

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
Université des Sciences et de la Technologie Houari Boumediene
Faculté de CHIMIE



MEMOIRE

Présenté en vue de l'obtention du diplôme de MAGISTER

En Chimie

Spécialité : Chimie Physique et Théorique

Par

Yasmine ASSES

Sujet

*Modélisation et Étude Structurale de Ligands du Récepteur
périphérique de la Cholécystokinine : Screening Virtuel*

Soutenu le 11 / 11 / 2006, devant le jury composé de:

B. MAUCHE	Professeur	USTHB	Président
S. KELLOU-TAIRI	Maître de conférences	USTHB	Directrice de thèse
Y. BADJAH-H.-A.	Professeur	USTHB	Examineur
B. MAIGRET	Directeur de recherches au CNRS	Nancy	Examineur
Y. AKACEM	Maître de conférences	USTHB	Examinatrice

A mes parents

Remerciements

Les travaux présentés dans ce mémoire de magister ont été menés au laboratoire de Physico-Chimie Théorique et Chimie Informatique, sous la direction scientifique de Madame Safia TAIRI-KELLOU, Maître de conférences à l'USTHB et au sein de l'équipe de Dynamique des Assemblages Membranaires (eDAM, UMR 7565 UHP) de Nancy, dirigée par le Docteur Bernard MAIGRET (DR1 au CNRS). Ces travaux s'insèrent dans le cadre de l'accord programme TASSILI CMEP (Comité Mixte d'Evaluation et de prospection de la coopération interuniversitaire Franco-Algérienne). Nous tenons à remercier les autorités du CMEP pour avoir financé les stages de recherche au sein de l'eDAM de Nancy.

J'exprime ma sincère gratitude à Monsieur le Professeur Boubekur MAOUCHE pour m'avoir accueillie dans le laboratoire. Je le remercie pour le soutien qu'il m'a témoigné et pour l'honneur qu'il me fait en acceptant de juger mes travaux et de présider le jury.

J'adresse mes remerciements les plus attentionnés à Madame Safia TAIRI-KELLOU pour m'avoir initiée à la recherche dans un cadre studieux et, pour m'avoir donné l'opportunité de découvrir un domaine passionnant : la biochimie théorique. Sa confiance, sa disponibilité, ses conseils avisés ainsi que sa rigueur scientifique ont été salutaires à l'aboutissement de ce travail.

Je tiens à remercier tout particulièrement le Docteur Bernard MAIGRET, Directeur de recherche au CNRS, pour m'avoir accueillie au sein de son équipe et d'avoir mis à ma disposition les moyens nécessaires à l'avancement de ces travaux. Je lui exprime ma profonde reconnaissance pour l'aide précieuse et l'intérêt qu'il a accordé en jugeant ce présent travail malgré ses nombreuses occupations.

J'adresse mes remerciements les plus respectueux à Monsieur Yacine BADJAH-HADJ-AHMED, Professeur à la faculté de chimie de l'USTHB pour l'honneur qu'il me fait en acceptant de juger mes travaux.

Je remercie profondément Madame Yamina AKACEM, Maître de conférences à la faculté de chimie de l'USTHB, d'avoir accepté de faire partie du jury. Je suis particulièrement ravie par cet honneur.

Mes remerciements les plus chaleureux s'adressent à Madame Salima EL AOUI, Chargée de cours à la Faculté des Sciences Biologiques, pour sa collaboration et son agréable sympathie.

J'associe à ces remerciements l'ensemble de mes enseignants et le personnel de la faculté de chimie de l'USTHB. Que Monsieur le Professeur Chaabane CHELGHOUM trouve ici, l'expression de mes vifs remerciements pour sa disponibilité.

Je tiens à remercier également l'ensemble de mes collègues du laboratoire de Physico-Chimie Théorique et Chimie Informatique, plus particulièrement Amel Toumi-Maouche, Amel Kadari, Karima Deramchi, Souhila Terrachet-Bouaziz et Yamina Belmiloud pour leur sincère sympathie et leur soutien durant ce parcours.

Ce travail étant le fruit d'une collaboration, je ne saurais oublier d'adresser ma gratitude la plus sincère à Madame Mercedes MARTIN-MARTINEZ de Madrid pour m'avoir permis d'utiliser son modèle de récepteur périphérique de la cholécystokinine et également à tous les membres de l'eDAM pour l'immense soutien, scientifique et humain, qu'ils m'ont apporté durant mes stages ; la liste est longue... mais, merci à Alex, Jean-Paul, Matthieu et Vincent.

Je remercie Monsieur Alain CARTIER pour sa bienveillance lors de mes séjours à Nancy.

L'amitié à toute épreuve et la constante présence de mes amies : Chichette, Dalila, Karima, Radia et Sadjia méritent un regard bien particulier et une attention de ma part dignes de ce qu'elles m'ont témoigné.

Enfin, une immense pensée revient à ma famille au complet, en particulier mes parents, ma sœur et mon frère pour leur soutien sans concession et leurs inlassables encouragements au cours de mes études. Aucun mot au monde ne saurait exprimer l'étendue de ma reconnaissance envers vous.

SOMMAIRE

ABREVIATIONS

INTRODUCTION GENERALE	6
-----------------------------	---

CHAPITRE I *RESUME DE L'ETAT ACTUEL DES CONNAISSANCES SUR LA CHOLECYSTOKININE ET SES RECEPTEURS*

I- INTRODUCTION.....	10
II- LA CHOLECYSTOKININE	10
III- LES RECEPTEURS DE LA CHOLECYSTOKININE.....	11
IV- PRINCIPAUX ROLES PHYSIOLOGIQUES	12
1- Actions sur le système gastro-intestinal	12
2- Actions sur le système nerveux central et périphérique	13
A) La satiété	13
B) L'analgésie opioïde	13
C) Les troubles psychiques et l'anxiété.....	13
D) La mémoire et l'apprentissage	14
E) Conclusion.....	14
V- DESCRIPTION DES PHARMACOPHORES IMPORTANTS.....	14
VI- RESUME DE L'ETAT ACTUEL DES CONNAISSANCES SUR LES SITES DE LIAISONS DES ANTAGONISTES.....	17
VII- CONCLUSION	20
REFERENCES BIBLIOGRAPHIQUES	21

CHAPITRE II *MODELISATION MOLECULAIRE ET DRUG DESIGN*

I- INTRODUCTION.....	27
II- LA MECANIQUE MOLECULAIRE.....	28
A- Bases et principes	28
B- Fonction énergie potentielle	28
C- Méthodes d'exploration de l'hypersurface d'un système moléculaire.....	30
C1 – Algorithmes de minimisation	30
C2 - Rotation des angles de torsion	33
D- Méthodes statistiques	33

<i>D1 – Dynamique Moléculaire</i>	33
<i>D2 - Monté Carlo</i>	35
<i>D3 - Recuit Simulé</i>	36
E- Conclusion.....	36
III- METHODES DU DRUG DESIGN	37
A- QSAR (Quantitative Structure-Activity Relationships)	38
B- Le Criblage à Haut Débit ou High Throughput Screening (HTS)	39
B1- Le criblage réel.....	39
B2- Le criblage virtuel	40
<i>REFERENCES BIBLIOGRAPHIQUES</i>	50

CHAPITRE III *METHODOLOGIE*

I– INTRODUCTION.....	54
II- PRESENTATION DU LOGICIEL UTILISE.....	56
A - <i>Préparation des ligands</i>	56
A1- Création de bases de données moléculaires (Create)	57
A2- Filtrage de bases de données (Filter).....	59
A3- Manipulation de la base de données (Handle)	59
B- <i>Préparation de la cible</i>	60
C- <i>Entonnoir de criblage (Screening Funnel)</i>	61
III- PRESENTATION DE LA CIBLE BIOLOGIQUE.....	62
IV- PRESENTATION DE LA CHIMIOTHEQUE	64
A- Recherche des molécules par sous structure	65
B- Recherche conformationnelle.....	70
C- Comparaison des surfaces ligand/récepteur avec MSSH	72
C1- Calcul des fingerprints.....	72
C2- Comparaison des fingerprints.....	72
C3- Filtrage de ces molécules	73
D- Préparation des fichiers pour le docking moléculaire GOLD	73
E- Calcul des énergies d'interactions ligand/récepteur	73
F- Calcul des énergies libres de liaisons $\Delta G_{\text{binding}}$	74
<i>REFERENCES BIBLIOGRAPHIQUES</i>	75

CHAPITRE IV *RESULTATS, STRATEGIES ET DISCUSSION*

I – INTRODUCTION	78
------------------------	----

II – RESULTATS ET ANALYSE	78
A- MSSH.....	78
B- GOLD.....	80
C- Discover	82
III- DISCUSSION	83
IV– STRATEGIES ET DISCUSSION	92
A- Comparaison MSSH/GOLD	92
B- Comparaison GOLD/Discover.....	94
V- CONCLUSION.....	95
<i>REFERENCES BIBLIOGRAPHIQUES</i>	96
CONCLUSION GENERALE ET PERSPECTIVES	99
ANNEXES	101

