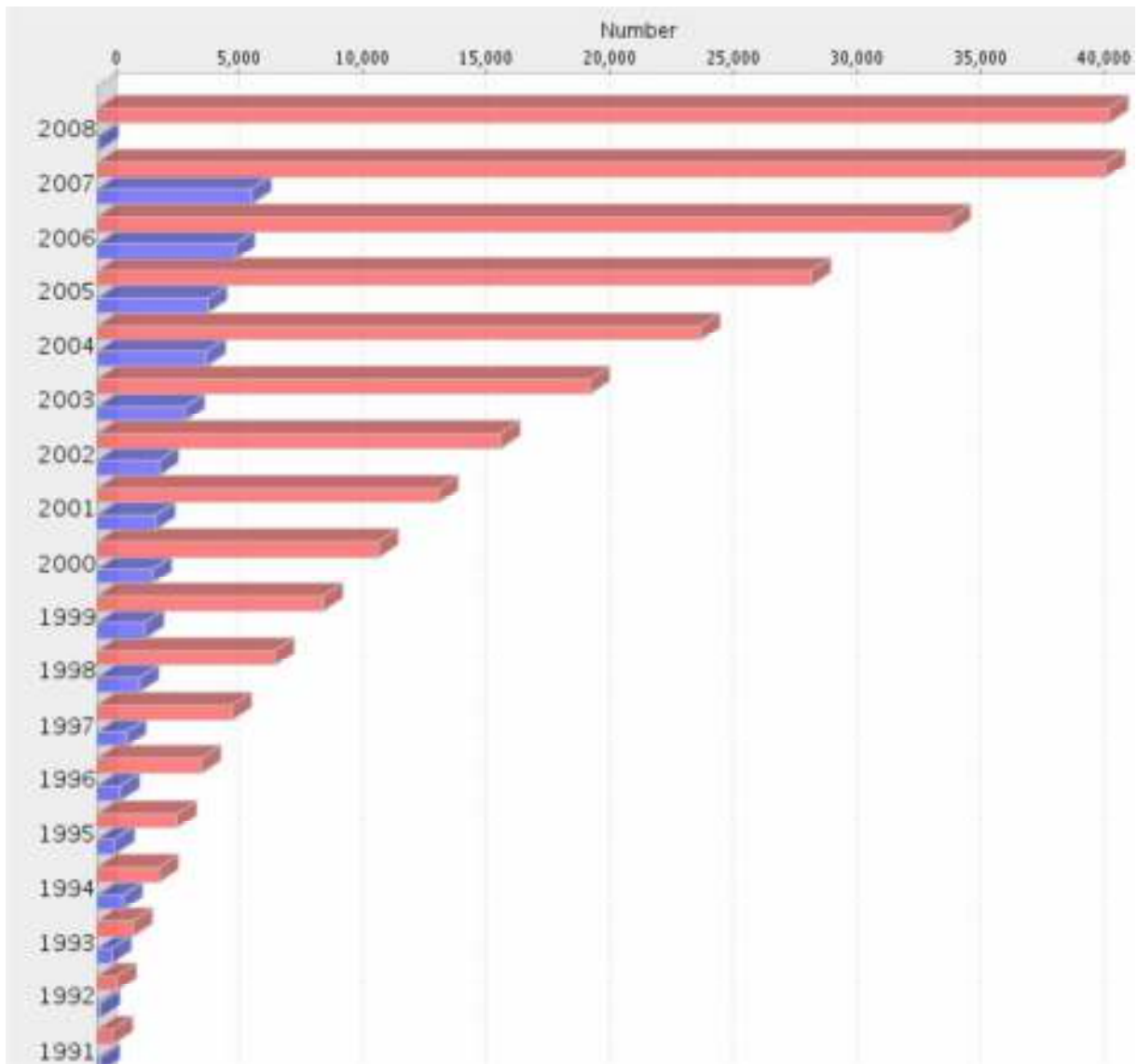


Figure 2 : Evolution du nombre de structures 3D Rx contenues dans la Protein Data Bank. En bleu, les nouvelles structures déposées dans l'année considérée et en rouge le nombre total de structures. Données de janvier 2008.

En 1995, Jones et Thornton [31], marquent une nouvelle étape dans les études sur les interactions protéine-protéine en réalisant un travail très complet sur une banque de données de 32 dimères. La composition en acides aminés et en atomes, la forme, les structures secondaires, les liaisons hydrogènes, les ponts salins, la segmentation, la complémentarité des interfaces furent analysés. Il ressort de cette étude que, bien que l'interface soit plus hydrophobe que le reste de la surface, certains acides aminés chargés et polaires (arginine, méthionine) montrent une certaine affinité pour l'interface. D'autres travaux ont permis de confirmer que

les acides aminés hydrophiles interviennent de manière spécifique dans les interactions protéiques. Larsen et al [32], ont réalisé une étude plus 'visuelle' d'une banque de 136 homodimères et ont montré que seulement un tiers des complexes possèdent un coeur hydrophobe alors que les autres deux tiers possèdent des interfaces où l'on retrouve des zones hydrophobes et des zones hydrophiles (avec des liaisons hydrogènes et présence de molécules d'eau).

Récemment, les projets de séquençage menés sur différents génomes [33-35], ont permis d'obtenir une meilleure caractérisation des différents gènes constitutifs d'un organisme. L'objectif majeur est de mieux comprendre les fonctions de ces gènes, leurs régulations et leurs interactions, en s'intéressant plus particulièrement au produit de leur transcription, les ARNs, ou de leur traduction, les protéines.

L'ensemble des processus cellulaires est contrôlé par l'association et la dissociation des protéines. Ainsi, les interactions protéines-protéines sont au coeur de différentes fonctions essentielles, comme la réplication de l'ADN, l'organisation du cytosquelette, le contrôle du cycle cellulaire, ou encore la réponse à certains signaux extra-cellulaires. Leur caractérisation à l'échelle d'un organisme constitue une étape majeure pour une meilleure compréhension du rôle des protéines, et plus généralement du fonctionnement des cellules.

5.1. Mise en évidence des interactions protéine-protéine

Différentes méthodes peuvent être utilisées pour étudier ponctuellement certaines interactions protéines-protéines parmi lesquelles les techniques de chromatographie d'affinité, de co-immunoprécipitation ou encore de *phage-display* [36]. De plus, des techniques d'analyses ont été proposées pour étendre l'identification des interactions protéiques à grande échelle, parmi lesquelles le système double-hybride de la levure [37], ou encore l'analyse par spectrométrie de masse de complexes purifiés [38].

5.2. Le rôle du solvant

Le rôle de l'eau qui reste le solvant biologique par excellence, fut lui aussi de plus en plus pris en compte. Janin [39], montre que, malgré un échange très rapide des molécules d'eau entre l'interface et le milieu aqueux (moins d'une milliseconde pour l'eau au centre de l'interface), les interactions avec les molécules d'eau sont aussi importantes pour la stabilité des protéines. En 2005, le groupe de Janin [40], propose de diviser les interfaces en deux types : les interfaces 'mouillées' et les interfaces 'sèches' où les molécules d'eau sont situées tout autour de l'interface. Ce sont ces dernières qui correspondraient plutôt aux interactions spécifiques. De même, les fragments de protéines qui ne sont pas suffisamment hydratés à l'état monomérique seraient des sites préférentiels pour les interactions intermoléculaires [41] et la conservation des molécules d'eau dans les interfaces a été étudiée [42].

Toutefois, il reste un problème majeur à l'étude de ce type d'interaction eau-protéine : pour pouvoir situer avec précision la position d'une molécule d'eau, il est nécessaire de posséder des structures cristallines aux rayons-X hautement résolues (de l'ordre de 2Å ou moins) ce qui ne représente pas la majorité des structures connues.

5.3. Les changements de conformation

En 2000, Sundberg et Mariuzza [43], ont publié une revue sur les changements de conformation subis par une protéine lors de son association à une autre protéine. Ces changements sont étudiés par de nombreux auteurs [44-45] et connus depuis longtemps (ex. : allostérie enzymatique) mais les avis divergent quant à leur importance. Sundberg et Mariuzza⁸⁴ ont montré que certains atomes bougent d'une distance allant jusque 3Å. Heifetz et Eisenstein [46], ont pour leur part analysé les différences de flexibilité des chaînes latérales des vingt acides aminés naturels et ont montré que les chaînes latérales de la lysine et de l'arginine sont les plus flexibles. Ces résultats pourraient permettre d'expliquer certains comportements des protéines comme leur faculté d'interagir avec plusieurs partenaires et la

reconnaissance du système immunitaire. Les mouvements effectués par les protéines lors de l'association pourraient aussi permettre d'amener un plus grand nombre de résidus hydrophobes à l'interface en vue de la stabiliser [47]. Gunasekaran et Nussinov [48] ont essayé de différencier les interfaces rigides de celles subissant des changements de conformation. Il apparaît que le tryptophane, les interactions entre résidus hydrophobes, entre résidus aromatiques, et entre résidus polaires et aromatiques sont associées à des grands changements de conformation.

Néanmoins, pour étudier les changements de conformation lors de la liaison, il faut posséder les structures 3D à la fois des monomères et du complexe (the Database of Macromolecular Movements) [49], ce qui n'est pas toujours le cas. De plus, il faut remarquer que les protéines en solution présentent naturellement de faibles déformations de leur structure [50]. Ces deux derniers points montrent que ce problème est encore loin d'être résolu mais le développement des techniques par Résonance Magnétique Nucléaire (RMN), qui permettent l'observation de ces changements de conformation en solution, devrait permettre une compréhension plus poussée du phénomène.

5.4. Identification des interactions protéines-protéines à grande échelle.

Depuis quelques années, différents projets ont été menés pour caractériser les interactions protéines-protéines à grande échelle, et pour différents organismes. Les premières études, menées chez *S. cerevisiae* par des techniques de double hybride, ont ainsi permis l'identification de 957 interactions potentielles impliquant 1004 protéines [19], puis 4549 interactions impliquant 3278 protéines [20]. Ces études ont été suivies par d'autres expériences sur ce même organisme, réalisées cette fois-ci par des techniques de spectrométrie de masse [21,22,51]. D'autres cribles à grande échelle, basés sur la technique du double hybride, ont permis d'identifier un grand nombre d'interactions protéines-protéines chez d'autres organismes tels que *H. Pylori* [52], la mouche *D. melanogaster* [53], ou le vers *C. elegans* [54]. Plus récemment,

les identifications d'interactions protéines-protéine ont pu être décrites chez l'homme, à partir de cribles à grande échelles utilisant la technique du double hybride [55-57].

À partir des données issues d'analyses à grande échelle, les premières cartes d'interactions protéines-protéines ou interactomes ont pu être proposées. Généralement, ces cartes d'interactions sont représentées sous la forme d'un graphe dont chaque noeud représente une protéine, et chaque arête reliant deux noeuds une interaction entre deux protéines. La **Figure 3**, présente une partie de l'interactome humain impliquant 401 protéines et 911 interactions [57]. En examinant cet interactome, nous pouvons constater que la majorité des protéines n'interagissent qu'avec un nombre limité de partenaires.

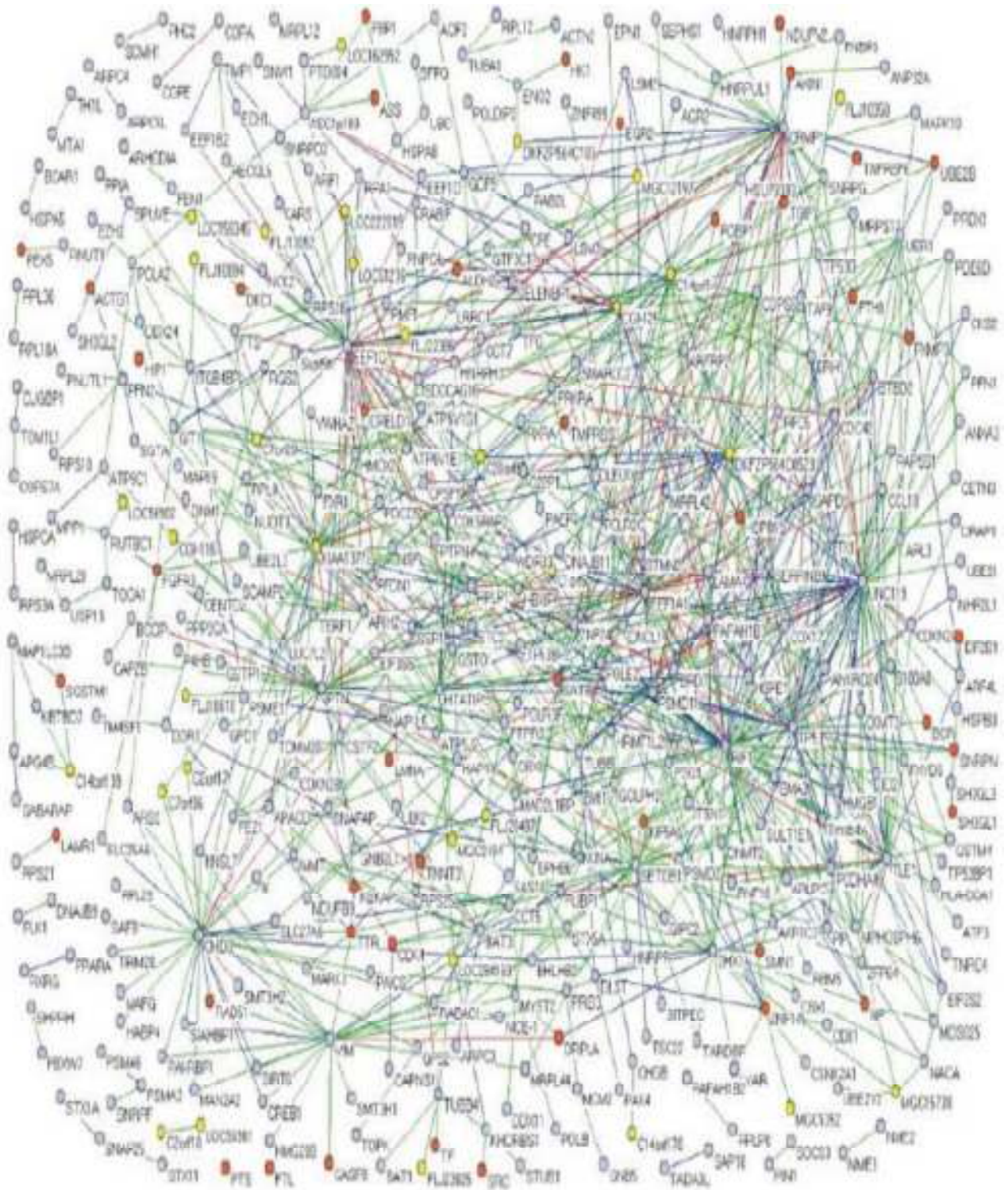


Figure 3 : Carte d'interaction protéine-protéine impliquant 911 interactions caractérisées chez l'homme. Les cercles oranges représentent les protéines impliquées dans des maladies génétiques (d'après la base de données OMIM). Les cercles bleus représentent les protéines annotées par la banque « *Gene Ontology* »(GO) (Ashburner et al., 2000). Les protéines n'ayant pas été annotées selon les

bases de données OMIM ou GO sont représentées par un cercle jaune. Les arêtes reliant les différents noeuds modélisent les interactions entre protéines. Ces arêtes sont colorées en fonction de leur score de confiance (score de confiance croissant du vert au violet en passant par le bleu et le rouge).

5.5. Les bases de données d'interactions protéine-protéine

Pour collecter, classer et rendre accessible ces différentes données expérimentales issues de l'analyse des interactions protéines-protéines, différentes bases de données ont été développées. Le **Tableau 1** en mentionne 7 parmi les plus utilisées.

Par ailleurs, elles sont couplées à différents outils d'analyse de séquences, de domaines ou de prédictions d'interaction extrêmement utiles pour accéder à un état des connaissances complet pour une protéine d'intérêt. Ces différences de traitements font que ces bases de données peuvent contenir des nombres très différents d'interactions. À titre d'exemple, la base de données BIND contient 188517 interactions, et la base de données DIP n'en possède que 56186.

Sigle	Base de donnée-description	taille	URL
BioGRID	Biological General Repository for Interaction Datasets.	198791	http://www.thebiogrid.org/
BIND	Biomolecular Interaction Network Database	188517	http://www.bind.ca/
DIP	Database of Interacting Proteins	56186	http://www.dip.doe-mbi.ucla.edu
IntAct	Interaction Act	155533	http://www.ebi.ac.uk/intact/site/index.jsf
MINT	Molecular INTeraction database	103808	http://mint.bio.uniroma2.it/mint/
HPRD	The Humain Protein Reference	38167	http://www.hprd.org

	Database		
MIPS	Munich Information Center of proteins sequences	15488	http://mips.gsf.de/denne/proj/impact/

Tableau 1. Bases de données d'interactions protéine-protéine

5.6. Pourquoi porter autant d'intérêt aux interactions avec les protéines ?

L'ambitieux objectif de l'étude des protéines est d'élucider la structure, les interactions et les fonctions de toutes les protéines d'une cellule et/ou d'un organisme entier [58]. Les objectifs sont multiples : meilleure connaissance des réseaux et processus cellulaires, meilleure compréhension des maladies, production de nouveaux médicaments...

Au cours des dix dernières années, plus de 550 génomes bactériens [59] et plus de 70 génomes eucaryotiques [60-64], ont été entièrement séquencés. Il est clair que, même si les gènes compris dans ces génomes ne sont pas clairement définis, nous avons maintenant accès à des séquences codant pour des dizaines de milliers de protéines. Dès lors, une question fondamentale se pose : quelles sont les fonctions biologiques de ces protéines ?

Pour exprimer leur fonction biologique, une grande partie des protéines a besoin d'interagir avec un partenaire car peu de protéines ont une fonction sous forme de monomère. Les sites d'interaction jouent donc des rôles essentiels en biologie. Ceux-ci concernent notamment la régulation des réseaux métaboliques, la reconnaissance immune, la réplication de l'ADN, la progression à travers le cycle cellulaire, la transduction de signaux, la synthèse protéique et la régulation des différentes enzymes et hormones.

La compréhension des interactions entre protéines et la prédiction des sites impliqués ouvrent donc de nombreuses voies dont deux principales qui dépendent l'une de l'autre : la compréhension de la biologie cellulaire et la production de

nouveaux médicaments. D'un point de vue fondamental, la connaissance des différents réseaux moléculaires d'interaction présente un intérêt scientifique énorme. Si on connaissait la fonction et les différents partenaires d'interaction d'une protéine, les nombreuses cascades cellulaires qui permettent aux organismes de vivre et de se développer pourraient être expliquées et certaines interrogations pourraient être levées.

Un corollaire à la connaissance des mécanismes d'interaction biologique est la compréhension des mécanismes suivis par certaines maladies. En effet, une partie des maladies existantes sont dues à une interaction anormale entre protéines, à la perte d'une interaction donnée ou à l'agrégation de plusieurs protéines [65]. De plus, la détection des molécules impliquées dans une nouvelle maladie serait plus rapidement réalisée. D'un point de vue pratique, la production de médicaments serait facilitée par une identification plus aisée des cibles pharmaceutiques potentielles. Une fois les cibles pharmaceutiques détectées avec précision grâce aux nouvelles connaissances, la production de peptides imitant le partenaire de l'interaction permettrait par exemple de bloquer les effets néfastes du système impliqué. Si, actuellement, il est encore impossible de prédire avec certitude la position dans la séquence des sites d'interaction, il est possible de s'aider des données et techniques déjà mises au point. Golemis et al [66], proposent une démarche en cinq étapes permettant de puiser le maximum des informations actuelles pour générer des médicaments utiles :

1. Identifier les cibles potentielles en se basant sur l'évidence des fonctions biologiques dans des conditions cliniques semblables à celles qui seront adressées et en se basant sur l'évidence que les fonctions de la cible peuvent être manipulées de manière productive.
2. À partir des données existantes, trouver un réseau physique et fonctionnel d'interactions pour la cible.
3. Déterminer la bibliographie des mutations des interactions protéine-protéine impliquant la cible.

4. Développer une stratégie discriminante pour identifier les inhibitions spécifiques/non-spécifiques de la protéine cible, des protéines en interaction et du réseau de contrôle.
5. Etre prêt à exploiter des découvertes très intéressantes...

Références

- [1] Word, J.M.; Lovell, S.C.; Richardson, J.S.; Richardson, D.C. Asparagine and glutamine: using hydrogen atom contacts in the choice of sidechain amide orientation *J. Mol. Biol.* (1999) 285, 1733-1745.
- [2] Maréchal, Y. The hydrogen bond and the water molecule *Elsevier Science & Technology*, (2007).
- [3] Marti, D.N. et Bosshard, H.R. Inverse electrostatic effect : electrostatic repulsion in the unfolded state stabilizes a leucine zipper. *Biochemistry*, (2004) 43, 12436–12447.
- [4] Nooren, I.M.A. et Thornton, J.M. Structural characterisation and functional significance of transient protein-protein interactions. *J Mol Biol*, (2003) 325 (5), 991–1018.
- [5] Stasiak, A., Di Capua, E. et Koller, T. Elongation of duplex DNA by recA protein. *J Mol Biol*, (1981) 151, 557–564.
- [6] Chothia, C. Hydrophobic bonding and accessible surface area in proteins. *Nature*, (1974) 248, 338–339.
- [7] Chothia, C. et Janin, J. Principles of protein-protein recognition. *Nature*, (1975) 256, 705–708.
- [8] Lo Conte, L., Chothia, C. et Janin, J. The atomic structure of protein-protein recognition sites. *J Mol Biol*, (1999) 285, 2177–2198.
- [9] Jones, S. et Thornton, J.M. Principles of protein-protein interactions. *Proc Natl Acad Sci U S A*, (1996), 93, 13–20.
- [10] Glaser, F., Steinberg, D. M., Vakser, I. A., and Ben-Tal, N. Residue frequencies and pairing preferences at protein-protein interfaces. *Proteins* (2001) 43, 89–102.
- [11] McCoy, A. J., Chandana Epa, V., and Colman, P. M. Electrostatic complementarity at protein/protein interfaces. *J Mol Biol* (1997) 268, 570-584.
- [12] Sheinerman, F. B., Norel, R., and Honig, B. Electrostatic aspects of protein-protein interactions. *Curr Opin Struct Biol* (2000) 10, 153-159.
- [13] Carugo, O., and Argos, P. Protein-protein crystal-packing contacts. *Protein Sci* (1997) 6, 2261-2263.

- [14] Larsen, T. A., Olson, A. J., and Goodsell, D. S. Morphology of protein-protein interfaces. *Structure* (1998) *6*, 421-427.
- [15] Ellis, R. J. *Curr Opin Struct Biol*, (2001) *11*,114-9.
- [16] Minton, A. P. *Curr Opin Struct Biol*, (2000) *10*,34-9.
- [17] Demchenko,A.P. Recognition between flexible protein molecules : induced and assisted folding. *Journal of Molecular Recognition*, (2001) *14* , 42–61.
- [18] Goh,C.S., Milburn,D. et Gerstein,M. Conformational changes associated with protein-protein interactions. *Curr. Opin. Struct. Biol.*, (2004) *14*, 104–109.
- [19] Uetz, P., Giot, L., Cagney, G., Mansfield, T. A., Judson, R. S., Knight, J. R., Lockshon, D., Narayan, V., Srinivasan, M., Pochart, P., *et al.* A comprehensive analysis of proteinprotein interactions in *Saccharomyces cerevisiae*. *Nature* (2000) *403*, 623-627.
- [20] Ito, T., Chiba, T., Ozawa, R., Yoshida, M., Hattori, M., and Sakaki, Y. A comprehensive two-hybrid analysis to explore the yeast protein interactome. *Proc Natl Acad Sci U S A* (2001) *98*, 4569-4574.
- [21] Gavin, A. C., Bosche, M., Krause, R., Grandi, P., Marzioch, M., Bauer, A., Schultz, J., Rick, J. M., Michon, A. M., Cruciat, C. M., *et al.* Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* (2002) *415*, 141-147.
- [22] Ho, Y., Gruhler, A., Heilbut, A., Bader, G. D., Moore, L., Adams, S. L., Millar, A., Taylor, P., Bennett, K., Boutilier, K., *et al.* Systematic identification of protein complexes in *Saccharomyces cerevisiae* by mass spectrometry. *Nature* (2002) *415*, 180-183.
- [23] Aloy, P., Ceulemans, H., Stark, A., and Russell, R. B. The relationship between sequence and interaction divergence in proteins. *J Mol Biol* (2003) *332*, 989-998.
- [45] Hart, G. T., Ramani, A. K., & Marcotte, E. M. How complete are current yeast and human protein interaction networks? *Genome Biol.* (2006) *7*, 120.
- [25] Arkin, M. R. & Wells, J. A. Small-molecule inhibitors of protein-protein interactions: progressing towards the dream. *Nat. Rev. Drug Discov.* (2004) *3*, 301–317.

- [26] González-Ruiz, D. & Gohlke, H. Targeting protein-protein interactions with small molecules: challenges and perspectives for computational binding epitope detection and ligand finding. *Curr. Med. Chem.* (2006) 13, 2607–2625.
- [27] Chothia, C. & Janin, J. Principles of protein-protein recognition. *Nature* (1975) 256, 705-708.
- [28] Li, H., Li, J., Tan, S.H. & Ng, S.K. Discovery of binding motif pairs from protein complex structural data and protein interaction sequence data. *Pac. Symp. Biocomput.* (2004) 312-323.
- [29] Aytuna, A.S., Gursoy, A. & Keskin, O. Prediction of protein-protein interactions by combining structure and sequence conservation in protein interfaces. *Bioinformatics.* (2005) 21, 2850-2855.
- [30] Grabowski, M., Joachimiak, A., Otwinowski, Z. & Minor, W. Structural genomics: keeping up with expanding knowledge of the protein universe. *Curr. Opin. Struct. Biol.* (2007) 17, 347-353.
- [31] Jones, S. & Thornton, J.M. Protein-protein interactions: a review of protein dimer structures. *Prog. Biophys. Mol. Biol.* (1995) 63, 31-65.
- [32] Larsen, T.A., Olson, A.J. & Goodsell, D.S. Morphology of protein-protein interfaces. *Structure.* (1998) 6, 421-427.
- [33] Adams, M. D., Celniker, S. E., Holt, R. A., Evans, C. A., Gocayne, J. D., Amanatides, P. G., Scherer, S. E., Li, P. W., Hoskins, R. A., Galle, R. F., *et al.* The genome sequence of *Drosophila melanogaster*. *Science* (2000) 287, 2185-2195.
- [34] Blattner, F. R., Plunkett, G., 3rd, Bloch, C. A., Perna, N. T., Burland, V., Riley, M., Collado- Vides, J., Glasner, J. D., Rode, C. K., Mayhew, G. F., *et al.* The complete genome sequence of *Escherichia coli* K-12. *Science* (1997) 277, 1453-1474.
- [35] Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., *et al.* Initial sequencing and analysis of the human genome. *Nature* (2001) 409, 860-921.
- [36] Phizicky, E. M., and Fields, S. Protein-protein interactions: methods for detection and analysis. *Microbiol Rev* (1995) 59, 94-123.
- [37] Fields, S., and Song, O. A novel genetic system to detect protein-protein interactions. *Nature* (1989) 340, 245-246.

- [38] Mann, M., Hendrickson, R. C., and Pandey, A. Analysis of proteins and proteomes by mass spectrometry. *Annu Rev Biochem* (2001) 70, 437-473.
- [39] Janin, J. Wet and dry interfaces: the role of solvent in protein-protein and protein-DNA recognition. *Structure. Fold. Des* (1999) 7, 277-279.
- [40] Rodier, F., Bahadur, R.P., Chakrabarti, P. & Janin, J. Hydration of protein-protein interfaces. *Proteins* (2005) 60, 36-45.
- [41] Fernandez, A. & Scheraga, H.A. Insufficiently dehydrated hydrogen bonds as determinants of protein interactions. *Proc. Natl. Acad. Sci. U. S. A* (2003) 100, 113-118.
- [42] Reichmann, D., Phillip, Y., Carmi, A. & Schreiber, G. On the contribution of water-mediated interactions to protein-complex stability. *Biochemistry* (2008) 47, 1051-1060.
- [43] Sundberg, E.J. & Mariuzza, R.A. Luxury accommodations: the expanding role of structural plasticity in protein-protein interactions. *Structure. Fold. Des* (2000) 8, 137-142.
- [44] Janin, J. & Chothia, C. The structure of protein-protein recognition sites. *J. Biol. Chem.* (1990) 265, 16027-16030.
- [45] Ming, D. & Wall, M.E. Interactions in native binding sites cause a large change in protein dynamics. *J. Mol. Biol.* (2006) 358, 213-223.
- [46] Heifetz, A. & Eisenstein, M. Effect of local shape modifications of molecular surfaces on rigid-body protein-protein docking. *Protein Eng* (2003) 16, 179-185.
- [47] Xu, D., Lin, S.L. & Nussinov, R. Protein binding versus protein folding: the role of hydrophilic bridges in protein associations. *J. Mol. Biol.* (1997) 265, 68-84.
- [48] Gunasekaran, K. & Nussinov, R. How different are structurally flexible and rigid binding sites? Sequence and structural features discriminating proteins that do and do not undergo conformational change upon ligand binding. *J. Mol. Biol.* (2007) 365, 257-273.
- [49] Goh, C.S., Milburn, D. & Gerstein, M. Conformational changes associated with protein-protein interactions. *Curr. Opin. Struct. Biol.* (2004) 14, 104-109.
- [50] Wodak, S.J. & Janin, J. Structural basis of macromolecular recognition. *Adv. Protein Chem.* (2002) 61, 9-73.

- [51] Krogan, N. J., Cagney, G., Yu, H., Zhong, G., Guo, X., Ignatchenko, A., Li, J., Pu, S., Datta, N., Tikuisis, A. P., *et al.* Global landscape of protein complexes in the yeast *Saccharomyces cerevisiae*. *Nature* (2006) 440, 637-643.
- [52] Rain, J. C., Selig, L., De Reuse, H., Battaglia, V., Reverdy, C., Simon, S., Lenzen, G., Petel, F., Wojcik, J., Schachter, V., *et al.* The protein-protein interaction map of *Helicobacter pylori*. *Nature* (2001) 409, 211-215.
- [53] Giot, L., Bader, J. S., Brouwer, C., Chaudhuri, A., Kuang, B., Li, Y., Hao, Y. L., Ooi, C. E., Godwin, B., Vitols, E., *et al.* A protein interaction map of *Drosophila melanogaster*. *Science* (2003) 302, 1727-1736.
- [54] Li, S., Armstrong, C. M., Bertin, N., Ge, H., Milstein, S., Boxem, M., Vidalain, P. O., Han, J. D., Chesneau, A., Hao, T., *et al.* A map of the interactome network of the metazoan *C. elegans*. *Science* (2004) 303, 540-543.
- [55] Lim, J., Hao, T., Shaw, C., Patel, A. J., Szabo, G., Rual, J. F., Fisk, C. J., Li, N., Smolyar, A., Hill, D. E., *et al.* A protein-protein interaction network for human inherited ataxias and disorders of Purkinje cell degeneration. *Cell* (2006) 125, 801-814.
- [56] Rual, J. F., Venkatesan, K., Hao, T., Hirozane-Kishikawa, T., Dricot, A., Li, N., Berriz, G. F., Gibbons, F. D., Dreze, M., Ayivi-Guedehoussou, N., *et al.* Towards a proteome-scale map of the human protein-protein interaction network. *Nature* (2005) 437, 1173-1178.
- [57] Stelzl, U., Worm, U., Lalowski, M., Haenig, C., Brembeck, F. H., Goehler, H., Stroedicke, M., Zenkner, M., Schoenherr, A., Koeppen, S., *et al.* A human protein-protein interaction network: a resource for annotating the proteome. *Cell* (2005) 122, 957-968.
- [58] Bock, J.R. & Gough, D.A. Predicting protein-protein interactions from primary structure. *Bioinformatics*. (2001) 17, 455-460.
- [59] Binnewies, T.T. *et al.* Ten years of bacterial genome sequencing: comparative genomics-based discoveries. *Funct. Integr. Genomics* (2006) 6, 165-185.
- [60] Goffeau, A. *et al.* Life with 6000 genes. *Science* (1996) 274, , 563-567.
- [61] Genome sequence of the nematode *C. elegans*: a platform for investigating biology. *Science* (1998) 282, 2012-2018.

[62] Finishing the euchromatic sequence of the human genome. *Nature* (2004) 431, 931-945.

[63] Lander,E.S. et al. Initial sequencing and analysis of the human genome. *Nature* (2001) 409, 860-921.

[64] Venter,J.C. et al. The sequence of the human genome. *Science* (2001) 291, 1304-1351.

[65] Stites,W.E. Protein-Protein Interactions: Interface Structure, Binding Thermodynamics, and Mutational Analysis. *Chem. Rev.* (1997) 97, 1233-1250.

[66] Golemis,E.A., Tew,K.D. & Dadke,D. Protein interaction-targeted drug discovery: evaluating critical issues. *Biotechniques* (2002) 32, 636-638.

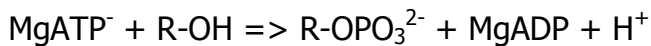
CHAPITRE II :

PROTEINES KINASES ET PEPTIDES

Les protéine-kinases sont des composants majeurs des voies de signalisation qui permettent aux cellules de répondre de manière adaptée à leur environnement.

La phosphorylation des protéines représente un mécanisme de régulation contrôlant pratiquement tous les aspects de la vie cellulaire [1]. Cette réaction covalente est assurée par les protéines kinases, qui participent, au sein des réseaux, à l'intégration de signaux essentiels pour le contrôle de la division, de la différenciation et de la mort des cellules [2]. Les protéines kinases régulent l'activité transcriptionnelle, l'apoptose, la progression dans le cycle cellulaire et la prolifération, les réarrangements du cytosquelette, la différenciation, la mobilité cellulaire, etc [1].

Les kinases catalysent la réaction suivante :



Selon la nature du groupement phosphorylé sur la protéine (R-OH), ces enzymes sont classées en protéines sérine/thréonine kinases et en protéines tyrosine kinases. L'étude du génome humain a conduit à l'identification de 518 gènes de protéine kinase qui correspondent à environ 2% du génome humain. La famille des protéines kinases, incluant 385 sérine/thréonine kinases, 90 tyrosine kinases et 43 protéines proches des tyrosines kinases. Elle est, en taille, la seconde famille d'enzyme après celle des protéases [3].

En raison de leur grande diversité d'actions, les protéines kinases doivent être rigoureusement régulées : l'activité aberrante de ces enzymes génère des maladies comme le cancer ou le diabète, mais aussi à des dysfonctionnements nerveux, cardiovasculaire, inflammatoires ou auto-immunitaires [3].

1. La famille des protéines kinases comme cibles thérapeutiques et analyse de leurs principaux inhibiteurs

La plupart des mécanismes de régulation intracellulaires impliquent des processus de phosphorylation/déphosphorylation par les kinases et les phosphatases [4]. La première phosphorylation de protéine a été identifiée dans le milieu des

années 50 par Krebs et Fischer [5]. L'existence de protéines kinases et de protéines phosphatases chez les bactéries a seulement été mise en évidence à la fin des années 1970 [6]. Les kinases représentent un groupe ubiquitaire d'enzymes qui sont impliquées dans un grand nombre de processus cellulaires. Par définition, le mot kinase est appliqué aux enzymes qui catalysent le transfert du groupement phosphate terminal de l'ATP vers un substrat qui peut être une petite molécule, un lipide ou une protéine. La plupart des kinases sont impliquées dans les différentes voies de transduction du signal. D'autres kinases ont un rôle important dans le métabolisme des glucides, des lipides, des nucléotides, des vitamines, etc..., d'autres encore participent à la régulation des gènes ou encore à la contraction musculaire. Du fait de leurs rôles dans les processus cellulaires, les kinases font partie des enzymes les plus étudiées au niveau structural, biochimique ainsi qu'au niveau cellulaire.

Les protéines kinases, plus particulièrement, représentent une des plus grandes familles de protéines, ce sont des enzymes comptant plus de 500 membres dans le génome humain [7]. Ces protéines kinases (PK) sont impliquées dans la plupart des voies de transduction du signal régulant le fonctionnement cellulaire sous tous ses aspects. La plupart des cancers sont associés au dérèglement de protéines kinases, tels que des mutations génétiques qui conduisent à une sur-expression ou à une activation constitutive de ces protéines. Elles sont également impliquées dans de nombreuses autres pathologies. La compréhension des mécanismes de contrôle de ces PK ainsi que la découverte d'inhibiteurs de PK sont devenus des points-clé pour la recherche de nouveaux agents thérapeutiques.

La fonction catalytique commune des protéines kinases consiste en la phosphorylation covalente d'une protéine-substrat qui se fait par le transfert du groupement phosphate de l'ATP à un résidu thréonine, sérine ou tyrosine du substrat [8]. Cette fonction est assurée par le domaine catalytique des PK dont la structure et en particulier les aminoacides-clé, sont très conservés [9]. Pour qu'un résidu Ser, Thr ou Tyr soit phosphorylé, il est nécessaire qu'il appartienne à une séquence de la protéine-substrat bien particulière, qui assure une conformation déterminée à cette

partie de la protéine-substrat, compatible avec le site catalytique de la protéine kinase.

Autrement dit, tous les résidus Ser, Thr et Tyr ne peuvent pas être phosphorylés, cependant une protéine-substrat peut avoir 1 à 100 sites de phosphorylation. Les PK peuvent être régulées par des protéines activatrices, des protéines inhibitrices ou des cofacteurs. L'activité des PK peut également être contrôlée par phosphorylation, faite par elles-mêmes (autophosphorylation) ou par d'autres protéines kinases de sorte qu'il se crée des cascades de protéines kinases qui amplifient les signaux membranaires. Cette activité peut être également contrôlée simplement par la localisation des PK dans la cellule.

2. Classification des protéines kinases

Le classement des kinases peut s'effectuer en fonction de la structure, de la séquence des protéines [10]. Les PK peuvent être classées selon leur localisation dans la cellule. Ils existent des PK cytosoliques et des PK membranaires qui présentent la propriété de récepteurs d'enzymes. Les PK peuvent également être classées en fonction de l'acide aminé qu'elles phosphorylent. Il existe principalement deux grandes classes de protéines kinases, les 'Tyrosine Kinases' (TK) et les 'Sérine/Thréonine Kinases' [11].

3. Protéines kinases A

Les protéines kinases représentent une des plus grandes familles encodées par le génome humain. Elles fonctionnent comme des interrupteurs moléculaires qui activent ou désactivent la plupart des processus biologiques tels que la mémoire, la différenciation, la division cellulaire, le métabolisme et la mort cellulaire. De nombreuses maladies, et notamment le cancer, sont liées à des déficiences de protéines kinases spécifiques. Ainsi, les protéines kinases jouent un rôle critique pour la survie et la régulation de chaque cellule.

Les protéines kinases AMPc (adénosine monophosphatase cyclique)-dépendantes (PKA) découvertes en 1968 et cristallisées en 1991 servent de prototype pour comprendre la structure, la fonction et la régulation de cette famille d'enzymes ainsi que son évolution à partir de kinases procaryotes. En l'absence de l'AMPc, la PKA existe sous forme inactive où deux sous-unités régulatrices (R sur le schéma) verrouillent deux sous-unités catalytiques (C sur le schéma). La libération des sous-unités catalytiques et donc l'activation de l'enzyme PKA est déclenchée et dépendante de la fixation de l'AMPc aux sous-unités régulatrices.

Les sous-unités C de la PKA ont pour fonction de phosphoryler les protéines, l'ATP étant le donneur de phosphate et d'énergie. La phosphorylation des enzymes-clés par ce mécanisme permet d'activer ou d'inhiber les réactions catalysées et donc de réguler les voies métaboliques.

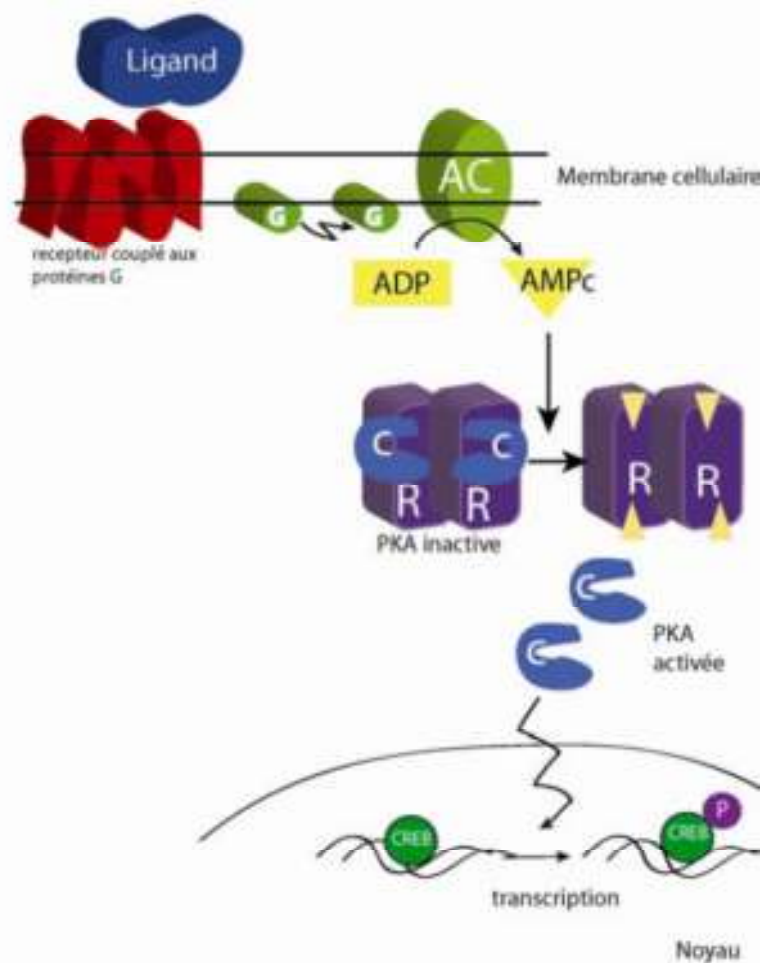


Figure 4 : Activation des protéines kinases A

La protéine kinase A (PKA) fut la première protéine kinase identifiée comme étant impliquée dans l'entrée du VHC (virus de l'hépatite C) [12].