ABBREVIATIONS

1D: Unidimensionnel

2D: Bidimensionnel

3D: Tridimensionnel

AC: Adénylcyclase

AMBER: Assisted Model Building with Energy Refinement

AMPC: Adénosine Monophosphate cyclique

ATP : Adénosine Triphosphate

BAPPL: Binding Affinity Prediction of Protein-Ligand

CADD: Computer Assisted Drug Design

CCDC: Cambridge Crystallographic Data Centre

CCK: Cholécystokinine

CCK₁: Récepteur périphérique de la cholécystokinine

CCK₂: Récepteur central de la cholécystokinine

CCK-8: Octapeptide de la cholécystokinine

CCK-8S: Octapeptide de la cholécystokinine Sulfatée

CFF91: Consistent Force Field 91

CHARMM: Chemistry at HARvard Molecular Mechanics

CoMFA: Comparative Molecular Field Approach

CORINA: COoRdINAtes

CPU: Central Processing Unit

CVFF : Consistent Valence Force Field

DM: Dynamique Moléculaire

eDAM : équipe Dynamique des Assemblages Membranaires

E_i: Energie d'interaction

eX : Boucle extracellulaire n° :X, avec X = 1,2,3.

fs: femtoseconde

GDP: Guanosine Diphosphate

GOLD: Genetic Optimization for Ligand Docking

GTP: Guanosine Triphosphate

HTS: High Throughput Screening

IC₅₀ : Concentration nécessaire à inhiber 50% de l'activité biologique d'un composé.

kDa: kilodalton

MC: Monté Carlo

MM: Molecular Mechanic (Mécanique Moléculaire)

MSSH: Molecular Simulation using Spherical Harmonics

NAMD: NANoscale Molecular Dynamics

nM: nanomole

ns: nanoseconde

OMEGA: Optimized Molecular Ensemble Generation Application

PC: Personal Computer

PDB: Protein Data Bank.

QSAR: Quantitative Structure-Activity Relationships

RCPG(s): Récepteurs Couplés aux Proteines G(stimulatrices)

RMN: Résonance Magnétique Nucléaire

RMSD: Root Mean Square Deviation

SMS: Sting Millenium Suite

TMX: Hélice transmembranaire n°X, avec X = I, II, III, IV, V, VI, VII.

VMD: Visual Molecular Dynamics

VSM-G: Virtual Screening Manager for Grids

ΔG_{bind} : Energie libre de liaison

INTRODUCTION GENERALE

De nos jours, la conception de nouvelles molécules thérapeutiques n'est plus le fruit du hasard mais repose sur des investissements lourds en labeurs, temps et argent. De plus, les critères, de plus en plus drastiques, d'acceptation de nouveaux médicaments compliquent considérablement le processus de recherche et de découverte d'un médicament. Par conséquent, l'optimisation des étapes du développement de substances thérapeutiques et l'interconnexion de différentes disciplines (biologie, pharmacochimie, chimie théorique et informatique) représentent une solution alternative pour diminuer les coûts et le temps de création de nouveaux médicaments.

Au cours du processus de recherche du médicament, le criblage qui s'apparente à une phase de repérage ou de tri systématique des banques de composés, est une technique permettant d'améliorer l'identification des composés prometteurs et ensuite leur optimisation en réduisant les coûts et en augmentant la probabilité de découverte de médicaments.

Le criblage à haut débit existe sous deux types qui diffèrent par les moyens employés : le criblage réel et le criblage *in silico* ou virtuel.

Le criblage réel ou expérimental, consiste à identifier parmi des centaines de milliers de molécules celles qui sont pourvues de propriétés biologiques les plus intéressantes. L'identification se fait au moyen d'un test *in vitro* qui consiste à mettre en présence la substance à tester et un système biochimique, puis à mesurer l'importance de la réaction éventuelle. Ces opérations qui évoluent avec les techniques de la robotique se déroulent de nos jours sans intervention humaine et sont pilotées par des logiciels adaptés au type d'analyse à effectuer. Cependant, ce procédé de criblage réel à haut débit est limité aussi bien par le nombre de composés possibles à tester que par le coût des tests qui est estimé à 1 dollar par molécule.

Pour remédier au problème du coût d'argent et de temps, les sociétés pharmaceutiques ont de plus en plus recours à l'outil informatique qui est de moins en moins cher et de plus en plus performant.

Ce type de criblage *in silico* ou virtuel mis en œuvre par la société Syn:tem, une société biopharmaceutique créée en 1995 et spécialisée dans le développement de nouvelles

molécules thérapeutiques pour les maladies du système nerveux central, consiste à rechercher dans une banque de composés ou une chimiothèque, des molécules complémentaires d'un site actif de la cible biologique, susceptibles de l'activer ou de l'inhiber.

Cette technique relativement récente dont il est question dans nos recherches contribue, *par une évaluation rationnelle de l'activité biologique des molécules*, à épargner le coût que nécessiterait une expérimentation "aveugle". En effet, grâce aux progrès en biologie moléculaire et en informatique, les chercheurs parviennent à simuler l'action de substances thérapeutiques, d'où la naissance des techniques du «drug design». Dans la quête des candidats-médicaments, le drug design utilisant les méthodes de la modélisation moléculaire comprend deux types d'approche : celle du «ligand-based » et celle du «receptor-based ». L'approche «ligand-based » utilise des techniques de calculs basées sur la connaissance du ligand, comme le QSAR (Quantitative Structure Activity Relationships) qui consiste à disposer d'une liste de molécules présentant une activité sur une cible donnée et ensuite à élaborer des modèles statistiques permettant de prédire l'activité d'une nouvelle molécule. L'autre approche : «receptor-based », basée sur la connaissance de la cible biologique, utilise les techniques du criblage virtuel à haut débit.

Les travaux menés et présentés dans ce mémoire, entrent dans le cadre d'une contribution à l'élucidation du mode d'interaction de la cholécystokinine (CCK), une hormone neuropeptidique impliquée dans un grand nombre d'actions physiologiques. Le but de nos recherches consiste dans une première partie, à rechercher des antagonistes du récepteur périphérique de la CCK en criblant virtuellement une base de données judicieusement choisie. La deuxième partie de nos travaux consiste à valider la stratégie employée dans le cadre de l'outil logistique effectuant le criblage.

A la suite d'une introduction générale présentant l'objectif des travaux menés, le présent mémoire comporte quatre chapitres suivis d'une conclusion générale et de trois annexes.

Le premier chapitre vise à introduire la cible moléculaire de nos recherches qui est la cholécystokinine. Largement décrite dans des travaux antérieurs⁽¹⁾, nous nous sommes limités à actualiser les connaissances, en résumant notre recherche bibliographique concernant les travaux les plus récents, portant sur les sites d'interactions relatifs au récepteur périphérique de la CCK.

Dans le second chapitre, nous décrivons les méthodes de la modélisation moléculaire, et plus particulièrement celles du drug design que nous avons largement utilisées dans le cadre de nos recherches.

En ce qui concerne le troisième chapitre, il est question de présenter la méthodologie que nous avons utilisée. La plate-forme que nous avons employée, Virtual Screening Manager for Grids (VSM-G), y est décrite ainsi que le détail des étapes de la stratégie adoptée.

Le quatrième chapitre est consacré à la présentation et à la discussion des résultats obtenus. Nous détaillons les résultats acquis, à l'issue de chaque étape du criblage, et nous présentons la stratégie employée pour la validation du protocole de criblage. Les résultats de la corrélation, entre les différents filtres du criblage effectué, sont également discutés⁽²⁾.

La conclusion générale porte sur les principaux apports de nos travaux. Nous présentons également les perspectives que nos résultats laissent entrevoir afin d'améliorer le processus de criblage virtuel de chimiothèques de grande taille.

Nous reportons en annexe 1 un tableau résumant les propriétés structurales et physico-chimiques essentielles des acides aminés.

Les annexes 2 et 3 contiennent les détails de nos résultats de docking moléculaire.

REFERENCES BIBLIOGRAPHIQUES

- 1- Taïri-Kellou, S., Thèse de Doctorat d'état, Conception Assistée par Ordinateur des Complexes ligand/récepteur de la Cholécystokinine, Alger, USTHB, N° d'ordre : 04/2002-E/CH.
- 2- Asses, Y., Beautrait, A., Toumi-Maouche, A., Maouche B., Taïri-Kellou, S. et Maigret, B., Criblage virtuel sur la cible biologique périphérique de la cholécystokinine, XXXIIème Congrès du CHITEL, Tunis (Tunisie); 1-6 Septembre 2006. (Communication par affiche).

CHAPITRE I

RESUME DE L'ETAT ACTUEL DES CONNAISSANCES SUR LA CHOLECYSTOKININE ET SES RECEPTEURS

I- INTRODUCTION

La cholécystokinine, appelée également CCK, est une hormone neuropeptidique connue pour ses actions régulatrices dans le système digestif et le système nerveux central des mammifères. L'appellation de cette substance provient de l'étymologie grecque signifiant hormone contractant la vésicule biliaire (kholé : bile, kustis : vessie ou sac, kinine : circuler, mouvoir). La CCK étant largement décrite dans un travail antérieur⁽¹⁾, nous nous limiterons dans ce chapitre à l'actualisation des connaissances relatives à cette hormone endogène et ses récepteurs. Une partie de ce chapitre sera consacrée à la description des sites de liaisons révélant les résidus cruciaux, du récepteur périphérique de la CCK, capables d'interagir avec les ligands de cette hormone. Nous dresserons, ensuite, un récapitulatif des différentes familles antagonistes dont la structure des composés a servi de base à la construction de notre chimiothèque.

II- LA CHOLECYSTOKININE

La CCK, connue pour son implication dans un grand nombre d'actions physiologiques et physiopathologiques, a toujours suscité un intérêt particulier dans les laboratoires de recherche et cela en vue de développer des ligands capables de pallier à son dysfonctionnement. Cette hormone endogène existe sous différentes formes moléculaires dérivant de l'hydrolyse d'une préprohormone de 115 acides aminés lors de son processus de maturation. Ces différentes formes moléculaires sont notées CCK-n, avec n le nombre d'acides aminés qui peut varier de 4 à 83. L'activité biologique est fonction de la longueur de la chaîne peptidique. Il faut noter que la sulfatation de la tyrosine et l'amidation de la phénylalanine procurent l'activité biologique optimale aux différents polypeptides endogènes (fig.1).

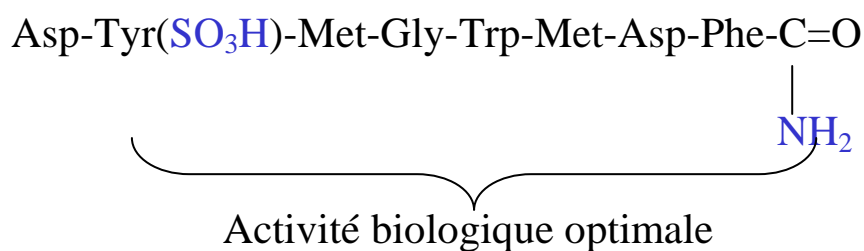


Figure 1 : Représentation de la structure primaire de la CCK-8S

La CCK est le neuropeptide le plus abondant du cortex cérébral (1 à 2mg dans le cerveau humain), elle y est principalement produite et localisée. Cette hormone se situe également au

niveau de l'hippocampe, les bulbes olfactifs, le septum, l'amygdale, l'hypothalamus et l'épine dorsale de la moelle épinière. Au niveau du système gastrique, la CCK est produite par les cellules I endocrines de la muqueuse se trouvant dans le duodénum, le jéjunum et l'intestin proximal.

III- LES RECEPTEURS DE LA CHOLECYSTOKININE

La caractérisation et la localisation des différents sites de liaison de la CCK sont possibles grâce à l'utilisation de ligands, agonistes et antagonistes. Nous pouvons citer, à titre d'exemple, le cas de la CCK-8 qui a permis grâce à sa sulfatation de distinguer entre 2 types de récepteurs de la CCK. Le premier récepteur caractérisé est le récepteur CCK₁, majoritairement localisé dans le système nerveux périphérique. Il représente une cible potentielle pour le traitement de nombreuses pathologies liées à la prise alimentaire et à la digestion. Le second type de récepteur, nommé CCK₂, est prépondérant dans le cortex cérébral. Cependant, par rapport au CCK₁, le récepteur CCK₂ prédomine aussi dans le pancréas de certains mammifères supérieurs.

Le clonage des récepteurs de la CCK a permis de localiser leurs gènes. Ainsi le gène du récepteur CCK₁ est situé sur le chromosome 4 et celui du récepteur CCK₂ se trouve sur le chromosome 11. La détermination de la structure secondaire des récepteurs CCK₁ et CCK₂ a été déduite par l'analyse du profil d'hydrophobicité de leur séquence primaire. Notons que ces 2 récepteurs membranaires de nature glycoprotéinique, et d'une quarantaine de kilodaltons chacun, présentent 50% d'homologie au niveau de leur structure primaire. On trouve aussi une homologie supérieure à 80% entre 2 récepteurs de même type CCK₁ ou CCK₂ appartenant à des espèces différentes.

Les récepteurs, partageant une structure secondaire déduite par l'analyse de leur structure primaire et de leur profil d'hydrophobicité, sont des protéines membranaires appartenant à la famille des récepteurs couplés aux protéines G (RCPGs) présentant :

- 7 domaines transmembranaires structurés en hélices α reliés par 3 boucles intracellulaires et 3 autres extracellulaires.
- une extrémité N-terminale extracellulaire et une extrémité C-terminale intracellulaire.
- 2 cystéines conservées situées dans la 1^{ière} et la 2^{ième} boucle extracellulaires, et formant le pont disulfure.

La structure 3D expérimentale de ces récepteurs membranaires demeure encore inaccessible en raison des difficultés rencontrées lors de leur cristallisation.

La famille, de ces récepteurs couplés aux protéines G dont la fonction est illustrée dans la figure 2, est largement présente dans le corps humain et joue un rôle physiologique important.

Les protéines G sont des protéines membranaires de couplage constituées de 3 sous unités ($\alpha\beta\gamma$), la sous unité α fixe la guanosine diphosphate, GDP, à l'état inactif. Le récepteur activé par l'hormone, induisant son changement conformationnel, interagit avec une protéine G stimulatrice. Un GTP remplace alors le GDP de la sous unité α , lui permettant ainsi de se séparer de $\beta\gamma$ et d'activer une enzyme membranaire, l'adénylcyclase (AC). Cette dernière permet, à partir de l'adénosine triphosphate (ATP), de catalyser la synthèse d'un second messenger, l'adénosine monophosphate cyclique (AMPc).

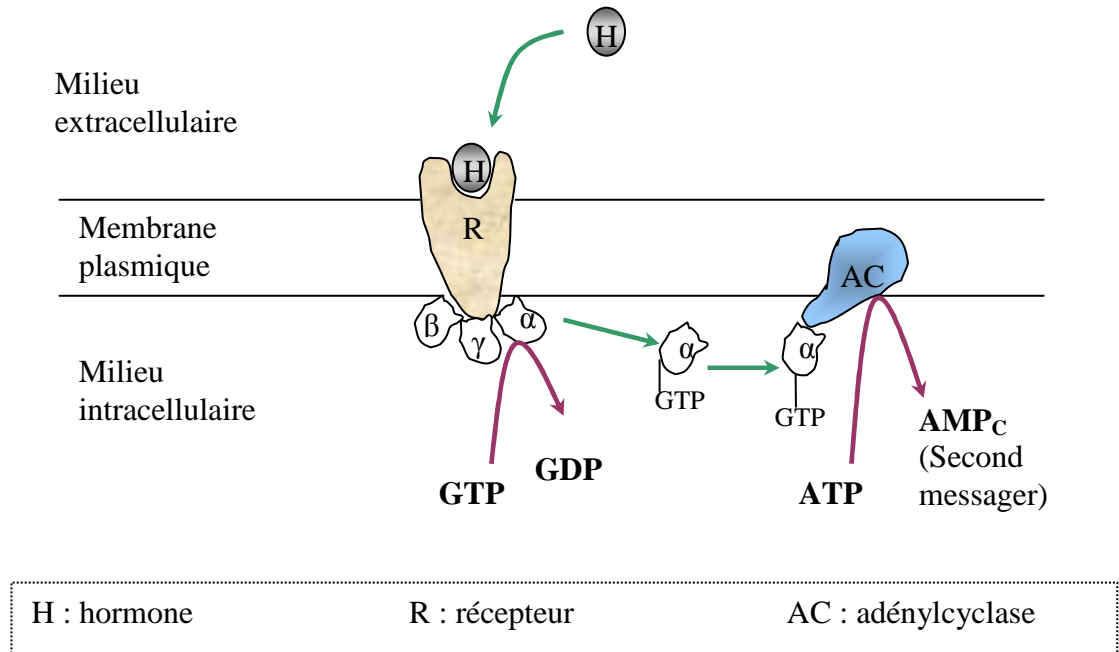


Figure 2 : Illustration du mode d'action des RCPGs.

Il faut noter que sur le marché, plus de 30% des médicaments sont des modulateurs de la fonction des RCPGs.

IV- PRINCIPAUX ROLES PHYSIOLOGIQUES

Ce peptide endogène agit en tant qu'hormone dans le système gastro-intestinal et en tant que neurotransmetteur-neuromodulateur dans le système nerveux central et périphérique.