4. Mécanismes d'activation des protéines kinases

En première approximation, les sérine/thréonine ou tyrosine kinases ont les mêmes repliements de leur domaine catalytique. Il consiste en :

1) un lobe N-terminal, essentiellement formé d'un feuillet- β à cinq bandes et d'une hélice α .

2) un lobe C-terminal, essentiellement composé d'hélices α et de la région flexible (segment d'activation) impliquée dans la régulation de l'accès au substrat.

La fente catalytique est à l'interface des deux lobes. L'activation nécessite à la fois l'accès du substrat à la fente catalytique et un positionnement de l'ATP favorable à son transfert. De nombreuses protéines kinases doivent être phosphorylées sur un résidu thréonine ou tyrosine du "segment d'activation" (nommé aussi "boucle d'activation").

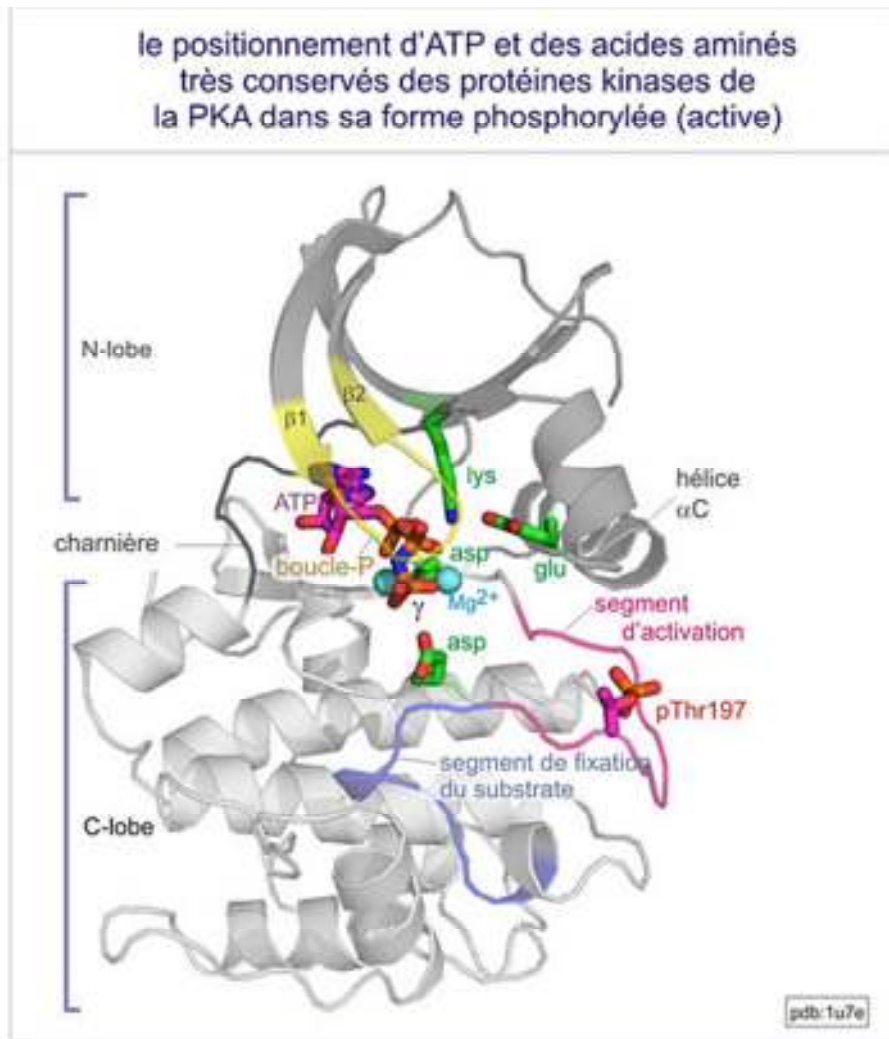


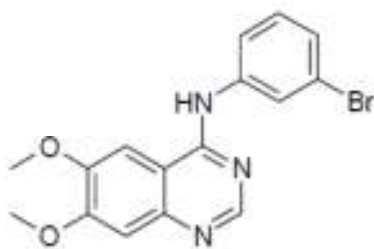
Figure 5: Motifs conservés des protéines kinases

5. Les inhibiteurs de protéines kinases

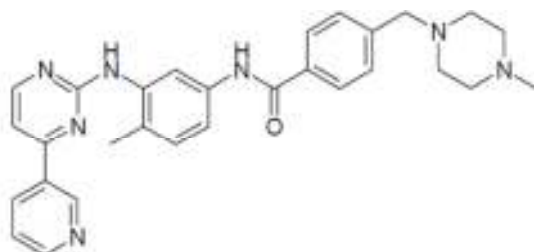
Les protéines kinases jouent un rôle central dans la transduction du signal, elles régulent de nombreux aspects de la fonction cellulaire. Leur dysfonctionnement est impliqué dans de nombreuses maladies, c'est pourquoi elles sont devenues des cibles particulièrement intéressantes pour la recherche de nouveaux agents thérapeutiques.

5.1 Ligands compétitifs de l'ATP

La plupart des inhibiteurs de protéine kinases sont des inhibiteurs compétitifs de l'ATP qui présentent une structure chimique proche du noyau adénine. On trouve de nombreux composés hétérocycliques azotés tels que des quinazolines, des pyrimidines, des pyrrolopyrimidines, des pyrrolopyridines, des pyrazolopyrimidines, etc... Parmi ces inhibiteurs compétitifs de l'ATP, nombreux sont ceux qui présentent une faible sélectivité du fait du haut degré de conservation du site de liaison à l'ATP à l'intérieur de la famille des protéines kinases. Bien que certains inhibiteurs compétitifs de l'ATP soient peu sélectifs, d'autres, au contraire, sont des inhibiteurs puissants et sélectifs comme par exemple l'Iressa (composé **1**) dérivé de quinazoline inhibant l'activité TK avec une IC_{50} de 23-79 nM [7,13]. Ce composé est utilisé dans le traitement du cancer des poumons. Dans la classe des phenylamino-pyrimidines, le Gleevec (composé **2**) est un exemple représentatif. C'est un inhibiteur de la protéine kinase Abl ($IC_{50} = 38$ nM) utilisé dans le traitement de la leucémie myéloïde chronique (CML) [7].



Composé 1 : Iressa

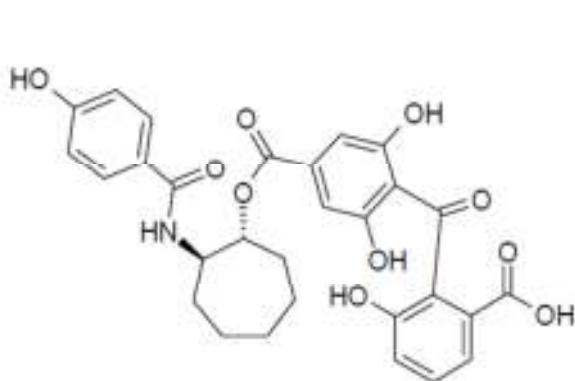


composé 2 : Gleevec

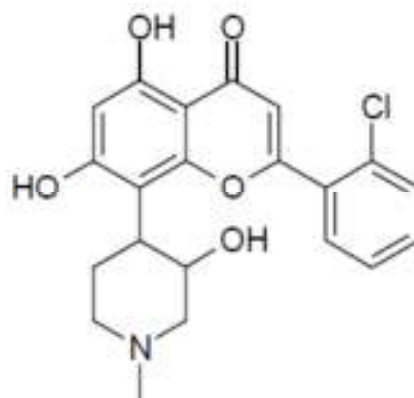
Il existe de nombreuses autres classes d'inhibiteurs de protéine kinases. On peut encore citer le balanol (composé **3**) qui est un composé naturel inhibant des Ser/Thr kinases spécifiques (la plupart étant des PKC) et qui est inactif vis à vis des Tyr kinases. Le flavopiridol (composé **4**) est un flavonoïde qui inhibe la plupart des kinases dépendantes des cyclines (Cdk1, 2...), et qui stoppe la croissance de nombreuses cellules tumorales. Il pourrait également jouer un rôle dans le traitement du virus HIV [7]. Les dérivés de la staurosporine (alcaloïde microbien inhibiteur

puissant mais non spécifique des protéines kinases) sont des inhibiteurs puissants et plus sélectifs que la staurosporine.

Les inhibiteurs compétitifs de l'ATP sont testés sur un nombre limité de Tyr et Ser/Thr kinases, ainsi ils sont souvent abandonnés lors des essais cliniques à cause de leur toxicité [14].



Composé 3 : balanol



composé 4 : flavopiridol

5.2 Ligands compétitifs du substrat

Les inhibiteurs entrant en compétition avec le substrat endogène de la kinase sont moins toxiques puisqu'ils se lient à des domaines bien moins conservés que le domaine de liaison de l'ATP et donc interagissent moins avec d'autres protéines. Les premiers inhibiteurs sélectifs de Tyr kinases découverts font partie de la famille des tyrphostines. Ce sont des inhibiteurs entrant en compétition avec le substrat [14]. Ces composés possèdent un motif catéchol mimant le groupement phénolique de la Tyr. Ce motif porte des groupements hydroxyyles qui ne sont pas très stables métaboliquement.

D'une manière générale, la plupart des inhibiteurs compétitifs du substrat sont des peptides basés sur la séquence peptidique du substrat dans lesquels le résidu phosphorylable est remplacé par un autre acide aminé non phosphorylable comme l'alanine (dans le cas des Ser kinases) ou encore des dérivés de tyrosine (1- ou 2-

naphtylalanine, p-fluoro-, p-guanidino-, p-chlorophénylalanine, D-tyrosine, 3-iodotyrosine, tétrafluorotyrosine..., dans le cas des Tyr kinases) [15,16]. De nombreux inhibiteurs de ce type ont été synthétisés et, pour certains d'entre eux, une très bonne affinité de l'inhibiteur pour la kinase a été déterminée (K_i de l'ordre du nanomolaire) [16].

Pour la protéine kinase A par exemple, de nombreux inhibiteurs basés sur la structure du substrat ont été proposés [17]. Les peptides les plus puissants sont souvent longs (~17 aminoacides). Le plus petit et le plus puissant d'entre eux est un nanopeptide (composé **26**, Figure 8) avec un K_i de 0,036 μM . Des peptides dérivés de la gastrine contenant un résidu tétrafluorotyrosine ont été décrits comme inhibiteurs de l'activité TK du récepteur de l'insuline [18].

Les inhibiteurs peptidiques souffrent malheureusement de nombreuses limitations telles qu'une faible biodisponibilité par voie orale (mauvaise absorption), une faible stabilité métabolique face aux peptidases et protéases présentent au niveau gastro-intestinal. Cependant ils constituent de bons chefs de file pour la découverte d'inhibiteurs entrant en compétition avec le substrat et ils restent des outils inestimables pour les études du mécanisme d'action des enzymes.

6. Inhiber la protéine kinase ; faut-il cibler son état actif ou inactif ?

L'analyse structurale de plusieurs sérine/thréonine kinases révèle qu'à l'état inactif, elles présentent une grande variabilité. A l'état actif au contraire, ces différentes structures convergent vers un cadre commun d'organisation, en ce qui concerne l'orientation correcte des acides aminés du site catalytique et du segment d'activation. Ceci a des conséquences importantes pour la mise au point d'inhibiteurs kinasiques à but thérapeutique (anti-cancéreux par exemple). Pour ce qui est des protéines kinases inactives, des inhibiteurs hautement sélectifs ont été obtenus (qui ciblent le site catalytique). En revanche ce n'est pas le cas pour les kinases dans leur conformation active car leurs sites catalytiques se ressemblent énormément. Cependant, un important inconvénient du ciblage des kinases inactives vient du fait

qu'elles échappent « facilement » à l'inhibiteur par de fréquentes mutations. Du point de vue de leur fonctionnement, les mutations ponctuelles sont aisément tolérées comme ceci est reflété par la grande variabilité de la structure du domaine catalytique des kinases inactives. Ces mutations qui n'ont pas de conséquences fonctionnelles majeures, sont le plus souvent à l'origine d'une perte considérable d'affinité pour l'inhibiteur.

Les protéines kinases sont des enzymes qui phosphorylent un substrat spécifique de la protéine par transfert d'un phosphate de l'ATP (adénosine triphosphate) vers le substrat. L'ATP représente une part importante de l'activité cellulaire car elle stocke et transporte l'énergie à l'intérieur de la cellule. Les protéines kinases sont classées d'après le résidu aminoacide qu'elles phosphorylent. Elles sont importantes car elles régulent les voies de transduction du signal cellulaire et interviennent dans les mécanismes de cascade de transduction du signal. Ces mécanismes impliquent des processus de phosphorylation pour la régulation des activités intracellulaires. Ces activités consistent en la synthèse et la dégradation du glycogène, la synthèse des acides gras, l'oxydation du pyruvate en acétyl-CoA, la formation de triglycérides et la régulation de la glycolyse et de la gluconéogénèse. Un changement du niveau de phosphorylation peut avoir des effets significatifs sur l'activité cellulaire, comme par exemple des désordres dans la croissance cellulaire, la régulation et la différenciation [19]. Puisque les protéines kinases régulent ces activités, des aberrations dans l'activité des kinases peuvent conduire à des maladies comme le cancer [20]. Depuis la mise en évidence de leur influence sur les maladies, les protéines kinases sont devenues des cibles privilégiées de la recherche médicamenteuse.

L'idée qu'on pouvait bloquer une protéine avec un antagoniste remonte à 1905 [21]. Un antagoniste inhibe les effets biologiques après liaison d'un ligand sur un site, tandis qu'un agoniste induit un effet biologique similaire à celui du ligand naturel. En pharmacologie moderne, cette idée a donné les inhibiteurs de protéine kinase. Un inhibiteur peut empêcher l'activité de l'une des 500 molécules cibles [21]. Par inhibition de l'activité de la kinase lorsqu'une protéine est sur-exprimée, il est

possible de réparer le mécanisme de cascade de signaux et ainsi de guérir la maladie [22].

Les petites molécules sont de petites séries d'acides aminés que l'on peut utiliser comme inhibiteurs à cause de leur capacité de liaison. Si une petite molécule se lie à une protéine kinase, elle peut empêcher la phosphorylation et par conséquent inhiber l'activité de la protéine [23]. La découverte de nouveaux médicaments, c'est-à-dire la découverte de petites molécules viables, destinées à l'inhibition, nécessite de trouver quelles sont les petites molécules capables de liaison avec la protéine kinase. Une technique nouvelle de découverte de petites molécules viables nécessite un criblage virtuel associé à une évaluation de l'activité kinasique [23].

On utilise le criblage virtuel pour déterminer quelles sont les petites molécules qui, sur le plan théorique, ont la capacité de se lier à une protéine donnée. Il existe des bases de données relatives à des milliers de petites molécules, que l'on utilise pour évaluer les capacités des petites molécules qui y sont enregistrées [24]. Un ligand peut s'évaluer en considérant la structure de la kinase, sa distribution de charge, la minimisation de l'énergie par empilement aromatique, la réduction de la répulsion stérique et les simulations dynamiques [25]. En déterminant la structure de la petite molécule et de la protéine, et en considérant leur capacité d'interaction stérique et électrostatique, on peut prédire si un ligand va se lier lors d'un essai [25]. On appelle assemblage dirigé cette approche qui fait appel à l'information chimique pour diriger l'orientation de la petite molécule dans le site actif [26].

En général, le ligand est capable de devenir un médicament s'il se lie sur certains sites et s'il est donc inhibiteur d'activité [27]. Les sites cibles pour la petite molécule peuvent être le site de liaison de l'ATP, des sites substrats ou des sites allostériques. La plupart des petites molécules sert à inhiber la fonction kinase par liaison vers le site de liaison de l'ATP, entre les deux lobes de la protéine kinase [22]. C'est cependant la liaison des ligands au site de liaison du substrat qui devient actuellement une approche prédominante. Si une petite molécule se lie au site de liaison du substrat, alors le substrat ne peut plus se lier, la phosphorylation ne peut pas se produire et l'activité est donc inhibée.

Depuis ces percées technologiques, les peptides, au même titre que les protéines, sont considérés comme des molécules thérapeutiques d'avenir. En 2004, plus de 20 % des médicaments appartenant au top 200 des ventes étaient à base de protéines ou de peptides, avec des chiffres d'affaires atteignant les 40 milliards \$ [28], soit environ 10 % du chiffre global de l'industrie pharmaceutique. En 2004, entre 600 et 700 peptides étaient en stade de développement et plus de 150 étaient à différents stades de phase clinique [28]. On les trouve dans différents domaines tels que le traitement de certaines formes de cancers, du SIDA, de l'ostéoporose, de maladies neuro dégénératives [29]. Les peptides offrent plusieurs avantages par rapport aux petites molécules qui constituent les médicaments traditionnels. Le premier avantage est que, représentant souvent la plus petite partie fonctionnelle d'une protéine, ils offrent une efficacité, une sélectivité et une spécificité que les concepteurs de petites molécules ont du mal à atteindre [30]. On peut classer les peptides thérapeutiques en trois catégories [31] :

- Les peptides naturels, aussi appelés peptides bioactifs qui sont soit produits par l'organisme, soit dérivés de protéines [32].
- Les peptides issus de bibliothèques génétiques [33].
- Les peptides issus de bibliothèques chimiques produites par synthèse en phase solide [34].

On peut également citer les peptides vecteurs qui ne sont pas à proprement parler des molécules médicamenteuses mais qui servent à transporter des molécules thérapeutiques [35]. Parmi ces peptides transporteurs, on retrouve par exemple la Pénétratine dérivée de l'homéodomaine d'Antennapédia, conçue pour la délivrance de protéines bioactives [36] ou le Transportan, peptide chimérique issu de la Galanine et du Mastoparan, conçu pour transporter des ARN interférants [37], etc. Finalement, une dernière voie de recherche mettant en oeuvre des peptides est celle qui les utilise non pas comme produits finis mais comme outils pour concevoir des peptidomimétiques, petites molécules qui miment les propriétés bioactives. On peut dater la découverte du premier peptide bioactif à plus de 80 ans. C'est en effet en 1923 que Banting et Macleod a eu le Prix Nobel de Médecine pour leur découverte de

l'insuline dont on connaît le rôle dans le traitement du diabète. À cette époque, et jusqu'à les années soixante, la synthèse de peptides par voie chimique classique prenait jusqu'à plusieurs mois, rendant impossible toute exploitation industrielle. Ce n'est qu'à partir de 1963, année du développement par Merrifield de la méthode de synthèse peptidique par voie solide [38] pour lequel il reçut le Prix Nobel de Chimie en 1984, et aussi grâce à la mise au point de techniques de séparation/purification performantes telles que l'HPLC (Chromatographie Liquide Haute Performance) [39] ou la SPE (extraction de phase solide) [40], que l'utilisation industrielle de peptides non issus du milieu naturel est devenue envisageable [41].

Les peptides offrent plusieurs avantages par rapport aux petites molécules qui constituent les médicaments traditionnels. Le premier avantage est que, représentant souvent la plus petite partie fonctionnelle d'une protéine, ils offrent une efficacité, une sélectivité et une spécificité que les concepteurs de petites molécules ont du mal à atteindre [30]. À titre d'exemple, on peut citer les travaux récemment menés sur NAP, un peptide de 8 résidus, le plus petit fragment actif de l'ADNP (Activity Dependent Neuroprotective Protein) qui montre une activité neuroprotectrice [42]. Un deuxième avantage, est que les produits de dégradation des peptides sont des acides aminés, ce qui élimine grandement les risques de toxicité [43].

La modélisation de peptides thérapeutique peut se diviser en deux voies : basée sur le ligand ou basée sur la structure de la cible. La modélisation basée sur la structure de la cible nécessite évidemment d'identifier cette cible puis d'en déterminer la structure ensuite d'identifier le site d'interaction, l'idéal étant de réaliser une co-cristallisation de la cible avec son ligand naturel, et de cribler des banques de données pour la recherche de molécules qui s'adaptent au mieux ; chimiquement et structurellement au site actif [44]. La modélisation basée sur le ligand est parfois une voie obligatoire, c'est notamment le cas lorsque la structure de la cible est inaccessible.

Beaucoup de récepteurs importants sont des protéines membranaires qui sont difficiles à cristalliser. De nombreux groupes pharmaceutiques comptent sur le développement de peptides, comme nouvelles voies thérapeutiques. Ceci a été rendu possible grâce aux avancées technologiques des dernières décennies. Comparés aux

petites molécules, les peptides présentent des qualités indéniables en termes d'activité. Toutefois, l'utilisation des peptides en médecine se heurte à différents problèmes liés à leur nature chimique ou biologique, principalement dus à leur sensibilité aux dégradations biologiques et à leur faible capacité à traverser les barrières membranaires. De nombreux groupes ont mis au point des stratégies permettant soit de nouvelles voies d'administration non orales, soit des mécanismes de protection pour rendre les peptides moins bio-vulnérables.

7. Peptides a activité biologique

En général, les peptides biologiquement actifs (ou peptides fonctionnels ou bioactifs) sont de petite taille, ils comportent entre 3 et 10 résidus d'acides aminés dans leur structure [45,46]. Ces petits peptides sont aussi appelés oligopeptides [47]. Cependant, il existe des exceptions concernant la taille des peptides bioactifs, par exemple :

- Le caséinomacropéptide (un peptide issu de la κ -caséine induisant la production de cholécystokinine, hormone liée à l'état de satiété) est composé de 64 acides aminés [46].

- Les peptides cationiques antibactériens qui ont une taille variant entre 11 et 50 résidus d'acides aminés [48-50].

- Certaines hormones peptidiques telles que l'insuline (51 résidus acides aminés), et l'hormone adrénocorticotrope ; 39 résidus.

8. Rôles attribués aux peptides biologiquement actifs

Depuis près de 30 ans, les études sur les peptides biologiquement actifs issus des aliments n'ont cessé de susciter un intérêt croissant. Ces recherches ont conduit à la mise en évidence de nombreux types d'activités physiologiques liés aux peptides (dont certains peuvent combiner plusieurs de ces bio fonctionnalités) :

- Rôle sur le système digestif [46,51] (effet anti sécrétoire).

- Rôle antithrombotique [52] (effet sur la coagulation).

- Rôle antihypertenseur [53].
- Rôle immunomodulateur [54].
- Rôle antimicrobien [48].
- Rôle analgésique [54,55].
- Rôles dans la réponse allergique [56].

Ces études ont montré que les peptides bioactifs étaient partiellement résistants à l'hydrolyse et capables d'exercer un effet soit localement dans le tube digestif, soit à distance dans l'organisme, après diffusion dans l'organisme via la circulation sanguine [46].

Une attention toute particulière a été portée sur les di- et tripeptides car, du fait de leur petite taille, ils seraient plus facilement résorbés au niveau du tube digestif [51].

9. LES PEPTIDES ANTIHYPERTENSEURS

Les peptides peuvent être présents naturellement dans les aliments, produits au cours du processus d'élaboration ou formés *in vivo* par hydrolyse enzymatique via les enzymes digestives retrouvées dans le tractus gastro-intestinal. Leur principale fonction est l'apport de l'azote et des acides aminés essentiels à la synthèse des protéines retrouvées dans les tissus et les organes. La plupart des composés biologiquement actifs identifiés jusqu'à maintenant, sont des petits peptides (ou oligopeptides), dont la taille varie de 2 à 10 résidus acides aminés.

Leur résorption au niveau de l'appareil digestif est dépendante de leur taille ainsi que de leurs propriétés hydrophiles/hydrophobes. En 1979, Oshima et al. étaient les premiers à décrire le potentiel thérapeutique que pouvait présenter la gélatine (aliment disposant d'une faible valeur nutritionnelle) dans le traitement de l'hypertension. Ces premiers résultats ont ouvert la voie à d'autres recherches en science des aliments. Certains programmes scientifiques, dont le but était de produire des peptides pharmacologiquement actifs, ont ainsi vu le jour. Ces projets se sont concrétisés par la synthèse de peptides analogues endogènes et/ou le

retraitement des sous-produits obtenus au cours du processus d'élaboration d'aliments. En plus de l'effet bénéfique que pouvait offrir cette approche, elle permettait aussi de recycler certains déchets alimentaires en limitant la pollution/contamination par rejet et en les convertissant en suppléments diététiques [58-61].

10. L'activité biologique

Trois valeurs permettent de caractériser l'activité biologique : le K_i , l' IC_{50} et l' EC_{50} . Le K_i représente, dans le cas d'une enzyme par exemple et surtout dans les conditions idéales dites michaeliennes, la constante d'inhibition. La concentration d'inhibition à 50 % de l'inhibition totale, notée IC_{50} (appelée aussi "Time Dependant Inhibition" en anglais) permet de caractériser l'activité d'une molécule sur l'activité biologique d'une cible. Les courbes dose-effet permettent, quant à elles, de définir l' EC_{50} , concentration effective à 50 % de l'effet total. Il est essentiel lors de la mesure de ces valeurs de se placer dans des conditions où la cible est un facteur limitant, ce qui revient à dire que le phénomène dynamique associé à la cible doit être une fonction linéaire de la concentration de la cible.

Références

- [1] Cohen P. The origins of the protein phosphorylation. *Nature Cell Biol* (2002) 4, 127-130.
- [2] Hunter T & Karin M. The regulation of transcription by phosphorylation. *Cell*, (1992) 70, 375-387.
- [3] Manning G, Whyte DB, Martinez R, Hunter T, Sudarsanam S. The protein kinase complement of the human genome. *Science*, (2002) 298, 1912-1934.
- [4] Hunter T. Protein Kinases and Phosphatases: The Ying and Yang of Protein Phosphorylation and Signalling. *Cell*, (1995) 80, 225-236.
- [5] Krebs E., Fischer E.H. The phosphorylase b to a converting enzyme of rabbit skeletal muscle. *Biochim. Biophys. Acta.*, (1956) 20, 150-157.
- [6] Wang J.Y., Koshland D.E.J. Evidence for protein kinase activities in the prokaryote *Salmonella typhimurium*. *J. Biol. Chem.* (1978) 253, 7605-7608.
- [7] Noble M. E. M., Endicott J. A., Johnson L. N. Protein Kinase Inhibitors: Insights into Drug Design from Structure. *Science*, (2004) 3003, 1800-1805.
- [8] Engh R.A., Bossemeyer D. The Protein Kinase Activity Modulation Sites: Mechanisms for Cellular Regulation-Targets for Therapeutic Intervention. *Advan. Enzyme Regul.* (2001) 41, 121-149.
- [9] Hanks S.K., Quinn A.M. Protein kinase catalytic domain sequence database, identification of conserved features of primary structure and classification of family members. *Methods Enzymol.* (1991) 200, 38-62.
- [10] Cheek S., Zhang H., Grishin N.V. Sequence and Structure Classification of Kinases. *J. Mol. Biol.* (2002) 320, 855-881.
- [11] Gies J.P. Les mécanismes de transduction des signaux membranaires. Seconde étape : Phosphorylation des protéines et ou intervention des calcioprotéines. Bases de pharmacologie moléculaire. Ed. ellipse, pp 105.
- [12] Farquhar MJ, Harris HJ, Diskar M, *et al.* Protein kinase A-dependent step(s) in hepatitis C virus entry and infectivity. *J. Virol* ; (2008) 82, 8797-8811.
- [13] Garcia-Echeverria C., Traxler P., Evans D. B. ATP Site-Directed Competitive and Irreversible Inhibitors of Protein Kinases. *Med. Res. Rev.* (2000) 20, 28-57.

- [14] Levitzki A. Protein Kinase Inhibitors as a Therapeutic Modality. *Acc. Chem. Res.*, (2003) 36, 462-469.
- [15] Lou Q, Leftwich M. E., McKay R. T., Salmon S. E., Rychetsky L., Lam K. S. Potent pseudosubstrate-based inhibitors for p60(c-src) protein tyrosine kinase. *Cancer Res.* (1997) 57, 1877-1881.
- [16] Al-Obeidi F. A., Wu J. J., Lam K. S. Protein tyrosine kinases: structure, substrate specificity, and drug discovery. *Biopolymers*, (1998) 47, 197-223.
- [17] Walsh D., Glass D. B. Utilization of the inhibitor protein of adenosine cyclic monophosphate-dependent protein kinase, and peptides derived from it, as tools to study adenosine cyclic monophosphate-mediated cellular processes. *Methods Enzymol.*, (1991) 201, 304-316.
- [18] Yuan C. J., Jakes S., Elliott S. Graves D. J. A rationale for the design of an inhibitor of tyrosyl kinase. *J. Biol. Chem.* (1990) 265, 16205-16209.
- [19] Balis, Frank. Evolution of Anticancer Drug Discovery and the Role of Cell Based Screening. *Journal of the National Cancer Institute.* (2002) 94, 78-79.
- [20] Cohen, P. Protein Kinases- the major drug targets of the twenty-first century? *Nat Rev Drug Discovery.* (2002) 309-15.
- [21] Drews, Jurgen. Drug Discovery: A Historical Perspective. *Science. Drug Discovery.* (2000) 287, 17.
- [22] Noble, Martin et al. Protein Kinase Inhibitors: Insights into Drug Design from Structure. *Science.* (2004) 303, 19.
- [23] Jorgensen, William. The Many Roles of Computation in Drug Discovery. *Science.* (2004) 303, 19.
- [24] Gray, Nathanael et al. Exploiting Chemical Libraries, Structure, and Genomics in the Search for Kinase Inhibitors. *Science.* (1998) 281.
- [25] Teague, SJ. Implications of Protein Flexibility for Drug Discovery. *Nat Rev Drug Disc.* (2003) 527, 41.
- [26] Fradera, X et al. Guided-docking approaches to structure based design and screening. *Curr Top med Chem.* (2004) 687-700.

- [27] Shah, Neil et al. Overriding Imatinib. Resistance with a Novel Kinase Inhibitor. *Science*. (2004) 305.
- [28] McGee P. First successes turn tide for peptide therapeutics. *Drug Discovery Dev.* (2005) 4.
- [29] Lien S. & Lowman H.B. Therapeutic peptides. *Trends Biotechnol.* (2003) 21, 556-562.
- [30] Hummel G., Reineke U. & Reimer U. Translating peptides into small molecules. *Mol. Biosyst.* (2006) 2, 499-508.
- [31] Sato N, Meijer L, Skaltsounis L, Greengard P, Brivanlou AH. Maintenance of pluripotency in human and mouse embryonic stem cells through activation on Wnt signalling by a pharmacological GSK-3-specific inhibitor. *Nature Med*, (2004) 10, 55-63.
- [32] Watt P.M., Screening for peptide drugs from the natural repertoire of bio diverse protein folds. *Nature Biotech.* (2006) 24, 177-183.
- [33] Sergeeva A. et al., Display technologies: Application for the discovery of drug and gene delivery agents. *Adv. Drug Delivery Rev.* (2006) 15, 1622-1654.
- [34] Shin D.S., Kim D.H., Chung W.J. & Lee Y.S., Combinatorial solid phase peptide synthesis and bioassays. *J. Biochem. Mol. Biol.* (2005) 38, 517-525.
- [35] Mae M. & Langel U., Cell-penetrating peptides as vectors for peptide, protein and oligonucleotide delivery. *Curr. Opin. Pharmacol.* (2006) 6, 509-514.
- [36] Derossi D., Joliot A.H., Chassaing G. & Prochiantz A., The third helix of the Antennapedia homeodomain translocates through biological membranes. *J. Biol. Chem.* (1994) 269, 10444-10450.
- [37] Pooga M., Hallbrink M., Zorko M. & Langel U., Cell penetration by transportan. *FASEB J.* (1998) 12, 67-77.
- [38] Merrifield B., Solid phase peptide synthesis. I. The synthesis of a tetrapeptide. *J. Am. Chem. Soc.* (1963) 85, 2149-2154.

- [39] Chen H. & Horvath C.S., High-speed highperformanceliquid chromatography of peptides andproteins. *J. Chrom. A*, (1995) 705, 3-20.
- [40] Kamysz W. et al., Fast and efficient purification of synthetic peptides by solid-phase extraction. *Acta Chrom.* (2004) 14, 180-186.
- [41] Patch J.A. & Barron A.E., Mimicry of bioactivepeptides via non-natural, sequence-specificpeptidomimetic oligomers.*Curr.Opin. Chem. Biol.* (2002) 6, 872-877.
- [42] Gozes I. & Spivak-Pohis I., Neurotrophic effects ofthe peptide NAP: a novel neuroprotective drug candidate.*Curr. Alzheimer Res.* (2006) 3, 197-199.
- [43] Loffet A., Peptides as drugs: is there a market? *J.Peptide Sci.* (2002) 8, 1-7.
- [44] Anderson A.C., The process of structure-based drugdesign. *Chem. Biol.* (2003) 10, 787-797.
- [45] SEKI E., OSAJIMA K., MATSUFUJI H., MATSUI T., OSAJIMA Y. Angiotensin I converting enzyme inhibitory activity of the short chain peptides derived from various food proteins. (1996) 43, 839-840.
- [46] TOME D. Peptides fonctionnels. *Danone World Newsletter*, (1998) n°17.
- [47] GONZÁLES DE LLANO D., HERRAIZ T., POLO M.C. Peptides. *Handbook of food analysis, vol. 1/* ed. by L.M.L. Nollet. Marcel Dekker Inc. (1996) 229-276.
- [48] HANCOCK R.E.W. Peptide antibiotics. *The Lancet*, (1997) 349, 418-422.
- [49] HANCOCK R.E.W., ROZEK A. Role of membranes in the activities of antimicrobial cationic peptides. *FEMS Microbiol. Let.* (2002) 206, 143-149.
- [50] POWERS J.P.S., HANCOCK R.E.W. The relationship between peptide structure and antibacterial activity. *Peptides*, (2004).
- [51] DZIUBA J., MINKIEWICZ P., NAŁĘCZ D., IWANIAK A. Database of biologically active peptide sequences. *Nahrung*, (1999) 43, 190-195.
- [52] RUTHERFURD K.J., GILL H.S. Peptides affecting coagulation. *Brit. J. Nutr.*, (2000) 84, S99-S102.

- [53] YAMAMOTO N. Antihypertensive peptides derived from food proteins. *Biopolymers*, (1997) 43, 129-134.
- [54] MEISEL H. Biochemical properties of regulatory peptides derived from milk proteins. *Biopolymers*, (1997) 43, 119-128.
- [55] ZIOUDROU C., STREATY R.A., KLEEW.A. Opioid peptides derived from food proteins. *J. Biol. Chem.* (1979) 254, 2446-2449.
- [56] NYBERG F., SANDERSON K., GLÄMSTA E.L. The hemorphins: a new class of opioid peptides derived from blood protein hemoglobin. *Biopolymers*, (1997) 43, 147-156.
- [57] GALLO R.L., MURAKAMI M., OHTAKE T., ZAIYOU M. Biology and clinical relevance of naturally occurring antimicrobial peptides. *J. Allergy Clin. Immunol.* (2002) 110, 823-831.
- [58] FUJII M., MATSUMURA N., MITO K., SHIMIZU T., KUWAHARA M., SUGANO S., KARAKI H. Antihypertensive effects of peptides in autolysate of bonito bowels on spontaneously hypertensive rats. *Biosci. Biotechnol. Biochem.* (1993) 57, 2186-2188.
- [59a] MATSUMURA N., FUJII M., TAKEDA Y., SHIMIZU T. Isolation and characterization of angiotensin I-converting enzyme inhibitory peptides derived from bonito bowels. *Biosci. Biotechnol. Biochem.*, (1993b) 57, 1743-1744.
- [59b] MATSUMURA N., FUJII M., TAKEDA Y., SUGITA K., SHIMIZU T. Angiotensin I-converting enzyme inhibitory peptides derived from bonito bowels autolysate. *Biosci. Biotechnol. Biochem.* (1993a) 57, 695-697.
- [60] HYUN C.K., SHIN H.K. Utilization of bovine blood plasma proteins for the production of angiotensin I converting enzyme inhibitory peptides. *Process. Biochem.*, (2000) 36, 65-71.
- [61] HERNÁNDEZ-LEDESMA B., RECIO I., RAMOS M., AMIGO L. Preparation of ovine and caprine β -lactoglobulin hydrolysates with ACE-inhibitory activity. Identification of active peptides from caprine β -lactoglobulin hydrolysed with thermolysin. *Int. Dairy J.* (2002) 12, 805-812.

CHAPITRE III :

METHODES ET LOGICIELS UTILISES

1. Introduction

Les techniques de la modélisation moléculaire ont débuté dans les années 1970 sous l'impulsion de quelques pionniers tous américains tels que Peter Kollman, William Jorgensen, Martin Karplus ou Norman Allinger. Cette technique présente l'intérêt d'allier une potentialité prédictive à une approche de phénomènes subtils au niveau atomique par ailleurs très difficiles à analyser avec des outils classiques. Un autre intérêt non négligeable est la possibilité d'effectuer une étude préparatoire, en général peu coûteuse, d'un phénomène d'intérêt afin de planifier au mieux une analyse expérimentale qui elle sera coûteuse, cela dans le but d'éviter au maximum les gaspillages ou les mauvaises directions.

À cause de l'augmentation continue de la puissance des processeurs, à coût quasiment constant, pratiquement tout le monde dispose sur son bureau d'une puissante machine de calcul largement suffisante pour effectuer des calculs routiniers de minimisation, *drug design* ou calculs de dynamique moléculaire sur des protéines ou de petits systèmes. L'arrivée massive, *via* internet, de solutions logicielles *open-source* soit entièrement gratuites, soit à des prix très bas (100 ou 200 €) mais extrêmement performantes (plates-formes Linux par exemple) rend cette technique extrêmement compétitive.

2. Bases de la mécanique moléculaire

L'ensemble des techniques utilisées est fondé sur une large simplification de la structure de la matière : les atomes ne sont plus que des masses sphériques pesant le poids atomique de l'élément en question et les liaisons ne sont plus que des ressorts (figure 1). Les molécules se présentent comme un assemblage de masses et de ressorts et le problème peut alors être résolu en utilisant les lois bien connues de la mécanique classique, d'où le nom de « mécanique moléculaire ». À ce point, il est crucial de mettre en évidence une importante limitation de la méthode : il n'y a pas d'électrons explicites, même s'il y a des charges partielles sur les sphères. S'il n'y a pas d'électrons, alors il n'y a pas d'orbitales et donc la notion même de liaison

disparaît ! Dans ces conditions, on ne peut ni rompre, ni créer de liaison et donc on n'a pas accès à la réactivité et toute étude sur les réactions est impossible. Donc si on est intéressés par les chemins réactionnels, alors il faut utiliser la mécanique quantique. Par contre, si ce qui nous intéresse c'est la structure, l'espace conformationnel, la dynamique ou l'interaction non covalente dans le cadre des biomolécules, alors la mécanique moléculaire est l'outil de choix [1].

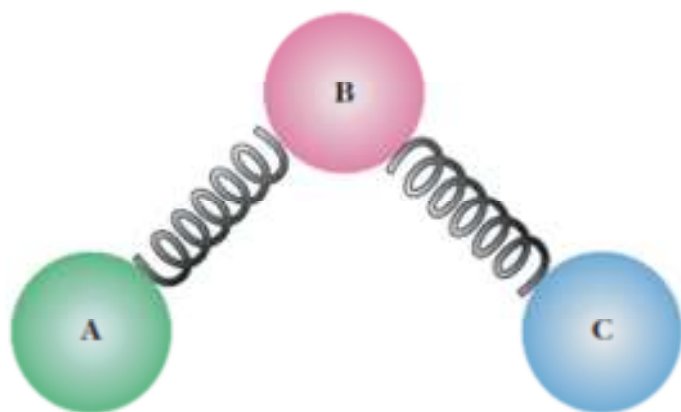


Figure 6 : principe de la mécanique moléculaire

3. Bases de la dynamique moléculaire

La dynamique moléculaire est un complément simple mais qui offre des perspectives infinies à la mécanique moléculaire : à l'assemblage de masses et de ressorts de la mécanique moléculaire, on ajoute des vitesses (ou vélocités) aléatoires sur les sphères (figure 2). On se trouve alors en présence d'un assemblage de masses, de ressorts et de vitesses et donc le problème peut être résolu en intégrant la seconde équation de mouvement de Newton par petites étapes de temps Δt .