1- Actions sur le système gastro-intestinal :

La CCK coordonne l'évacuation biliaire par la régulation de la contraction des muscles lisses de la vésicule biliaire ainsi que les muscles de la paroi intestinale. La vidange gastrique est

ralentie, induisant un relâchement de l'estomac proximal, la contraction du sphincter pylorique et ensuite l'inhibition de la sécrétion d'acide gastrique.

La CCK joue un double rôle dans le pancréas. Au niveau du pancréas exocrine, elle stimule la sécrétion d'enzymes digestives, hydrolysant les hydrates de carbone, les graisses, les protéines et les acides nucléiques. Au niveau du pancréas endocrine, elle stimule la sécrétion de glucagon, d'insuline, de somatostatine et d'autres polypeptides pancréatiques.

La CCK maintient la croissance normale du pancréas. En effet, la diminution du taux de CCK dans le sang ou l'administration d'un antagoniste induit une atrophie du pancréas. Inversement, un excès de CCK dans le sang peut développer des nodules précancéreux du pancréas⁽²⁾.

2- Actions sur le système nerveux central et périphérique

A) La satiété

La CCK peut présenter des actions différentes, par son excès ou par son déficit, sur la prise alimentaire. En effet, l'administration à fortes doses par voie centrale ou périphérique de la CCK réduit la prise alimentaire. Par ailleurs, les agonistes de la CCK pourraient agir dans le traitement des désordres de l'appétit, comme l'anorexie et la boulimie nerveuses^(1, 3).

La leptine, hormone sécrétée par les tissus adipeux, inhibe avec la CCK la prise alimentaire en agissant sur la sécrétion de l'insuline et du glucagon.

B) L'analgésie opioïde

La participation de la CCK dans l'analgésie fut découverte en 1980. L'administration de ce peptide atténue l'amplitude et la durée de la réponse analgésique à la morphine. Ceci implique que les antagonistes de la CCK vont dans le sens des opiacés dans le traitement de la douleur⁽⁴⁻⁵⁾.

C) Les troubles psychiques et l'anxiété

L'administration de la CCK provoque un comportement anxieux, des attaques de panique surtout dans le cas des patients atteints de troubles psychiques. Les antagonistes du récepteur CCK₂ présentent un réel potentiel dans le traitement des maladies neuropsychiatriques telles que la schizophrénie⁽⁶⁾.

D) La mémoire et l'apprentissage

Il est entièrement logique de trouver la CCK impliquée dans les diverses fonctions du cerveau étant donné l'abondance importante de ce peptide dans le cortex cérébral et l'hippocampe. Des expériences faites sur des rongeurs affirment qu'une grande concentration de la CCK serait impliquée dans le processus de mémorisation et d'apprentissage^(7,8).

E) Conclusion

Il ressort de ce qui précède qu'une anomalie dans le métabolisme de ce peptide engendrerait de nombreux troubles physiologiques. Sa distribution ubiquitaire dans l'organisme et son activité à faible dose ($\mu\text{g}/\text{kg}$) font d'elle une cible de choix dans le développement de nouvelles molécules thérapeutiques. Ce qui a suscité l'intérêt des firmes pharmaceutiques et des groupes de recherche pour la conception d'agonistes ou d'antagonistes de ce peptide.

V- DESCRIPTION DES PHARMACOPHORES IMPORTANTS

La variété des effets physiologiques de la CCK et son implication dans plusieurs désordres pathologiques, ont mené à la découverte de nombreux ligands, potentiellement sélectifs sur les récepteurs CCK₁ et CCK₂. Le grand potentiel thérapeutique, révélé *in vitro*, de certains agonistes et antagonistes a conduit à de nombreux essais sur l'animal ; ce qui a généré des candidats-médicaments très prometteurs. Dans ce paragraphe, nous relaterons brièvement les antagonistes découverts avant 2002 puisqu'ils ont été largement décrits dans des travaux antérieurs^(1 et références incluses).

Compte tenu du fait que notre recherche porte sur le criblage de molécules sur la cible périphérique de la CCK et en utilisant une cavité antagoniste, nous nous limiterons ici à présenter uniquement les ligands décrits dans la littérature comme des antagonistes sur le récepteur CCK₁.

Les premiers antagonistes aux récepteurs de la CCK découverts étaient des peptides issus des modifications structurales et remplacements de certains acides aminés de la structure de la CCK-8, CCK-7 et CCK-4^(1 et références incluses, 9, 10). Cependant, malgré leur bon potentiel de sélectivité envers l'un ou l'autre des récepteurs, les antagonistes peptidiques sont biodégradables et ne présentent pas de caractéristiques "franchement" agonistes ou antagonistes. Les développements thérapeutiques se sont essentiellement axés sur la recherche de ligands nonpeptidiques et se sont soldés par la découverte de l'asperlicine⁽¹¹⁾ en tant qu'antagoniste sélectif du récepteur CCK₁. Diverses manipulations effectuées sur la structure de l'asperlicine ont conduit à la découverte des

dérivés de la benzodiazépine, tels que le dévazépide^(12, 13) et la quinazolidinone⁽¹⁴⁾ qui représentent actuellement deux familles de pharmacophores puissants et hautement sélectifs.

La découverte des dérivés du diphenylpyrazolidinone, tels que LY294920 et LY288513, résulte de l'optimisation d'une tête de série issue d'un criblage aléatoire⁽¹⁵⁾. Ces dérivés possèdent une affinité de l'ordre de la nanomole sur les récepteurs de la CCK.

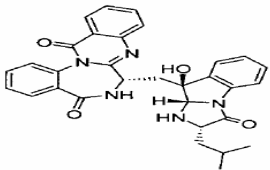
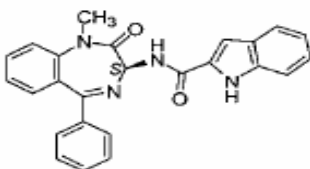
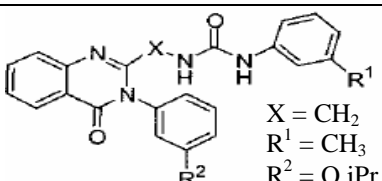
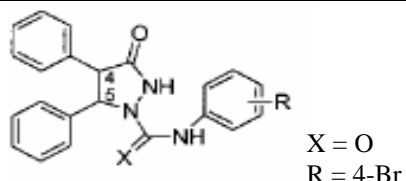
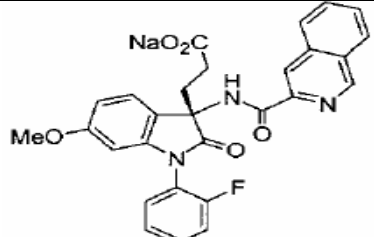
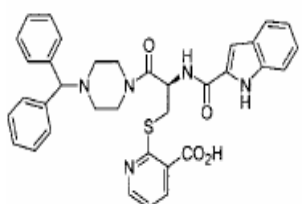
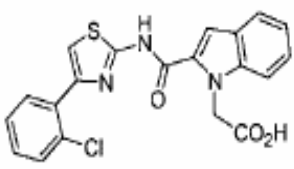
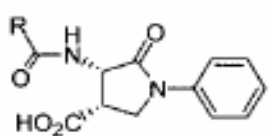
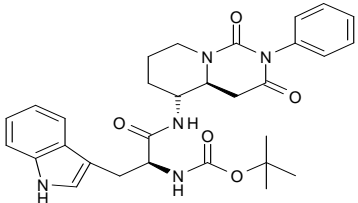
Les dérivés du 1,3,3-trisubstitué indol-2-one, nommés AG-041R et T-O632, résultent également d'un programme de criblage aléatoire^(16,17). Il en est de même pour l'antagoniste sélectif TP-680⁽¹⁸⁾ et pour le composé SR-27897 appelé également Lintript⁽¹⁹⁾, ce dernier a été obtenu par optimisation d'un composé issu du criblage d'une large chimiothèque des laboratoires Sanofi.

L'emploi du 1, 3, 5-trisubstitué pyrrolidinone a servi de base à la découverte de l'antagoniste, de la CCK sur le récepteur CCK₁, le SC-50998⁽²⁰⁾.

Notons que la structure de la CCK-4 a été le point de départ pour la recherche d'antagonistes à l'institut de chimie médicinale de Madrid. Le remplacement des résidus Met31-Asp32 par le squelette 5-amino-1,3-dioxoperhydroprido[1,2-c]pyrimidine effectué sur le CCK-4, a conduit à des séries hautement sélectives d'antagonistes CCK₁^(9,10). Dans le cadre de notre recherche, nous avons utilisé la cavité (le site de liaison) de la série la plus récente : les dérivés pyridopyrimidiques⁽¹⁰⁾ pour docker nos composés.

Nous récapitulons ces antagonistes non-peptidiques sélectifs sur le récepteur CCK₁, avec leur activité, dans le tableau suivant :

Tableau 1 : Structure 2D des antagonistes sélectifs sur le récepteur CCK₁

Nom	Structure 2D	IC ₅₀ / CCK ₁ (nM)
Asperlicine		1,40
Devazepide		0,08
Dérivé de la quinazolidinone	 X = CH ₂ R ¹ = CH ₃ R ² = O iPr	1,64
Dérivé du diphenylpyrazolidinone LY288513	 X = O R = 4-Br	20,50
Dérivé du 1,3,3-trisubstitué indol-2-one T-0632		0,24
TP-680		1,20
SR-27897 Lintitrip		0,58
Dérivé du 1,3,5-trisubstitué pyrrolidinone SC-50998		16,00
Dérivé pyridopyrimidique		5,20

VI- RESUME DE L ETAT ACTUEL DES CONNAISSANCES SUR LES SITES DE LIAISONS DES LIGANDS

La caractérisation des interactions entre les pharmacophores de la CCK et leurs partenaires dans les récepteurs, aidera certainement la recherche avec une approche plus rationnelle pour concevoir de nouvelles molécules, ou modifier celles qui existent déjà par l'optimisation de leurs propriétés. Pour cela, la combinaison de l'outil informatique, par l'exploitation des méthodes de la modélisation moléculaire, et de l'outil biologique, par les techniques de la mutagenèse dirigée, a déjà fait ses preuves.

A la suite d'une recherche bibliographique de plusieurs travaux, nous avons listé les sites de liaisons de la CCK et de quelques uns de ses agonistes et antagonistes avec le récepteur membranaire périphérique CCK₁.

Nous présentons dans le tableau 2, les acides aminés du récepteur périphérique qui interagissent avec les acides aminés de la CCK-8S⁽²¹⁻²⁶⁾. Nous donnons également les régions où se situent les acides aminés appartenant au récepteur CCK₁.

"TM" est l'abréviation de la région transmembranaire et "e" l'abréviation des boucles extracellulaires.

Tableau 2: Résidus de CCK₁ interagissant avec CCK-8S

AA de CCK ₁	AA de CCK-8S	Localisation de AA _{CCK1}
Trp39	Arg1	Région N-term
Gln40	Asp2	Région N-term
Leu50	Met/Nle	TM I
Ileu51	Met/Nle	TM I
Leu53	Met/Nle	TM I
Cys94	Met/Nle	TM I
Lys105		e1
Asp106	Arg1	e1
Phe107	Trp9	e1
Lys115		e1
Thr117		TM III
Thr118		TM III
Met121	Met/Nle	TM III
Ser124		TM III
Val125	Phe9	TM III
Phe170		TM V
Met195	Tyr3	e2
Arg197	Tyr3(SO ₃ H)	e2
Phe218	Phe9	TM V

Trp326	Phe9	TMVI
Ile329	Phe9	TMVI
Phe330	Phe9	TMVI
Asn333	Amide du C-Ter	TMVI
Arg336	Asp8	TMVI
Ser342		e3
Ser348	Met7	e3
Ile352	Met/Nle, Phe	TMVII
Leu356	Met/Nle, Phe	TMVII
Tyr360	Met/Nle, Phe	TMVII

A titre d'illustration, nous donnons la figure secondaire suivante, représentant la localisation de ces interactions pertinentes.

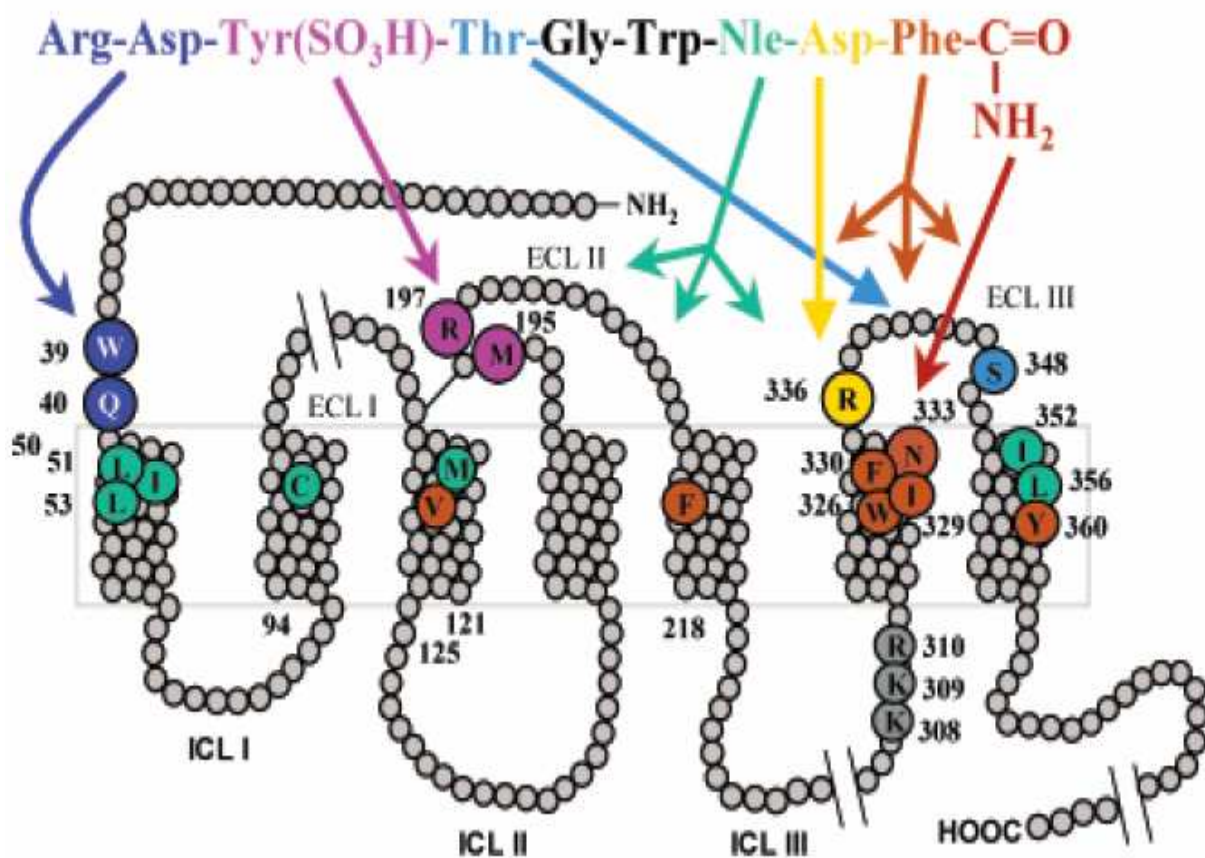


Figure 3 : Représentation schématique du site de liaison du complexe CCK/CCK₁⁽²⁶⁾

Nous donnons ci-dessous le site de liaison de l'agoniste non-peptidique SR-146,131 au sein du récepteur CCK₁⁽²¹⁾.

Tableau 3 : site de liaison SR-146,131/ CCK₁⁽²¹⁾

AA de CCK ₁	Localisation de AA _{CCK1}
Arg336	TMVI
Asn333	TMVI
Phe330	TMVI
Ile329	TMVI
Ile352	TMVII
Leu356	TMVII
Met121	TMIII
Cys94	TMI

Concernant les sites de liaisons antagoniste/CCK₁, décrits dans la littérature^(10, 19, 26), nous reportons dans les tableaux 4-a, 4-b et 4-c les résidus qui se sont avérés pertinents.

a - Site de liaisons dérivé pyridopyrimidique /CCK₁⁽¹⁰⁾

Tableau 4-a : Site de liaison dérivé pyridopyrimidique/CCK₁⁽¹⁰⁾

AA de CCK ₁	Localisation de AA _{CCK1}
Arg336	TMVI
Asn333	TMVI
Phe330	TMVI
Ile329	TMVI
Ile352	TMVII
Val125	TMIII
Leu214	TMV
Gly122	TMIII
Pro177	TMIV
Thr118	TMIII

b - Site de liaisons SR-27897/CCK₁⁽¹⁹⁾

Tableau 4-b : Site de liaison SR-27897/CCK₁⁽¹⁹⁾

AA de CCK ₁	Localisation de AA _{CCK1}
Lys115	e1
Tyr179	e2
Phe185	e2
Lys187	e2
Phe198	e2
Leu214	TMV
Asn333	TMVI

c - Sites de liaisons SR-27897/CCK₁⁽²⁶⁾

Tableau 4-c : Site de liaison SR-27897/CCK₁⁽²⁶⁾

AA de CCK ₁	Localisation de AA _{CCK1}
Arg336	TMVI
Ser348	e3
Asn333	TMVI
Phe330	TMVI
Val125	TMIII
Leu214	TMV
Leu46	TMI
Leu50	TMI
Ile352	TMVII
Leu356	TMVII

Nous avons reporté en bleu les résidus, du site de liaison antagoniste/ CCK₁, issus des 3 travaux^(10, 19, 26). En rouge, sont indiqués les résidus communs aux 2 études^(10, 26). L'analyse de ces résultats montre que l'étude effectuée par l'équipe de Gouldson et al.⁽¹⁹⁾, prévoit un site moins hydrophobe que celle des deux autres études^(10, 25).