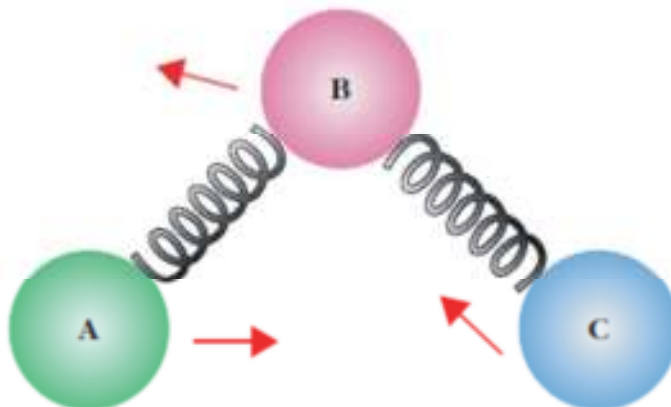


Figure7 : principe de la dynamique moléculaire

Les vitesses sont appliquées de façon aléatoire sur les atomes (sphères) selon la température choisie pour l'expérience. La technique globale est assez simple mais souffre de deux paramètres très défavorables. Le premier est précisément Δt : en effet, si on veut intégrer de façon fiable une collection de mouvements, souvent très disparates comme dans le cas de la dynamique d'une biomolécule, il est nécessaire d'intégrer plus vite que le plus rapide de tous les mouvements. Malheureusement, les systèmes biologiques comprennent un grand nombre d'atomes d'hydrogènes. Comme cet atome est le plus léger de toute la classification, c'est évidemment lui qui a les mouvements les plus rapides. Les vibrations les plus rapides sont donc issues des liaisons de type X-H pour les quelles l'ordre de grandeur est de l'ordre de 10^{-14} secondes ! Si l'on veut intégrer précisément ce type de systèmes organiques, il faut que Δt soit 10^{-15} secondes (1 fs) afin d'avoir environ 10 points pour simuler une sinusoïde (vibration) ! Dans ces conditions, la simple simulation d'une dynamique d'une protéine quelconque pendant 1 ns (10^{-9} secondes) implique 1 million d'intégrations de l'équation de Newton pour un système pouvant compter plusieurs dizaines de milliers d'atomes soit plusieurs millions ou milliards d'interactions non liées à calculer pour chaque intégration, ce qui représente finalement beaucoup de calculs. Malgré la puissance des calculateurs actuels, cette approche est très exigeante et nécessite le plus souvent l'utilisation de nombreux processeurs en parallèle.

Pour donner un exemple, un système comprenant une sérine/thréonine kinase classique avec l'eau nécessaire pour hydrater parfaitement la protéine et les ions nécessaires pour neutraliser cette dernière représente environ 45 000 à 50 000 atomes. Simuler la dynamique de ce système pendant 2 ns demande 1 semaine de calculs sur 8 processeurs récents.

3.1. Influence de la température

Les vitesses sont appliquées au hasard sur les sphères mais de façon à ce que leur répartition suive la courbe de distribution de Maxwell-Boltzman à la température choisie (figure 3). Si l'on choisit une température basse, par exemple 100 K (courbe bleue), alors tous les atomes auront à peu près la même vitesse : à l'extrémité de chaque ressort, on aura donc des masses ayant des vitesses très proches et la structure globale sera peu perturbée. Si on choisit une température élevée, par exemple 1 000 K (courbe rouge), alors les atomes auront en majorité des vitesses très différentes (aires grises sous la courbe rouge). Ainsi aux extrémités d'un ressort, il y aura de fortes chances de trouver des atomes avec des vitesses très différentes et la structure atomique globale sera très perturbée, ce qui permettra une exploration efficace de l'espace conformationnel.

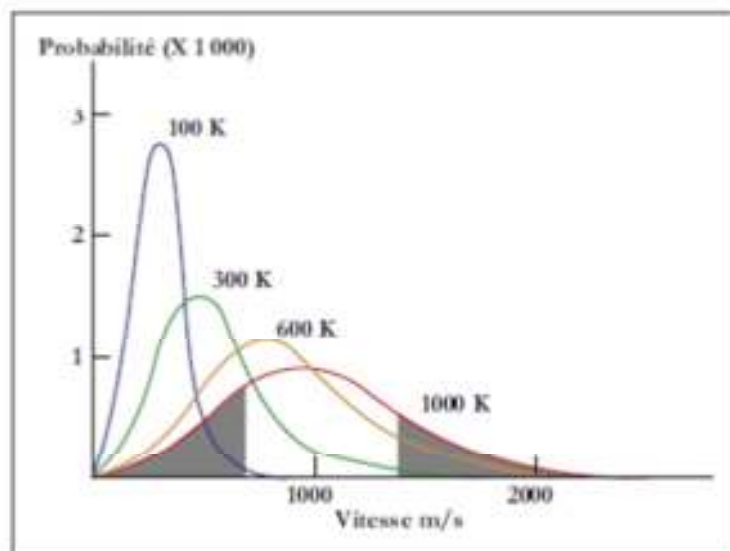


Figure 8 : distributions de Maxwell- Boltzman à différentes températures

La dynamique moléculaire est en fait à peu près insensible à la température : modifier la température ne modifie pas le problème physique. La seule conséquence d'une température élevée est la facilité à passer les barrières rotationnelles ou conformationnelles. Donc, si ce qui nous intéresse est l'exploration d'un espace conformationnel, alors il faut chauffer (1 000 à 1 200 K) ; et si, par contre, ce qui nous intéresse est la dynamique d'une protéine d'intérêt dans des conditions physiologiques, alors il faut se placer à la température physiologique, soit environ 310 K, ce qui est relativement « froid » [1].

À ce point une des questions les plus importantes est : combien de temps doit durer une simulation ? La réponse est entièrement dépendante de l'échelle de temps des phénomènes intéressants que l'on veut simuler : en gros, une simulation doit durer au moins 10 fois le temps nécessaire à la survenue d'un de ces événements. Il faut donc garder en mémoire les quelques éléments suivants :

- si la barrière énergétique qu'il faut passer pour que survienne l'événement à simuler est de l'ordre de 1 kcal.mole^{-1} , alors l'échelle de temps du phénomène est d'environ $1,2 \text{ ps}^{-1}$ et donc la simulation doit durer au moins 12 ps ;
- si la barrière énergétique est de l'ordre de 5 kcal.mole^{-1} , alors l'échelle de temps est d'environ $1,5 \text{ ns}^{-1}$ et donc la simulation doit durer au moins 15 ns ;
- si la barrière énergétique est de l'ordre de $10 \text{ kcal.mole}^{-1}$, alors l'échelle de temps est plus grande que 1 ms et donc la simulation sera extrêmement difficile !
- il faut se rappeler que l'agitation thermique à 300 K est de l'ordre de $2,5 \text{ kcal.mole}^{-1}$.

3.2. Problème du solvant

Si l'on s'intéresse aux petites molécules organiques, dans la plupart des cas et à cause de la prédominance des contraintes géométriques, la présence d'un solvant explicite ne présente pas toujours un grand intérêt. Ce n'est plus le cas pour les systèmes biologiques où la présence d'un solvant (en général l'eau) a une importance capitale sur le comportement des molécules étudiées. Dans beaucoup de logiciels disponibles il est prévu une méthode de solvation implicite. En fait le

meilleur solvant est toujours un solvant explicite mais sa présence implique l'addition de très nombreux atomes au système, ce qui alourdit considérablement la charge de calculs. Ainsi par exemple, une protéine d'environ 400 acides aminés représente environ un peu moins de 7 000 atomes mais, pour la solvater correctement, il faut rajouter un peu plus de 13 000 molécules d'eau, soit environ 40 000 atomes. La solvataion est donc un problème crucial dans le cadre des biomolécules.

3.3. Intérêt de la dynamique moléculaire en amont du docking

La faiblesse majeure des algorithmes de docking est l'absence ou le peu de prise en compte de la flexibilité de la protéine lors de l'opération d'arrimage entre le ligand et la protéine [2]. Ceci ne permet donc pas une co-adaptation optimale du ligand et du récepteur. La dynamique moléculaire est capable de traiter la flexibilité de la protéine. Par conséquent, le couplage des deux puissantes techniques que sont le docking et la dynamique moléculaire doit théoriquement augmenter le pouvoir prédictif de notre modèle.

4. Utilisation de la mécanique et de la dynamique moléculaires pour la découverte et la conception de nouvelles molécules

Les techniques de modélisation moléculaire sont très largement utilisées par l'industrie pharmaceutique et de très nombreux laboratoires académiques dans le monde. En effet, la plupart des médicaments agissent par interaction directe avec des macromolécules biologiques (protéines, ADN ou ARN, polysaccharides, etc.) ou des complexes de ces macromolécules. Dans la plupart des cas, l'interaction responsable de l'activité de la molécule est due uniquement à des interactions à distance sans échange de liaison (interaction non covalente). Dans la plupart des cas, une bonne affinité est le résultat d'une excellente complémentarité de forme et de fonctions chimiques entre le ligand et sa cible que l'on appelle le site actif. La découverte et le développement d'un nouveau médicament sont devenus une entreprise longue et extrêmement coûteuse d'où l'idée d'essayer de rationaliser et d'accélérer le processus.

En effet, tout nouveau composé doit à la fois présenter l'activité désirée avec un minimum d'effets secondaires mais, en plus, il doit être plus efficace que ses concurrents. La découverte de nouveaux médicaments passe par deux étapes clés : la découverte de *hits* et l'identification d'une série de *leads*. Les *hits* sont des molécules qui présentent l'activité biologique recherchée à des degrés divers. Avec les moyens actuels cette étape est assez souvent atteinte. La seconde consiste en l'identification d'une série homogène de molécules ayant des éléments structuraux en commun dont de légères modifications entraînent des variations dans l'activité biologique, les *leads*. On procède ensuite à une optimisation des *leads* jusqu'à obtenir un candidat médicament possédant à la fois l'efficacité, la sélectivité, l'absence de toxicité et les propriétés physicochimiques aptes à le faire atteindre sa cible *in vivo*.

Les techniques de modélisation moléculaire de par leur faible coût global et leur rapidité de mise en œuvre jouent de plus en plus un rôle moteur tout au long de ce processus. Un autre aspect important de ces techniques, qui est bien pris en compte dans la recherche industrielle, est la capacité de ces méthodes à éliminer très vite les très mauvaises solutions. En restreignant souvent sévèrement les pistes possibles, elles permettent des économies très perceptibles de temps et de moyens. Nous allons illustrer cette technique par quelques exemples couvrant les situations habituelles auxquelles on est confronté dans la conception de nouvelles molécules. Deux cas se présentent : on connaît la structure 3D de la biomolécule cible ou bien on ne la connaît pas. La stratégie sera très différente dans les deux cas. Nous allons examiner ces deux situations avec une protéine comme biomolécule. Un acide nucléique (ADN ou ARN) donnera plus ou moins les mêmes résultats mais l'approche est plus complexe dans ce cas.

4.1. Cas 1 : la structure 3D est connue

Deux possibilités se présentent alors : un ligand est présent dans le site actif ou le site actif est bien connu pour d'autres raisons, ou alors aucun ligand ni site actif ne sont connus.

-Site actif connu ou ligand co-cristallisé avec la protéine

Dans ce cas, la situation est la plus satisfaisante et on peut procéder par modélisation moléculaire. La méthode la plus utilisée actuellement est le docking moléculaire (amarrage). Il s'agit d'une technique très simple consistant en fait à positionner un ligand potentiel dans le site actif et à estimer son affinité pour la protéine via un champ de forces de mécanique moléculaire. On peut opérer en utilisant des ligands provenant de protéines proches ou homologues, de bases de données contenant des listes de molécules diverses ou de bases de molécules purement virtuelles imaginées au laboratoire. Le principe est simple puisqu'il suffit d'essayer de multiples positions de chaque molécule dans le site de façon à obtenir pour chacune d'elle la ou les meilleures « poses » (étape de *docking*) puis on effectue une classification énergétique de toutes ces poses avec différents paramètres de mesure (étape de *scoring*) afin d'établir une liste d'affinité des molécules testées pour la protéine d'intérêt. Les molécules les mieux classées seront évidemment les premières candidates à la synthèse dans le cas de bases virtuelles ou seront testées en priorité dans le cas de bases réelles de molécules.

Ce type d'approche appelle plusieurs commentaires. Il existe de nombreuses bases de molécules. La plupart des grandes firmes pharmaceutiques possèdent en propre des bases comptant plusieurs centaines de milliers de molécules propriétaires qui ont été accumulées au cours du temps ou récupérées *via* le rachat de petites bibliothèques. Mais il existe aussi des bases en libre accès dont la plus connue est la base zinc qui contient actuellement plus de 3 millions de molécules.

Le NCI met également à la disposition une base générale contenant environ 256 000 molécules mais aussi une base de diversité contenant 2 000 molécules représentatives choisies pour leur diversité chimique maximale.

Il existe actuellement un grand nombre de logiciels spécialisés dans le *docking*, tels que Dock, Autodock, Gold, Hex, Slide, Glide, Affinity, FlexX, etc. Il existe également plus d'une dizaine de fonctions de *scoring* qui sont intégrées dans la plupart des logiciels pratiquant cette approche. Chacune a ses avantages et inconvénients mais aucune n'est parfaite.

Une technique consiste à essayer de « mouler » le site actif (en fait certaines parties fonctionnelles) avec des molécules proposées, au contact de la surface du site ; il s'agit de la méthode de design. Dans cette technique, on examine soigneusement les résidus présents à la surface de la poche du site actif, en repérant précisément leurs propriétés : charges, groupes donneurs ou accepteurs de liaisons hydrogènes, résidus hydrophobes, aromatiques, zones polaires, etc. Ensuite, en face de chacune de ces entités, on va positionner exactement un partenaire complémentaire. Par exemple, à une charge positive (ammonium), on va opposer une charge négative (carboxylate par exemple), à un donneur de liaison H on va opposer un accepteur, et ainsi de suite. Une fois cette opération effectuée, on est en présence d'une collection de petits éléments positionnés exactement dans l'espace et qui sont susceptibles d'interagir fortement avec le site actif.

Dans toutes ces opérations, il est évident qu'une bonne expérience de chimie de synthèse est un plus car cela évite de proposer des structures trop difficiles à préparer ou impossibles à synthétiser. En particulier, il faudra éviter de tomber dans le piège de la complexité en essayant à tout prix de prendre en compte toutes les particularités du site actif. Dans beaucoup de cas, un bon ligand fait preuve d'une remarquable économie de moyens pour arriver à ses fins : un petit nombre de contacts judicieusement placés sont largement suffisants plutôt qu'une multitude de contacts qui brouillent les cartes.

-Site actif de localisation inconnue

Dans ce dernier cas et malgré la présence d'une structure 3D la situation n'est pas simple ! En effet sans aucune idée sur la localisation du site actif il est extrêmement difficile d'avoir une idée raisonnable des sites d'actions d'un ligand vis-à-vis d'une protéine. Les molécules classiques sont en général de petites tailles (souvent de la taille de 1 à 3 acides aminés) et fortement fonctionnalisées : elles peuvent donc interagir avec un nombre considérable de poches ou éléments de surface, sans qu'il soit possible de favoriser *a priori* telle ou telle position.

Une des techniques possibles est de chercher les poches et cavités présentes dans la protéine et de les examiner une par une. Il existe plusieurs logiciels, dont certains gratuits comme Surfnet, qui pratiquent ce type de recherche. Par contre, même dans une protéine très moyenne, le nombre de cavités peut être considérable et, on a peu de chances d'arriver à un résultat fiable. On peut également s'aider de cartes de potentiels électrostatiques ou de lipophilie moléculaire qui donnent en général des renseignements très qualitatifs mais souvent précieux quand on a peu de chose sous la main. Les potentiels électrostatiques peuvent montrer les grandes directions d'interaction autour de la protéine et parfois les zones d'interaction en surface. Quant aux potentiels de lipophilie, ils permettent de repérer les poches hydrophobes dans la protéine (mais aussi parfois en surface) qui sont souvent l'indice d'un site actif.

4.2. Cas 2 : la structure 3D n'est pas connue

Là encore deux possibilités se présentent:

- il existe une protéine dont on connaît la structure 3D et qui est homologue de la protéine d'intérêt, dans ce cas on va pouvoir proposer un modèle de protéine par la technique dite de la reconstruction par homologie de séquence.
- on connaît une série de ligands de la protéine d'intérêt et dans ce cas on va opérer par modélisation indirecte.

5. Problèmes liés à l'approche par modélisation moléculaire

Un des principaux problèmes est d'arriver à gérer correctement l'efficacité d'un médicament et sa sélectivité. Si l'on prend comme exemple les kinases qui sont actuellement une des cibles privilégiées dans la lutte contre le cancer, celles-ci représentent une vaste famille (518 membres chez l'homme) d'une grande homogénéité au niveau du site actif. Dans tous les cas, ce site doit lier impérativement une molécule d'ATP. Si l'on cible spécifiquement ce site, on aura effectivement une molécule très active mais très peu sélective (cas de la staurosporine par exemple). Si l'on veut être spécifique pour une kinase particulière (kinase BCR-ABL par exemple), il faut cibler un site spécifique de la kinase d'intérêt. C'est le mode d'action de l'imatinib mesylate qui cible une poche lipophile accessoire située sur le bord du site actif. Cette poche est spécifique à cette kinase particulière et l'imatinib mesylate est effectivement (relativement) sélectif.

Une conception rationnelle des médicaments est essentiellement une approche interdisciplinaire. En effet il faut avoir une bonne connaissance des protéines d'intérêt (fonction, structure, expression, régulation) mais aussi des ligands naturels, ce qui suppose une étude profonde au niveau de la biochimie. À partir de là, les techniques de la modélisation moléculaire peuvent proposer une conception rationnelle de molécules nouvelles qui seront à leur tour testés et ces résultats après retour à l'étape modélisation devraient générer des candidats médicaments que la biologie validera.

6. Modélisation d'une protéine

L'énergie d'une protéine peut être obtenue par différentes approches : regroupées sur la figure 4.

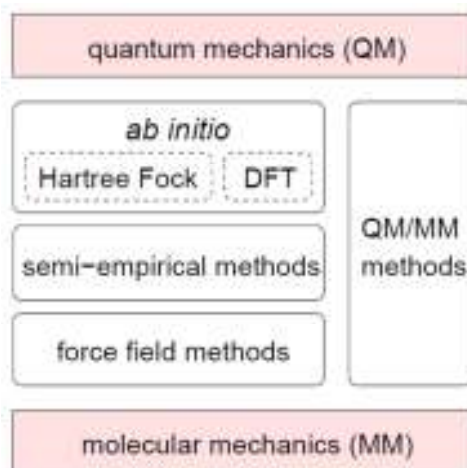


Figure 9 : Principales méthodes de calcul d'énergie utilisées en simulation moléculaire.

Les méthodes *ab initio* tentent de résoudre l'équation de Schrödinger dans laquelle on traite les électrons actifs du système. L'objectif de telles techniques est de déterminer l'énergie d'une molécule ainsi que la fonction d'onde ou la densité des électrons. Une méthode très utilisée est la méthode Hartree-Fock, pour laquelle chaque électron perçoit les autres électrons comme un champ moyen [3]. Les méthodes *ab initio* sont limitées à des systèmes comprenant peu d'électrons donc peu d'atomes.

On assiste depuis une quinzaine d'années au développement des méthodes DFT (density functional theory) [4, 5] pour les systèmes biologiques. La théorie de la fonctionnelle de la densité repose sur la détermination de l'état fondamental électronique par sa densité totale au lieu de la fonction d'onde.

Les méthodes semi-empiriques reposent sur les fondements de la mécanique quantique mais supposent un certain nombre d'approximations, notamment sur les calculs des termes non diagonaux dans l'hamiltonien. On peut citer les méthodes de type MNDO (modified neglect of differential overlap) [6], AM1 (Austin model 1) [7], et PM3 (parametric model number 3) [8], toutes les trois basées sur l'approximation NDDO (neglect of differential diatomic overlap) [9]. Les méthodes par champs de force ne considèrent, par contre, que les mouvements des noyaux.

Le calcul de l'énergie potentielle dépend d'une forme fonctionnelle ainsi que d'un jeu de paramètres empiriques. Les électrons ne sont plus explicitement pris en compte, donc la rupture ou la création de liaisons chimiques n'est plus possible. Le calcul par champ de force est très rapide comparé aux méthodes *ab initio* ou semi-empiriques, et permet d'aborder les propriétés statistiques ou de traiter des systèmes contenant plusieurs centaines ou plusieurs milliers d'atomes. Enfin, les méthodes QM/MM [10, 11] sont des méthodes hybrides qui traitent une partie du système quantiquement et l'autre partie par une approche classique. Dans ce travail, nous avons utilisé l'approche par champs de force qui est adaptée aux systèmes de grande taille.

6.1. Modélisation par champs de force

L'utilisation d'un champ de force est pertinente lorsqu'on veut réaliser des simulations de mécanique ou dynamique moléculaire pour des systèmes comprenant plusieurs centaines d'atomes. Un champ de force se définit comme une forme fonctionnelle et un ensemble de paramètres attribué à un type d'atome. Les champs de force les plus couramment utilisés pour les simulations de biomolécules sont : AMBER (assisted model building and energy refinement) [12], CHARMM (chemistry at harvard using molecular mechanics) [13], OPLS (optimized potentials for liquid simulations) [14], GROMOS 96 (Groningen molecular simulation package) [15] et ECEPP (empirical conformational energy program for peptides) [16].

6.2. Simulation de biomolécules par dynamique moléculaire

La dynamique moléculaire est la méthode la plus intuitive et la plus naturelle pour explorer la surface d'énergie potentielle d'une biomolécule. Historiquement, la première molécule d'intérêt biologique — l'inhibiteur de la trypsine pancréatique bovine (BPTI) — a été modélisée par dynamique moléculaire il y a moins de 30 ans [17] et cette technique de simulation est toujours largement utilisée actuellement [18]. La dynamique moléculaire cherche à obtenir une exploration de la surface

d'énergie potentielle, l'accumulation de statistiques et bien évidemment, la construction d'une dynamique réelle, par exemple, de repliement [18].

- Principales échelles de temps rencontrées en dynamique moléculaire

L'intégration temporelle se fait de manière discrète en dynamique moléculaire. Il convient donc de définir avec soin le pas de temps utilisé. Le tableau 1 résume quelques différents temps caractéristiques des systèmes étudiés. Avec un pas de temps trop court, la simulation ne peut pas converger et explorer correctement l'espace des phases. Réciproquement, un pas trop grand aboutit à une forte instabilité du système [19].

Mouvement	Temps (s)
Elongation d'une liaison	10^{-14}
Flexion d'une liaison	10^{-14}
Rotation des chaînes latérales de surface	$10^{-11} - 10^{-10}$
Rotation des chaînes latérales intérieures	$10^{-4} - 1$
Repliement de protéines	$10^{-6} - 10^2$

Tableau 1 : Principales échelles de temps rencontrées en dynamique moléculaire pour des systèmes biologiques. [19, 20]

Le défi des simulations de biomolécules est de modéliser des phénomènes lents comme le repliement d'une protéine tout en tenant compte de mouvements rapides comme la vibration d'une liaison. Les algorithmes de dynamique modifient les équations du mouvement de façon à geler les vibrations les plus rapides. La dynamique moléculaire peut actuellement simuler des biomolécules sur des temps de l'ordre de 10 ns ou des systèmes allant jusqu' à 10^4 – 10^6 atomes [18].

6.3. Le docking Moléculaire

Les interactions entre molécules sont à la base de la plupart des mécanismes biologiques. Les détails de ces interactions, au niveau moléculaire, sont donc d'un très grand intérêt et peuvent être étudiés par cristallographie aux rayons X ou résonance magnétique nucléaire (RMN). Malheureusement, ces techniques ne permettent pas de détailler la totalité de ces interactions étant donné le nombre incalculable de molécules différentes au sein d'une seule cellule.

Le docking moléculaire vise à prédire la structure d'un complexe moléculaire à partir des molécules isolées, ce qui est considérablement plus facile à mettre en oeuvre, moins cher et plus rapide que l'utilisation d'une des méthodes expérimentales mentionnées ci-dessus. Les logiciels de docking sont donc des outils très utiles en biologie, pharmacie et médecine, car la plupart des principes actifs sont de petites molécules (ligand) qui interagissent avec une cible biologique d'intérêt thérapeutique, généralement protéique (récepteur), afin d'influencer le mécanisme dans lequel cette protéine est impliquée.

La procédure typique à mettre en place pour prédire les interactions entre un ligand et son récepteur est présentée dans la Figure 5. La première étape est l'obtention d'une structure pour le récepteur, par cristallographie aux rayons X, par RMN ou par modélisation.

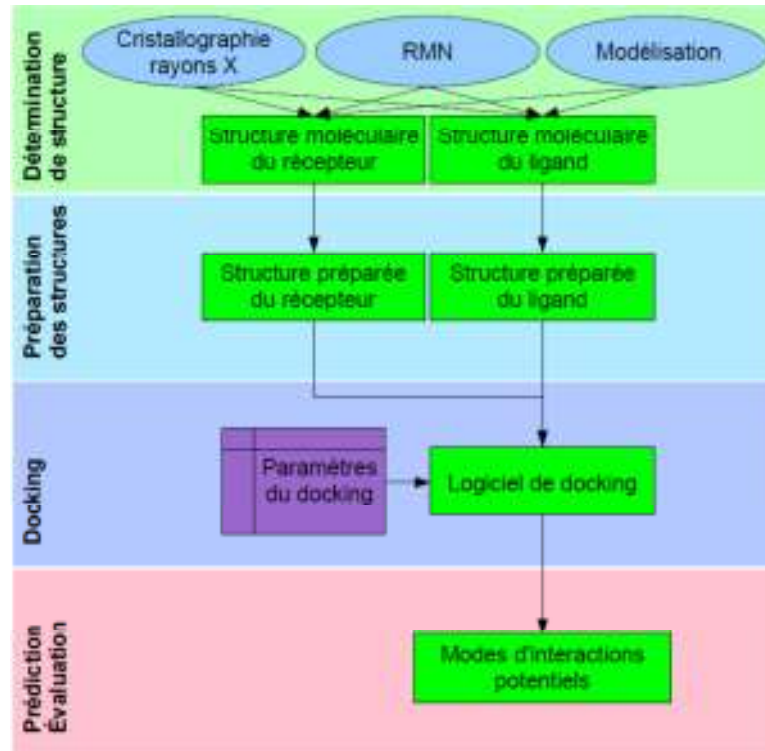


Figure 10 : Étapes typiques d'un docking moléculaire.

La conformation initiale du ligand est généralement modifiée et optimisée durant le docking. Une fois les deux structures préparées, les paramètres sont donnés au logiciel de docking, et ce dernier propose un ou plusieurs modes d'interactions potentiels qui peuvent être ensuite utilisés d'une multitude de façons différentes.

6.3.1 Les programmes d'assemblage moléculaire

Avant de détailler les divers programmes de docking existants, nous allons présenter les étapes de base de ces programmes. Il existe d'ailleurs plusieurs publications de synthèse très intéressantes sur le sujet [21-25]. Plus de 30 logiciels de docking sont actuellement disponibles [26]

Le principe des programmes de docking est d'essayer d'assembler deux structures macromoléculaires, protéines et/ou acides nucléiques, afin d'obtenir un complexe. Ce processus se déroule en plusieurs étapes :

- Recherche globale des assemblages potentiels :

Cette première étape est basée sur la complémentarité de forme entre les partenaires. En règle générale, la plus grosse molécule (le récepteur) est fixée dans l'espace tandis que la plus petite molécule (le ligand) explore l'espace des conformations autour du récepteur. Nous pouvons noter 3 grandes techniques pour effectuer cette recherche.

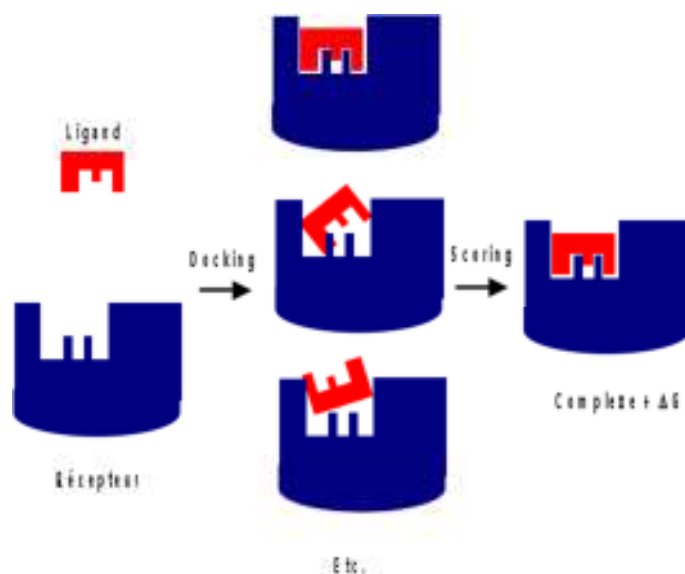
La plus largement utilisée est la transformée de Fourier rapide ou Fast Fourier Transform-FFT [27]. Avec cette méthode, la molécule est discrétisée en petit cube unitaire d'une grille. Cette technique a l'avantage de fournir des solutions très rapidement. De nombreux programmes de docking utilisent cette approche : comme ZDock, SmoothDock, MolFit, DOT, FTDock, 3D Dock, GRAMM ou encore BiGGER (voir tableau 1.3). Une autre approche possible est d'utiliser les harmoniques sphériques pour le calcul de la surface moléculaire et la recherche rapide de complexes potentiels : il s'agit de la transformée de Fourier polaire ou Polar FT [28] qui est utilisée dans le programme Hex.

- Reclassement des complexes trouvés :

Toutes ces techniques permettent d'obtenir de quelques centaines à quelques milliers d'assemblages potentiels. Il faut alors pouvoir discriminer ce type de complexe des solutions potentielles. Il est important de rajouter des informations comme le potentiel électrostatique à la surface des partenaires, la prise en compte de la formation de liaisons hydrogène ou des termes de désolvatation dans le but de reclasser les assemblages potentiels. Tous ces paramètres sont le plus souvent inclus dans des fonctions de score.

Deux autres phases cruciales du docking sont : l'incorporation d'informations (expérimentales ou issues de la littérature) et le traitement de la flexibilité.

Chaque programme peut néanmoins utiliser ces étapes de différentes manières, par exemple, en combinant la recherche de candidats potentiels avec l'utilisation de fonctions de score pour guider les recherches, ou en modélisant la flexibilité avant, après ou pendant cette même phase de recherche.



6.3.2 Les divers programmes de docking :

A travers cette partie, nous allons développer plus en détails les programmes et les techniques que nous avons présentées dans le paragraphe précédent. Nous nous focaliserons sur les programmes les plus remarquables. Le tableau 2 résume les différentes spécificités de chacun. Les programmes utilisant la technique de FFT ou assimilée : Comme expliqué dans la partie précédente, un grand nombre de programmes utilisent la FFT pour la première étape de la recherche de complexes.

Katchalski-Katzir et al. furent les premiers à employer cette technique pour l'amarrage moléculaire au travers d'un algorithme qui devint par la suite le programme MolFit [29,30]. Le récepteur et le ligand sont d'abord représentés sur une grille. On attribue ensuite un score à chaque voxel de la grille. Les modèles sont ensuite classés en fonction des résultats obtenus par la fonction de corrélation : plus

la complémentarité de surface est bonne, plus les valeurs sont élevées. Cette recherche est utilisée par de nombreux programmes comme ZDock, SmoothDock, DOT, FTDock, 3D Dock ou GRAMM. La différence entre ces programmes se fait alors sur d'autres étapes comme le reclassement des résultats grâce à la fonction de score ou l'incorporation de la flexibilité au niveau des chaînes latérales ou du squelette protéique.

Une approche un peu différente a été développée par Ritchie et Kemp avec le programme HEX [28]. La surface des molécules est représentée par des harmoniques sphériques. Alors que les méthodes précédentes accélèrent les recherches par translation, l'approche développée par D. Ritchie favorise les recherches par rotation en utilisant une transformée de Fourier à une dimension (les translations étant aussi prises en compte). Cette implémentation, ajoutée à partir de la version 3.1 de Hex, permet un gain de temps de 30 à 50% par rapport aux versions précédentes [31].

Les programmes présentés dans cette section diffèrent donc par leurs méthodologies et aussi par le temps de calcul associé à celles-ci. En effet, suivant le degré de précision demandé, les temps de calcul varient de quelques minutes à plusieurs jours (tableau 2). Néanmoins, les moyens de calcul se développent et l'utilisation de clusters ou de grilles d'ordinateurs permettent de réduire considérablement ces temps. De plus, de nouvelles stratégies utilisant le potentiel des cartes graphiques commencent à se développer [25].

Programme	Recherche	Fonction de score	Temps	Références
RosettaDock	MC	Vdw + désolvatation + électro.	Semaine	Wang et al.(2005)
SmoothDock	FFT	Forme + désolvatation+ électro.	heures	Camacho et Gatchell(2003)
Molfit	FFT	Forme+ non polaire+ électro.	heures	Berchanski et Eisenstein(2003)
ZDock	FFT	Forme+ électro.+ désolvatation	heures	Chen et al.(2003a)
RosettaDock	MC	Vdw+ désolvatation+ électro.	semaine	Gray et al.(2003)
HADDOCK	MC-DM	Champ de force	jours	Dominguez et al. (2003)
DOT	FFT	Forme+ électro.	minutes	Mandell et al. (2001)
HEX	PFT	Forme+ électro.	minutes	Ritchi et Kemp (2000)
PatchDock	géométrie	Forme+ électro.	minutes	Schneidman-Duhovny et al. (2005b)

Programme	Recherche	Fonction de score	Temps	Références
FTDOCK	FFT + DM	Forme + électro.	jours	Smithet al.(2005b)
3D-Dock	FFT	Forme+ électro.	heures	Aloys et al. (2001)
ATTRACT	Vecteurs propres	Potentiel réduit	heures	ZAcharias (2003)
GRAMM	FFT	Forme+ électro.	heures	Tovchigrechko et al.(2002)

Tableau 2 : présentation des programmes de docking les plus utilisés

MC : Monte-Carlo ; FFT : Fast Fourier Transform ; DM : dynamique moléculaire ; PFT : Polar Fourier Transform.

Electro. : prise en charge des interactions électrostatiques

7. Logiciels utilisés

7.1. Hyperchem 8

Hyperchem 8.0.3 est un logiciel de modélisation moléculaire développé par Autodesk, INC, et distribué par hypercube, INC (Ontario, Canada) [32]. C'est un logiciel sophistiqué, qui est connu pour sa qualité, sa flexibilité et sa facilité d'usage. Unissant l'animation et la visualisation 3D, hyperchem peut faire des calculs de mécanique, de dynamique moléculaire et même quantiques.

7.2. Hex 6.3

Hex est un logiciel de docking moléculaire développé par Dave Ritchie [31]. Il utilise une représentation des surfaces moléculaires des protéines par l'utilisation d'harmoniques sphériques.

Références

- [1] M. Laguerre, Modélisation moléculaire et conception de nouvelles molécules. *Bull Cancer*. (2007) 94, 7.
- [2] Rester, R. Dock around the Clock – Current Status of Small Molecule Docking and Scoring *QSAR Comb. Sci.* (2006), 25, 605-615.
- [3] C. J. Cramer. *Essentials of Computational Chemistry : theories and models*. Wiley, Chichester, England, 2nd edition, (2004).
- [4] P. Hohenberg and W. Kohn. Inhomogeneous Electron Gas. *Physical Review B*, (1964) 136, 864–887.
- [5] W. Kohn and L. J. Sham. Self-consistent Equations Including Exchange and Correlation Effects. *Physical Review A*, (1965) 140, 1133–1138.
- [6] M. J. S. Dewar and W. Thiel. Ground States of Molecules, 38. The MNDO Method. Approximations and Parameters. *Journal of American Chemical Society*, (1977) 99, 4899– 4907.
- [7] M. J. S. Dewar, E. G. Zoebisch, E. F. Healy, and J. J. P. Stewart. AM1: A New General Purpose Quantum Mechanical Model. *Journal of American Chemical Society*, (1985) 107, 3902–3909.
- [8] J. J. P. Stewart. Optimization of Parameters for Semi-Empirical Methods. I-Method. *Journal of Computational Chemistry*. (1989) 10, 209–220.
- [9] J.A. Pople, D. L. Beveridge, and P. A. Dobosh. Approximate Self-consistent Molecular Orbital Theory V. Intermediate Neglect of Differential Overlap. *Journal of Chemical Physics*. (1967) 47, 2026.
- [10] J. Gao. Methods and Applications of Combined Quantum Mechanical and Molecular Mechanical Potentials. In K. B. Lipkowitz and D. B. Boyd, editors, *Reviews in Computational Chemistry*. (1996) 7, 119–185.
- [11] P. Amara and M. Field. Combined Quantum Mechanics and Molecular Mechanics Potentials. In P. V. R. Schleyer, editor, *Encyclopedia of Computational Chemistry*. (1998) 1, 431–437.
- [12] S. J. Weiner, P. A. Kollman, D. A. Case, U. C. Singh, C. Ghio, G. Alagona, Jr. S. Profeta, and P. K. Weiner. A new force field for molecular mechanical simulation of

nucleic acids and proteins. *Journal of American Chemical Society*, (1984) 106, 765–784.

[13] B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus. CHARMM: A Program for Macromolecular Energy, Minimization, and Dynamics Calculations. *Journal of Computational Chemistry*. (1983) 4, 187–217.

[14] W. L. Jorgensen and J. Tirado-Rives. The OPLS potential functions for proteins. Energy minimizations for crystals of cyclic peptides and crambin. *Journal of American Chemical Society*, (1988) 110, 1657–1666.

[15] W. F. van Gunsteren and H. J. C. Berendsen. *Computer Simulation of Molecular Dynamics: Methodology, Applications and Perspectives in Chemistry*. *Angewandte Chemie International Edition in English*. (1990) 29, 992–1023.

[16] F. A. Momany, R. F. McGuire, A. W. Burgess, and H. A. Scheraga. Energy Parameters in Polypeptides VII. Geometric Parameters, Partial Charges, Non-bonded Interactions, Hydrogen Bond Interactions and Intrinsic Torsional Potentials for Naturally Occurring Amino Acids. *Journal of Physical Chemistry*. (1975) 79, 2361–2381.

[17] J. A. McCammon, B. R. Gelin, and M. Karplus. Dynamics of folded proteins. *Nature*, (1977) 267, 585–590.

[18] M. Karplus and J. A. McCammon. Molecular dynamics simulations of biomolecules. *Nature Structural Biology*, (2002) 9, 646–652.

[19] A. R. Leach. *Molecular modelling: principles and applications*. Pearson Education, Harlow, 2nd edition, (2001).

[20] T. Schlick. *Molecular Modeling and Simulation: an interdisciplinary guide*. Springer-Verlag, New-York, USA, (2002).

[21] Halperin, I., Ma, B., Wolfson, H. et Nussinov, R. Principles of docking : An overview of search algorithms and a guide to scoring functions. *Proteins : Struct. Func. Genet.* (2002) 47, 409–443.

[22] Smith, G. R. et Sternberg, M. J. E. Prediction of protein-protein interactions by docking methods. *Curr. Op. Struct. Biol.* (2002) 12, 28–35.

[23] Vajda, S. et Camacho, C. J. Protein-protein docking : is the glass half-full or halfempty ? *Trends in Biotechnology*. (2004) 22, 110–116.

- [24] Gray, J. J. High-resolution protein-protein docking. *Curr. Op. Struct. Biol.* (2006) 16,183–193.
- [25] Ritchie, D. W. Recent progress and future directions in protein-protein docking. *Curr Protein Pept Sci*, (2008) 9,1–15.
- [26] SOUSA S. F., FERNANDES P. A., RAMOS M. J., Protein-ligand docking: current status and future challenges., *Proteins.* (2006), 65, 15-26.
- [27] Harrison, R. W., Kourinov, I. V. et Andrews, L. C. The Fourier-Green's function and the rapid evaluation of molecular potentials. *Protein Eng.* (1994) 7, 359–369.
- [28] Ritchie, D. W. et Kemp, G. J. L. Protein docking using spherical polar Fourier correlations. *Proteins : Struct. Func. Genet.* (2000) 39, 178–194.
- [29] Katchalski-Katzir, E., Shariv, I., Eisenstein, M., Friesem, A. A. et Aflalo, C. Molecular surface recognition : Determination of geometric fit between proteins and their ligands by correlation techniques. *Proc. Natl. Acad. Sci.* (1992) 89, 2195–2199.
- [30] Berchanski, A., Shapira, B. et Eisenstein, M. Hydrophobic complementarity in protein-protein docking. *Proteins : Struct. Func. Bioinf.* (2004) 56,130–142.
- [31] Ritchie, D. W. Evaluation of protein docking predictions using Hex 3.1 in CAPRI rounds 1 and 2. *Proteins.* (2003) 52, 98–106.
- [32] Hyperchem (Molecular Modelling System) Hypercube, Inc. USA (2007).

CHAPITRE IV :

RESULTATS ET DISCUSSION

1. Introduction

La modélisation moléculaire [1] est un outil complémentaire à l'expérimentation de plus, elle donne des grandeurs qui sont au-delà des possibilités expérimentales et ceci dans des conditions parfaitement contrôlées de pureté, de pression, de température, etc.

Par le rôle qu'elles jouent dans de nombreuses voies de signalisation, les protéines kinases A, contribuent à de nombreuses maladies comme le cancer, l'inflammation, les diabètes, les maladies cardiovasculaires, ainsi que les maladies infectieuses.

Ce travail consiste à étudier l'interaction de la protéine kinase A avec des fragments de protéine tel que le tripeptide feG et ses analogues.

Les peptides, au même titre que les protéines, sont considérés comme des molécules thérapeutiques. De manière générale, les peptides biologiquement actifs (ou peptides fonctionnels ou bioactifs) sont de petite taille, ils comportent entre 3 et 10 résidus acides aminés dans leur structure [2,3]. Ces petits peptides sont aussi appelés oligopeptides [4].

2. Préparation du récepteur

La structure tridimensionnelle de la protéine kinase A (PKA) a été obtenue par diffraction aux rayons X à haute résolution (2.20Å). Un modèle a été déposé dans la base de données **Bookhaven Protein Data Bank (code d'accès 1YDR)**. Il représente la co-cristallisation de la PKA réalisée en présence de l'inhibiteur : 1-(5-isoquinolinesulfonyl)-2-methylpiperazine. La protéine PKA est constituée de 350 AA, elle comporte 15 hélices α et 14 feuillets β (figure 1).

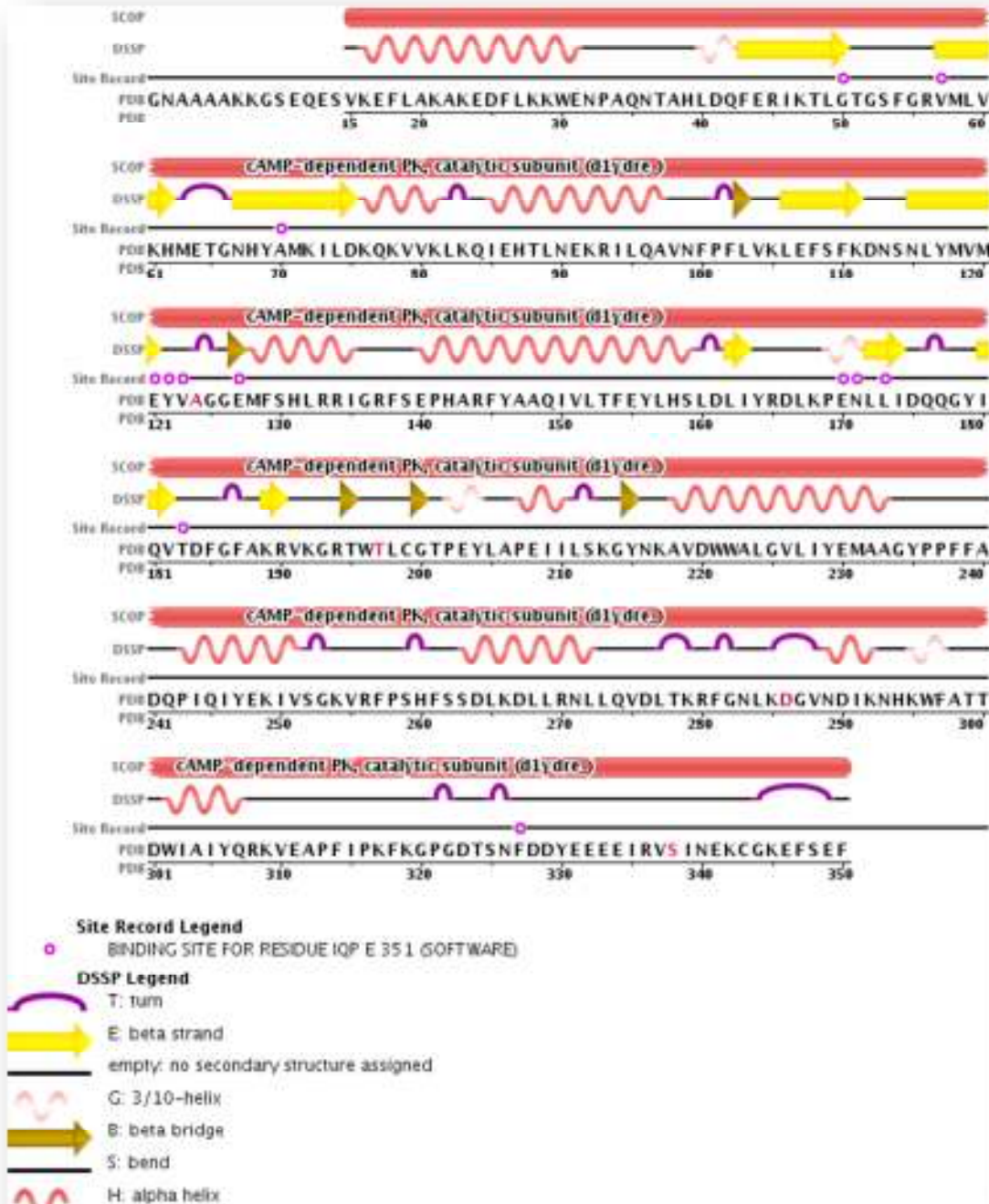


Figure11 : Structure secondaire de 1YDR

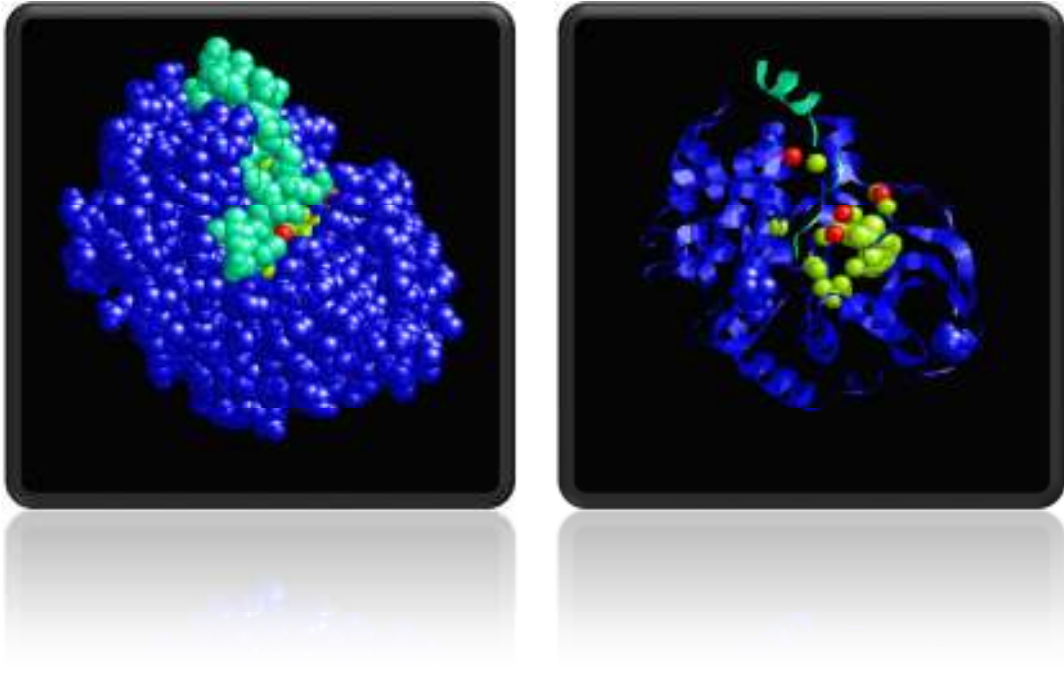


Figure 12 : Structure de la protéine kinase A (1YDR)

Par le souci d'améliorer la vitesse de simulations, nous n'avons tenu compte, dans notre modèle que des derniers 273 AA de la PKA. En outre toutes les molécules d'eau ont été éliminées de la structure. Le ligand utilisé lors de la co-cristallisation, a été extrait du site actif de la PKA pour obtenir ainsi l'apo-protéine, qui après minimisation est prête à recevoir les tripeptides à étudier.

3. Construction des peptides

Les peptides ont été construits en utilisant le logiciel Hyperchem version 8 [5]. Tout d'abord, les molécules sont construites à l'aide du module Database: c'est la partie la plus importante du système, elle regroupe des données de base de plus de 10.000 molécules tel que : les acides aminés, les acides nucléiques et les sucres. Des modifications se font via le mode Build et Add. Hydrogens (ajouter des hydrogènes), Set Charge (pour fixer une charge), Constrain (ajout des contraintes) avec le module Build.

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 forces à appliquer.

L'objectif d'un champ de forces 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 forces, son application au système d'une part, et sa bonne qualité prédictive des propriétés mesurées lors des simulations, d'autre part, influencent directement la validité des résultats. De là, vient la nécessité d'utiliser le champ de forces le plus performant possible et qui soit adapté au type de molécules à étudier.

4. Minimisation de l'énergie

Les méthodes de minimisation ont deux rôles principaux. Le premier est celui d'auxiliaire d'autres méthodes, par exemple pour effectuer une analyse et avant d'entreprendre le calcul d'une trajectoire de dynamique moléculaire, il peut s'avérer nécessaire d'optimiser une structure, c'est à dire d'éliminer toutes les interactions stériques fortement défavorables [6]. Le second rôle est celui de l'exploration dans l'espace conformationnel. Cela permet d'avoir une estimation de la barrière de potentiel entre deux conformations. L'objectif des méthodes de minimisation est extrêmement simple : étant donné une fonction d'une ou plusieurs variables indépendantes, il s'agit de trouver la variable pour laquelle la fonction est la plus faible.

La minimisation de l'énergie d'une molécule est faite en deux étapes. Tout d'abord, une équation décrivant l'énergie du système est exprimée en fonction des coordonnées des atomes. Puis cette équation sera définie et évaluée pour une conformation donnée. Ensuite, la conformation est ajustée à la valeur minimale de cette équation. Un minimum peut être obtenu après un ajustement, ou par contre il peut demander plusieurs milliers d'itérations, dépendant de la nature de l'algorithme, du type d'équation et de la taille de la molécule.

En général, l'algorithme de la plus grande pente ou « steepest descent » est utilisé au début, pendant 100 à 200 pas. Après, l'algorithme du gradient conjugué ou « conjugate gradient » peut être appliqué pour compléter la minimisation jusqu'à la convergence. La convergence dans l'algorithme « steepest descent » est lente. L'algorithme du gradient conjugué est la méthode de choix pour les grands systèmes moléculaires, puisqu'il présente l'avantage de posséder à une convergence rapide.

Au niveau de notre étude, la minimisation a été effectuée sur la protéine kinase A seule, les peptides seuls et les complexes PKA-tripeptides.

Les techniques de minimisation présentent cependant quelques limites qui sont dues à une convergence trop rapide du système. En voici quelques unes:

- Atteinte rapide du minimum le plus proche.
- Pas de réarrangements très importants.
- Exploration conformationnelle faible.
- Barrières d'énergie non franchies.
- L'énergie chute très rapidement (10^7 à -1200 cal/mol).

5. 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 et le développement d'algorithmes performants, rendent possible l'étude d'assemblages supramoléculaires de complexité croissante par les méthodes de chimie théorique [7]. Le principe de la dynamique moléculaire est de fournir de l'énergie cinétique au système pour franchir des barrières énergétiques.

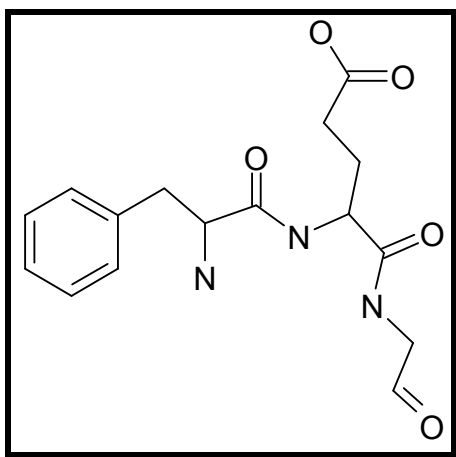
La dynamique moléculaire a été effectuée sur la PKA seule, les peptides seuls et les complexes PKA-peptide formés.

6. Docking des peptides : construction des complexes

L'étape suivante, après la construction des peptides, est le positionnement de ces molécules dans le site actif de la protéine kinase A. Pour ce faire, nous avons utilisé le module Docking à l'aide logiciel Hex6.3. 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.

7. Résultats et discussion

7.1. Optimisation des peptides



D-Phe-D-Glu-L-Gly (feG)

La construction et l'optimisation (par mécanique moléculaire) de toutes les structures peptidiques ont été exécutées par le logiciel hyperchem 8.

	Séquences peptidiques
Peptides biologiquement actifs	feG, FEG, WEG, yeG
Peptides biologiquement inactif	wdG, weG

F: Phe (L- phemylalanine) **f:** (D- phenylalanine)
E: Glu (L- acide glutamique) **e:** (D- acide glutamique)
G: Gly (L-Glycine) **g:** (D- Glycine)
W: Trp (L- Tryptophane) **w:** (D- Tryptophane)
D: Asp (L- acide aspartique) **d:** (D- acide aspartique)
Y: Tyr (L- Tyrosine) **y:** (D- Tyrosine)

E (Kcal/mol)	E(liaison)	E(valence)	E(dièdre)	E(vdW)	E(totale)
Tripeptides actifs					
feG	2.027	3.142	-0.920	10.673	14.949
FEG	2.035	3.742	-0.830	11.959	16.993
WEG	2.276	21.439	-0.512	9.319	32.661
yeG	2.031	3.060	-1.324	11.892	15.696
Tripeptides inactifs					
wdG	14.414	32.906	2.848	13.273	63.433
weG	15.190	34.352	13.833	6.302	70.217

Tableau1 : Résultats de minimisation moléculaire des tripeptides

7.2. Optimisation de la protéine kinase A

L'optimisation de la géométrie de la protéine kinase A a été effectuée à l'aide du champ de forces Amber 99, la chaîne principale (backbone) a été maintenue rigide, tandis que les chaînes latérales demeurent souples. Cette approximation permet aux chaînes latérales des protéines de trouver plus facilement la position dans laquelle les interactions sont les plus favorables.

AMBER 99 a été développé par Peter Albert Kollman et *col* [8]. L'équation utilisée peut être simplifiée par l'équation suivante :

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

Avec :

k_r : constante de force de la liaison.

r : longueur de liaison.

r_{eq} : longueur de la liaison à l'équilibre.

k_{θ} : constante de force de déformation d'angle de valence.

θ : angle de valence.

θ_{eq} : angle de valence à l'équilibre.

V_n : constante de force.

n : périodicité de $A_{i,n}$.

γ : angle de phase.

R_{ij} : La distance entre les deux atomes.

A_{ij} et B_{ij} constantes de Van Der Waals.

q_i, q_j : sont les charges portées par les atomes,

ϵ : la constante diélectrique du milieu.

Les énergies de liaison, de flexion et de torsion sont représentées par une simple expression harmonique diagonale. L'énergie d'interaction de Van Der Waals est représentée par un potentiel 6-12, et l'énergie électrostatique est modélisée par des interactions Coulombiennes de charges centrées sur les atomes.

7.3. Dynamique moléculaire de la protéine kinase

En général, lors de la modélisation d'un système moléculaire biologique de type protéine-ligand, les molécules d'eau rencontrées issues du solvant dans le milieu naturel ne sont pas prises en compte, car on considère en premier lieu qu'elles n'ont aucun rôle significatif sur la liaison ligand-récepteur. Si on détaille cette approximation, on constate que l'inclusion d'une molécule d'eau (environ 1.4 Å de rayon) dans le complexe moléculaire modélisé dans le vide provoquerait deux effets antagonistes :

- 1) Un effet favorable sur l'enthalpie de liaison, dû à la formation très probable de liaisons hydrogène supplémentaires (l'eau étant une très petite molécule pouvant agir aussi bien comme donneur que comme récepteur, et imposant peu de contraintes stériques).
- 2) Un effet défavorable au niveau entropique, causé par l'extraction du solvant, ce qui réduit le nombre de degrés de liberté de la molécule d'eau.

On estime en général que ces deux effets ont chacun un impact peu important, et que d'autre part ils se compensent. Leur somme est donc considérée négligeable par rapport à l'énergie d'interaction totale entre le ligand et le récepteur biologique, et la suppression des molécules d'eau est généralement opérée comme approximation acceptable lors de la modélisation.

Des contraintes au niveau de la chaîne principale (backbone) ont été appliquées à la protéine étudiée. Ces contraintes maintiennent le squelette polypeptidique rigide, permettant aux distances de coordination de rester stables.

La dynamique commence par une initialisation du système : à $t=0$ on a $r(t)=0$, c'est à dire la structure initiale, précédemment minimisée. Puis on a une montée en température jusqu'à 300 K avec un time step (pas d'intégration) 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 500 picosecondes. La figure 1 schématise la variation potentielle en fonction du temps.

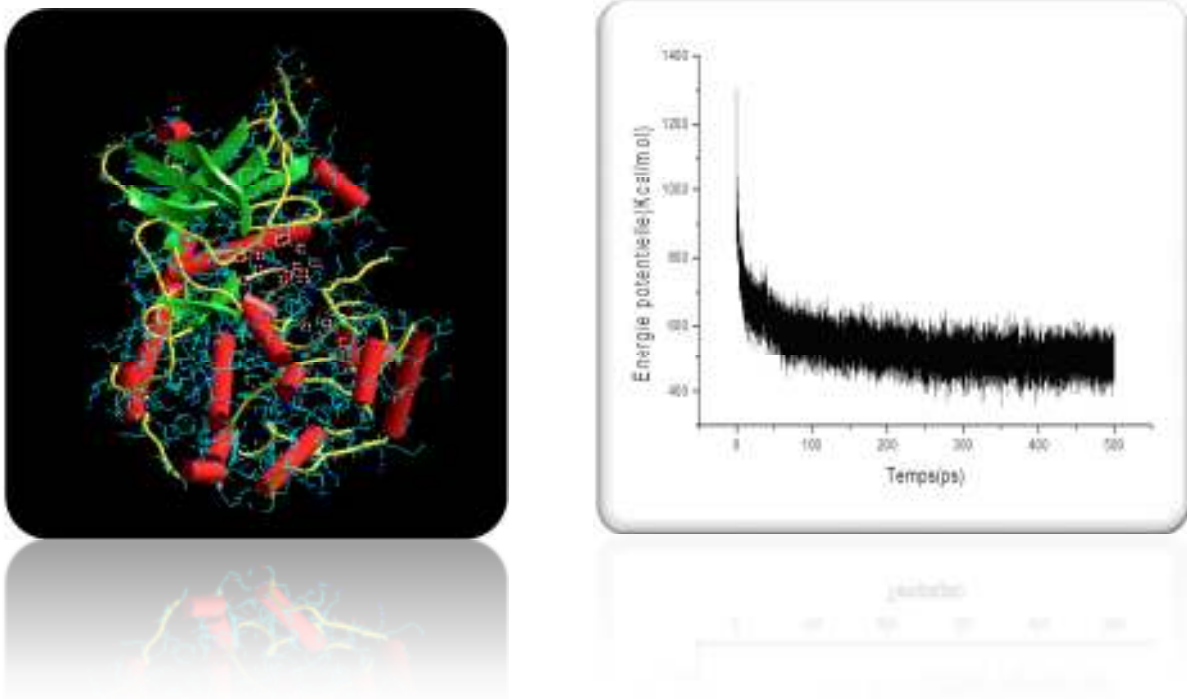


Figure 13 : l'énergie potentielle de la protéine kinase seule en fonction du temps.

Energie de VDW de (1ydr) = -181.243 Kcal/mol

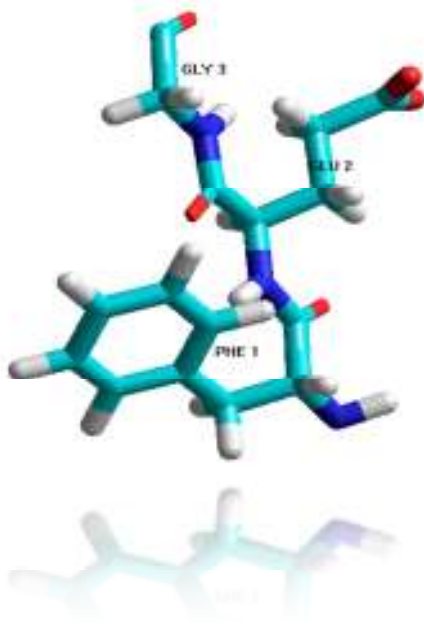
Energie Potentielle (1ydr) = 243.546 Kcal/mol

On a utilisé le programme ORIGIN5.0 [9] pour dessiner les cartes d'énergie. *Origin* est un logiciel de traitement et d'analyse de données scientifiques pour environnement Microsoft Windows développé par OriginLab. Il permet notamment de tracer des graphes 2D et 3D et possède des fonctions d'analyse de pics. Il est aussi capable d'importer des fichiers de divers formats tels que Excel, Mathematica ou SigmaPlot, et d'exporter les graphes sous format JPEG, GIF, Tiff etc.

7.4. Dynamique des peptides

- Tripeptides biologiquement actifs

1. Tripeptide feG



Tripeptide feG

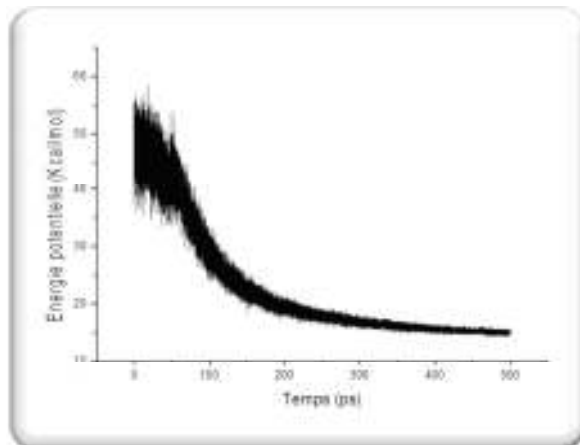


Figure 14 : l'énergie potentielle de feG en fonction du temps

2. Tripeptide FEG :

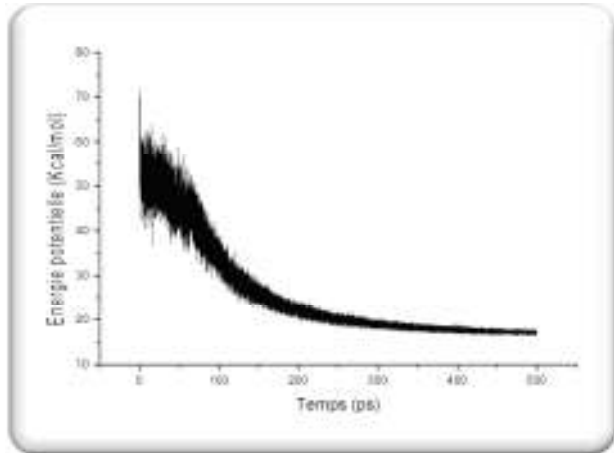
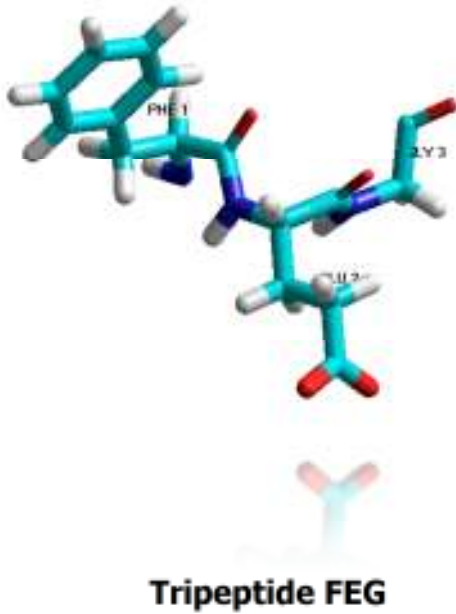


Figure 15 : l'énergie potentielle de FEG en fonction du temps

3. Tripeptide WEG :

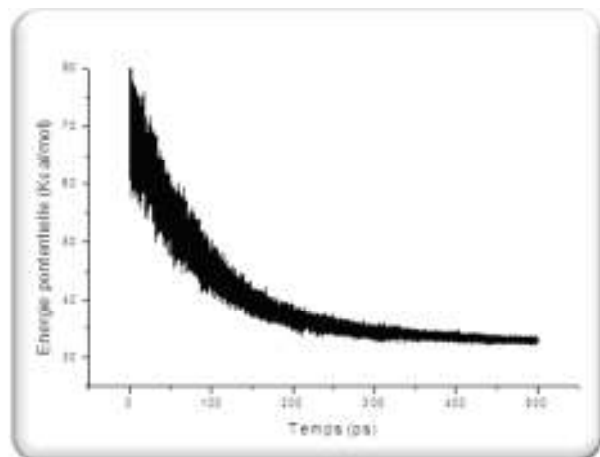
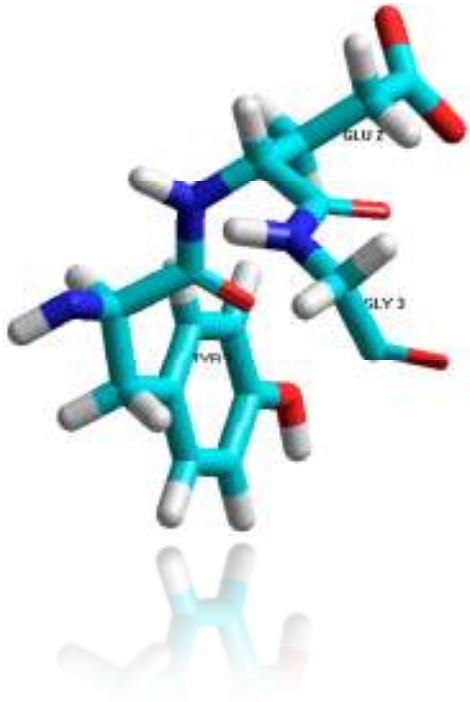


Figure 16 : l'énergie potentielle de WEG en fonction du temps.

4. Tripeptide yeG :



Tripeptide yeG

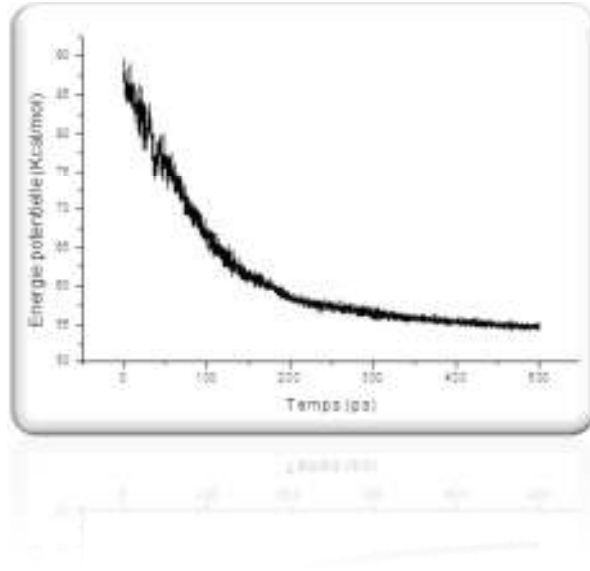


Figure 17 : l'énergie potentielle de yeG en fonction du temps.

- **Tripeptides biologiquement inactifs**

1. Tripeptide wdG



Tripeptide wdG

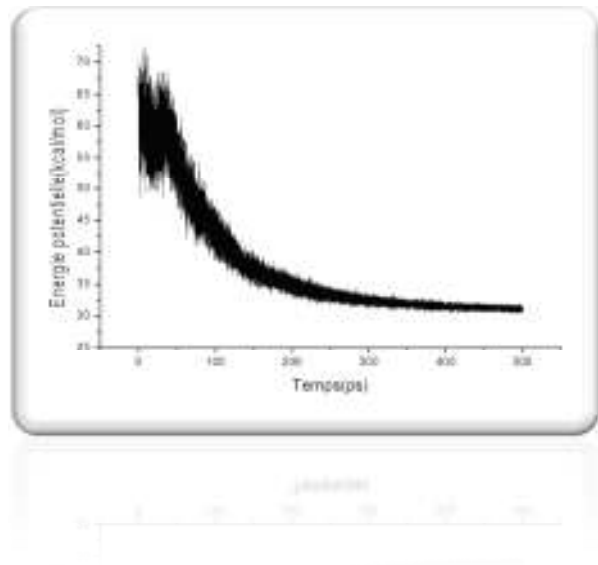


Figure 18: l'énergie potentielle de wdG en fonction du temps.

2. Tripeptide weG



Tripeptide weG

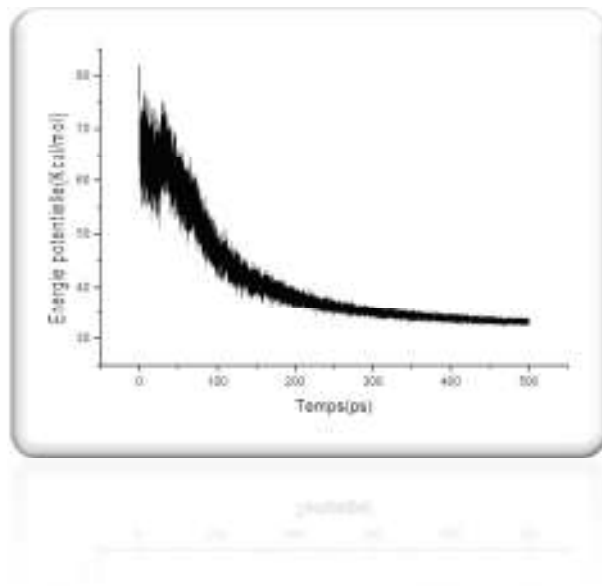


Figure 19 : l'énergie potentielle de weG en fonction du temps.

7.5. Formation des complexes PKA- tripeptide

Nous avons, dans un premier temps, modélisé le complexe protéine-tripeptide à l'aide du programme Hex 6.3 [10]. Les résultats obtenus ont été classés grâce à une fonction de score tenant compte de l'énergie potentielle d'interaction de l'assemblage.

La contribution principale dans le domaine de l'amarrage de protéines est d'explorer et de démontrer l'utilité de la méthode utilisant les coordonnées sphériques polaires ; nommé: *polaire sphérique de Fourier* (ou "spherical polar Fourier," SPF) [11]. L'idée fondamentale est de représenter la forme et les propriétés électrostatiques des protéines (ou d'autres biomolécules telles que le ADN et le ARN) en tant qu'expansions de Fourier.

Ces idées ont été implémentées dans le programme d'amarrage Hex 6.3 [10]. Pendant un calcul d'amarrage, le programme Hex peut évaluer plusieurs millions d'orientations par seconde sur un ordinateur individuel ordinaire (ou "personal computer," PC). Avec plus de 12.000 téléchargements, Hex est utilisé à travers le monde aussi bien dans le milieu universitaire qu'en industrie et a été cité dans plus de 200 publications scientifiques.

1. Complexe PKA-feG

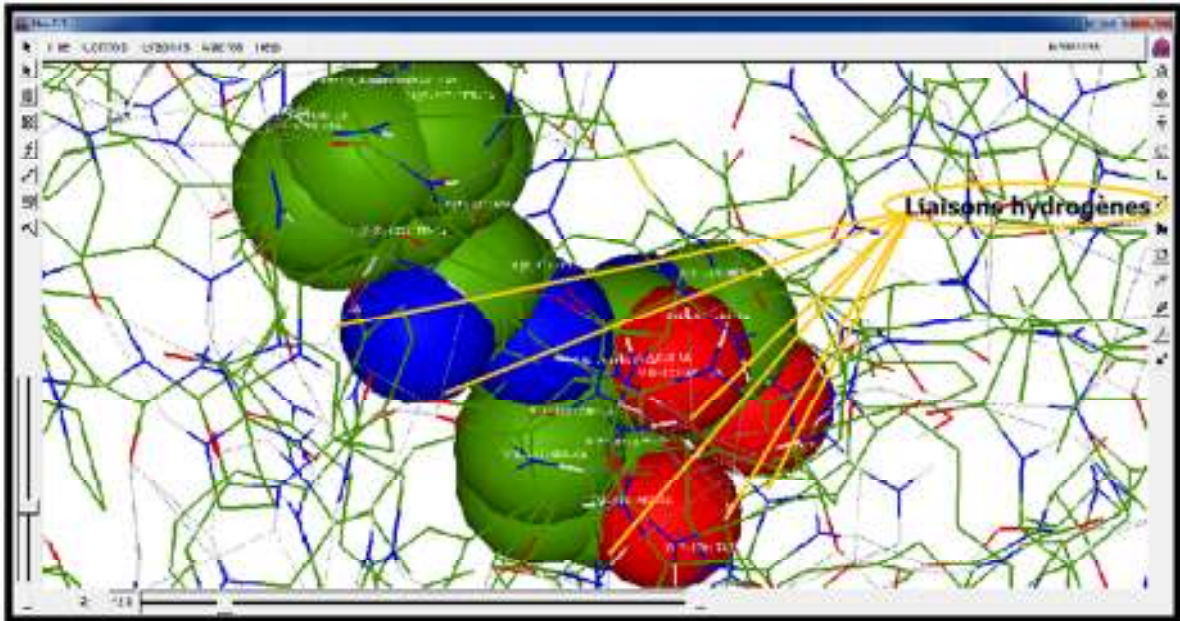
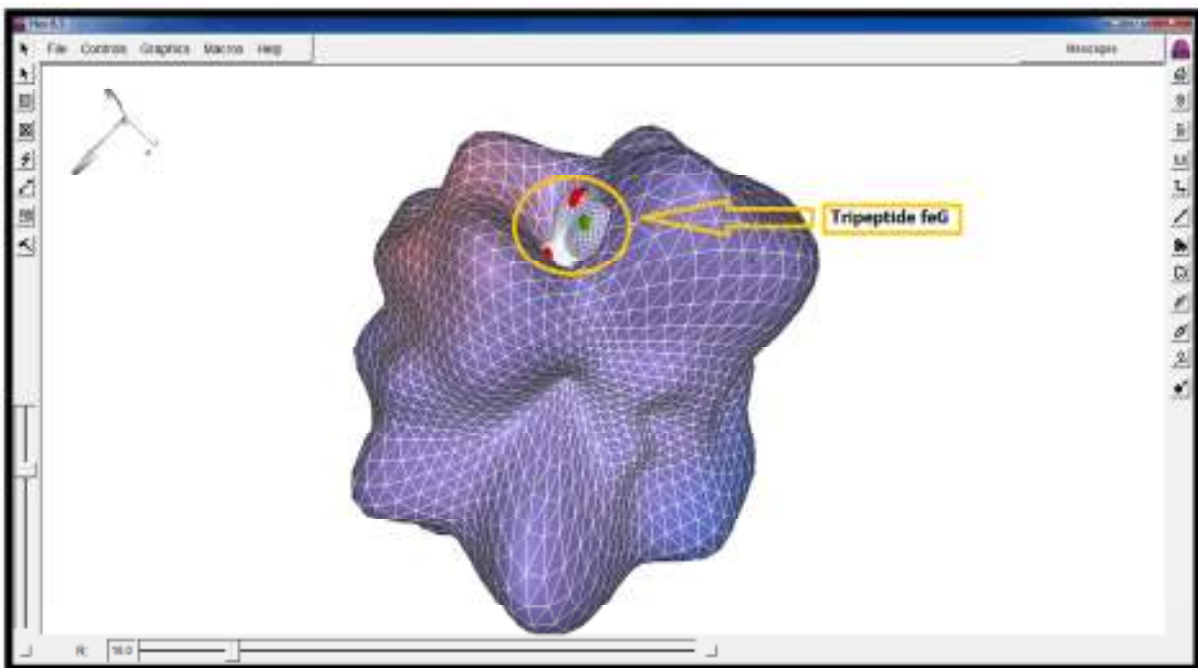


Figure 20 : les liaisons hydrogènes entre PKA et feG



2. Complexe PKA-FEG

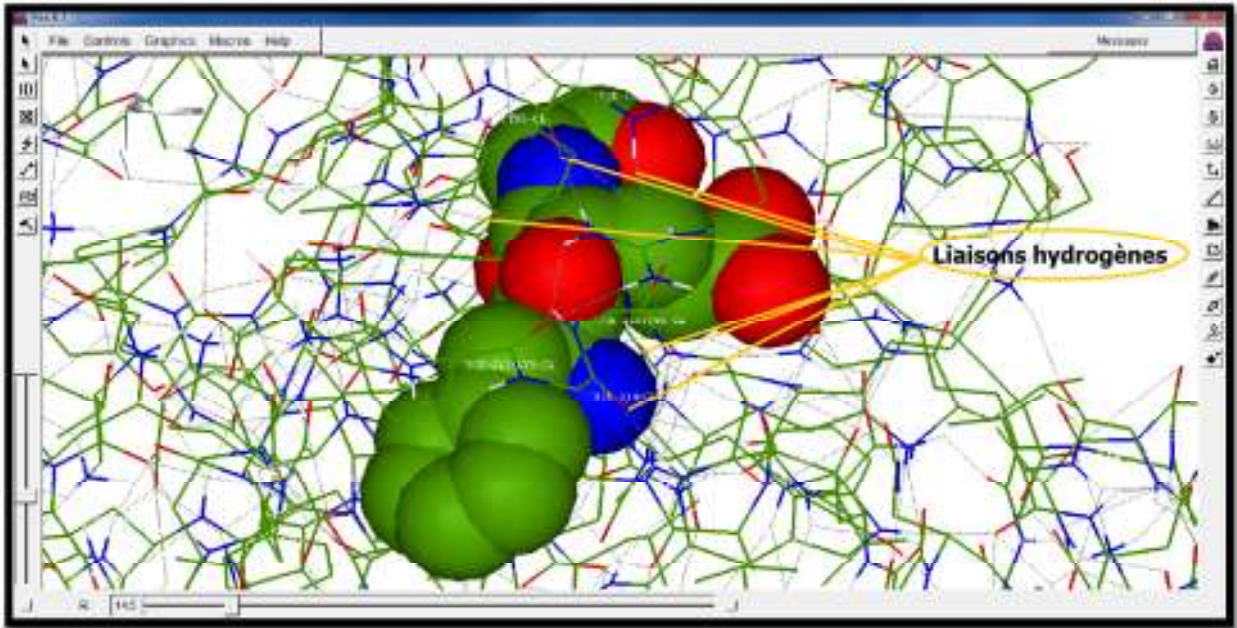
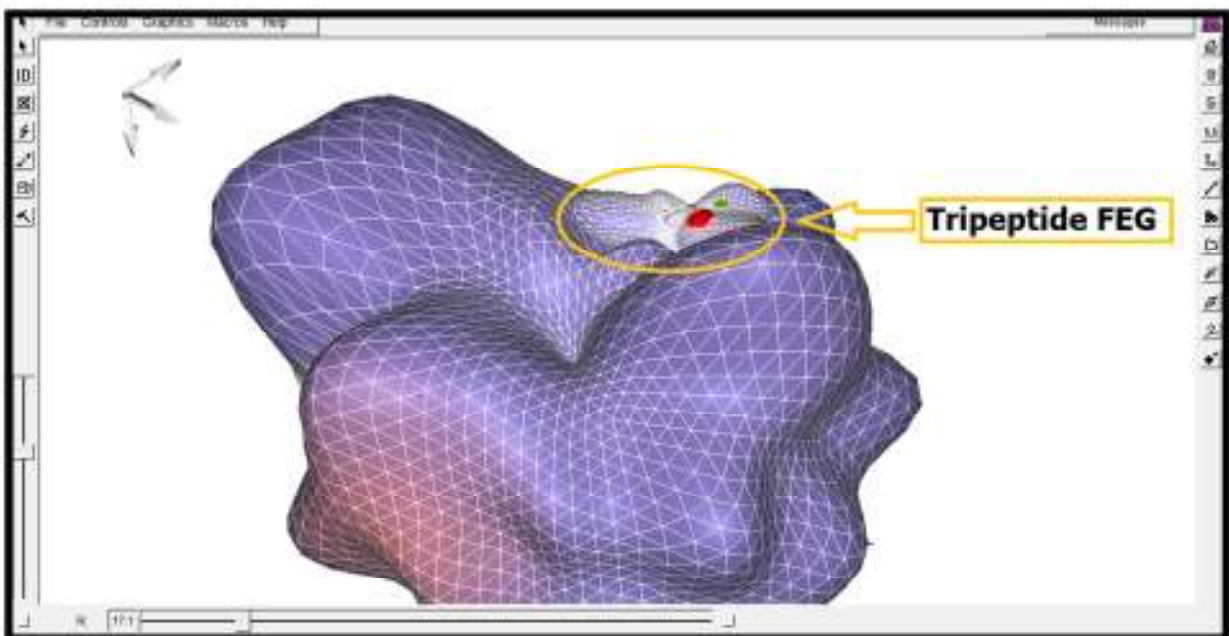