VII- CONCLUSION

Cette revue bibliographique et ces informations de la littérature, concernant aussi bien les ligands potentiels et les sites de liaison, sont à la base des recherches effectuées dans le cadre de ce manuscrit. En effet, pour élaborer notre chimiothèque, nous avons pris en considération, les groupements les plus fréquemment rencontrés dans les séries de ligands antagonistes, et en particulier, sur le récepteur CCK₁⁽²⁷⁻⁴⁵⁾.

Nous avons également comparé et analysé nos résultats, issus des différents dockings, aux informations rapportées par les travaux les plus récents de la littérature. Cette comparaison concerne particulièrement les résidus jugés pertinents dans les interactions antagoniste/CCK₁^(10,19,26).

REFERENCES BIBLIOGRAPHIQUES

- 1- Taïri-Kellou, S., Thèse de Doctorat d'état, Conception Assistée par Ordinateur des Complexes ligand/récepteur de la Cholécystokinine, Alger, USTHB, N° d'ordre : 04/2002-E/CH.
- 2- Noble, F., Wank, S. A., Crawley, J. N., Bradwejn, J., Seroogy, K. B., Hamon, M. et Roques, B. P., International union of pharmacology. XXI. Structure, distribution, and functions of cholécystokinine receptors. *Pharmacological Reviews*, 51, Issue 4, 745-781, **1999**.
- 3- Rehfeld, J. F., *Cholecystokinine*, *Best Practice & Research*, 18, 569-586, **2004**.
- 4- Faris, P. L., Komisaruk, B. R., Watkins, L. R. et Meyer, D. J., Evidence for the neuropeptide cholecystokinine as an antagonist of opiate analgesia. *Science*, 219, 310-312, **1983**.
- 5- Agnes, R. S., Lee, Y. S., Davis, P., Ma, S., Badghisi, H., Porreca, F., Lai, J. et Hruby, V. J., Structure-Activity Relationships of Bifunctional Peptides Based on Overlapping Pharmacophores at Opioid and Cholecystokinine Receptors. *J. Med. Chem.*, 49, 2868 – 2875, **2006**.
- 6- Tachikawa, H., Harada, S., Kawanissi, Y., Okubo, T. et Shiraishi, H., Novel polymorphism in the promoter and coding regions of the human cholecystokinine B receptor gene: An association analysis with schizophrenia. *Am J Med Genet*, 88, 700–704, **1999**.
- 7- Flood, J. F., Smith, G. F. et Morley, J.E., Modulation of memory processing by cholecystokinine: dependence on the vagus nerve. *Science*, 236, 832-834, **1987**.
- 8- Dauge, V. et Lena, I., CCK in anxiety, and cognitive processes. *Neurosci. Biobehav. Rev.*, 22, 815-825, **1998**.
- 9- Martín-Martínez, M., Bartolomé-Nebreda, J.M., Gómez-Monterrey, I., González-Muñiz, R., García-López, M.T., Ballaz, S., Barber, A., Fortuño, A., Del Río, J. et Herranz, R. Synthesis and stereochemical structure-activity relationships of 1,3-dioxoperhydropyrido[1,2-c]pyrimidine derivatives: potent and selective cholecystokinine-A receptor antagonists. *J. Med. Chem.*, 40, 3402-3407, **1997**.
- 10- Martín-Martínez, M., Marty, A., Jourdan, M., Escrieut, C., Archer, E., González-Muñiz, R., García-López, M.T., Maigret, B., Herranz, R. et Fourmy, D. Combination of molecular modeling, site directed mutagenesis and SAR studies to delineate the

- binding site of pyridopyrimidine antagonists on the human CCK₁ receptor. *J. Med. Chem.*, 48, 4842-4850, **2005**.
- 11- Goëtz MA, Lopez M, Monaghan RL, Chang RSL, Lotti VJ, Chen TB. Asperlicin, a novel non-peptidal cholecystokinin antagonist from *Aspergillus alliaceus*. Fermentation, isolation and biological properties. *J. Antibiot.*, 38, 1633–1637, **1985**.
- 12- Evans, B. E., Bock, M. G., Rittle, K. E., DiPardo, R. M., Whitter, W. L., Veber, D. F., Anderson, P. S. et Freidinger, R. M., Design of potent, orally effective, nonpeptidal antagonists of the peptide hormone cholécystokinine. *Proc. Natl. Acad. Sci. USA*, 83, 4918–4922, **1986**.
- 13- Evans, B. E., Rittle, K. E., Bock, M. G., DiPardo, R. M., Freidinger, R. M., Whitter, W. L., Lundell, G. F., Veber, D. F., Anderson, P. S., Chang. R. S.L., Lotti, V.J., Cerino, D. J., Chen, T. B., Kling, P. J., Kunkel, K. A., Springer, J. P. et Hirshfield, J., Methods for drug discovery: Development of potent, selective orally effective cholecystokinin antagonists. *J. Med. Chem.*, 31, 2235–2246, **1988**.
- 14- Padia, J. K., Field, M., Hinton, J., Meecham, K., Pablo, J., Pinnock, R., Roth, B. D., Singh, L., Suman-Chauhan, N., Trivedi, B. K. et Webdale, L., Novel nonpeptide CCK-B antagonists: Design and development of quinazolinone derivatives as potent, selective, and orally active CCK-B antagonists. *J. Med. Chem.*, 41, 1042–1049, **1998**.
- 15- Howbert, J. J., Lobb, K. L., Britton, T. C., Mason, N. R. et Bruns, R. F., Diphenylpyrazolidinone and benzodiazepine cholecystokinin antagonists: A case of convergent evolution in medicinal chemistry. *Bioorg. Med. Chem. Lett.*, 5, 875–880, **1993**.
- 16- Kitamura, H., Kato, A. et Esaki, T., AG-041R, a novel indoline-2-one derivative, induces systemic cartilage hyperplasia in rats, *Eur J Pharmacol*, 418, 225–230, **2001**.
- 17- Taniguchi, H., Yazaki, N., Endo, T. et Nagasaki, M., Pharmacological profile of T-0632, a novel potent and selective CCKA receptor antagonist, in vitro. *Eur. J. Pharmacol.*, 304, 147–154, **1996**.
- 18- Tachibana, I., Kanagawa, K., Yamamoto, Y. et Otsuki, M., Pharmacological profile of a new serine derivative cholecystokinin receptor antagonist TP-680 on pancreatic, biliary and gastric function. *J. Pharmacol. Exp. Ther.*, 279, 1404–1412, **1996**.
- 19- Gouldson, P., Legoux, P., Carillon, C., Delpech, B., Le Fur, G., Ferrara, P. et Shire, D., The agonist SR 146131 and the antagonist SR 27897 occupy different sites on the human CCK₁ receptor. *Eur. J. Pharmacol.*, 400, 185-194, **2000**.

- 20- Flynn, D. L., Villamil, C. I., Becker, D. P., Gullikson, G. W., Moumami, C., Yang, D-C., 1,3,4-trisubstituted pyrrolidinones as scaffolds for construction of peptidomimetic cholecystokinin antagonists. *Bioorg. Med. Chem. Lett.*, 2, 1251–1256, **1992**.
- 21- Escricuet, C., Gigoux, V., Archer, E., Verrier, S., Maigret, B., Behrendt, R., Moroder, L., Bignon, E., Silvente-Poirot, S., Pradayrol, L. et Fourmy, D. The biologically crucial C-terminus of cholecystokinin and the non-peptide agonist SR-146,131 share a common binding site in the human CCK₁ receptor. Evidence for a crucial role of Met-121 in the activation process, *J. Biol. Chem.*, 277, 7546-7555, **2002**.
- 22- Kennedy, K., Gigoux, V., Escricuet, C., Maigret, B., Martinez, J., Moroder, L., Fréhel, D., Gully, D., Vaysse, N. et Fourmy, D. Identification of two amino acids of the human cholecystokinin-A receptor that interact with the N-terminal moiety of cholécystokinine. *The Journal of Biological Chemistry*, 5, 2920-2926, **1997**.
- 23- Gigoux, V., Escricuet, C., Silvente-Poirot, S., Maigret, B., Gouilleux, L., Fehrentz, J-A., Gully, D., Moroder, L., Vaysse, N. et Fourmy, D. Met-195 of the cholecystokinin-A receptor interacts with the sulfated tyrosine of cholecystokinin and is crucial for receptor transition to high affinity state. *The Journal of Biological Chemistry*, 23, 14380-14386, **1998**.
- 24- Gigoux, V., Maigret, B., Escricuet, C., Silvente-Poirot, S., Bouisson, M., Fehrentz, J-A., Moroder, L., Gully, D., Martinez, J., Vaysse, N. et Fourmy, D. Arginine 197 of the cholecystokinin-A receptor binding site interacts with the sulfate of the peptide agonist cholecystokinin. *Protein Sci.*, 8, 2347-2354, **1999**.
- 25- Gigoux, V., Escricuet, C., Fehrentz, J-A., Poirot, S., Maigret, B., Moroder, L., Gully, D., Martinez, J., Vaysse, N. et Fourmy, D. Arginine 336 and asparagine 333 of the human cholecystokinin-A receptor binding site interact with the penultimate aspartic acid and the C-terminal amide of cholecystokinin. *J. Biol. Chem.*, 274, 20547-20464, **1999**.
- 26- Archer-Lahlou, E., Tikhonova, I., Escricuet, C., Dufresne, M., Seva, C., Pradayrol, L., Moroder, L., Maigret, B. et Fourmy, D., Modeled structure of a G-protein-coupled receptor: the cholecystokinin-1 receptor. *J. Med. Chem.*, 48, 180-191, **2005**.
- 27- Latorre, M., Bartolomé-Nebreda, J.M., Garcia-Lopez, M.T., Gonzalez-Muñiz, R., Herranz R., Del Río, J. et Cenarruzabeitia, E., Pharmacological Study of IQM 97,423, a Potent and Selective CCK₁ Receptor Antagonist with Protective Effect in Experimental Acute Pancreatitis. *Pharmacology*, 72, 68–76, **2004**.