Figure 21 : les liaisons hydrogènes entre PKA et FEG



3. Complexe PKA-WEG

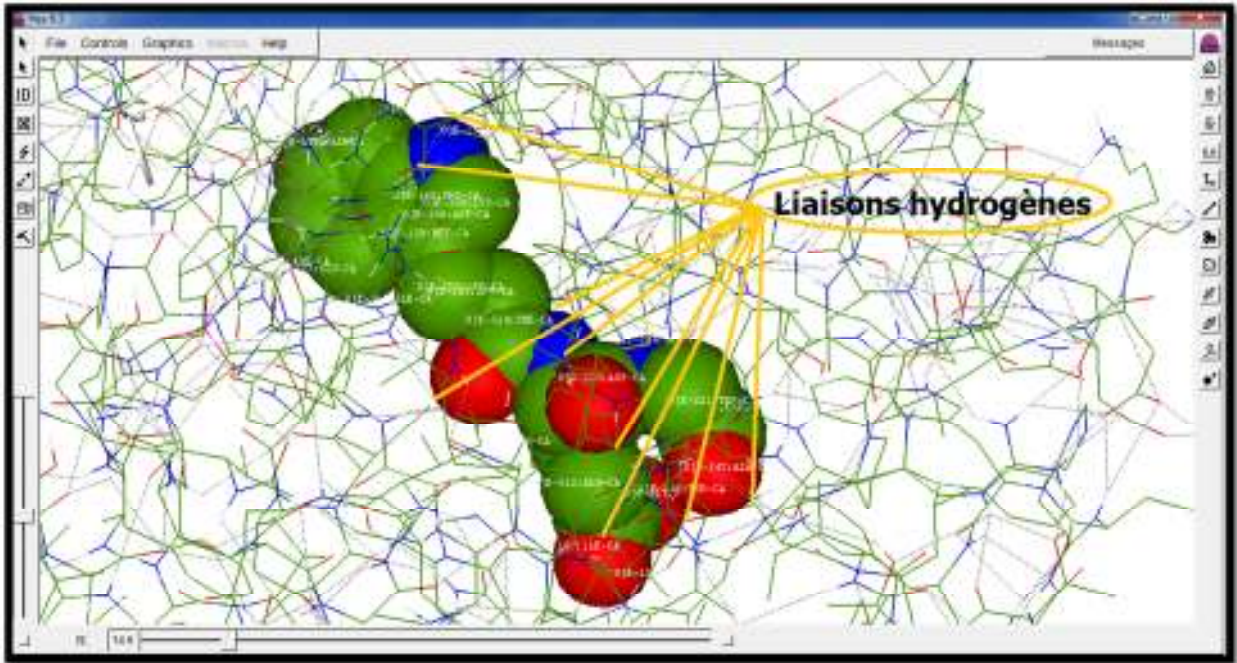
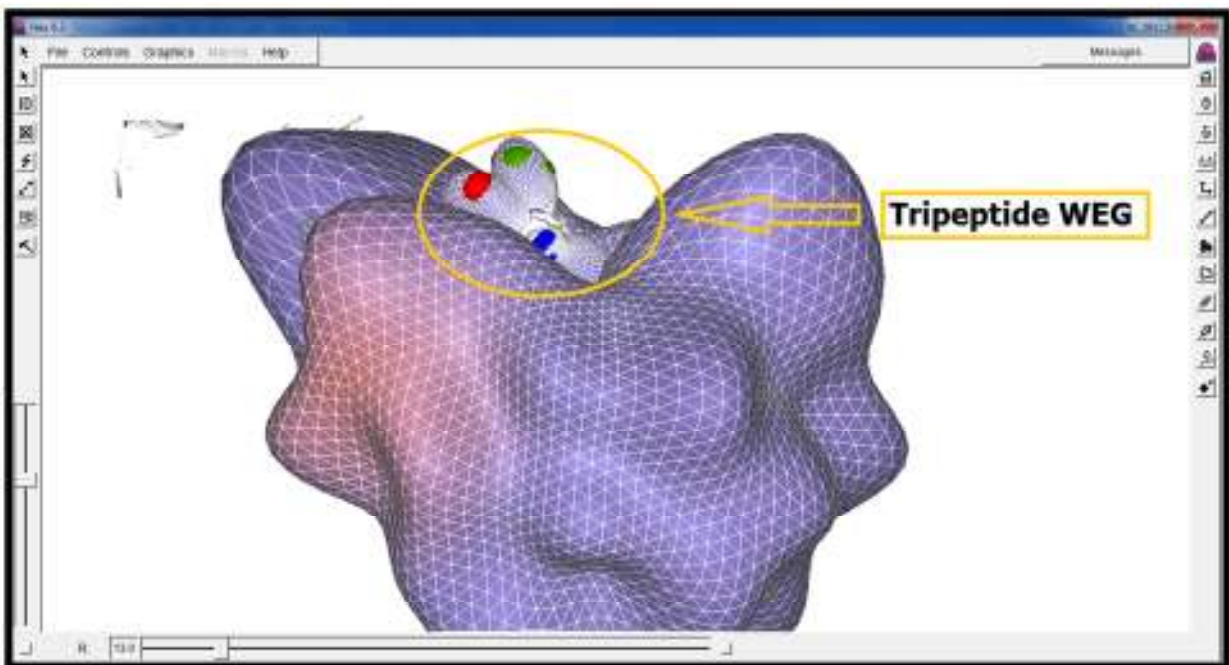


Figure 22: les liaisons hydrogènes entre PKA et WEG



4. Complexe PKA-yeG

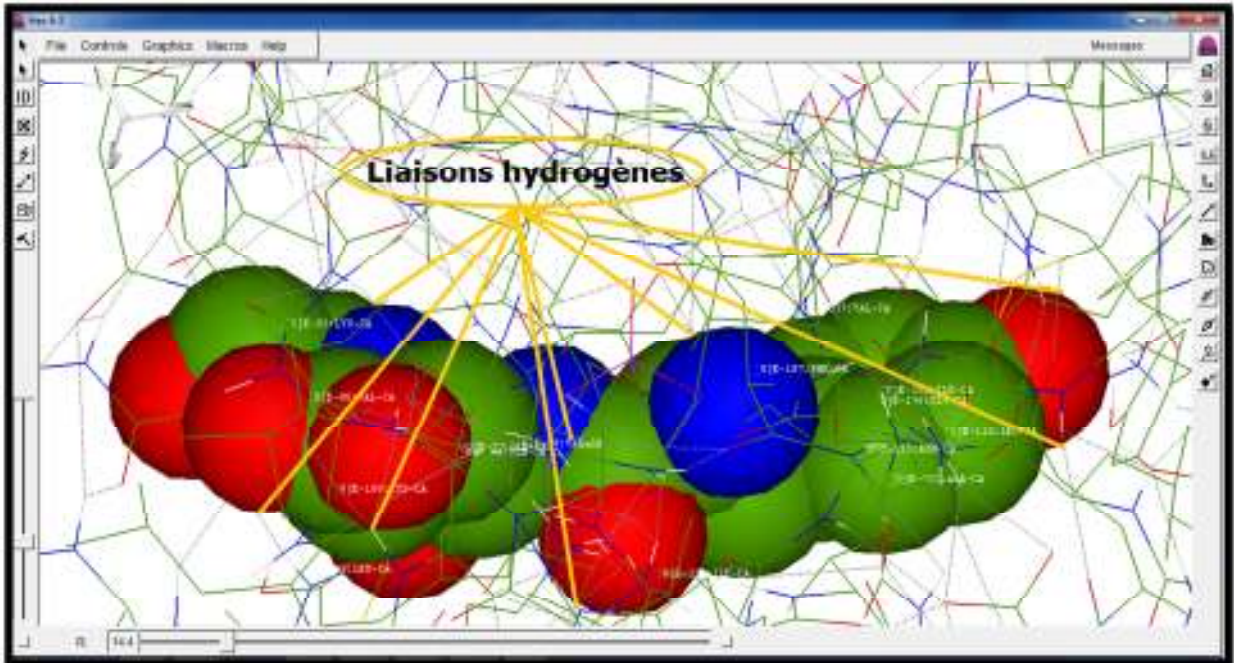
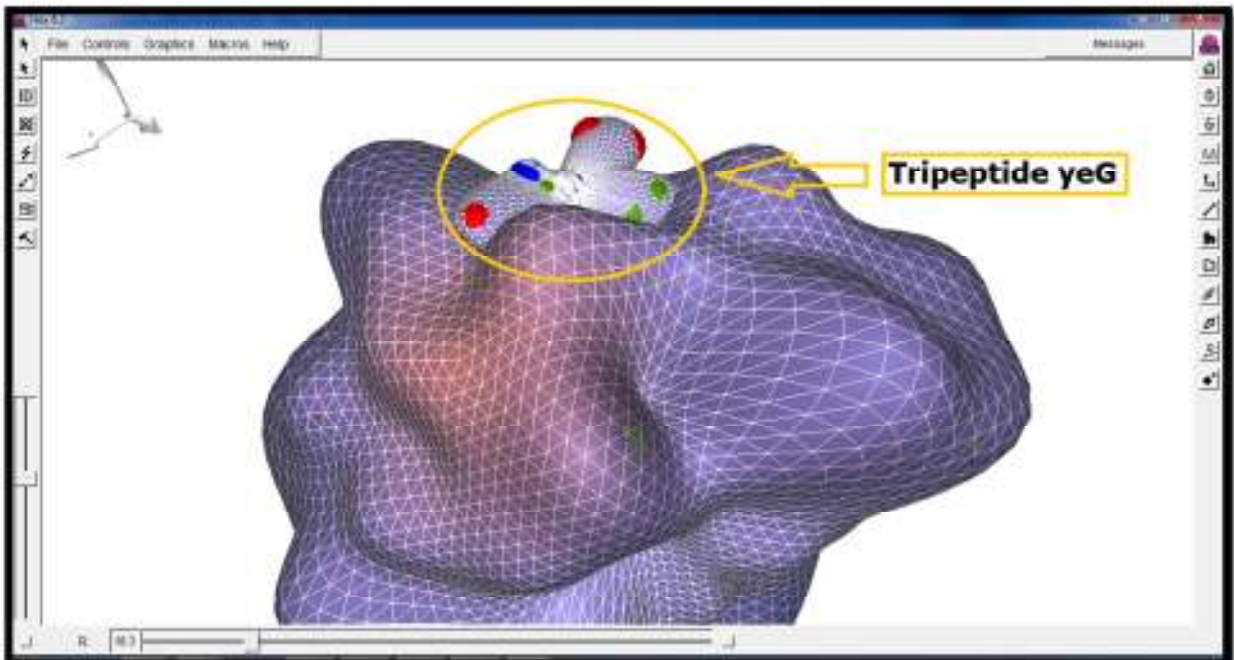


Figure 23: les liaisons hydrogènes entre PKA et yeG



5. Complexe PKA-wdG

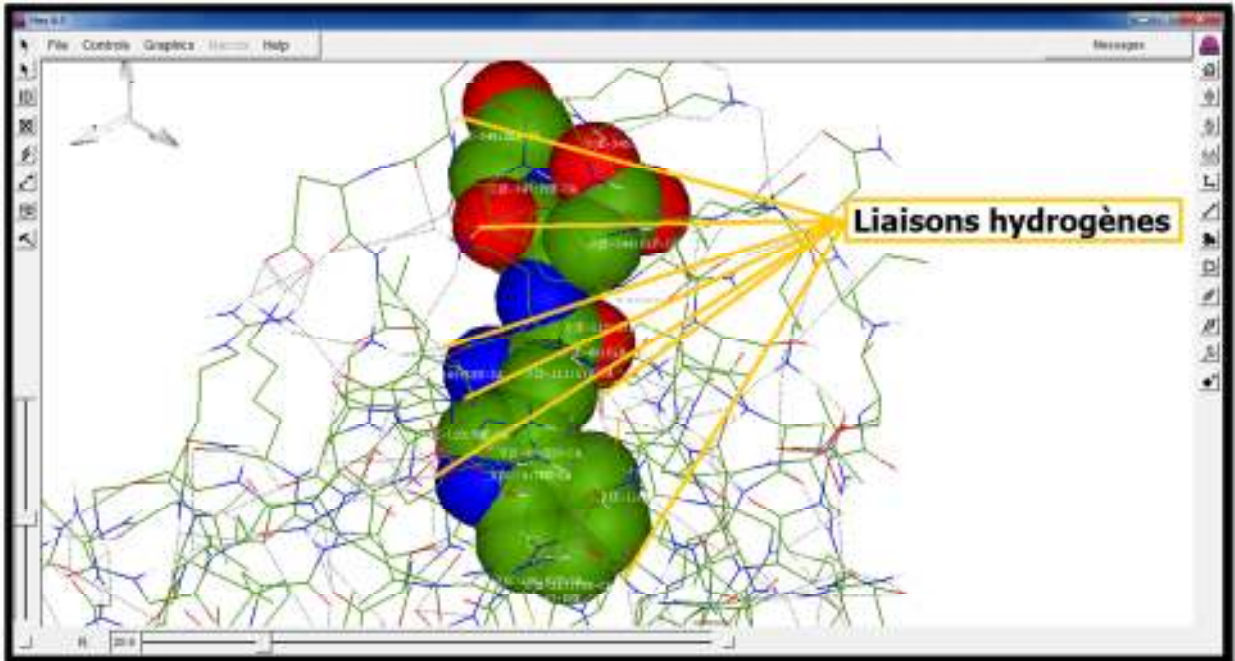
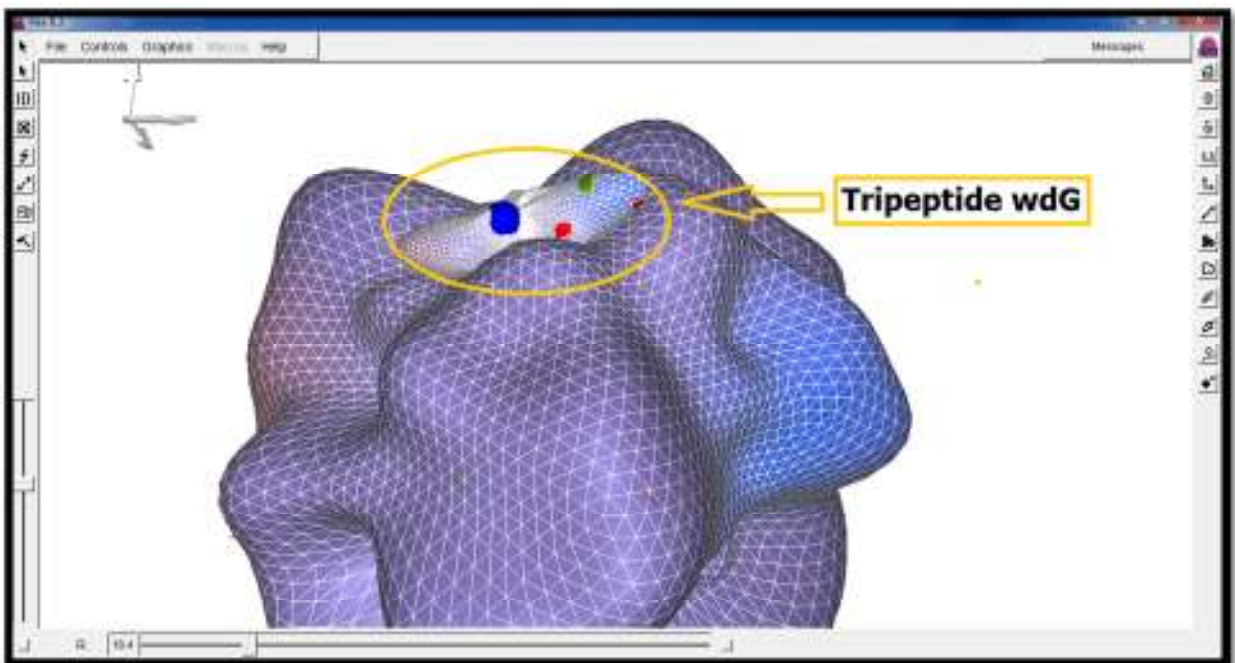


Figure 24: les liaisons hydrogènes entre PKA et wdG



6. Complexe PKA-weG

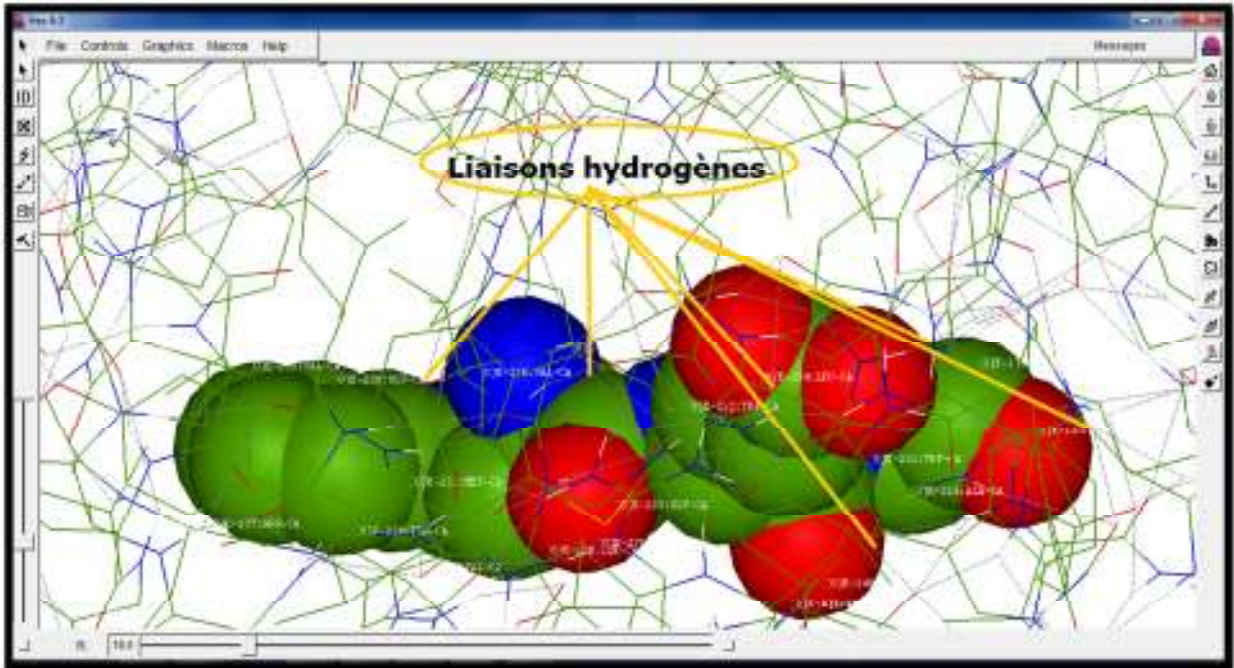
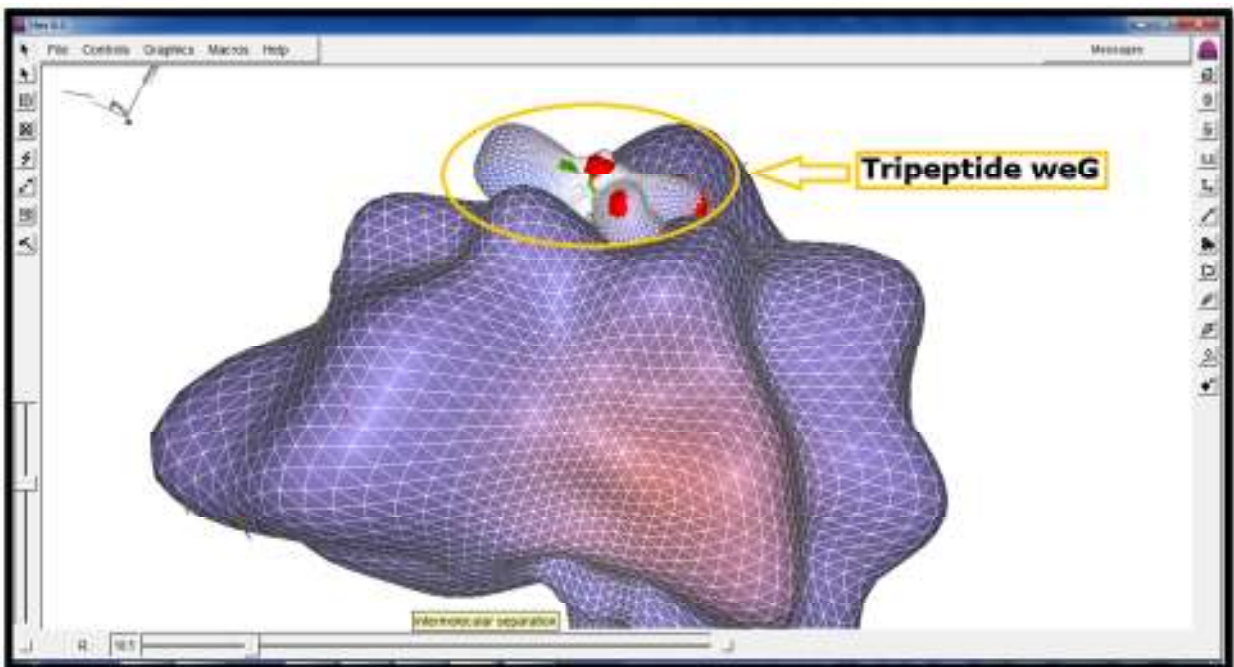


Figure 25: les liaisons hydrogènes entre PKA et weG



D'après les interactions des liaisons hydrogènes qui s'établissent entre les acides aminés de la protéine kinase A avec les différents tripeptides, on peut prédire les acides aminés qui forment le site actif ; ce qui aboutit aux résidus suivants : **Lys 76, Lys 111, Leu 116, Met 118, Phe 347, Lys 345, Phe 350.**

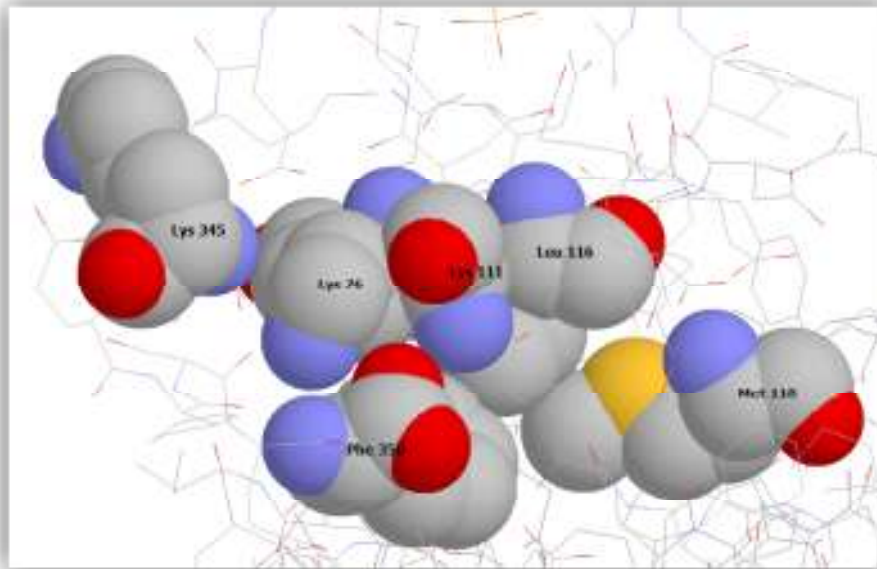


Figure 26: acides aminés du site actif

7.6. Energies 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 protéine kinase A-tripeptides est gouvernée, on fait souvent référence à une décomposition de l'énergie potentielle totale de la molécule en plusieurs termes (Energie électrostatique, Energie de VDW,...).

Les énergies d'interaction entre les différents tripeptides étudiés et la protéine kinase A sont obtenues à l'aide de la relation suivante :

$$E_{\text{interaction}} = E_{\text{potentielle (complexe 1YDR-tripeptide)}} - [E_{\text{potentielle (1YDR seule)}} + E_{\text{potentielle (tripeptide)}}]$$

complexe	E totale (KJ/mol)	E électrostatique (KJ/mol)	% d'inhibition
Tripeptides biologiquement actifs			
feG	-253.7	-41.4	47
FEG	-262.5	-35.9	53
WEG	-291.3	-45.2	52
yeG	-269.6	-41.7	50
Tripeptides biologiquement inactifs			
wdG	-261.9	-27.3	25
weG	-270.4	-34.9	25

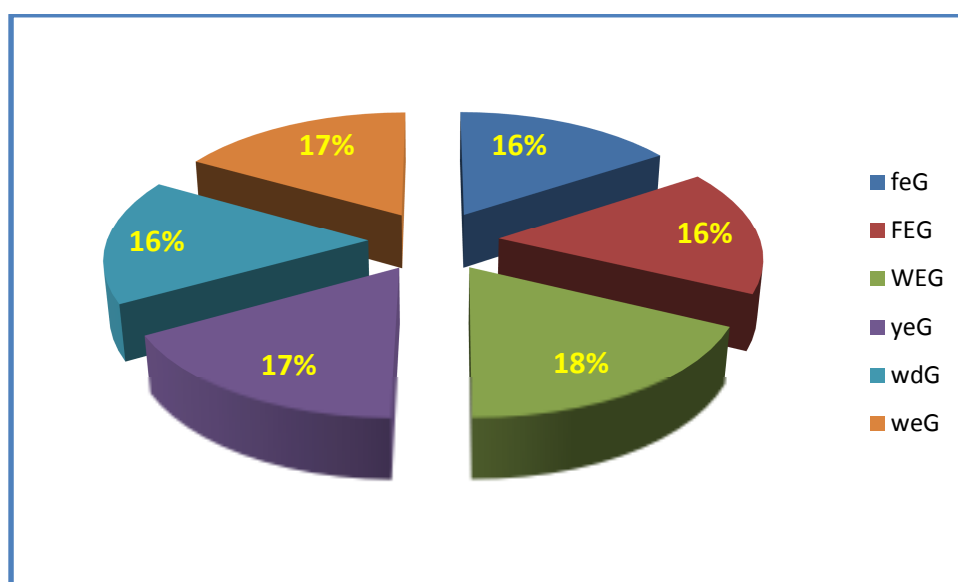


Figure 27: Pourcentage d'énergie d'interaction totale des complexes PKA- tripeptide

- **Classement des énergies d'interactions totales :**

Tripeptides biologiquement actifs: **E(WEG) < E (yeG) < E(FEG) < E(feG)**

Tripeptides biologiquement inactifs : **E(weG) < E(wdG)**

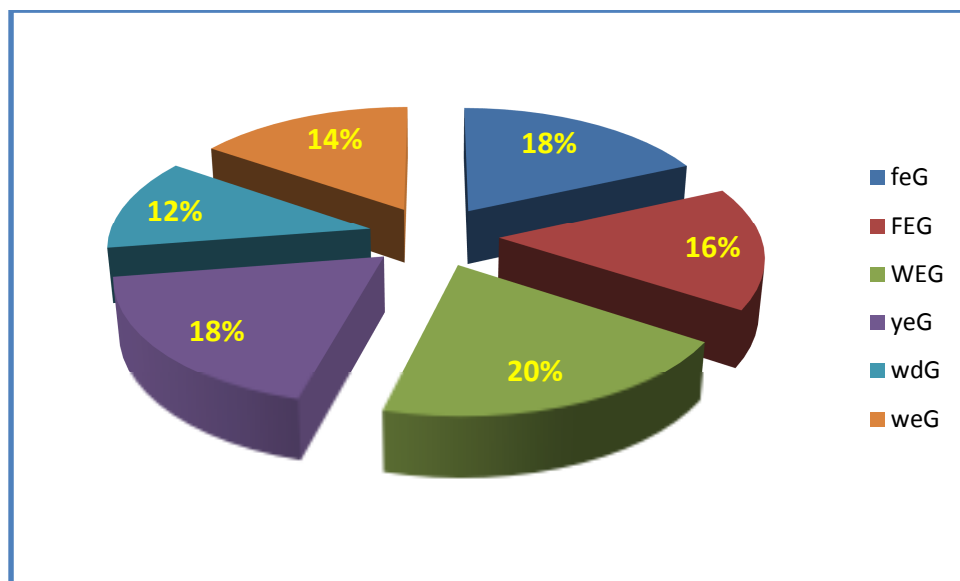


Figure 28: Pourcentage d'énergie d'interaction électrostatique des complexes PKA-tripeptide

- **Classement des énergies d'interactions électrostatiques:**

Tripeptides biologiquement actifs: **$E(\text{WEG}) < E(\text{yeG}) < E(\text{feG}) < E(\text{FEG})$**

Tripeptides biologiquement inactifs : **$E(\text{weG}) < E(\text{wdG})$**

L'analyse des résultats obtenus lors de la décomposition d'énergie nous permet d'énoncer quelques conclusions :

En premier lieu, il apparaît que l'énergie électrostatique pourrait mieux rendre compte de l'interaction entre les différents peptides étudiés et la PKA comparativement à l'énergie totale. Elle est à la base de la formation des liaisons hydrogènes dont l'importance est cruciale au fonctionnement des macromolécules.

En outre, le composé qui possède le plus bas niveau d'énergie d'interaction électrostatique est celui qui présente probablement la meilleure activité.

Pour le tripeptide WEG qui possède la plus basse énergie d'interaction totale (-291.3 KJ/mol) et la plus faible énergie d'interaction électrostatique (-45.2 KJ/mol), la présence du résidu L- tryptophane (acide aminé aromatique, apolaire et

hydrophobe) en position N- terminal favorise l'interaction avec le site actif de la protéine kinase A et engendre une complémentarité relativement importante.

Les tripeptides feG, FEG et yeG possédants également une énergie d'interaction électrostatique relativement faible (présence du résidu phenylalanine et tyrosine en N- terminal) présentent une activité inhibitrice pour la PKA.

Par contre, pour wdG et weG et avec une énergie électrostatique respectivement égale à - 27.3 KJ/mol et -34.9 KJ/mol, on remarque que la présence du résidu D- tryptophane en position N- terminal défavorise l'inhibition de la PKA.

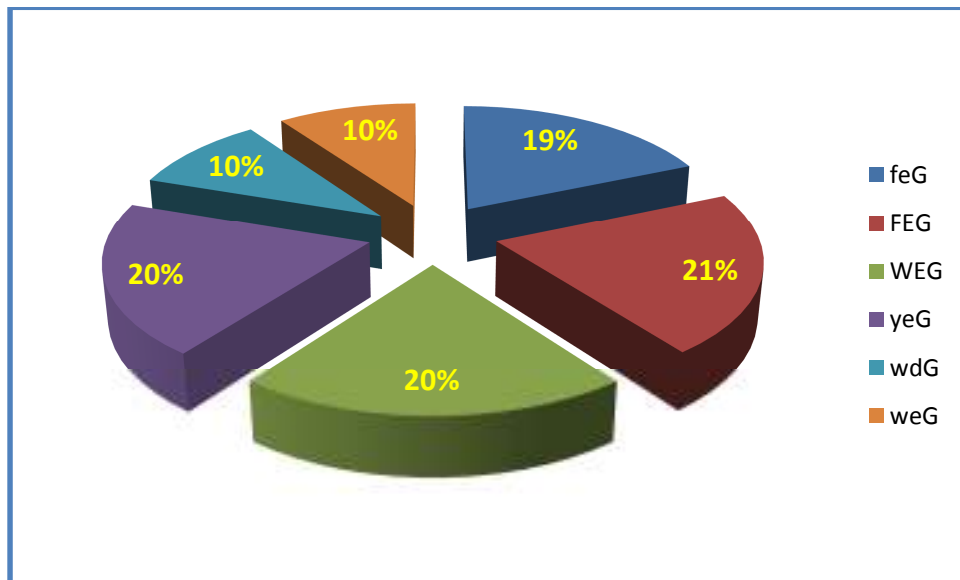


Figure 29: Représentation des valeurs expérimentales / pourcentage d'inhibition de tripeptides

- **Classement des pourcentages d'inhibition:**

Tripeptides biologiquement actifs: $\%(FEG) > \%(WEG) > \%(yeG) \quad \%(feG)$

Tripeptides biologiquement inactifs : $\%(weG) = \%(wdG)$

Comparant les valeurs expérimentales des pourcentages d'inhibition des différents tripeptides, on note que les tripeptides WEG, yeG, FEG et feG présentent une bonne inhibition de la PKA avec des pourcentages égaux à 52, 50, 53 et 47 respectivement.

7.7. Dimension de la cavité et les tripeptides étudiés

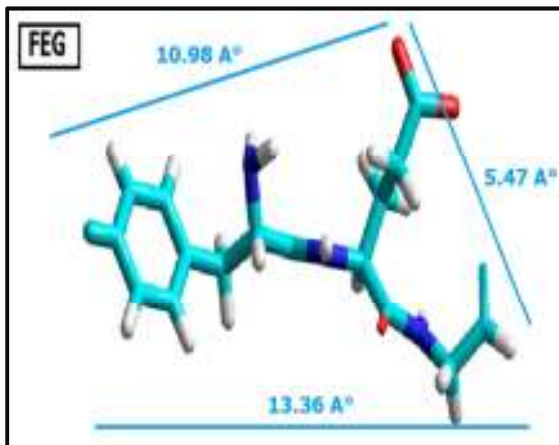


Figure 30: dimension du tripeptide FEG

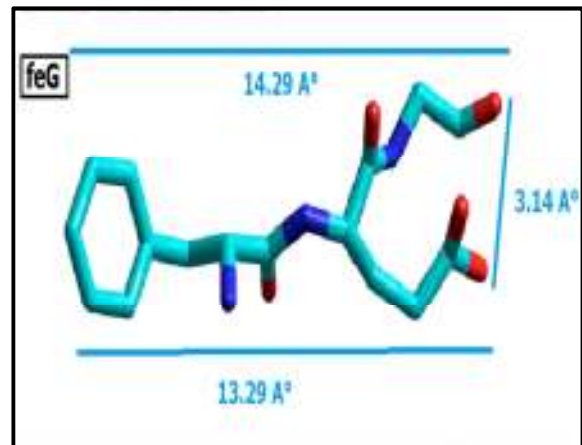


Figure 31: dimension du tripeptide feG

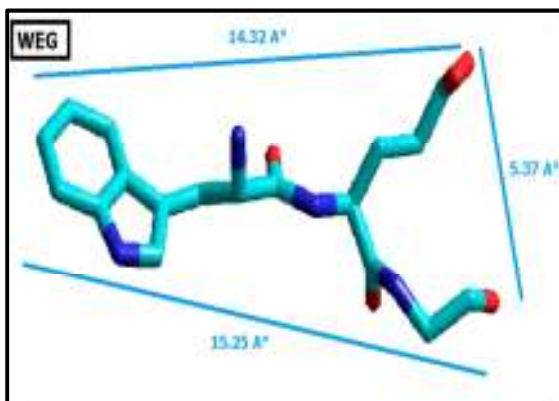


Figure 32: dimension du tripeptide WEG

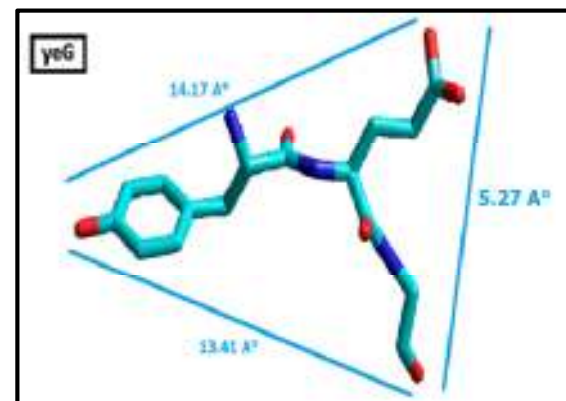


Figure 33: dimension du tripeptide yeG

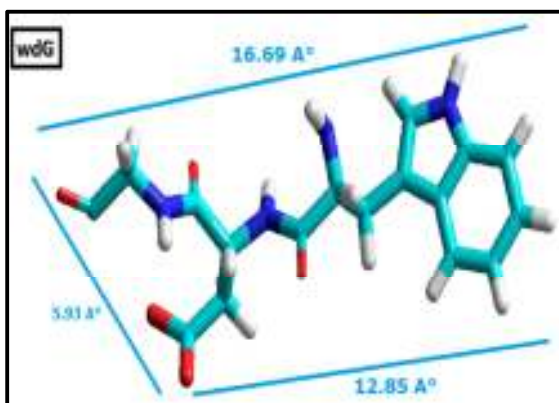


Figure 34: dimension du tripeptide wdG

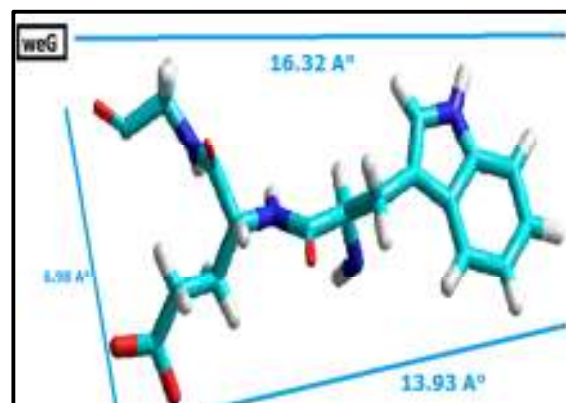


Figure 35: dimension du tripeptide weG

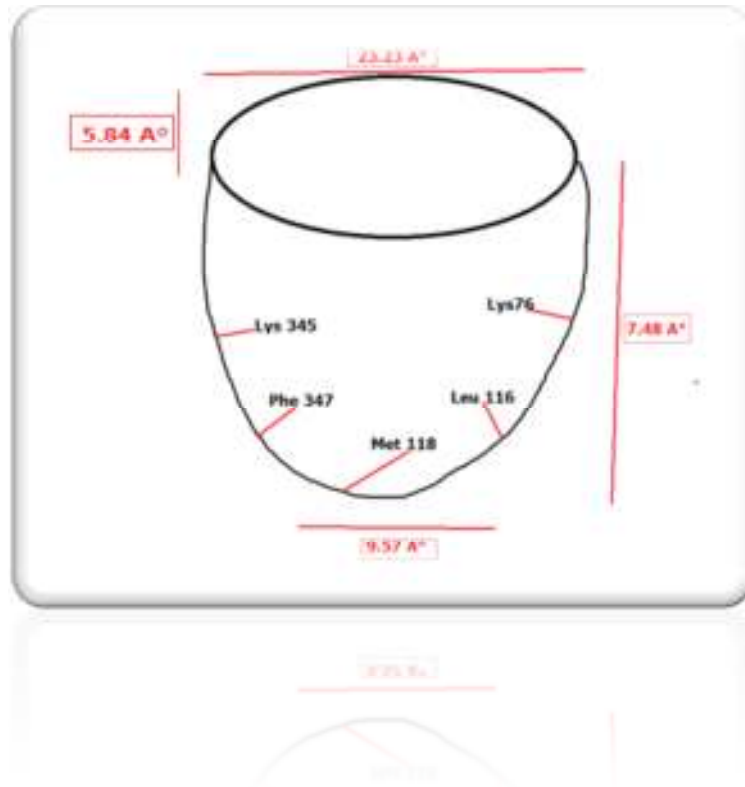


Figure 36: Représentation de la cavité de site actif

On ne peut discuter de complémentarité que dans l'intervalle de dimensions de la poche du site actif [12,16] possédant dans notre cas une largeur de 5.84 Å et une profondeur de 7.48 Å.

En comparant les dimensions des différents tripeptides biologiquement actifs avec celles de la cavité du site actif, on peut dire que WEG, yeG, FEG et feG présentent une meilleure complémentarité par rapport aux tripeptides biologiquement inactifs.

Pour conclure, en se basant principalement sur l'énergie d'interaction électrostatique et les dimensions de la cavité de la protéine kinase A ; la présence des résidus : L- tryptophane, L/ D phénylalanine et D- tyrosine en position N-terminal dans les tripeptides favorisent leurs interaction avec la protéine kinase A et par conséquent présentent une meilleure inhibition de la protéine.

Références

- [1] H.Dugas, Principes de base en modélisation moléculaire, Aspects théoriques et pratique, Librairie de l'université de Montréal, (1996), Chapitre 3
- [2] SEKI E., OSAJIMA K., MATSUFUJI H., MATSUI T., OSAJIMA Y. Angiotensin I converting enzyme inhibitory activity of the short chain peptides derived from various food proteins. (1996) 43, 839-840.
- [3] TOME D. Peptides fonctionnels. *Danone World Newsletter*, (1998) 17.
- [4] GONZÁLES DE LLANO D., HERRAIZ T., POLO M.C. Peptides. *Handbook of food analysis, vol. 1/* ed. by L.M.L. Nollet. Marcel Dekker Inc. (1996), 229-276.
- [5] Hyperchem (Molecular Modelling System) Hypercube, Inc. USA (2007).
- [6] J. Gibart, J. Garnier, Cours de modélisation moléculaire, seconde édition, DRED-DS2, ATF « Chimie et Physico- Chimie des Grandes Molécules », (1993).
- [7] Christophe Chipot, *Centre National de la Recherche Scientifique*, les méthodes numériques de la dynamique moléculaire, (2003).
- [8] 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, programme Amber 6, University of California, San Francisco. (1999).
- [9] Drug Discovery & Development magazine. (2008) 11, 30-32.
- [10] Ritchie, D. W. et Kemp, G. J. L. Protein docking using spherical polar Fourier correlations. *Proteins : Struct. Func. Genet.* (2000) 39, 178–194.
- [11]. David W. Ritchie, Algorithmes Haute-Performance pour la Reconnaissance de Formes Moléculaires, Laboratoire Lorrain de Recherche en Informatique et ses Applications – UMR 7503, (2011).
- [12] I. Abdelli , S. Sari. & S. Ghalem, Molecular Modelling of Green Tea Catechins Inhibitors of Human Brain BACE1, *IBScientific Journal of Science*, (2008) 3,27-29.
- [13] S. Sari, W. Medjahed, J. Kajima Mulengi and S. Ghalem, Approach of the Interaction Enzyme-Substrate by Molecular Modelling, *Asian Journal of Chemistry.* (2007) 19,2117-2124.

[14] I. Abdelli, S. Sari, C. Ziani-chérif and S. Ghalem, Theoretical Study of the Inhibition of β -Secretase. Asian Journal of Chemistry. (2007) 19, 5407-5416.

[15] Wassila Soufi, Meriem Merad, Faïza Boukli, Saïd Ghalem. The Complementarity Effect for Cdc25 Phosphatase Inhibitors. Advances in Molecular Imaging, (2011) 1, 17-23.

[16] Mohamed Beghalia, Hocine Allali, Saïd Ghalem¹, Aïssa Belouatek, Abdelhamid Sari. Molecular Modeling of Chemicals Products Inhibitors of Growth Struvite Crystal. Advances in Molecular Imaging. (2011) 1, 33-42.

CONCLUSION GENERALE

L'objectif de notre travail consiste en l'étude des interactions protéine kinase A-tripeptides (feG et analogues) par modélisation moléculaire. 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 :

Dans le premier chapitre nous avons exposé quelques généralités sur les interactions protéine-protéine jouant un rôle clé dans l'organisation des systèmes biologiques. Elles permettent la régulation de certains processus biologiques, la transmission des signaux, ou encore la catalyse de diverses réactions biochimiques. Comprendre et prédire la manière dont les protéines interagissent devrait permettre, à terme, d'interférer avec le processus biologique dans lequel elles sont impliquées. C'est un phénomène complexe qui implique plusieurs types d'interactions au niveau moléculaire. L'étude par modélisation moléculaire devrait permettre une meilleure compréhension des interactions protéine-protéine et aider à l'interprétation et la rationalisation des données expérimentales.

Au niveau du deuxième chapitre, nous avons donné un aperçu sur les protéines kinases et les tripeptides. En raison de leur grande diversité d'actions, les protéines kinases doivent être rigoureusement régulées : l'activité aberrante de ces enzymes génère des maladies comme le cancer ou le diabète, mais aussi à des dysfonctionnements nerveux, cardiovasculaire, inflammatoires ou auto-immunitaires.

Si une petite molécule se lie à une protéine kinase, elle peut empêcher la phosphorylation et par conséquent inhiber l'activité de la protéine. La découverte de nouveaux médicaments, c'est-à-dire la découverte de petites molécules fiables, destinées à l'inhibition, nécessite de trouver quelles sont les petites molécules capables de se lier avec la protéine kinase.

Les peptides offrent plusieurs avantages par rapport aux petites molécules qui constituent les médicaments traditionnels. Le premier avantage est que, représentant

souvent la plus petite partie fonctionnelle d'une protéine, ils offrent une efficacité, une sélectivité et une spécificité que les concepteurs de petites molécules ont du mal à atteindre. En général, les peptides biologiquement actifs (ou peptides fonctionnels ou bioactifs) sont de petite taille, ils comportent entre 3 et 10 résidus acides aminés dans leur structure.

Le chapitre III est réservé à la description des différentes méthodes de la modélisation moléculaire retenues tel que la mécanique moléculaire, la dynamique moléculaire) et 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 sur la validité des résultats. D'où la nécessité du choix du champ de force adapté au type de molécules à étudier.

Le chapitre IV est consacré aux résultats obtenus ainsi qu'à la discussion et l'interprétation de ces derniers.

Dans le but de déterminer les interactions protéine-peptides, nous avons effectué des calculs de dynamique moléculaire à partir des structures des complexes pour rechercher la conformation la plus stable correspondante à l'énergie la plus faible qu'adopte le complexe formé. Afin de générer les trajectoires des différents systèmes, le champ de force AMBER a été exécuté implémenté dans le logiciel HyperChem 8.

La discussion engagée est basée sur les énergies d'interactions des complexes obtenues à l'aide du logiciel HEX 6.3. L'exploitation de quelques données expérimentales comme le pourcentage d'inhibition pour les tripeptides retenus pour cette étude (feG, FEG, WEG, yeG, wdG et weG) et en se basant sur l'énergie

d'interaction, il ressort que le tripeptide WEG présente une meilleure inhibition de la PKA avec une énergie d'interaction totale égale à -291.3 KJ/mol, une énergie d'interaction électrostatique de -45.2 KJ/mol et avec un pourcentage d'inhibition égal à 52%.

L'examen de la cavité enzymatique confirme que la structure du tripeptide WEG, avec la présence et la position des résidus L- tryptophane et acide L- glutamique forment de fortes interactions de type hydrogène avec les acides aminés de la cavité de la protéine kinase A.

Il est à noter que la reconnaissance moléculaire des complexes protéines - peptides est hautement difficile, et que notre contribution à la compréhension des processus reste certainement à compléter par l'utilisation des logiciels beaucoup plus performants (comme le logiciel Amber conçu pour la dynamique moléculaire et le logiciel Autodock pour le docking moléculaire) que ceux que nous avons utilisés, et en tenant compte de l'effet de solvant qui a été négligé lors de notre approche.

Liste des figures

Figure 1 : Nombre de structures protéiques par intervalle de résolution donné.

Figure 2 : Evolution du nombre de structures (Données de janvier 2008).

Figure 3 : Carte d'interaction protéine-protéine impliquant 911 interactions caractérisées chez l'homme.

Figure 4 : Activation des protéines kinases A.

Figure 5: Motifs conservés des protéines kinases.

Figure 6 : principe de la mécanique moléculaire.

Figure 7 : principe de la dynamique moléculaire.

Figure 8 : distributions de Maxwell- Boltzman à différentes températures.

Figure 9 : Principales méthodes de calcul d'énergie utilisées en simulation moléculaire.

Figure 10 : Étapes typiques d'un docking moléculaire.

Figure 11 : Structure secondaire de 1YDR.

Figure 12 : Structure de la protéine kinase A (1YDR).

Figure 13 : l'énergie potentielle de la protéine kinase seule en fonction du temps.

Figure 14 : l'énergie potentielle de feG en fonction du temps.

Figure 15 : l'énergie potentielle de FEG en fonction du temps.

Figure 16 : l'énergie potentielle de WEG en fonction du temps.

Figure 17 : l'énergie potentielle de yeG en fonction du temps.

Figure 18: l'énergie potentielle de wdG en fonction du temps.

Figure 19 : l'énergie potentielle de weG en fonction du temps.

Figure 20 : les liaisons hydrogènes entre PKA et feG.

Figure 21 : les liaisons hydrogènes entre PKA et FEG.

Figure 22: les liaisons hydrogènes entre PKA et WEG.

Figure 23: les liaisons hydrogènes entre PKA et yeG.

Figure 24: les liaisons hydrogènes entre PKA et wdG.

Figure 25: les liaisons hydrogènes entre PKA et weG.

Figure 26: acides aminés du site actif.

Figure 27: Pourcentage d'énergie d'interaction totale des complexes PKA- tripeptide.

Figure 28: Pourcentage d'énergie d'interaction électrostatique des complexes PKA tripeptide.

Figure 29: Représentation des valeurs expérimentales / pourcentage d'inhibition de tripeptides.

Figure 30: dimension du trpeptide FEG.

Figure 31: dimension du trpeptide feG.

Figure 32: dimension du trpeptide WEG.

Figure 33: dimension du trpeptide yeG.

Figure 34: dimension du trpeptide wdG.

Figure 35: dimension du trpeptide weG.

Figure 36: Représentation de la cavité de site actif.



PUBLICATIONS INTERNATIONALES



Approach of the Interaction Enzyme-Substrate by Molecular Modelling

S. SARI, W. MEDJAHED, J. KAJIMA MULENGI and S. GHALEM*

Laboratory of Organics Chemistry, Analyses and Natural Substances, Cosna

Department of Chemistry, Faculty of Sciences

Aboubakr Belkaid University, Tlemcen, Algeria

Fax: 043 21 58 86; E-mail: s_ghalem2002@yahoo.fr; sari_souad@yahoo.fr

In the biological systems, the chemical reactions of the metabolism proceed seldom spontaneously, but they are generally catalysed by particular proteins which one calls enzymes. Whole of these biological functions pass by a reaction of complexation while utilizing certain atoms belonging to a given number of amino acids representing the active site. On the level of this active site, the behaviour of the atoms is relatively different as for the approach from the enzyme and the substrate. At the time of this approach, there are interactions enzyme-substrate which control the stability of a E-S complex leading to the end product according to the specificity of the enzyme. It is on this level that resides the difficulties of comprehension of the interactions between enzyme and substrate. On the basis of this idea we propose to elucidate this mechanism by molecular modelling. We are particularly interested in studying the interaction of the phospholipase A₂ (PLA₂) with aziridines structures by molecular modelling and molecular dynamics.

Key Words: Interactions enzyme-substrate, Phospholipase A₂, Aziridines, Molecular modelling.

INTRODUCTION

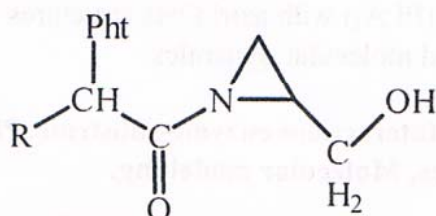
Phospholipase A₂ (PLA₂) is one of the most intensively studied membrane proteins which hydrolyze phospholipids at the sn-2 position to form fatty acid and lysophospholipid products¹. These are small proteins and the 3-D structures are known to high resolution for several species². Phospholipase A₂ proteins are of high pharmaceutical concern since they are responsible for the release of arachidonic acid from membranes and since the subsequent conversion of this fatty acid to leukotrienes and prostaglandins is part of the inflammatory response. Other functions of PLA₂ have also been documented in the literature, such as signal transduction, host defence, blood coagulation and membrane remodelling. On the basis of their structures and enzymatic characteristics, PLA₂ secreted have been recently classified into 10 groups³.

Aziridines are three-membered heterocycles containing nitrogen. The inherent reactivity of aziridines is due, in large part, to ring strain energy (SE) of 26.7 kcal/mol for the parent, unsubstituted aziridine⁴. Aziridines can be found in natural products such as mitomycin, porfiromycin and mitiromycin, which are potent antitumor and antibody agents^{5,6}.

The aziridines were the synthetic targets in the synthesis since the discovery of Gabriel⁷ in year 1888, of the smallest heterocycle containing the nitrogen. The majority of the methods of preparation of the aziridines imitate significant developments of the syntheses of epoxies^{8,9} or cyclopropanes¹⁰. The aziridines are for a long time a source of interest for the chemists, due to the multiplicity of their use in organic chemistry¹¹. They are known as well to be synthons significant in organic synthesis as to compose part of the skeleton of various biologically active natural products.

EXPERIMENTAL

The construction and the optimization of all the molecules of aziridines has been carried out by program EMO (Energy of a Molecule). The conformational analysis of all structures was carried out by using the option SCAN, which allows: A sweeping of the surface of energy and to eliminate the conformers the least probably stable.



N-acylhydroxymethylaziridine

R= H (Ref. 12), R= CH₂Ph (Ref. 12), R= CH₂PhOH and R= -CH₂COOH

The results obtained are presented in Table-1.

TABLE-1
RESULTS OBTAINED USING PROGRAM EMO

Steric energy (KJ/mol)	Stretch.	Bending	Torsion	VDW	Electr.	Total
Az-H	10.60	568.53	32.68	40.03	-28.46	623.366
Az-CH ₂ Ph	13.61	573.31	7.07	74.13	-34.97	633.155
Az-CH ₂ COOH	13.22	581.61	36.42	65.96	-70.16	627.054
Az-CH ₂ PhOH	14.07	576.61	14.66	72.80	-38.30	639.825

Once the optimized structures, it is necessary to allot to each atom an electronic partial load. A calculation of load was carried out using software SPARTAN with the level of optimization B3LYP/6-31G*, NPA.

The optimization of the geometry of the PLA₂ was carried out using the force field Amber99 (Assisted Model Building with Energy Refinement), the principal chain (backbone) was maintained rigid, while the side chains remain flexible. This approximation allows the side chains proteins to more easily find the position in which the interactions are most favourable. AMBER99 was developed by Case *et al.*¹³. The equation used can be simplified by the following equation:

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

The torsion, stretching and bending energies are represented by a simple diagonal harmonic expression. The energy of interaction of van der Waals is represented by a potential 6-12 and electrostatic energy is modelled by Coulombiennes interactions of loads centered on the atoms.

The following step is the positioning of the ligands in the active site of the PLA₂. using the program Dock in ChemOffice. Once that the receiving complex ligand- is formed, this one will adapt the most stable conformation, *i.e.* with the lowest energy level.

In present studies, a minimization of the geometry and a molecular dynamics calculation were carried out on the PLA₂ alone, the substrates and complexes PLA₂-aziridines.

RESULTS AND DISCUSSION

The first step is downloading of the PLA₂ from the Bookhaven Protein Data Bank (access code is 1 kvo). It crystallizes in the hexamer forms; each monomer consists of 124 amino acids, 2 calcium atoms and the inhibitor OAP which was used for the co-crystallization (Fig. 1).

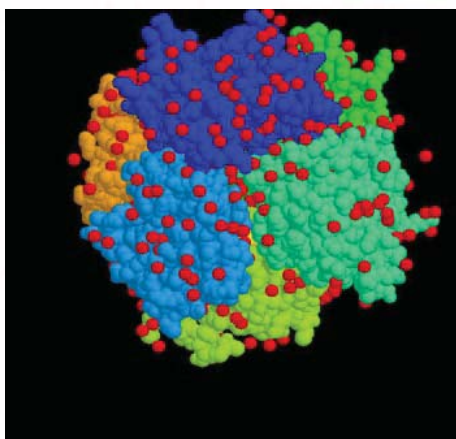


Fig. 1. Hexamer form of the PLA₂

The molecular modelling have a reducing effect, thus in present model that first 70 amino acids of the monomeric form as well as only one calcium atom present by eliminating all the molecules from water and the OAP (Fig. 2).

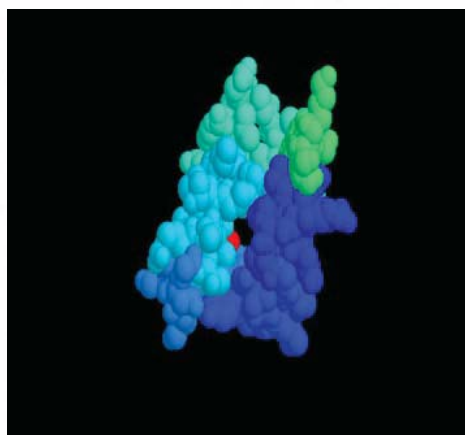


Fig. 2. Monomer chain, 70 amino acids and 1 atom of calcium

This step is followed by the optimization of the geometry and a calculation of molecular dynamics of all the starting structures.

Molecular dynamics of the PLA₂: Constraints on the level of the principal chain (backbone) were applied to the PLA₂ alone like with the various studied complexes. These constraints maintains the skeleton of rigid polypeptide, allowing the distances coordination between calcium and the complex to remain stable (Figs. 3 and 4).

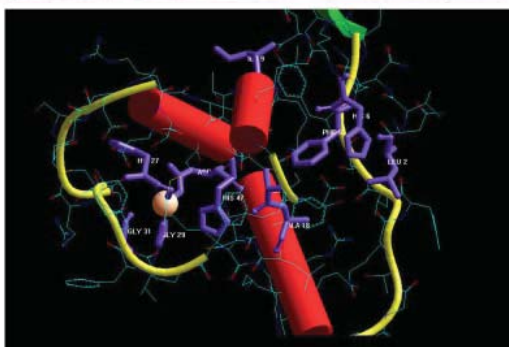


Fig. 3. Residues of PLA₂ constituting the active site

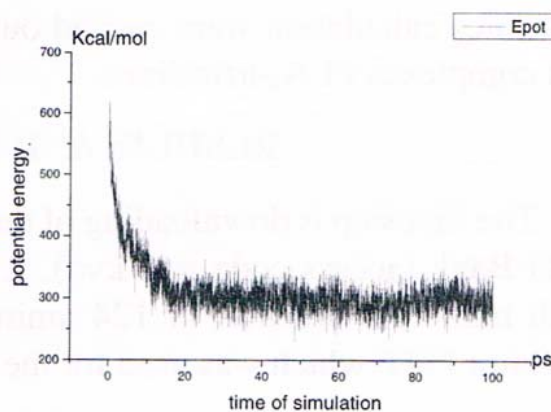


Fig. 4. Variation of the potential energy of the PLA₂ according to time

Molecular dynamics calculations of complexes PLA₂-aziridines are given in Figs. 5 to 12.

Energy of interaction: With the object of establishing correlations between energies of interaction and the activities and to consider by which type of forces the nature of the interactions PLA₂-ligands is controlled, one often refers to a decomposition of the total potential energy of the molecule in several terms (electrostatic energy, VDW energy...) (Table-6). Energies of interaction between the various studied substrates and the PLA₂ is obtained using the following relation:

$$E_{\text{interaction}} = (E_{\text{total potential energy of the complex}}) - (E_{\text{total potential energy of PLA}_2} + E_{\text{total potential energy of substrate}})$$

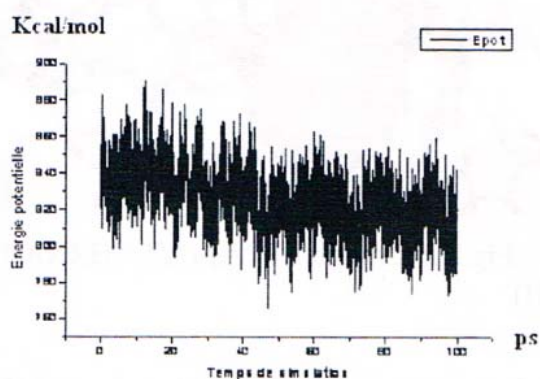


Fig. 5. Variation of the potential energy of the complex PLA₂-(AZ-H) according to time

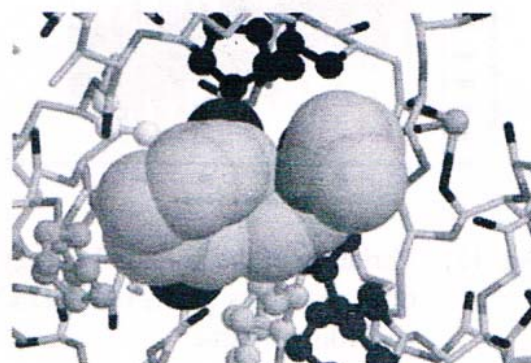


Fig. 6. Complex PLA₂-(AZ-H)

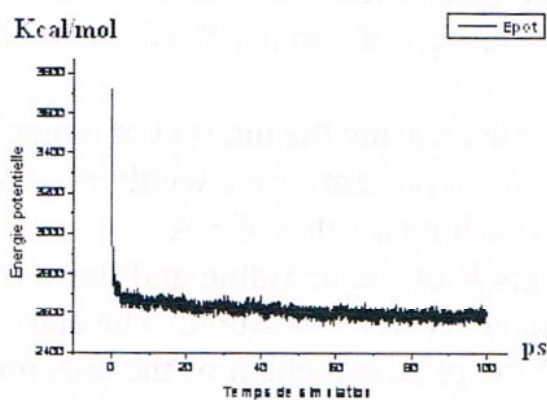


Fig. 7. Variation of the potential energy of the complex PLA₂-(AZ-CH₂Ph) according to time



Fig. 8. Complex PLA₂-(AZ-CH₂Ph)

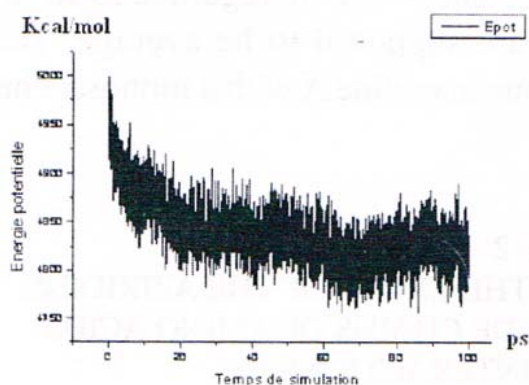


Fig. 9. Variation of the potential energy of the complex PLA₂-(AZ-CH₂PhOH) according to time

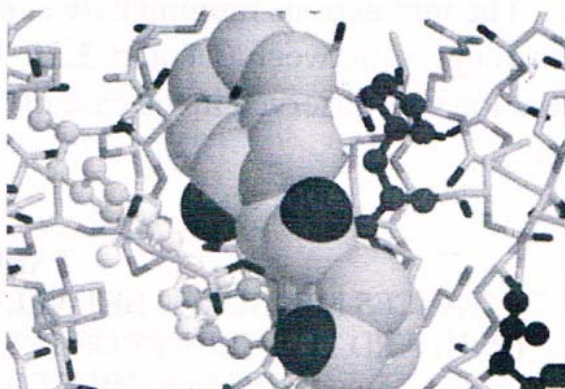


Fig. 10 Complex PLA₂-(AZ-CH₂PhOH)

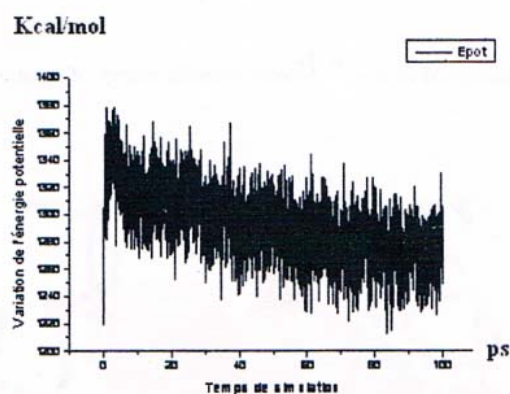


Fig. 11. Variation of the potential energy of the complex $\text{PLA}_2\text{-(AZ-CH}_2\text{COOH)}$ according to time

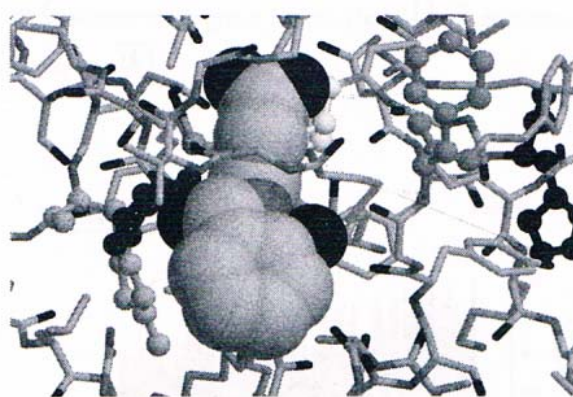


Fig. 12. Complex $\text{PLA}_2\text{-(AZ-CH}_2\text{COOH)}$

In the light of the results obtained, it arises that the introduction of the cumbersome groupings generates a conformational rearrangement within the cavity of the active site, which probably will increase complimentary activity.

The case of R=H presents a characteristic because of perfect linearity, not presenting any steric obstruction; its energy of van der Waals is equal to -92.29 Kcal/mol.

It could be discuss that increasing or decreasing the interval of dimensions of the cavity of the active site having in present case a width of 10 \AA and a depth of 14 \AA which narrows by reaching a value of 8 \AA .

The distances between the groupings R of the aziridine and those of the side chains of the constituent amino acids were measured. The active site and possibly of other groupings of the principal chain of the enzyme responsible of the interactions which can generate a conformation favourable significant activity. The results obtained are reported in Tables 2-5. The measured distances vary between 3.46 and 19.95 \AA for the whole of the studied complexes.

The interactions ranging between 2.5 and 3.1 \AA are regarded as strong and ranging between 3.1 and 3.55 \AA are supposed to be average. This situation was observed in the case of Concanavale A with a monosaccharide¹⁴.

TABLE-2
DISTANCES MEASURED BETWEEN THE GROUP OF THE AZIRIDINE (R=H) AND THE GROUPS OF THE SIDE CHAINS OF AMINO ACIDS RESPONSIBLE FOR INTERACTIONS

Distances (\AA)	Phe23	His6	Phe5	His47	Tyr21
R = H	4.25	6.20	6.45	8.75	11.00

TABLE-3
DISTANCES MEASURED BETWEEN THE GROUP OF THE AZIRIDINE
(R=CH₂Ph) AND THE GROUPS OF THE SIDE CHAINS OF AMINO ACIDS
RESPONSIBLE FOR INTERACTIONS

Distances (Å)	His27	Pro36	His47	Tyr51	Tyr66
R = CH ₂ Ph	9.67	12.69	4.37	9.30	8.80

TABLE-4
DISTANCES MEASURED BETWEEN THE GROUP OF THE AZIRIDINE
(R=CH₂PhOH) AND THE GROUPS OF THE SIDE CHAINS OF AMINO
ACIDS RESPONSIBLE FOR INTERACTIONS

Distances (Å)	Phe5	Tyr21	His27	His47	Tyr24
R = CH ₂ PhOH	5.08	6.94	3.46	5.18	13.65

TABLE-5
DISTANCES MEASURED BETWEEN THE GROUP OF THE AZIRIDINE
(R=CH₂COOH) AND THE GROUPS OF THE SIDE CHAINS OF AMINO
ACIDS RESPONSIBLE FOR INTERACTIONS

Distances (Å)	His6	Tyr21	Phe23	His27	His47
R = CH ₂ COOH	19.97	5.71	7.69	7.38	12.08

TABLE-6
VALUES OBTAINED OF INTERACTIONS ENERGY IN Kcal/mol

R	E _{total pot} complex E-S	E _{total. pot} S	VDW Energy Complex E-S	VDW Energy S	VDW Inter. Energy	Total inter. Energy
H	785.2017	236.4357	-273.860	-0.3260	-92.2908	305.2199
CH ₂ Ph	2552.6130	438.2930	-194.122	5.9238	-18.8028	1828.5954
CH ₂ PhOH	4775.0330	247.4404	-163.452	3.0965	14.6945	4284.0465
CH ₂ COOH	1248.8980	224.5801	-126.280	4.6933	50.2697	759.2937

Vdw energy of PLA₂ = -181.243 Kcal/mol, Pot. Energy of PLA₂ = 243.5461 Kcal/mol, E-S: complex Enzyme-substrate, S: substrate

Tyrosin having a pK_a side chain (phenol) of 10.07 Å is strongly deprotonated, possibly being able to have interactions cation- π , e.g. the distances separating Ca²⁺ and the groups of the aziridines are:

R = H (13.23 Å), R = CH₂Ph (5.25 Å), R = CH₂PhOH (4.57 Å) and
R = CH₂COOH (3.25 Å)

Another approach is that of exploitation of the hydrophobic interactions which remain among the principal interactions which are established between the residues preparing the hydrophobic cavity of the enzyme and the hydrophobic part of the ligands during the introduction inside the active site, in order to avoid the exposure to the solvent.

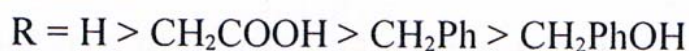
Conclusions

This contribution which consists with the elucidation of the mechanisms of interaction between the enzyme and of the inhibitors in order to establish a correlation between the activity and the energy of interaction of the complexes advances the following assumptions while waiting for the biological tests of these aziridines to validate present model.

In present work, because of the presence of the phthalimido grouping the interaction of van der Waals and cation- π , distances separating the groups from the aziridines and those of the principal amino acids of the side chain of the enzyme, the classification by increasing activity is according to:



while being based on the energy of total interaction, the order is slightly disturbed and as follows:



REFERENCES

1. D.L. Scott, S.P. White, Z. Otwinowski, W. Yuan, M.H. Gelb and P.B. Sliger, *Science*, **250**, 1541 (1990).
2. B.W. Dijkstra, K.H. Kalk, W.G.J. Hol and J. Drenth, *J. Mol. Bio.*, **147**, 93 (1981).
3. E.A. Dennis, *TIBS*, **22**, 1 (1997).
4. R.D. Bach and O. Dmitrenko, *J. Org. Chem.*, **67**, 3884 (2002).
5. H.M.I. Osborn, *J. Tetrahedron Asym.*, **8**, 1693 (1997).
6. Y.Wang and S. Kohn, *J. Am. Chem. Soc.*, **124**, 4666 (2002).
7. S. Gabriel, *Chem. Ber.*, **21**, 1049 (1888).
8. a) W.H. Pearson, B.N. Lian and S.C. Bergmeier, in eds.: A.R. Katritzky, C.W. Reese and E.F. Scriven, *Comprehensive Heterocyclic Chemistry II*, Vol. 1A, p. 1 (1996); b) D. Tanner, *Angew. Chem. Int. Ed.*, **33**, 599 (1994).
9. I. Erden, in eds.: A.R. Katritzky, C.W. Reese and E.F. Scriven, *Comprehensive Heterocyclic Chemistry II*, Vol. 1A, p. 97 (1996).
10. T. Aratani, E.N. Jacobsen, A. Pfaltz and H. Yamamoto, *Springer*, **3**, 1451 (1999).
11. R.H. Garrett and M.C. Grisham, *Biochemistry*, Brooks & Cole, edn. 3 (2004).
12. W. Medjahed, A.T. Zatla, J.K. Mulengi, F.Z.B. Ahmed and H. Merzouk, *Tetrahedron Lett.*, **45**, 1211(2004).
13. 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 and P.A. Kollman, *Programme Amber 6*, University of California, San Francisco (1999).
14. A. Imberty, K.D. Hardman, J.P. Carver and S. Pérez, *Glycobiology*, **1**, 631(1991).

(Received: 1 May 2006;

Accepted: 26 October 2006)

AJC-5225

Theoretical Study of the Inhibition of β -Secretase

I. ABDELLI, S. SARI, C. ZIANI-CHÉRIF and S. GHALEM*

Laboratory of Organics Chemistry, Analyses and Natural Substances (COSNA)
Department of Chemistry, Faculty of Sciences Aboubakr Belkaid University
Tlemcen, Algeria

E-mail: s_ghalem2002@yahoo.fr; i_abdelli@yahoo.fr

The brain of the human being suffering of the Alzheimer disease undergoes modifications caused by enzymes *e.g.*, β -secretase and γ -secretase. These enzymes produce fragments of amyloid β . The inhibition of the cleavage of this amyloid β by enzymes β and γ -secretase prevents the amyloid forming deposits plates and can improve the efficiency of the system of the elimination. Our research objective is to study the inhibition of the β -secretase by the methods of molecular modelling.

Key Words: Alzheimer disease, β -Secretase inhibitors, Molecular modelling.

INTRODUCTION

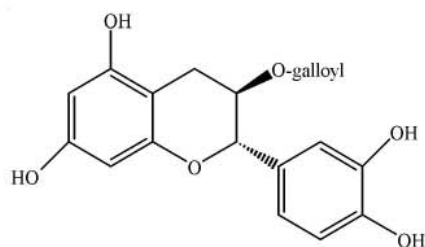
Alzheimer disease is a degenerative disease of the brain from which there is no recovery. The disease attacks nerve cells in all parts of the cortex of the brain, as well as some surrounding structures, thereby impairing a person's abilities to govern emotions, recognize errors and patterns, coordinate movement and memory. At the last, an affected person loses all memory and mental functioning¹.

β -Secretase is an enzyme that appears to be directly involved in the early development of Alzheimer disease. β -Secretase^{2,3} is a protease (an enzyme that catalyzes the splitting of interior peptide bonds in a protein).

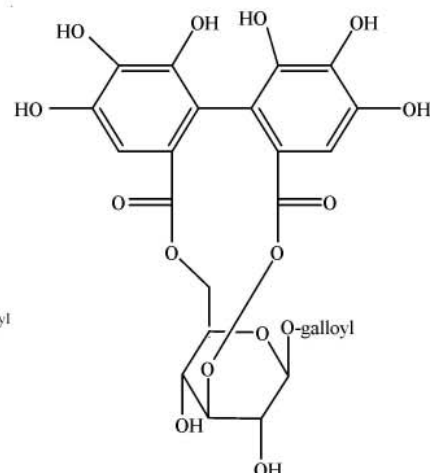
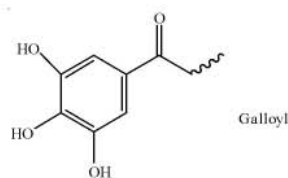
The discovery of this enzyme was made in 1999⁴⁻⁷ and may lead to the development of drugs designed (inhibitors) to block this enzyme and hopefully block the progression of Alzheimer disease.

The inhibitors used in present work were divided into 2 categories A and B on the basis of their geometries and their nature.

Category A: The inhibitors of this category are natural substances extracted the green tea⁸, their structure is represented below:

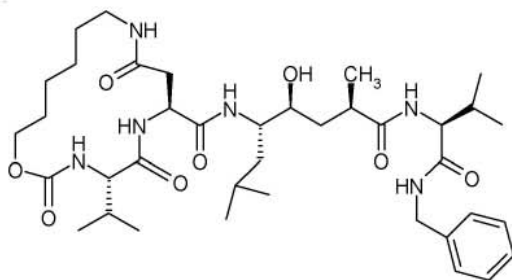


Inhibitor 1: (Catechin gallate)

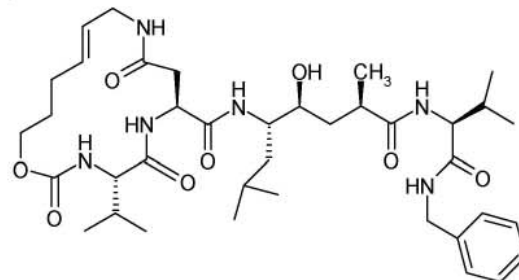


Inhibitor 2: (Corilagin)

Category B: The inhibitors of this category represent the cycloamide-urethane derivatives⁹, they are obtained by synthesis. They have the following structures:



Inhibitor 3

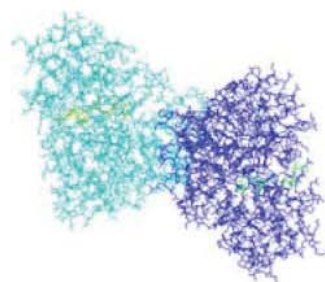
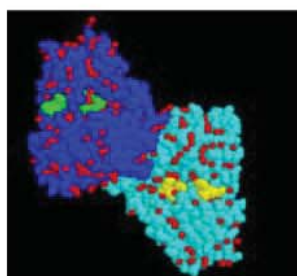


Inhibitor 4

THEORETICAL APPROACH

The downloading of β -secretase was made from the data base Bookhaven Protein Data Bank (access code 1FKN) it is co-crystallized with inhibitor OM99-2.

The three-dimensional structure of β -secretase was obtained by diffraction in X-ray with a resolution (1.90 Å). It is noticed that β -secretase crystallizes in the dimer forms (Fig. 1).

Fig. 1. The dimerous form of β -secretase

With the reducing effect of molecular modelling, we simplified the model of the enzyme and retained only one monomer (Fig. 2). This monomer comprises the amino acids forming the active site (Fig. 3).

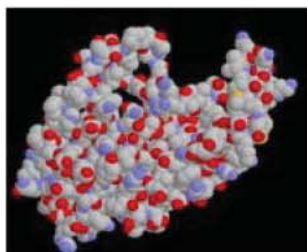


Fig. 2. Simplified model of β -secretase

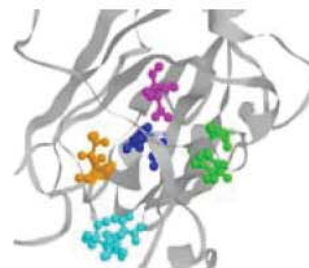


Fig. 3. Amino acids of the active site (Asp32, Tyr198, Asp228, Thr232, Arg235)

The optimization of the geometry¹⁰ of the β -secretase was carried out using the field of forces Amber99.

The main chain was maintained rigid, whereas the lateral chains remain flexible. This approximation allows the lateral chains of proteins to more easily find the position in which the interactions are the most favourable.

The value of the energy of optimization is: $E_{\text{Opt}} = -1933.20$ Kcal/mol.

The optimization of all the inhibitors was made by program EMO¹¹⁻¹⁴. The results obtained are collected in Table-1.

TABLE-1
RESULTS OBTAINED USING PROGRAM EMO

Steric energy (KJ/mol)	$E_{\text{Stretching}}$	E_{Bending}	E_{Torsion}	E_{VdW}	$E_{\text{Electrostatic}}$	E_{Steric}
Inhib-1	7.59	44.14	-76.12	108.33	-55.73	27.61
Inhib-2	27.38	119.49	159.12	252.69	-102.46	456.22
Inhib-3	22.65	134.49	7.79	169.29	-65.09	269.15
Inhib-4	22.28	162.54	65.69	205.46	-109.28	346.70

After optimization of all the inhibitors, we made a calculation of load to attribute to each atom an electronic partial load, for it we used the software AMPAC.

RESULTS AND DISCUSSION

Molecular dynamics of the β -secretase (Fig. 4): This stage is necessary to eliminate conformations the least favourable and to lead to the most stable conformation.

We began dynamics with an initialization of the system: with $t = 0$, we have $R(T) = 0$, *i.e.* the initial structure, previously minimized. Then we

heated the system up to 300 K during 1000 steps with a step of integrations of 1fs.

In 300 K, there is an equilibration. Speeds are straightened to keep the constant temperature (there is exchange between the kinetic energy and the potential energy). Then, there is production of conformations. The time of simulation of the molecular dynamics is 100 picoseconds.