- 28- Bellier, B. et Garbay, C. How a single inversion of configuration leads to a reversal of the binding mode: proposal of a novel arrangement of CCK₂ ligands in their receptor, and contribution to the development of peptidomimetic or non-peptide CCK₂ ligands. *European Journal of Medicinal Chemistry*, 38, 671- 686, **2003**.
- 29- De la Figuera, N., Martin-Martinez, M., Herranz, R., Garcia-Lopez , M. T., Latorre, M., Cenarruzabeitia, E., Del Rio J. et Gonzaleiz-Muniz R., Highly constrained dipetoid analogues containing a type II'β-turn mimic as novel and selective CCK-A receptor ligands. *Bioorganic & Medicinal Chemistry Letters*, 9, 43-48, **1999**.
- 30- Gibson, S. E., Guillo, N., Jones, J. O., Buck, I. M., Kalindjian, S. B., Roberts S. et Tozer, M. J., CCK₂ receptor antagonists containing the conformationally constrained phenylalanine derivatives, including the new amino acid Xic. *Eur. J. Med. Chem.* , 37, 379–389, **2002**.
- 31- Gupta, S. P. Quantitative Structure-Activity Relationship Studies on Cholecystokinin Antagonists. *Current Pharmaceutical Design*, 8, 111-124, **2002**.
- 32- Low, C. M. R., Buck, I. M., Cooke, T., Cushnir, J. R., Kalindjian, S. B., Kotecha, A., Pether, M. J., Nigel, P. S., Vinter, J. G. et Wright, L., Scaffold Hopping with Molecular Field Points: Identification of a Cholecystokinin-2 (CCK₂) Receptor Pharmacophore and Its Use in the Design of a Prototypical Series of Pyrrole- and Imidazole-Based CCK₂ Antagonists. *J. Med. Chem.*, 48, 6790-6802, **2005**.
- 33- Agrawal, V. K., Sharma, R. et Khadikar, P. V., Quantitative Structure–Activity Relationship Studies on 5-Phenyl-3-ureido-1,5-benzodiazepine as Cholecystokinin-A Receptor Antagonists. *Bioorganic & Medicinal Chemistry*, 10, 3571–3581, **2002**.
- 34- Sinha, J., Kurup, A., Paleti, A. et Gupta, S. P., Quantitative Structure-Activity Relationship Study on Some Nonpeptidal Cholecystokinin Antagonists. *Bioorganic & Medicinal Chemistry*, 7, 1127-1130, **1999**.
- 35- Varnavas, A., Lassiani, L., Valenta, V., Berti, F., Mennuni, L. et Makovec, F., Anthranilic Acid Derivatives: A New Class of Non-Peptide CCK₁ Receptor Antagonists. *Bioorganic & Medicinal Chemistry*, 11, 741–751, **2003**.
- 36- Varnavas, A., Lassiani, L., Valenta, V., Berti, F., Tontini, A., Mennuni, L., Makovec, F., Anthranilic acid based CCK₁ antagonists : the 2-indole moiety may represent a « needle » according to the recent homonymous concept. *European Journal of Medicinal Chemistry*, 39, 85-97, **2004**.

- 37- Varnavas, A., Lassiani, L., Valenta, V., Mennuni, L., Makovek, F., Hadjipavlou-Litina, D., Anthranilic acid based CCK₁ receptor antagonists: preliminary investigation on their second “touch point”. *European Journal of Medicinal Chemistry*, 40, 563-581, **2005**.
- 38- Varnavas, A., Lassiani, L., Valenta, V., Ciogli, A., Gasparrini, F., Mennuni, L. et Makovec, F., N-Terminal Anthranoyl-Phenylalanine Derivatives as CCK₁ Receptor Antagonists: The Final Approach. *Medicinal Chemistry*, 1, 501-517, **2005**.
- 39- Trivedi, B. K., Padia J. K., Holmes, A., Rose S., Wright, D. S., Hinton, J. P., Pritchard, M. C., Eden, J. M., Kneen, C., Webdale, L., Suman-Chauhan, N., Boden, P., Singh, L., Field, M. J., et Hill, D., Second Generation “Peptoid” CCK-B Receptor Antagonists: Identification and Development of *N*-(Adamantyloxycarbonyl)-*r*-methyl-(*R*)-tryptophan Derivative (CI-1015) with an Improved Pharmacokinetic Profile. *J. Med. Chem.*, 41, 38-45, **1998**.
- 40- Gupta, S.P., Nagappa, A.N., Quantitative Structure-Activity Relationship Studies on Cholecystokinine Antagonists. *Medicinal Chemistry Reviews*, 3, 349-350, **2004**.
- 41- Lattmann, E., Singh, H., Boonprakob, Y., Lattmann, P. et Sattayasai, J., Synthesis and evaluation of *N*-(3-oxo-2,3-dihydro-1*H*-pyrazol-4-yl)-1*H*-indole-carboxamides as cholecystokinine antagonists, *Journal of Pharmacy and Pharmacology*, 58, 393-401, **2006**.
- 42- González-Puga, C., García-Navarro, A., Escames, G., León, J., López-Cantarero, M., Ros, E. et Acuña-Castroviejo, D., Selective CCK-A but not CCK-B receptor antagonists inhibit HT-29 cell proliferation: synergism with pharmacological levels of melatonin. *Journal of Pineal Research*, 39, 243-250, **2005**.
- 43- Shajan, A.S. P., Massimo, D. et Christoph, B., CCK₁ Antagonists: Are They Ready for Clinical Use?. *Digestive Diseases*, 24, 70-82, **2006**.
- 44- Chopra, M. et Mishra, A. K., Ligand-Based Molecular Modeling Study on a Chemically Diverse Series of Cholecystokinine-B/Gastrin Receptor Antagonists: Generation of Predictive Model. *J. Chem. Inf. Model.*, 45, 1934-1942, **2005**.
- 45- McClure, K., Hack, M., Huang, L., Sehon, C., Morton, M., Li L., Barrett, T. D., Shankley, N. et Breitenbucher, J. G., Pyrazole CCK₁ receptor antagonists. Part 2: SAR studies by solid-phase library synthesis and determination of Free–Wilson additivity. *Bioorganic & Medicinal Chemistry Letters*, 16, 77–80, **2006**.

CHAPITRE II

MODELISATION MOLECULAIRE
ET DRUG DESIGN

I - INTRODUCTION

La modélisation, au sens large du terme, consiste à bâtir un modèle artificiel, le plus souvent mathématique, d'un phénomène réel pour l'utiliser ensuite à des fins prédictives ou de contrôle. Un modèle est le reflet de la réalité, reflet que l'on espère le plus précis possible mais en l'état actuel de la technologie, c'est à dire évolutive, il reste perfectible.

La modélisation moléculaire est l'élaboration des modèles dans le cadre de la chimie c'est-à-dire des systèmes d'intérêt chimique. Cette méthode utilisant l'outil informatique et les nombreuses possibilités de calcul théorique est appelée également conception assistée par ordinateur. De ce fait, elle est considérée aujourd'hui comme une nouvelle technique de recherche et de compréhension des phénomènes physico-chimiques des molécules. On peut dire qu'elle a fait de l'ordinateur un lieu d'expériences numériques en épargnant au chimiste le coût que nécessiterait une expérimentation aveugle.

L'objet de cette thèse est l'utilisation de la modélisation moléculaire pour la conception de futurs médicaments.