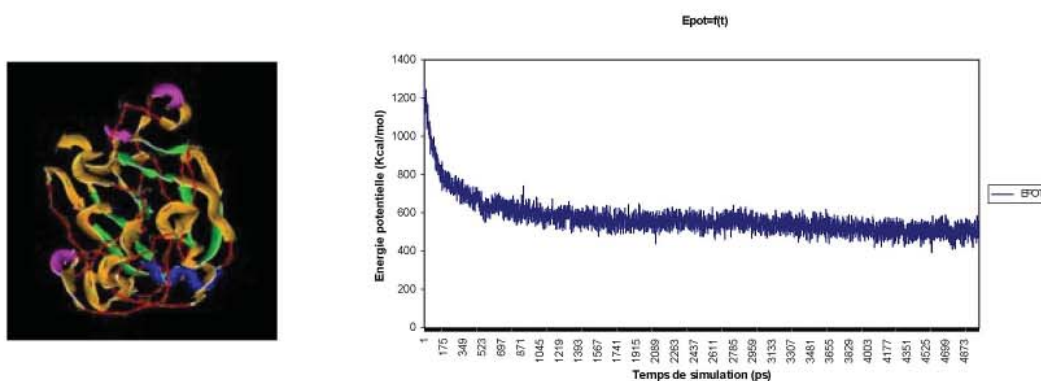


Fig. 4. Variation of the potential energy of the 1FKN only according to time

Molecular dynamics of the inhibitors:

A) Inhib-1:

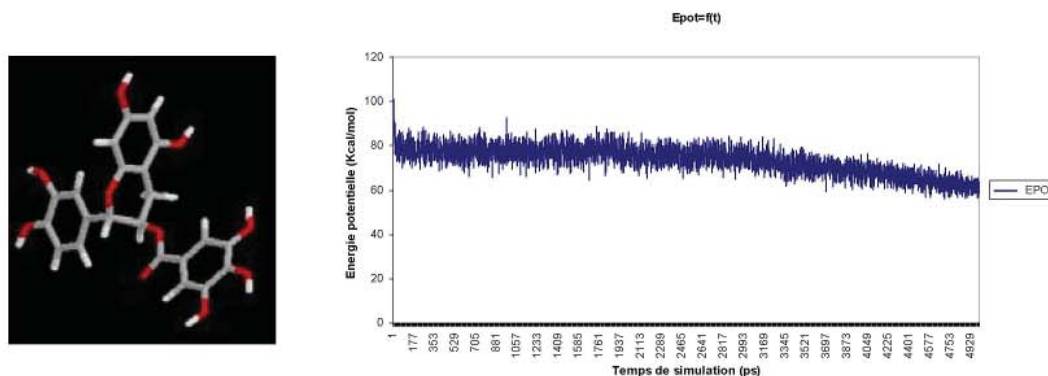


Fig. 5. Variation of the potential energy of the inhib-1 only according to time

B) Inhib-2:

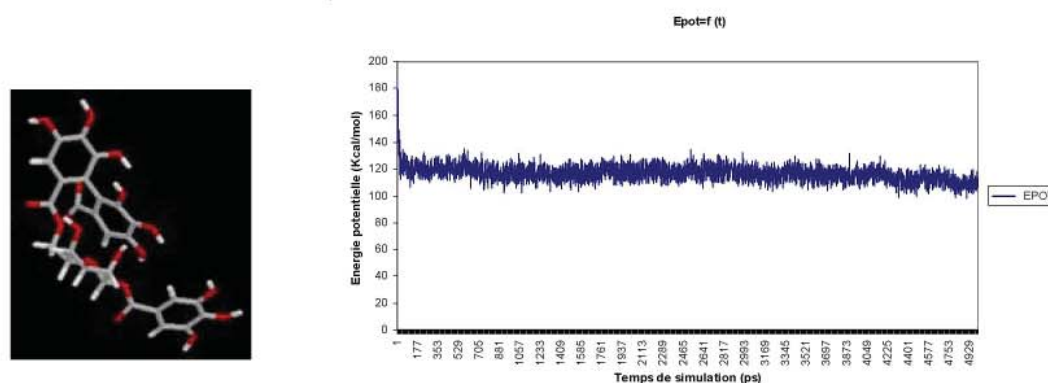


Fig. 6. Variation of the potential energy of the inhib-2 only according to time

C) Inhib-3:

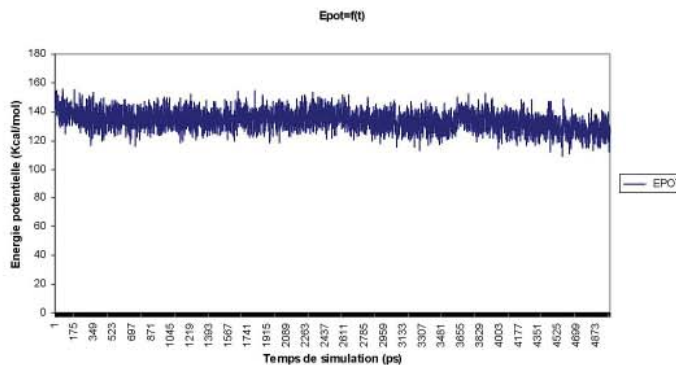
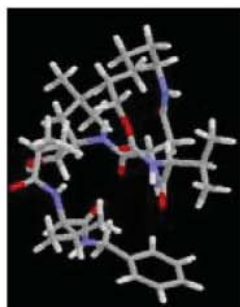


Fig. 7. Variation of the potential energy of the inhib-3 only according to time

D) Inhib-4:

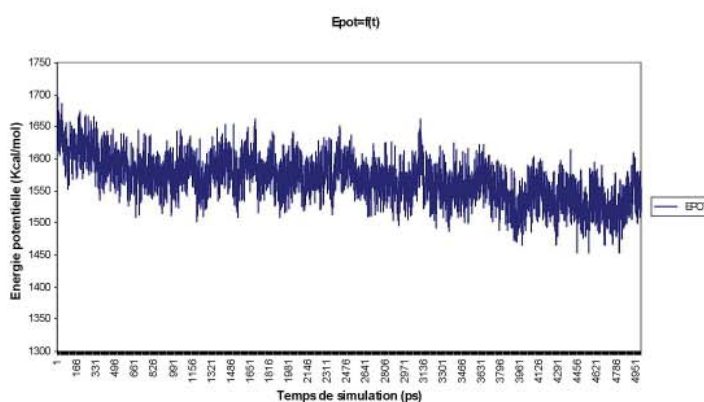
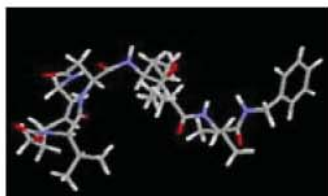


Fig. 8. Variation of the potential energy of the inhib-4 only according to time

Docking molecular: This step consists moved closer to two entities molecular that is the location of the inhibitors in the active site of β -secretase (forming of the complexes). We used the Dock program of the software ChemOffice (2000).

Once all the complexes were formed, we are going to make an optimization of the geometry and a calculation of molecular dynamics to look for the most stable conformation.

a) Complex 1FKN-Inhib-1:

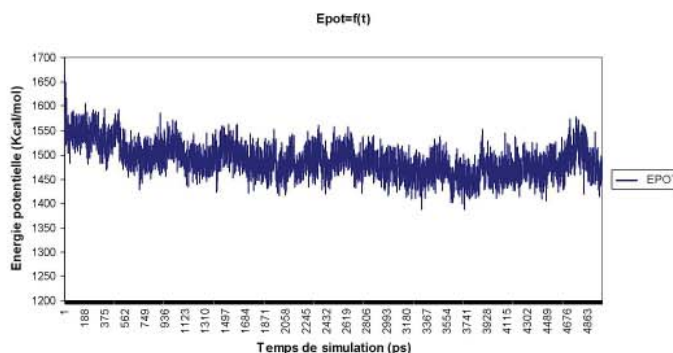
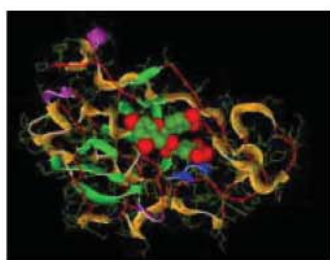


Fig. 9. Variation of the potential energy of complex 1 according to time

b) Complex 1FKN-Inhib-2:

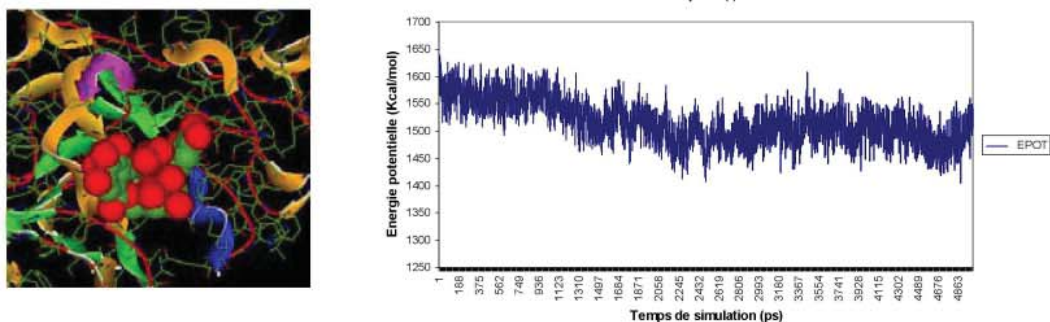


Fig. 10. Variation of the potential energy of complex 2 according to time

c) Complex 1FKN-Inhib-3:

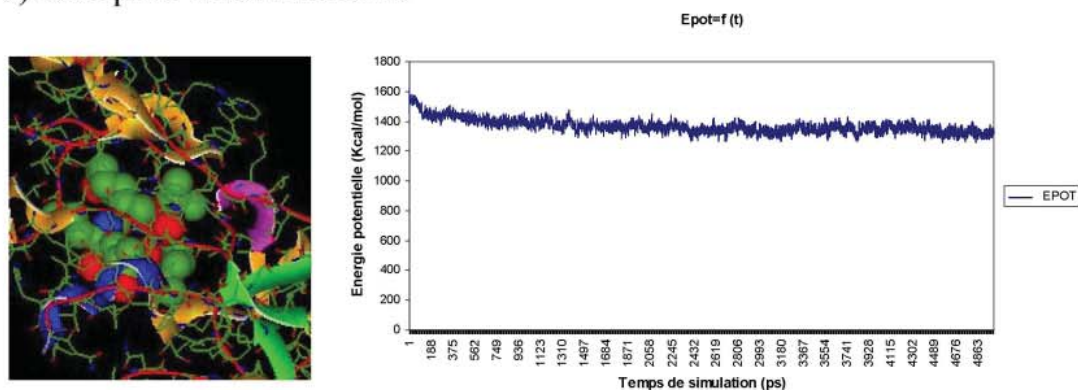


Fig. 11. Variation of the potential energy of complex 3 according to time

d) Complex 4 1FKN-Inhib-4:

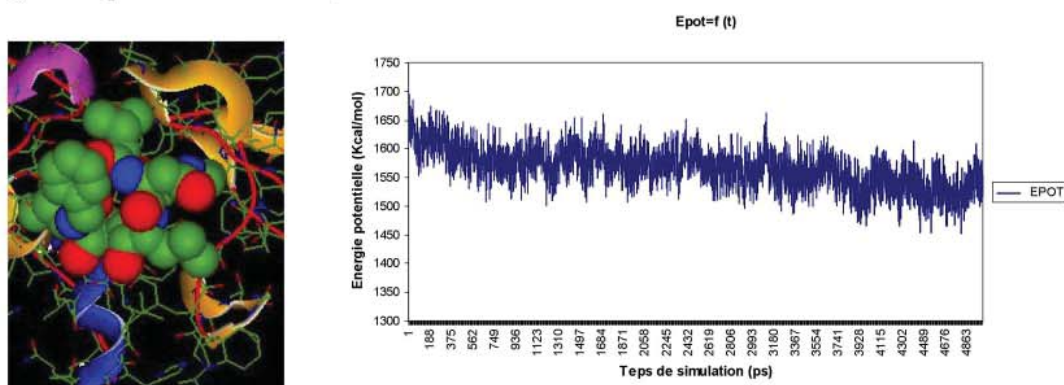


Fig. 12. Variation of the potential energy of complex 4 according to time

Distances separating the amino acids from the active site and the groupings of inhibitors:

TABLE-2
DISTANCE IN Å MEASURED BETWEEN THE GROUPINGS OF INHIBI
AND THE GROUPINGS OF THE SIDE CHAINS OF AMINO ACIDS
RESPONSIBLE FOR INTERACTION

	Asp32	Tyr198	Asp228	Thr232	Arg235
Inhib1	6.10	9.62	6.02	8.75	11.94

TABLE-3
DISTANCE IN Å MEASURED BETWEEN THE GROUPINGS OF INHIB-2
AND THE GROUPINGS OF THE SIDE CHAINS OF AMINO ACIDS
RESPONSIBLE FOR INTERACTION

	Asp32	Tyr198	Asp228	Thr232	Arg235
Inhib-2	6.94	7.63	12.37	2.99	2.42

TABLE-4
DISTANCE IN Å MEASURED BETWEEN THE GROUPINGS OF INHIB-3
AND THE GROUPINGS OF THE SIDE CHAINS OF AMINO ACIDS
RESPONSIBLE FOR INTERACTION

	Asp32	Tyr198	Asp228	Thr232	Arg235
Inhib-3	3.73	24.21	10.05	8.46	8.15

TABLE-5
DISTANCE IN Å MEASURED BETWEEN THE GROUPINGS OF INHIB-4
AND THE GROUPINGS OF THE SIDE CHAINS OF AMINO ACIDS
RESPONSIBLE FOR INTERACTION

	Asp32	Tyr198	Asp228	Thr232	Arg235
Inhib-4	3.02	14.04	6.84	8.54	2.86

The energies of interactions between the β -secretase and the various studied inhibitors are calculated from the following relation:

$$E (\text{interaction}) = [E \text{ pot (complex enzyme - inhibitor)}] - [E \text{ pot (enzyme) + E pot (inhibitor)}]$$

It is necessary to take into account also Van der Waals's interactions because it is the interactions between nondependent atoms that stabilize complex enzyme-inhibitor.

TABLE-6
RESULTS OF INTERACTIONS ENERGIES (Kcal/mol)

Inhibitors	(Kcal/mol)						Category
	$E_{\text{total pot of the complex E-I}}$	$E_{\text{total Pot of the inhibitor I}}$	$E_{\text{vdw (complex-inhibitor)}}$	$E_{\text{vdw (inhibitor)}}$	$E_{\text{Vdw inter.}}$	$E_{\text{total inter}}$	
Inhib-1	1474.05	60.56	-679.25	25.58	-39.72	867.90	A
Inhib-2	1541.54	111.26	-703.89	29.09	-67.87	884.78	A
Inhib-3	1336.00	118.77	-1048.2	21.40	-420.30	671.73	B
Inhib-4	1552.51	134.45	-738.10	31.21	-104.21	872.57	B

Potential energy of the β -secretase only is: $E_{\text{pot}} = 545.48$ (Kcal/mol).
Potential energy of the β -secretase only is: $E_{\text{Vdw}} = -665.11$ (Kcal/mol).

The values of K_i and IC_{50} are given below:

TABLE-7
VALUES OF K_i AND IC_{50}

Category A		Category B	
Inhibitors	IC_{50}	Inhibitors	K_i
Inhib-1	6.0×10^{-6} M	Inhib-3	25.1 nM
Inhib-2	1.6×10^{-4} M	Inhib-4	14.2 nM

In the light of the results obtained during present work, it emerges that the values obtained relating to energies of total interaction are of the same order of magnitude for both categories with the exception of the inhibitor 3.

The inhibitor 3 belonging in the category B constituted by 114 atoms, synthetic product possesses an important number of doublet and being able to interact with the residues of the active site, possesses an energy of (671.73 Kcal/mol) who can engender a conformation favourable to an important complementarity which from a consequent activity.

It is noted a peculiarity between the inhibitor 3 and the inhibitor 4, on the level of this last the presence of a double connection inserted in the cycle into 17 links which confers a relatively stable conformation of energy less low than the inhibitor 3; thus an interaction of atom nondependent is less important.

This structure presents a relatively cumbersome conformation from where a steric gene influencing the stability of the complex.

Distances measured between the groupings of the inhibitors and those of the side chains, by not considering that the most brought closer groupings show according to the inhibitor considered of the interactions variable of 2.42 to 24.21 Å.

The interactions ranging between 2.5 and 3.1 Å are regarded as strong and those ranging between 3.1 and 3.55 Å are supposed to be average. The interactions higher than 3.55 Å are weak¹⁵.

The examination of the cavity of the active site (Fig. 10) presents geometry of 8.65 Å of depth, an opening of 16.76 and 9 Å, this narrowed pocket reaching a width of 12.61 Å.



Fig. 13. Dimension of the cavity enzymatique

The category A, natural products extracts of the green tea, presents a geometry of association of cycle at 6 links and a less number of electronic doublet being able to interact with the residues of the active site.

By taking account of the various geometrical constraints, the approach according to inhibitors' considered raises difficulties being able to influence the complementarity and consequently the activity.

The exploitation of some experimental data in occurrence⁸ the IC_{50} for the structures of category A and while being based on energy of total interaction, shows an agreement acceptable for inhibitors 1 and 2. Indeed, inhibitor 1 (867.90 Kcal/mol) is more active ($IC_{50} = 6 \times 10^{-6}$ M) than inhibitor 2 (884.76 Kcal/mol) with $IC_{50} = 1.6 \times 10^{-4}$ M.

Concerning the category B, K_i (constants of inhibition)⁹ available and total energies of interaction, show the following order: inhibitor 3 (671.73 Kcal/mol) with $K_i = 25.1$ nM is more active than inhibitor 4 (872.57 Kcal/mol) with $K_i = 14.2$ nM.

The examination of the values of the energy of Van der Waal's interaction which is generally retained to explain the interactions between biological systems, show that in certain structures; mainly the category A whose geometry and characterized by a steric obstruction, the order of stability is relatively disturbed.

Inhibitor 1 (-39.72 Kcal/mol, $IC_{50} = 6 \times 10^{-6}$ M) normally more stable than inhibitor 2 (-67.87 Kcal/mol, $IC_{50} = 6 \times 10^{-4}$ M), is observed destabilizing taking into account the values of the energy of van der Waals. The same order is obtained by examining the values of the IC_{50} . Allotting this disorder in the classification to us to a competitive inhibition of with the geometry of the inhibitors.

This phenomenon is generated by a change of conformation of the enzyme (active site) which prevents the mutual connection of the substrate.

Another probable approach is that the connection of the inhibitor to the free enzyme with a site meadow of the active site can prevent the connection of the substrate by steric obstruction which is probably the present case.

Conclusions

This modest work consists in studying the inhibition of β -secretase which is an enzyme involved in the Alzheimer's disease by method of molecular modelling.

Within sight of the results obtained and experimental values recorded in the bibliography, it arises:

1st classification: while basing itself on the energy of total interaction.

Inhib-3 > Inhib-1 > Inhib-4 > Inhib-2

2nd classification: while basing itself on the energy of van der Waals's interaction.

Inhib-3 > Inhib-4 > Inhib-2 > Inhib-1

Thus according to these 2 classifications, we can say that the inhibitor 3 would be probably the best to slow down the evolution of studied pathology (Alzheimer's disease).

REFERENCES

1. C. Belisle and B. Rivard, The insanity of the Alzheimer type and other attacks cognitive. Ed. Formed, MI (1999).
2. A.K. Ghosh, L. Hong and J. Tang, *Curr. Med. Chem.*, **9**, 1135 (2002).
3. M.S. Wolfe, *J. Med. Chem.*, **44** (2001).
4. I. Hussain, D. Powell and et al, *Mol. Cell Neurosci.*, **14**, 419 (1999).
5. S. Sinha, J.P. Anderson, R. Barbour, G.S. Basi, R. Caccavello, D. Davis, M. Doan, H.F. Dovey, N. Frigon, J. Hong, K. Jacobson-Croak, N. Jewett, P. Keim, J. Knops, I. Lieberburg, M. Power, H. Tan, G. Tatsuno, J. Tung, D. Schenk, P. Seubert, S.M. Suomensari, S. Wang, D. Walker, J. Zhao, L. McConlogue and V. John, *Nature*, **402**, 537 (1999).
6. R. Vassar, B.D. Bennett, S. Babu-Khan, S. Kahn, E.A. Mendiaz, P. Denis, D.B. Teplow, S. Ross, P. Amarante, R. Loeloff, Y. Luo, S. Fisher, J. Fuller, S. Edenson, J. Lile, M.A. Jarosinski, A.L. Biere, E. Curran, T. Burgess, J.-C. Louis, F. Collins, J. Treanor, G. Rogers and M. Citron, *Science*, **286**, 735 (1999).
7. R. Yan, M.J. Bienkowski, M.E. Shuck, H. Miao, M.C. Tory, A.M. Pauley, J.R. Brashler, N.C. Stratman, W.R. Mathews, A.E. Buhl, D.B. Carter, A.G. Tomasselli, L.A. Parodi, R.L. Heinrikson and M.E. Gurney, *Nature*, **402**, 533 (1999).
8. S.Y. Jeon, K. Bae, Y.H. Seong and K.S. Song, *Bioorg. Med. Chem. Lett.*, **13**, 3905 (2003).
9. A.K. Ghosh, T. Devasamudram, L. Hong, C. De Zutter, X. Xu, V. Weerasena, G. Koelsch, G. Bilcer and J. Tang, *Bioorg. Med. Chem. Lett.*, **15**, 15 (2005)
10. D.A. Case, D.A. Pearlman, J.W. Caldwell, T.E. Cheatham III, W.S. Ross, C.L. Simmerlig, 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. Pitner, I. Massova, G.I. Seibel, U.C. Singh, P.K. Weiner and P. A. Kollman, Amber6, University California, San Francisco, USA (1999).
11. A. Bouraoui, M. Fathallah, B. Blaive and R. Gallo, *J. Chem. Soc. Perkin Trans. II*, 1211 (1990).
12. A. Bouraoui, M. Fathallah, F.M. Henni, B. Blaive and R. Gallo, in ed.: J.L. Rivail, Modeling of Molecular Structures and Properties. Proceeding of an International Meeting Nancy, France, 11-15, September, Studies in Physical and Theoretical Chemistry, Vol. 71, pp. 381-393 (1989).
13. A. Zinelabidine, A. Bouraoui, M. Fathallah, F.M. Henni, B. Blaive and R. Gallo, *J. Mol. Struct. (Theochem.)*, **286**, 267 (1993).
14. B. Blaive, G. Legsaï and R. Laï, *J. Mol. Struct.*, **354**, 245 (1995).
15. A. Imberty, K.D. Hardman, J.P. Carver and S. Pérez, *Glycobiology*, **1**, 631 (1991).

Theoretical study of corrosion inhibition by bipyrazolic derivatives

S. Sari.* , I. Abdelli. & S. Ghalem.*

Sari, Abdelli & Ghalem are with UNIVERSITY ABOU BEKR BELKAID - TLEMEN DEPARTEMENT OF CHEMISTRY. Laboratoire de Chimie Organique, Substances Naturelles et Analyses (COSNA)

Abstract

Corrosion is a field which touches all kinds of materials (metals, ceramics, polymers) in variable environments (aqueous medium, atmosphere, high temperatures). We can not stop corrosion but we can control it and slow down it by using inhibitors to prolong the lifespan of protected surface. The use of the organic inhibitors is one of the most practical methods for protection against corrosion in particular in acid medium. Corrosion inhibitors may be divided into three broad classes, namely oxidizing, precipitation and adsorption inhibitors [1]. Compounds containing nitrogen, oxygen, sulphur and phosphorous in the conjugated system have particularly been reported as efficient corrosion inhibitors [2, 3]. The choice of the inhibitors was based on the fact that these compounds contains π -electrons and heteroatoms such as N, O and S, which involve greater adsorption of the inhibitor molecules onto the surface of mild steel [4].

Citation: S. Sari. I. Abdelli.,& S. Ghalem. "Theoretical study of corrosion inhibition by bipyrazolic derivatives" IBScientific Journal Of Science Volume 3(2) pp 30,31 2008

Copyright: © S. Sari et al. This is an open-access article. Unrestricted use, distribution or reproduction in any medium of the article is permitted provided the original work is properly cited.

Article Submitted: 08 Oct 2007 **Accepted:** 09 March 2008

Competing Interests: The authors have declared that no competing interests exist.

* **Correspondence:** sari_souad@yahoo.fr s_ghalem2002@yahoo.fr

Long experience and voluminous research have been devoted to find the easiest way to slow down iron and steel attack in acid. The use of inhibitors tasks more and more attention until now. The efficiency of an organic compound as corrosion inhibitor depends not only on the characteristics of the environment in which it acts, the nature of the metal surface and electrochemical potential at the interface, but also on the structure of the inhibitor itself, which includes the number of adsorption active centres in the molecule, their charge density, the molecule size, the mode of adsorption, the formation of metallic complexes and the projected area of the inhibitor on the metallic surface [5-6].

The molecular weight of the inhibitor has a direct influence on its inhibition efficiency. The increase in molecular weight of the inhibitor is due to an increase in the length of the hydrocarbon chain of amines [7], nitriles [8] or mercaptants. The rise of the inhibition efficiency is due to the inductive effect of the methyl groups. The effect of the molecular configuration encountered in compounds like sulphides and aliphatic amines has also studied [9].

Pyrazolic compounds are good inhibitors of iron corrosion in sulphuric and hydrochloric acid [10]. Their inhibiting effect is closely related to the molecular structure containing two nitrogen atoms in the pyrazolic ring. The aim of this work is to study theoretically the inhibition of the bipyrazolic compounds [11], which are synthesised and tested as a corrosion inhibitors of Armco Iron [11], and the effect of the length of the chain between pyrazolic rings.

THEORETICAL APPROACH

Summary

Younes Mokrab

Postdoctoral Associate, Department of Biochemistry Oxford

Corrosion is an undesirable phenomenon which affects a wide range of industrial materials. Because it is caused by a complexity of factors inherent in the environment, a major focus has been on countering it using chemical compounds that decelerate the processes involved. In the current issue, Sari and Ghalem present a theoretical study for the use of the bipyrazolic class of chemical compounds as inhibitors for corrosion.

A conformational analysis, for all structures of inhibitors was performed using an option « scan » in the EMO program (Energy of MOlecule) and the found structures were optimized at DFT level using hybrid functional density B3LYP with 6-31G* basis set and POP=NPA (Gaussian 03). The structure of the studied bipyrazolic compounds are shown in figure 1.

RESULTS AND DISCUSSION

To investigate the influence of electronic structure on the efficiency of the compounds, some parameters such as the energies of molecular orbital, (high occupied molecular orbital energy), ELUMO (lowest unoccupied molecular orbital energy) and the dipole moment (μ) were carried out. The results are given in table 1. The energy of HOMO is often associated with the electron donating ability of a molecule. High value of E HOMO is likely to indicate a tendency of the molecule to donate electrons to appropriate acceptor molecules. Therefore the energy of LUMO indicates the ability of the molecule to accept electrons. The lower is the value of E LUMO, the more probable it is that the molecule accept electrons [12]. Consequently, less negative HOMO energy and the smaller energy gap (E LUMO - E HOMO) are reflected in a stronger chemisorptions bond and perhaps greater inhibition efficiency [13]. For the dipole moment (μ), lower value of (μ) will favour accumulation

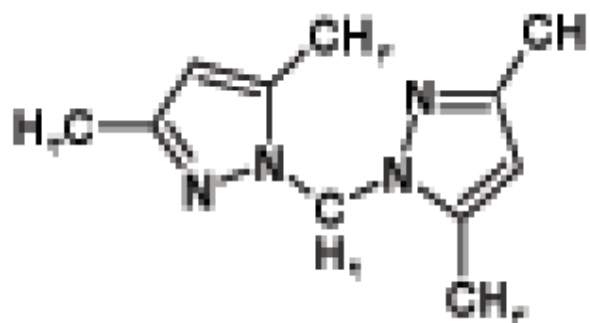


Figure 1 : Molecular structure of bipyrazole compounds studied. (A) Bis(3,5-dimethylpyrazolyl)methane (inh1).

of the inhibitor in the surface layer and therefore higher inhibition efficiency [14]. Meanwhile, several authors' state that the inhibition efficiency increases with increasing values of the dipole moment [15-16]. On the other hand, survey of literature reveals that several irregularities appeared in case of correlation of dipole moment with inhibitor efficiency [17-18]. The results obtained with theoretical calculation show that inh3 is the best inhibitor.

The inhibition efficiencies obtained can be explained by the length and the nature of the carbonic chain joining the two pyrazolic cycles. In fact, the bipyrazolic compounds differ by the number *n* of CH₂ groups between the pyrazolic cycles. The number *n* is equal to 1, 3 and 4 for the compounds 1, 2 and 3, respectively. It is clear that the maximum of inhibition is obtained for *n*=4 and on the both sides the efficiency decreases for *n*=1 and 3, respectively. This result is related to the high flexibility of the molecule with four CH₂; at *n*=1, the molecule loses its flexibility. The existence of two oxygens in compounds 4 and 5 has no effect and is accompanied by an augment in the efficiency.

Compound	E HOMO (eV)	E LUMO (eV)	ΔE (eV)	μ (D)
Inh1	-0.22968	0.01640	0.24608	0.1490
Inh2	-0.21848	0.02305	0.24153	2.5541
Inh3	-0.21592	0.02205	0.23797	3.5183
Inh4	-0.22307	0.02454	0.24761	4.7127
inh5	-0.22195	0.02096	0.24291	2.7986

Table1. Calculated quantum chemical indices

CONCLUSION

We can conclude that with higher alkyl chain length, the inhibition becomes high effective.

The band gap difference of 0.23797 eV, which is in favour of inh3, indicates that inh3 molecule can easily transfer an electron from HOMO level to LUMO level. However, we remark that this difference is principally supported by an increase of LUMO energy of inh1 indicating its lowest donor effect.

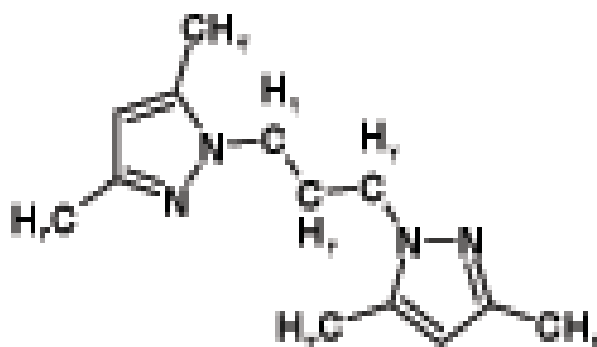


Figure 2: (B) 1,3-Bis (3,5-dimethylpyrazolyl) propane (inh2).

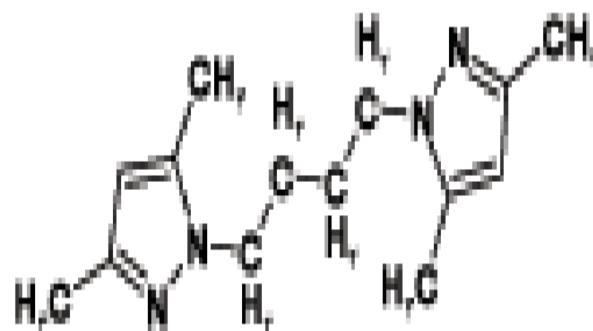


Figure 3: (C) 1,4- Bis(3,5-dimethylpyrazolyl) butane (inh3).

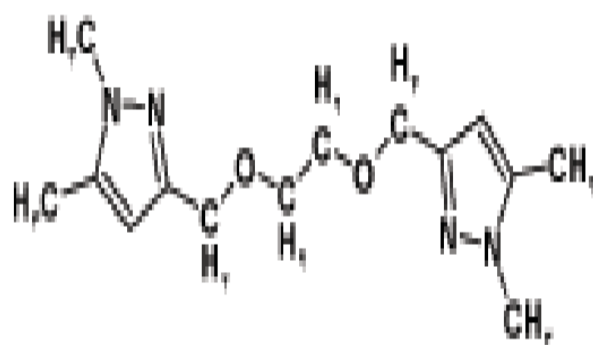


Figure 4: (D) 1,6-[Bis (1,5-dimethylpyrazol-3-yl)] 2,5-dioxahexane (inh4).

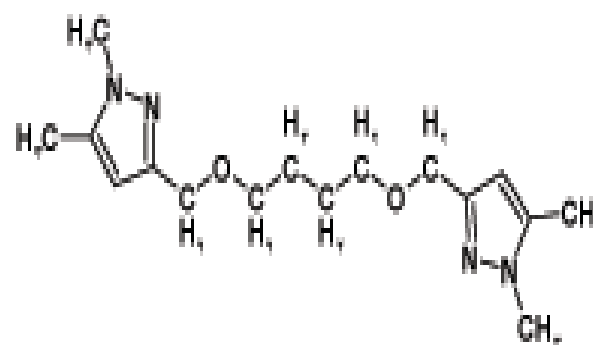


Figure 5: (E) 1,8-[Bis (1,5-dimethylpyrazol-3-yl)] 2,7-dioxaoctane (inh5).

REFERENCES

- [1] J.M. Sykes, 1990, Br. Corros. J, 25: 175.
- [2] L. Larabi, Y. Harek, M. Traisnel, A. Mansri, 2004, J. Appl. Electrochem, 34: 833.
- [3] H.I. Wong, H. Fan, J. Zheng, Mater. 2004, Chem. Phys. 77: 655.
- [4] L. Larabi, Y. Harek, O. Benali, S. Ghalem, 2005, Progress in Organic Coatings, 54: 256-262.
- [5] A.Aouniti, B. Hammouti, M. Brighli, S. Kertit, F. Berhili, S. El Kadiri, A. Ramdani, 1996, J. Chim. Phys, 93: 924.
- [6] B.Hammouti, R. Salghi, S. Kertit, 1998, J. Electrochem. Soc. Of India, 47: 31.
- [7] S. Sankarapavinasan, F. Pishapinadon, M.F. Ahmed, 1991, Corrosion. Sci, 32: 193.
- [8] V. Carassiti, F. Zucchi, G. Trabaneli, 1970, 3 SEIC, Ann. Univ. Ferrara, N.S. Sez, 5: 525.
- [9] D.M.J.Incorvia, 1985, 6 SEIC, Ann. Univ. Ferrara, 8 : 81.
- [10] W.A.Badawy, N.M. Hefny, S.S. El Egamy, 1990, Corros. Sci, 36 : 978.
- [11] F. Touhami , 2000, Corrosion Science, 42 : 929-940.

- [12] I.Lukovits, K. Palfi, E. Kalman, 1997, Corrosion, 53 : 915.
- [13] H.F.Finley, N.Hackerman, 1960, J. Electrochem, Soc, 107: 259.
- [14] N.Khalil, Electrochim. Acta, vol. V (48), pp 2635, 2003.
- [15] M. Langrenée, B. Mernari, N. Chaibi, M. Traisnel, H. Vezin, F. Bentiss, 200, Corros.Sci, 43 : 951.
- [16] M. A. Quraishi, R. Sardar, 2003, J. Appl. Electrochem, 33: 1163.
- [17] K.F. Khaled, N. KBabié-Samardzija, Hackerman, 2005, Electrochim. Acta, 50: 2515.
- [18] G. Bereket, E. Hur, C. Ogretir, 2002, J. Mol. Struct. (Theochem), 578: 79.

Molecular Modelling of Green Tea Catechins Inhibitors of Human Brain BACE1

I. Abdelli*, S. Sari. & S. Ghalem.*

Abdelli, Sari & Ghalem are with UNIVERSITY ABOU BEKR BELKAID - TLEMCEM DEPARTEMENT OF CHEMISTRY. Laboratoire de Chimie Organique, Substances Naturelles et Analyses (COSNA)

Abstract

The research and the synthesis of new chemical compounds are often associated by a molecular modelling study. The molecular modelling is an allowing technique, not only to represent the properties and the chemical reactions but also to treat the models of the structures in two or three dimensions. It implies the use of the theoretical methods of calculation (molecular mechanics, molecular dynamic, and quantum mechanics) allowing to determine the geometry, the configuration of the atoms of a molecule and to evaluate the properties physicochemical of the molecule. Our work consists to study by the methods of molecular modelling the derivatives of catechin which is a natural substance extracted the green tea. They are inhibitors of b-secretase (BACE1) enzyme implied in Alzheimer disease.

Citation: I. Abdelli., S. Sari. & S. Ghalem. "Molecular Modelling of Green Tea Catechins Inhibitors of Human Brain BACE1" IBSscientific Journal Of Science Volume 3(2) pp 27,29 2008

Copyright: © I. Abdelli et al. This is an open-access article. Unrestricted use, distribution or reproduction in any medium of the article is permitted provided the original work is properly cited.

Article Submitted: 08 Oct 2007 **Accepted:** 09 March 2008

Competing Interests: The authors have declared that no competing interests exist.

* **Correspondence:** i_abdelli@yahoo.fr s_ghalem2002@yahoo.fr

Alzheimer disease (AD) is the most prevalent neurodegenerative disease, affecting approximately 20 millions people worldwide. As a progressive disorder, it gradually robs the patient of cognitive function including memory and thinking ability, ultimately leading to death [1a-3]. AD account for 60% to 70% of cases of progressive cognitive impairment in elderly patients. It is more common among women than men by a ratio of 1.2 to 1.5. In the United State alone, the number of new cases per year has been estimated to 360000 equating 40 new cases every hour [4].

The accumulation of the β -amyloid peptide ($A\beta$) in the brain is the central event leading to the development of AD [5]. $A\beta$ is a 40/42-residue fragment of the brain transmembrane protein β -amyloid precursor protein (APP), released by two proteases known as the β - and γ -secretases (Figure 1). Currently, there is a great deal of interest in the development of inhibitor drugs against β -secretase for several reasons. The activity of this protease is an early step in the pathogenesis pathway common to all familial and sporadic forms of AD. The recent identification of β -secretase to be an aspartic protease has also raised the hope that its drug development can be as successful as that of HIV protease, another aspartic protease. Finally, the absence of phenotype in mice devoid of a functional β -secretase gene suggests that the inhibition of this protease during clinical treatment is physiologically tolerable [6-8].

Summary

Younes Mokrab

Postdoctoral Associate, Department of Biochemistry Oxford

Molecular modelling is becoming an increasingly powerful tool for modern 'rational' drug discovering and drug design. The last ten years saw a remarkable advancement in the knowledge of protein three-dimensional structures at the atomic level, notably for proteins playing essential biological functions such as enzymes, ion channels and transporters. Therefore, it has become possible to rationally target specific proteins with small chemical compounds, potentially identifying new lead drugs. In the present issue, Ghalem and Abdelli present a short illustration on this subject, studying the interaction of a key human enzyme, B-secretase with potential inhibitors from Green Tea.

IDENTIFICATION OF β - SECRETASE

The cloning of a new human aspartic protease identified to be β -secretase was independently reported by five laboratories [9-13]. There was remarkable agreement from these works in the properties of this new protease, including its presence in the brain, its activity in an acidic environment, its cellular co-localization with APP possibly in endosomes, the cleavage at the β -secretase site of APP, and the faster cleavage of Swedish mutant of APP than the native APP. This and other evidence led to the conclusion that this new protease was indeed the long sought β -secretase. Since the name β -secretase did not appropriately reflect its protease nature, several new names were given. The name memapsin 2 [13] is rooted in the "MEMbrane Aspartic Protease of the pepSIN family" and the number 2 distinguishes it from a homologous protease memapsin 1. This name conforms with the nomenclature of aspartic proteases recommended by the IUBMB Enzyme Nomenclature [14], e.g., pepsin, gastricsin, cathepsin etc. Other names used were BACE 1, an acronym based on " β -site APPcleaving enzyme" [9], ASP-2 [10, 12] and 'membrane-bound aspartic protease' [11]. The structures of β -secretase with amino acids of the active site are represented in the (Figures 2 and 3).

THEORETICAL APPROCH

In the course of searching for BACE1 (β -secretase) inhibitors from natural products, the ethyl acetate soluble fraction of green tea, witch was suspected to be rich in catechin content, showed potent inhibitory activity. The structures of the inhibitors [15] are represented in the (Figure 4).

- (1) : (+)-Catechin
- (2) : (-)-Catechin
- (3) : (+)-Epicatechin
- (4) : (-)-catechin gallate.

All calculations were carried with Gaussion 03W, version 6.0 [16]. A conformational analysis was performed, for inhibitors, on all rotatable bonds using an option "scan" in the EMO program, and then the found structures were minimized using Molecular Force Field EMO (MM2-

Allinger) and fully optimized at DFT level using hybrid B3LYP functional density with 6-31G* basis set and POP=NPA [17-20].

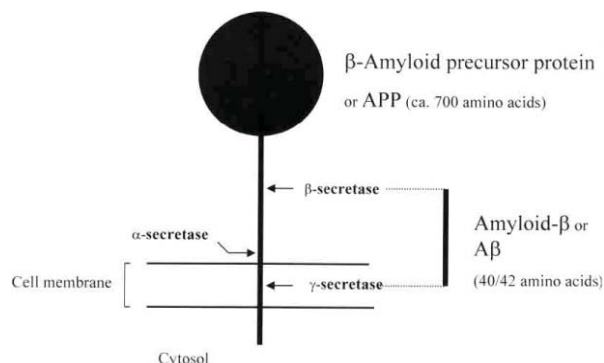


Figure 1 : Schematic presentation of the production of Aβ involving γ-, β -secretase (BACE1), the development of the Alzheimer's disease.

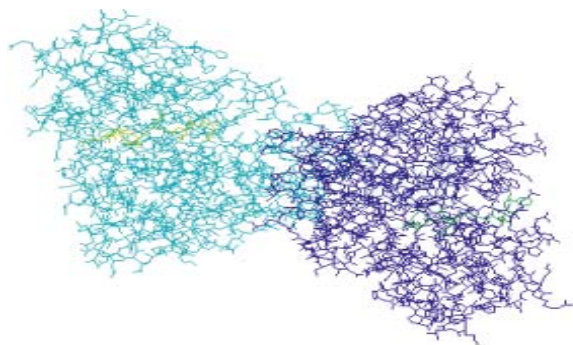


Figure 2: Structure of β-secretase (code pdb 1FKN)

Comparing these values of energy, we remark that the catechin gallate (which represent the lowest value of IC50= 6.0 10⁻⁶) has the low value of energy (-1676.1204571 eV). We can explain this difference in values by the structure of the inhibitors, we notices that all the inhibitors are cycles with 6 links which to differentiate by the groupings R1 and R2. Thus according to these resulted one can say the catechin gallate can present a better activity by contribution at the other inhibitors

CONCLUSION

This modest work consist in studying the inhibition of Bace1 which is an enzyme involved in the Alzheimer disease by methods of molecular modelling. Within sight of the results obtained and experimental values recorded in the bibliography, it arises that the catechin gallate would be probably the best to slow down the evolution of pathology (Alzheimer disease).

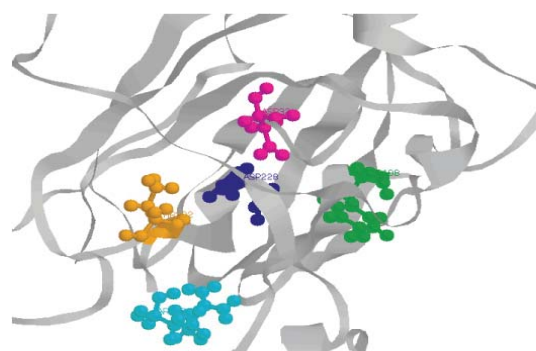


Figure 3: Amino acids of active site of β-secretase

RESULTS AND DISCUSSION

The IC₅₀ values are presented in the Table 1. To check the fundamental structure-activity relationships, the inhibitory activities of authentic catechins such as (+)-catechin (1), (-)-catechin (2), (+)-epicatechin (3), (-)-catechin gallate (4) were tested [15]. As a result 1, 2 and 3 showed relatively low activity regardless of their stereochemistry (Table 1). Therefore, at least, a pryrrogallol moiety in the catechin skeleton appeared to be essential for the stronger inhibitory activity (compound 4 IC₅₀=6.0 10⁻⁶). Of the purpose of our work which consists has to study this activity of various inhibitors by the theoretical methods one basing itself on the values of energy of activity.

INHIBITORS	IC ₅₀ (M)
(+) Catechin	3.5 10 ⁻⁵
(-) Catechin	3.0 10 ⁻⁵
(+) Epicatechin	2.8 10 ⁻⁵
(-) Catechin gallate	6.0 10 ⁻⁶

TABLE 1: IC₅₀ Values of catechins as BACE1 inhibitors

In the light of the results obtained during our work, it arises that the values obtained relating to energies Table2 are in the same order of magnitude for all inhibitors.

INHIBITORS	ENERGY DFT (eV)
(+) Catechin	-1031.041268
(-) Catechin	-1031.0400527
(+) Epicatechin	-1031.0387096
(-) Catechin gallate	-1676.1204571

TABLE 2: Energy Values of catechins as BACE1 inhibitors

REFERENCES

- [1a] J.L. Cummings, G. Cole, "Alzheimer Disease", 2002, JAMA, 287: 2335-8.
- [1b] DT. Villard, JC. Morris, "The diagnosis of Alzheimer's", 1998, Alz. Dis Re, 3: 142-52.
- [2] DJ. Selkoe, "Alzheimer's disease: genes, proteins, and therapy", 2001, Physiological Reviews, 81: 741-66.
- [3] DJ. Selkoe, P.J.R. Lansbury, "Biochemistry of Alzheimer's disease and prion diseases". 1999, In: G.J. Siegel, B.W. Agranoff, R.W. Albers, S.K. Fisher, M.D. Uhler 6th Ed, Basic Neurochemistry, Molecular, Cellular and Medical Aspects, Philadelphia, Lipincott-Raven, 949-68.
- [4] R. Brookmeyer, S. Gray, C. Kawas., "Projections of Alzheimer's disease in the United States and the public health impact of delaying disease onset", 1998, Am J public Health, 88: 1337-42.
- [5] D.J. Selkoe, Nature 399A, 1999, 23-31.
- [6] H. Cai, Y. Wang, D. McCarthy, H. Wen, D.R. Borchelt, D.L. Price, P.C. Wong, Nature Neuroscience, 2001, 4: 233-234.
- [7] Y. Luo, B. Bolon, S. Kahn, B.D. Bennett, S. Babu-Khan, P. Denis, W. Fan, H. Kha, J. Zhang, Y. Gong, L. Martin, J. Louis, Q. Yan, W.G. Richards, M. Citron, R. Vassar, Nature Neuroscience, 2001, 4: 231-232.
- [8] S.L. Roberds, J. Anderson, G. Basi, M.J. Bienkowski, D.G. Branstetter, K.S. Chen, S.B. Freedman, N.L. Frigon, D. Games, K. Hu, K. Johnson-Wood, K.E. Kappenman, T.T. Kawabe, I. Kola, R. Kuehn, M. Lee, W. Liu, R. Motter, N.E. Nichols, M. Power, D.W. Robertson, D. Schenk, M. Schoor, G.M. Shopp, M.E. Shuck, S. Sinha, K.A. Svensson, G. Tatsuno, H. Tintrup, J. Wijsman, S. Wright, L. McConlogue, Human Mol. Genetics, 2001, 10: 1317-24.
- [9] S.L. Roberds, J. Anderson, G. Basi, M.J. Bienkowski, D. G. Branstetter, K.S. Chen, S.B. Freedman, N.L. Frigon, D. Games, K. Hu, K. Johnson-Wood, K.E. Kappenman, T.T. Kawabe, I. Kola, R. Kuehn, M. Lee, W. Liu, R. Motter, N.E. Nichols, M. Power, D.W. Robertson, D. Schenk, M. Schoor, G.M. Shopp, M.E. Shuck, S. Sinha, K.A. Svensson, G. Tatsuno, H. Tintrup, J. Wijsman, S. Wright, L. McConlogue, 2001 Human Mol. Genetics, 10: 1317-24.
- [10] R. Yan, M.J. Bienkowski, M.E. Shuck, H. Miao, M.C. Tory, A.M. Pauley, J.R. Brashier, N.C. Stratman, W.R. Mathews, A.E. Buhl, D.B. Carter, A.G. Tomasselli, L.A. Parodi, R.L. Heinrikson, M.E. Gurney, 1999, Nature, 402: 533-7.
- [11] S. Sinha, J. P. Anderson, R. Barbour, G.S. Basi, R. Caccavello, D. Davis, M. Doan, H.F. Dovey, N. Frigon, J. Hong, K. Jacobson-Croak, N. Jewett, P. Keim, J. Knops, I. Lieberburg, M. Power, H. Tan, G. Tatsuno, J. Tung, D. Schenk, P. Seubert, S.M. Suomensari, S. Wang, D. Walker and V. John, 1999, Nature, 402: 537-40.
- [12] J. Hussain, D. Powell and D.R. Howlett, 1999, Mol. Cell Neurosci, 14: 419-27.
- [13] X. Lin, G. Koelsch, S. Wu, D. Downs, A. Dashti, J. Tang, 2000, Proc. Natl. Acad. Sci. USA, 97: 1456-60.
- [14] International Union of Biochemistry, Molecular Biology Enzyme Nomenclature, Academic Press, San Diego; also at website: www.chem.qmw.ac.uk/iubmb/enzyme/EC34/3423.html#23 1992.
- [15] So-Young. Jeon, KiHawan. Bae, Yeon-Hee. Seong and Kyung-Stik. Song, "Green Tea Catechins as a BACE1 (=Secretase) Inhibitor", 2003, Bioorganic& Medicinal Chemistry Letters, 13: 3905-3908.
- [16] M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, J.A.Jr. Montgomery, T. Vreven, K.N. Kudin, J.C. Burant, J.M. Millam, S.S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G.A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J.E. Knox, P.H. Hratchian, J.B. Cross, C. Adamo, J. Jaramillo, R. Gomperts, R.E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J.W. Ochterski, P.Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J.J. Dannenberg, V.G. Zakrzewski, S. Dapprich, D.A. Daniels, M.C. Strain, O. Farkas, D.K. Malick, A.D. Rabuck, K. Raghavachari, J.B. Foresman, J.V. Ortiz, Q. Cui, A.G.

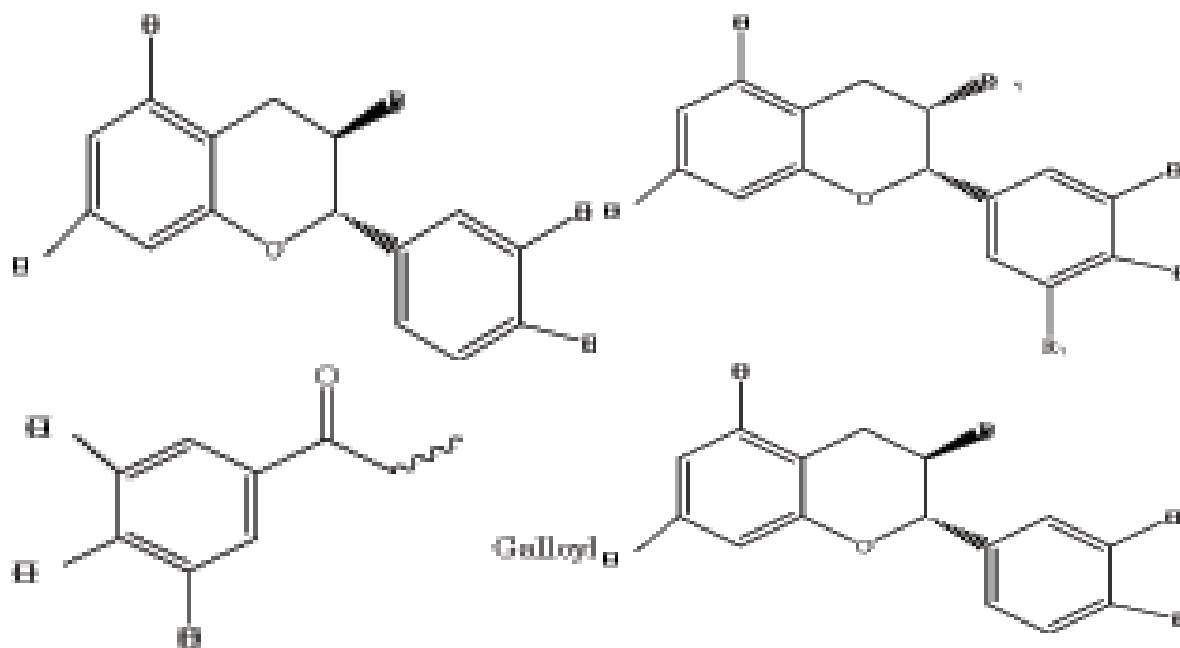


Figure 4: Structures of the various inhibitors. A: 1: R1=H, R2=H, 2 : Enantiomer of 7; B: 3: R1=H, R2=H

[11] S. Sinha, J. P. Anderson, R. Barbour, G.S. Basi, R. Caccavello, D. Davis, M. Doan, H.F. Dovey, N. Frigon, J. Hong, K. Jacobson-Croak, N. Jewett, P. Keim, J. Knops, I. Lieberburg, M. Power, H. Tan, G. Tatsuno, J. Tung, D. Schenk, P. Seubert, S.M. Suomensaari, S. Wang, D. Walker and V. John, 1999, *Nature*, 402 : 537-40.

[12] J. Hussain, D. Powell and D.R. Howlett, 1999, *Mol. Cell Neurosci*, 14: 419-27..

[13] X. Lin, G. Koelsch, S. Wu, D. Downs, A. Dashti, J. Tang, 2000, *Proc. Natl. Acad. Sci. USA*, 97: 1456-60.

[14] International Union of Biochemistry, *Molecular Biology Enzyme Nomenclature*, Academic Press, San Diego; also at website: www.chem.qmw.ac.uk/iubmb/enzyme/EC34/3423.html#23 1992.

[15] So-Young. Jeon, KiHawan. Bae, Yeon-Hee. Seong and Kyung-Stik. Song, "Green Tea Catechins as a BACE1 (–Secretase) Inhibitor", 2003, *Bioorganic& Medicinal Chemistry Letters*, 13: 3905-3908.

[16] M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, J.A.Jr. Montgomery, T. Vreven, K.N. Kudin, J.C. Burant, J.M. Millam, S.S. Iyengar, J. Tomasi, V.Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G.A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J.Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J.E. Knox, P.H. Hratchian, J.B. Cross, C. Adamo, J. Jaramillo, R. Gomperts, R.E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J.W. Ochterski, P.Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J.J. Dannenberg, V.G. Zakrzewski, S. Dapprich, D.A. Daniels, M.C. Strain, O. Farkas, D.K. Malick, A.D. Rabuck, K. Raghavachari, J.B. Foresman, J.V. Ortiz, Q. Cui, A.G. Baboul, S. Clifford, J. Cioslowski, B.B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R.L. Martin, D.J. Fox, T. Keith, M.A. Al-Laham, C.Y. Peng, A. Nanayakkara, M. Challacombe, P.M.W. Gill, B. Johnson, W. Chen, M.W. Wong, C. Gonzalez, J.A. Pople, Gaussian 03, Revision C.02; Gaussian: Wallingford, CT, 2004.

[17] A. Bouraoui, M. Fathallah, B. Blaive and R. Gallo, 1990, *J. Chem. Soc. Perkin Trans. II*, Vol. V (1211).

[18] A. Bouraoui, M. Fathallah, F. M. Henni, B. Blaive and R. Gallo, in ed.: J.L. Rivail, "Modelling of Molecular Structures and Properties", 1989, *Proceeding of an International Meeting Nancy, France*,

11,15, September, *Studies in Physical and Theoretical Chemistry*, 71: 381.

[19] A. Zinelabidine, A. Bouraoui, M. Fathallah, F. M. Henni, B. Blaive and R. Gallo, 1993, *J. Mol. Struct. (THEOCHEM)*, 286: 267.

[20] B. Blaive, G. Legsaí and R. Laí, 1995, *J. Mol. Struct.* 354: 245.

Quantum Chemical Studies on N-Donors Based-Pyrazole Compounds as Corrosion Inhibitors for Steel in Acidic Media

AHMED ATTAYIBAT[†], RACHID TOUZANI[‡]§, SMAAIL RADI[†], SGHIR EL KADIRI[‡]
SOUÂD SARI, IMÈNE ABDELLI and SAID GHALEM*

Laboratoire de Chimie Organique, Substances Naturelles et Analyses (COSNA)

Département de Chimie, Faculté des Sciences, Université de

Tlemcen, BP 119, Tlemcen 13000, Algeria

E-mail: s_ghalem2002@yahoo.fr

In this study, some possible relationship between the experimental inhibition corrosion in acidic media and the theoretical energy calculations for four series of compounds containing pyrazoles: (A) the first series of compounds, comprises only one pyrazole ring, (B) the second series has two pyrazoles (bipyrazole), (C) the third one contains one pyrazole and one pyridine (pyridyl pyrazole) and (D) the last series concerns tripodal pyrazoles, has presented. These sets of compounds have been tested for their corrosion inhibition properties of steel in low concentration of hydrochloric acid medium. For these compounds, DFT method studies have been performed. They were based on (E_{HOMO}), (E_{LUMO}) and the difference ($E_{LUMO}-E_{HOMO}$). The correlation between these two parameters (theoretical and experimental) was shown to be in good agreement for some cases.

Key Words: Pyrazole, Bipyrazole, Pyridylpyrazole, Tripods, Corrosion inhibitors, Quantum chemical calculations.

INTRODUCTION

Pyrazole-based ligands form a variety of coordination complexes with a number of metals ions, providing varying coordination geometry and nuclearity¹. The key feature of these heterocycles is their π -electron deficiency. Hence they behave as excellent π -acceptors which provide soft sites for metal coordination. On the other hand, the π -excessive five membered nitrogen heterocycle, pyrazole is a poorer π -acceptor. In fact, it is a better π -donor and hence acts as hard donor site. Recently, the systems based on nitrogen have known considerable interest in the designing of various

[†]Laboratoire de Chimie Organique, Macromoléculaire et Produits Naturels, Faculté des Sciences, Oujda, Morocco.

[‡]Laboratoire de Chimie Appliquée et Environnement, Faculté des Sciences, Oujda, Morocco.

[§]Faculté Pluridisciplinaire de Nador, Selouane, Morocco.

pyrazole ligands and their properties²⁻⁴. The derivatives containing pyrazole appeared to be good corrosion inhibitors of steel in acidic medium⁵⁻⁹. These organic molecules can be adsorbed on the metal surface forming a connection between the electron pair and the nitrogen atoms or the electronic orbital π cloud reducing the corrosive attack of metal in acidic medium^{10,11}. In addition, more recently, several corrosion inhibition semi-empirical studies have been realized with a view to establish a correlation between the experimental inhibition efficiency and the theoretical data¹²⁻¹⁵. The electronic properties of the inhibitors, the effect of energies E_{LUMO} and E_{HOMO} and the hydrophilic and hydrophobic nature of these organic compounds were brought back in these studies. In this study, a theoretical study of the electronic and energy properties E_{LUMO} and E_{HOMO} as well as the difference ($E_{LUMO}-E_{HOMO}$) of the constituents of four series of compounds containing pyrazole are reported. A comparison between the inhibiting efficiencies and theoretical calculations for each series of compounds will be presented.

METHOD OF CALCULATION

Theoretical calculations were carried out by using the Gaussian 03W software, version 6.0 on a PC (Pentium IV)¹⁶. The conformational analysis of the various geometries was performed with the help of the EMO program (Energy of Molecule) using the option SCAN. The minimization of the structures obtained was carried out using the EMO program (Champ de force MM2-Allinger) and finally we have used the DFT (DFT/B3LYP/6-31G* POP=NPA)¹⁷⁻²⁰.

RESULTS AND DISCUSSION

In order to obtain more information about a possible correlation between the corrosion inhibition efficiency of the ligands in acidic medium and the quantum theoretical chemical calculation parameters, four different series of compounds containing pyrazole (A, B, C and D) were selected (Fig. 1). These compounds were synthesized according to the literature methods. The first series (A) contains L1 and L2²¹; the second series (B) has two pyrazole groups (bipyrazole)²²; the third series (C) possesses one pyrazole and one pyridine (pyridylpyrazole)^{23,24} and the last series (D) contains tripodal compounds²⁵. The theoretical quantum calculations were performed taking compounds shown in Fig. 2 as references. Their corrosion inhibition efficiency (E (%)) of steel in 1 M HCl medium has been performed in our laboratory using 10^{-3} M as concentration of the ligand inhibitor, by two different methods: gravimetric and electrochemical. The experimental results are given on Table-1. When comparing the [E (%)] for the first series (A), the ligand L2 (having hydroxyl group) shows a better inhibition efficiency than the ligand L1 (having ester group), this may be due to the electron

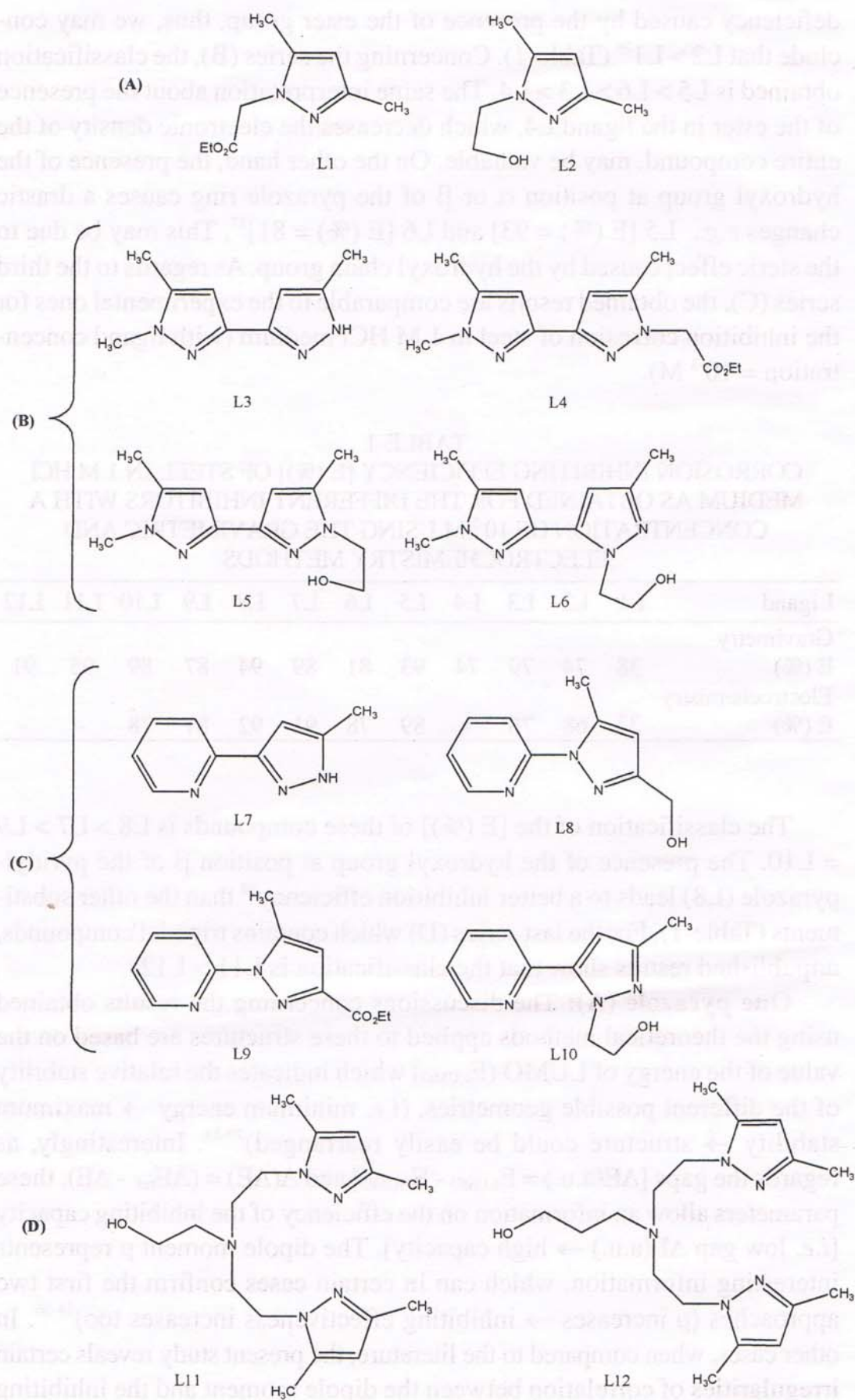
deficiency caused by the presence of the ester group, thus, we may conclude that $L2 > L1$ ²⁶ (Table-1). Concerning the series (B), the classification obtained is $L5 > L6 > L3 > L4$. The same interpretation about the presence of the ester in the ligand L4, which decreases the electronic density of the entire compound, may be valuable. On the other hand, the presence of the hydroxyl group at position α or β of the pyrazole ring causes a drastic changes *e.g.*, L5 [E (%) = 93] and L6 [E (%) = 81]²⁷. This may be due to the steric effect caused by the hydroxyl chain group. As regards to the third series (C), the obtained results are comparable to the experimental ones for the inhibition corrosion of steel in 1 M HCl medium (with ligand concentration = 10^{-3} M).

TABLE-1
CORROSION INHIBITING EFFICIENCY [E (%)] OF STEEL IN 1 M HCl
MEDIUM AS OBTAINED FOR THE DIFFERENT INHIBITORS WITH A
CONCENTRATION OF 10^{-3} M USING THE GRAVIMETRIC AND
ELECTROCHEMISTRY METHODS

Ligand	L1	L2	L3	L4	L5	L6	L7	L8	L9	L10	L11	L12
Gravimetry												
E (%)	38	74	79	74	93	81	89	94	87	89	95	91
Electrochemistry												
E (%)	32	68	76	-	89	78	91	92	81	78	-	-

The classification of the [E (%)] of these compounds is $L8 > L7 > L9 = L10$. The presence of the hydroxyl group at position β of the pyridyl-pyrazole (L8) leads to a better inhibition efficiency²⁸ than the other substituents (Table-1). For the last series (D) which contains tripodal compounds, unpublished results show that the classification is $L11 > L12$.

One pyrazole (A): The discussions concerning the results obtained using the theoretical methods applied to these structures are based on the value of the energy of LUMO (E_{LUMO}) which indicates the relative stability of the different possible geometries, (*i.e.* minimum energy \rightarrow maximum stability \rightarrow structure could be easily rearranged)²⁹⁻³³. Interestingly, as regards the gaps [$\Delta E(\text{a.u.}) = E_{LUMO} - E_{HOMO}$] and $\Delta(\Delta E) = (\Delta E_{\text{ref.}} - \Delta E)$, these parameters allow an information on the efficiency of the inhibiting capacity [*i.e.* low gap $\Delta E(\text{a.u.}) \rightarrow$ high capacity]. The dipole moment μ represents interesting information, which can in certain cases confirm the first two approaches (μ increases \rightarrow inhibiting effectiveness increases too)³⁴⁻³⁶. In other cases, when compared to the literature, the present study reveals certain irregularities of correlation between the dipole moment and the inhibiting capacity³⁷⁻³⁹.



On the basis of results obtained at the time of present work, the classification may be $L2 > L1$ (Table-2): Compared to Pyra-ref (Fig. 2). Compared to the effect of grafting of groupings on Pyra-ref and $\Delta E(\text{Pyra-ref}) > \Delta E(L1) > \Delta E(L2)$. For this series, it is concluded that a satisfactory correlation between the experimental studies and the theoretical calculations.

TABLE-2
THEORETICAL CALCULATION FOR THE SERIES (A)

	E_{HOMO} (a.u.)	E_{LUMO} (a.u.)	ΔE (a.u.)	μ (D)	$\Delta(\Delta E)$ (a.u.)
Pyra-ref	-0.24724	0.01398	0.26122	2.4676	–
L1	-0.23029	-0.01204	0.21825	3.6667	0.04297
L2	-0.11035	0.09177	0.20212	3.4014	0.05910

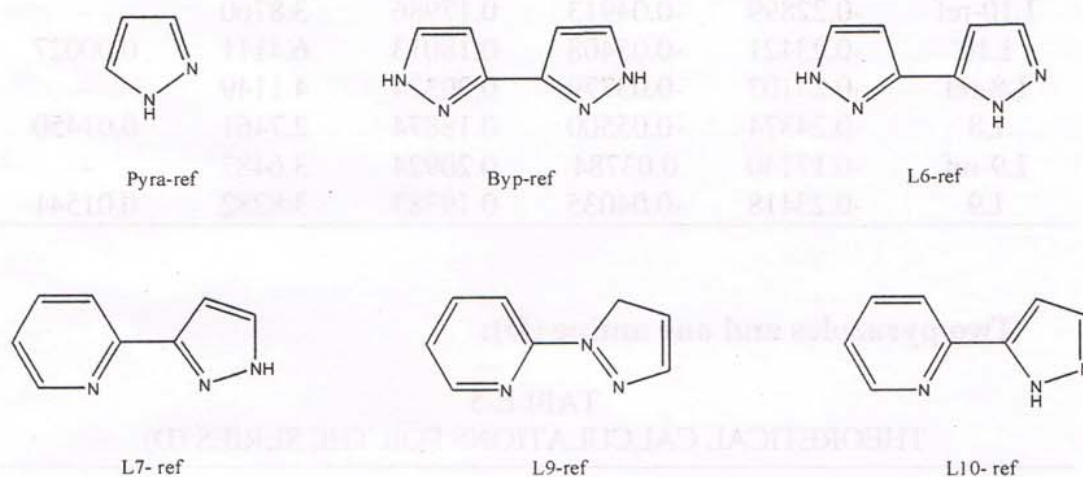


Fig. 2. References products

Two pyrazoles (bipyrazole) (B): From Table-3, it is concluded that the classification of the ligands from the lowest ΔE (a.u.) (good efficiency) to the highest is $L4 > L6 > L5 > L3$. The comparison of the experimental and the theoretical results shows that the correlation in this case is not quite satisfactory. Indeed, the compound L5 with a hydroxyl group should be more efficient than the compound L4.

TABLE-3
THEORETICAL CALCULATION OF SERIES (B)

	E_{HOMO} (a.u.)	E_{LUMO} (a.u.)	ΔE (a.u.)	μ (D)	$\Delta(\Delta E)$ (a.u.)
Byp-ref	-0.21883	-0.00243	0.21640	0.0009	–
L3	-0.20425	0.00631	0.21056	0.4760	0.00584
L4	-0.20356	-0.01074	0.19282	2.5076	0.02358
L5	-0.21113	-0.00588	0.20525	3.0433	0.01115
L6-ref	-0.21903	-0.01968	0.19935	4.9509	–
L6	-0.20956	-0.01136	0.19820	4.3932	0.00115

One pyrazole and one pyridine (pyridylpyrazole) (C): The theoretical calculations for the third series which contains one pyrazole and one pyridine (Table-4) show that the classification of the ligands in the same series is as follow: L9 > L8 > L10 > L7. The examination of the experimental and theoretical results leads to a better correlation for compound L8 than the rest of the series. This is not the case for the ligand L7 which exhibit a poorer correlation.

TABLE-4
THEORETICAL CALCULATION OF SERIES (C)

	E_{HOMO} (a.u.)	E_{LUMO} (a.u.)	ΔE (a.u.)	μ (D)	$\Delta(\Delta E)$ (a.u.)
L7-ref	-0.16140	0.04838	0.20978	1.7088	-
L7	-0.22421	-0.02881	0.19540	1.8955	0.01438
L10-ref	-0.22899	-0.04913	0.17986	3.8760	-
L10	-0.23421	-0.05408	0.18013	6.4111	0.00027
L8-ref	-0.24103	-0.03779	0.20324	4.1149	-
L8	-0.24374	-0.05500	0.18874	2.7461	0.01450
L9-ref	-0.17140	0.03784	0.20924	3.6487	-
L9	-0.23418	-0.04035	0.19383	3.8282	0.01541

Two pyrazoles and one amine (D):

TABLE-5
THEORETICAL CALCULATIONS FOR THE SERIES (D)

	E_{HOMO} (a.u.)	E_{LUMO} (a.u.)	ΔE (a.u.)	μ (D)
L11	-0.21445	0.01844	0.23289	1.5178
L12	-0.20378	0.02072	0.22450	4.1200

As can be observed from Table-5, the ligand L12 is better than the ligand L11 (L12 > L11). In addition, the correlation between the experimental results using the gravimetric method and the theoretical predictions are not in good agreement.

Conclusion

As a conclusion, the correlation between experimental and predicted results has been more limited than would have been desirable for almost the entire series under investigation. Thus, a general rule was not possible to establish according to the present results. Nevertheless, a quite satisfactory correlation was obtained for the ligand L8 belonging to the series (C). It is worthnoting that this type of molecule have been shown to exhibit antitumor properties which encourages us to project a similar study in order to correlate their biological activity with the theoretical calculations.

ACKNOWLEDGEMENT

The authors thank Professor Bruno Blaive, University of Marseille, France for providing the molecular mechanics program EMO.

REFERENCES

1. R. Mukherjee, *Coord. Chem. Rev.*, **203**, 151 (2000).
2. S. Trofimenko, *J. Am. Chem. Soc.*, **88**, 1842 (1966).
3. S. Trofimenko, *Chem. Rev.*, **93**, 943 (1993).
4. E.C. Constable and P.J. Steel, *Coord. Chem. Rev.*, **93**, 205 (1989).
5. A. El-Ouafi, B. Hammouti, H. Oudda, S. Kertit, R. Touzani and A. Ramdani, *Anti-Corros. Method. M.*, **49**, 199 (2002).
6. A. Dafali, B. Hammouti, R. Touzani, S. Kertit, A. Ramdani and K. El-Kacemi, *Anti-Corros. Method. M.*, **49**, 96 (2002).
7. M. Benabdellah, R. Touzani, A. Dafali, B. Hammouti and S. El Kadiri, *Mater. Lett.*, **61**, 1197 (2007).
8. K. Tebbji, H. Oudda, B. Hammouti, M. Benkaddour, M. El-Kodadi, F. Malek and A. Ramdani, *Appl. Surf. Sci.*, **241**, 326 (2005).
9. M. Benabdellah, R. Touzani, A. Aouniti, A. Dafali, S. El Kadiri, B. Hammouti and M. Benkaddour, *Mater. Chem. Phys.*, **105**, 373 (2007).
10. G. TrabANELLI, *Chemical Industries Corrosion Mechanism*, Marcel Dekker, New York, Vol. 3 (1987).
11. G. Schmitt, *Br. Corros. J.*, **19**, 165 (1984).
12. A.I. Altsybiera, S.Z. Levin and A.P. Dorokhov, *Third European Symposium of Corrosion Inhibitors*, University of Ferrara, Ferrara Italy, p. 577 (1971).
13. G. Lewis, *Corrosion*, **38**, 62 (1982).
14. J. Fang and J. Li, *J. Mol. Struct. (Theochem.)*, **593**, 197 (2002).
15. N. Khalil, *Electrochim. Acta*, **48**, 2635 (2003).
16. M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, J.A.Jr. Montgomery, T. Vreven, K.N. Kudin, J.C. Burant, J.M. Millam, S.S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G.A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J.E. Knox, P.H. Hratchian, J.B. Cross, C. Adamo, J. Jaramillo, R. Gomperts, R.E. Stratmann, O. Yazyev, A.J. Austin, R. Cammi, C. Pomelli, J.W. Ochterski, P.Y. Ayala, K. Morokuma, G.A. Voth, P. Salvador, J.J. Dannenberg, V.G. Zakrzewski, S. Dapprich, D.A. Daniels, M.C. Strain, O. Farkas, D.K. Malick, A.D. Rabuck, K. Raghavachari, J.B. Foresman, J.V. Ortiz, Q. Cui, A.G. Baboul, S. Clifford, J. Cioslowski, B.B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R.L. Martin, D.J. Fox, T. Keith, M.A. Al-Laham, C.Y. Peng, A. Nanayakkara, M. Challacombe, P.M.W. Gill, B. Johnson, W. Chen, M.W. Wong, C. Gonzalez and J.A. Pople, *Gaussian 03, Revision C.02; Gaussian: Wallingford. CT* (2004).
17. A. Bouraoui, M. Fathallah, B. Blaive and R. Gallo, *J. Chem. Soc. Perkin Trans. II*, 1211 (1990).
18. A. Bouraoui, M. Fathallah, F.M. Henni, B. Blaive and R. Gallo, in ed.: J.L. Rivail, *Modelling of Molecular Structures and Properties*, Proceeding of an International Meeting Nancy, France, 11-15, September, *Studies in Physical and Theoretical Chemistry*, Vol. 71, p. 381 (1989).
19. A. Zinelabidine, A. Bouraoui, M. Fathallah, F.M. Henni, B. Blaive and R. Gallo, *J. Mol. Struct. (Theochem.)*, **286**, 267 (1993).
20. B. Blaive, G. Legsaï and R. Lai, *J. Mol. Struct.*, **354**, 245 (1995).

21. W.G. Haanstra, W.L. Driessen, J. Reedijk, U. Turpeinen and R. Hämäläinen, *J. Chem. Soc., Dalton Trans.*, 2309 (1989).
22. A. Attayibat, S. Radi, Y. Lekchiri, A. Ramdani, B. Hacht, M. Morcellet, M. Bacquet and S. Willai, *J. Chem. Res.*, 655 (2006).
23. A. Satake and T. Nakata, *J. Am. Chem. Soc.*, **120**, 10391 (1998).
24. F. Malek, A. Ramdani and S. Radi, *J. Chem. Res.*, 640 (2004).
25. S. Radi, A. Attayibat, A. Ramdani, Y. Lekchiri, B. Hacht, M. Morcellet, M. Bacquet and S. Willai, *Bull. Korean Chem. Soc.*, **28**, 474 (2007).
26. M. Bouklah, B. Hammouti, M. Benkaddour, A. Attayibat and S. Radi, *Pigm. Resin Technol.*, **34**, 197 (2005).
27. M. Bouklah, A. Attayibat, B. Hammouti, A. Ramdani, S. Radi and M. Benkaddour, *Appl. Surf. Sci.*, **240**, 341 (2005).
28. K. Tebbji, H. Oudda, B. Hammouti, M. Benkaddour, M. El Kodadi and A. Ramdani, *Colloids and Surface A: Physiochem. Eng. Asp.*, **259**, 143 (2005).
29. M. Salazar-Villanueva, P.H.H. Tejada, U. Pal, J.F. Rivas-Silva, J.I.R. Mora and J.A. Ascencio, *J. Phys. Chem. A.*, **110**, 10274 (2006).
30. D. Zhang, G. Guo, C. Liu and R.Q. Zhang, *J. Phys. Chem. B.*, **110**, 23633 (2006).
31. M.-X. Chen, X.H. Yan and S.H. Wei, *J. Phys. Chem. A.*, **111**, 8659 (2007).
32. K. Ito, Z. Chen, C. Corminboeuf, C.S. Wannere, X.H. Zhang, Q.S. Li and P.V.R. Schleyer, *J. Am. Chem. Soc.*, **129**, 1510 (2007).
33. P. Gutta and D.J. Tantillo, *Angew. Chem. Int. Ed.*, **44**, 2719 (2005).
34. M. Lagrenée, B. Mernari, N. Chaibi, M. Traisnel, H. Vezin and F. Bentiss, *Corros. Sci.*, **43**, 951 (2001).
35. M.A. Quraishi and R. Sardar, *J. Appl. Electrochem.*, **33**, 1163 (2003).
36. L. Larabi, Y. Harek, O. Benali and S. Ghalem, *Prog. Org. Coat.*, **54**, 256 (2005).
37. K.F. Khaled, N. Babic-Samardzija and N. Hackerman, *Electrochim. Acta*, **50**, 2515 (2005).
38. G. Bereket, E. Hür and C. Ögretir, *J. Mol. Struct. (Theochem.)*, **578**, 79 (2002).
39. F. Touhami, A. Aouniti, Y. Abed, B. Hammouti, S. Kertit, A. Ramdani and K. El Kacemi, *Corros. Sci.*, **42**, 929 (2000).

(Received: 20 October 2007;

Accepted: 11 August 2008)

AJC-6746

**THE IUMRS INTERNATIONAL CONFERENCE IN ASIA 2008
(IUMRSICA 2008)**

9 — 13 DECEMBER 2008

NAGOYA, JAPAN

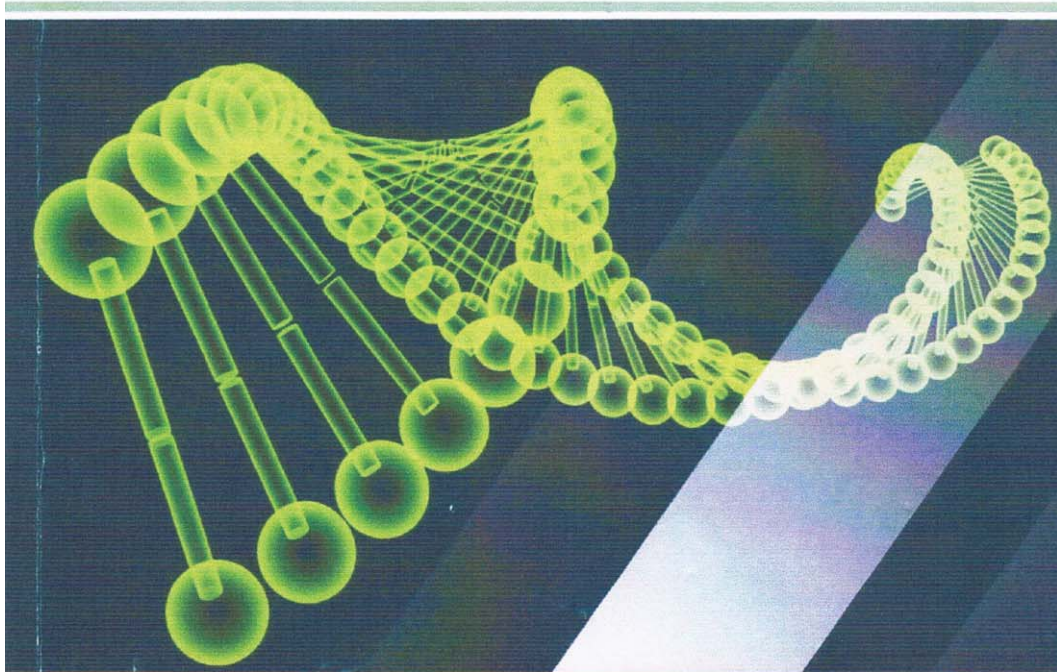
Contact:

Secretariat for IUMRS-ICA 2008, c/o Inter Group Corporation,
2-38-2 Meieki, Nakamura-ku, Nagoya 450-0002, Japan

E-mail: iumrsica2008@intergroup.co.jp;

Website: <http://www.iumrs-ica2008.jp/index.html>

EUE ÉDITIONS
UNIVERSITAIRES
EUROPÉENNES



Said Ghalem
Imene Abdelli
Souad Sari

Inhibition compétitive interaction enzyme- substrat

par modélisation moléculaire

ملخص

التفاعلات البروتين-البروتين لها أهمية كبيرة في المجالات البيولوجية. وتشارك بشكل كبير في تشكيل هياكل الجزيئات في إشارة و تنظيم المسارات الأيضية المختلفة. الخصوصية العالية لهذه التفاعلات يجعلها مثالية لأهداف العوامل العلاجية.

تحركات البروتينات، على وجه الخصوص، تمثل واحدة من أكبر عائلات البروتينات، وترتبط مع معظم السرطانات، مثل الطفرات الجينية التي تؤدي إلى التعبير عن مدى أو تفعيل التأسيسي من هذه البروتينات

هدفنا هو دراسة من قبل وسائل تثبيط النمذجة الجزيئية مختلف التفاعلات A المضادة للالتهابات. مع tripeptides بين بروتين الكيناز البروتين.

المصطلحات

Tripeptides , برنامج الرص. التفاعلات بروتين-بروتين. النمذجة الجزيئية.

Résumé :

Les interactions protéine-protéine ont un aspect essentiel des processus biologiques. Elles sont fortement impliquées dans la formation de structures macromoléculaires, dans la signalisation, dans la régulation et dans les différentes voies métaboliques. La haute spécificité de ces interactions fait d'elles des cibles idéales pour les agents thérapeutiques.

Les protéines kinases, plus particulièrement, représentent une des plus grandes familles de protéines, La plupart des cancers sont associés au dérèglement de protéines kinases, tels que des mutations génétiques qui conduisent à une sur-expression ou à une activation constitutive de ces protéines.

Notre objectif de travail est d'étudier par les méthodes de modélisation moléculaire l'inhibition de la protéine kinase A par des fragments de protéine (tripeptides) présentant une activité anti inflammatoire.

Mots clés : protéine kinase A, tr peptides, modélisation moléculaire, docking moléculaire.

Abstract:

The interactions protein-protein has an essential aspect of the biological processes. They are strongly implied in the formation of macromolecular structures, indication, the regulation and the various metabolic ways. The high specificity of these interactions makes they ideal targets for the therapeutic agents.

The proteins kinase, represent one of more large families of proteins. Most cancers are associated with the proteins kinases disordered state, such as genetic mutations which lead to an on-expression or an activation constitutive of these proteins.

Our objective of work is to study by the methods of molecular modeling the inhibition of the protein kinase A by protein fragments (three-peptides) presenting an anti inflammatory activity.

Key words: proteins kinases, three peptides, molecular modeling, molecular docking.