En 1894, le chimiste Emile Fischer compara la liaison entre un substrat et son enzyme, à celle d'une clé et sa serrure⁽¹⁾. En 1913, le médecin Paul Ehrlich généralisa ce concept aux interactions entre un médicament et sa cible biologique⁽¹⁾. Aujourd'hui, la connaissance de la structure tridimensionnelle d'un ligand dans la poche de liaison de son récepteur est une information pertinente dans la conception de meilleurs ligands de ce récepteur. Par analogie avec le concept d'Emile Fischer, cette conception consiste à faire les "copies" d'une clé.

L'utilisation fructueuse de la modélisation moléculaire dans la conception de nouveaux ligands a donné un nouvel essor à la recherche dans le domaine de la pharmacologie.

Pour aborder une étude par modélisation moléculaire, nous avons le choix entre deux types d'approches théoriques de base et qui sont :

La mécanique quantique et la mécanique moléculaire

Il faut noter que ces deux approches sont radicalement différentes. En effet,

- Les méthodes issues de la mécanique quantique sont basées sur la recherche de la meilleure répartition des électrons dans le champ engendré par les noyaux du système et cela en résolvant de manière approchée l'équation de Schrödinger : $\hat{H}\Psi = E \Psi$, où \hat{H} est l'opérateur hamiltonien associé à l'énergie E du système et Ψ la fonction d'onde décrivant le mouvement des particules (électrons +noyaux).

- La mécanique moléculaire ne tient pas compte explicitement des électrons. Elle traite les atomes et les électrons qui leur sont associés comme des masses ponctuelles chargées, obéissant aux lois de la mécanique classique.

Compte tenu que le formalisme des méthodes issues de la mécanique quantique a été largement développé dans plusieurs manuscrits de thèses soutenues à l'U.S.T.H.B⁽²⁻⁴⁾ et n'a pas été utilisé dans les travaux de cette thèse, nous nous limiterons ici à comparer les capacités de cette méthode avec celles issues de la mécanique moléculaire. Nous présentons brièvement dans ce chapitre, la mécanique moléculaire qui est à la base de l'élaboration de nos modèles. Nous présentons également les techniques utilisées dans le drug design.

II – LA MECANIQUE MOLECULAIRE

A- Bases et principes

Les méthodes de la mécanique moléculaire permettent de déterminer les propriétés structurales et thermodynamiques des systèmes étudiés. L'inconvénient majeur de ces méthodes est qu'elles ne sont pas adaptées à la description de systèmes impliqués dans la création ou la destruction de liaisons.

Basée sur l'approximation de Born-Oppenheimer qui permet de séparer les mouvements des électrons de ceux des noyaux, la mécanique moléculaire considère que les électrons des atomes du système adoptent une distribution optimale, ce qui permet de décrire l'énergie d'une molécule en fonction de la position des noyaux de ses atomes. De manière simplifiée, la mécanique moléculaire assimile les atomes d'une molécule à des masses ponctuelles chargées reliées par des ressorts.

B – Fonction énergie potentielle

L'énergie potentielle ou stérique de la molécule peut être donnée par la formule suivante :

$$E_S = \underbrace{\sum E_{liaisons} + \sum E_{angles} + \sum E_{dièdres} + \sum E_{dièdresimpropres}}_{E_{cov\ alente}} + \underbrace{\sum E_{vdW} + \sum E_{électrostatique}}_{E_{non\ cov\ alente}} \quad (II-1)$$

Cette énergie est exprimée sous la forme d'une somme de contributions associées aux écarts de la structure par rapport à des paramètres structuraux de référence. En d'autres termes

c'est une représentation empirique des aspects dynamiques et énergétiques des systèmes moléculaires limités à une classe de molécules ayant des ressemblances structurales et fonctionnelles. L'ensemble des fonctions ou des modèles mathématiques représentant cette énergie stérique est appelé champ de forces.

Différents champs de forces ont été développés en vue de modéliser les systèmes biologiques et non biologiques⁽⁵⁾.

Les variables de calcul sont les coordonnées internes du système : longueurs de liaison, angle de valence, angles dièdres, angles dièdres impropres et les distances entre les atomes non liés.

Nous donnons dans la figure suivante les termes communs constituant un champ de forces ou l'équation (II-1). Le détail de ces termes est donné dans un travail antérieur⁽⁶⁾.

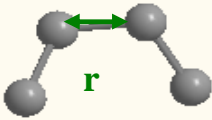
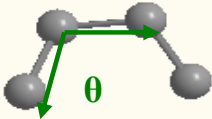
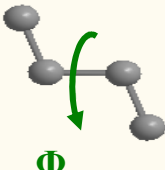
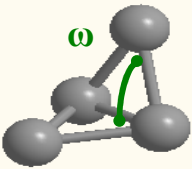
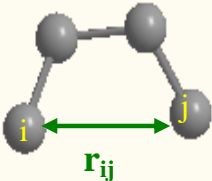

	Equations	Représentations
$E_{\text{cov alente}}$	$E_{\text{liaisons}} = \sum_{\text{liaisons}} \frac{1}{2} K_0 (r - r_0)^2$	
	$E_{\text{angles}} = \sum_{\text{angles}} \frac{1}{2} K_\theta (\theta - \theta_0)^2$	
	$E_{\text{dièdres}} = \sum_{\text{dièdres}} \sum_n K_\phi (1 + \cos(n(\phi - \phi_0)))$	
	$E_{\text{dièdresimpropres}} = \frac{1}{2} K_\omega \omega^2$	
$E_{\text{non cov alente}}$	« Lennard-Jones » $E_{\text{vdW}} = \sum_{\substack{ij, \text{atomes} \\ \text{non-liés}}} 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right]$	
	« Buckingham » $E_{\text{vdW}} = A_{ij}' e^{-B_{ij}' R_{ij}} - \frac{C_{ij}}{R_{ij}^6}$	
	$E_{\text{électrostatique}} = \sum_{\substack{ij, \text{atomes} \\ \text{non-liés}}} \frac{q_i q_j}{4\pi\epsilon_0 \epsilon_r r_{ij}}$	

Figure 1 : Illustration schématique d'un champ de forces

Notons que certains champs de force incluent, dans l'expression de l'énergie non covalente de la molécule, un terme décrivant les liaisons hydrogènes et qui est du type :

$$E_{HB} = \sum_{\text{paires}} \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{10}} \quad (\text{II-2})$$

Certains contiennent également des termes dus au couplage entre les variations de différentes coordonnées internes.

Il faut noter que le choix du champ de force est à faire sur la base des résultats déjà obtenus dans la littérature concernant son application aux systèmes moléculaires analogues au système étudié.

C - Méthodes d'exploration de l'hypersurface d'un système moléculaire

L'étape qui suit le choix du champ de force approprié au système étudié est la recherche de la géométrie la plus stable de ce dernier. Le but essentiel de la minimisation est de relaxer la structure initiale de la molécule afin d'éliminer les mauvais contacts interatomiques. Cette prospection consiste à déterminer les valeurs des variables géométriques pour lesquelles la valeur de la fonction d'énergie potentielle ou stérique est minimale. Le nombre important du degré de liberté de la molécule fait que la molécule fluctue autour d'un ensemble de conformations stables correspondant à différents *minima*. Parmi ces derniers un seul est global, et correspond à la plus basse valeur de l'énergie potentielle, tous les autres sont des *minima* locaux.

C1 – Algorithmes de minimisation

Les algorithmes de minimisation mesurent l'énergie le long de l'hypersurface énergétique d'un système de façon incrémentale, dans le but de déterminer des directions qui mènent à un minimum. Ces algorithmes sont basés sur le calcul des dérivées de la fonction d'énergie et utilisent des processus itératifs. En d'autres termes, à partir de l'énergie potentielle et du gradient d'un jeu de coordonnées atomiques $\left(g_i = -\frac{\partial E}{\partial x_i}\right)$, ces outils de minimisation engendrent un nouveau jeu de positions atomiques correspondant à une énergie potentielle plus basse. Cependant, ces algorithmes diffèrent tous dans la manière de descendre

le long de l'hypersurface. Nous détaillons ci-dessous les différents algorithmes les plus fréquemment utilisés :

La méthode du gradient simple ou de la pente abrupte (« steepest descent ») :

La recherche de l'énergie minimale pour une première position atomique passe par la détermination du scalaire α , donné par la relation itérative suivante :

$$x_{i+1} = x_i - \alpha g_i \quad (\text{II-3})$$

La recherche du *minimum* en un point s'effectue dans la direction opposée au gradient de la fonction énergie, dans le sens de la plus grande pente.

La méthode du steepest descent s'avère efficace lorsque la conformation de départ est loin du *minimum*, cependant elle converge lentement et adopte un comportement oscillatoire près du *minimum*. Pour cette raison, elle est utilisée en première étape dans un processus de minimisation.

La méthode du gradient conjugué (« Conjugate Gradients » ou méthodes Fletcher - Reeves) :

Tout comme la méthode précédente, la recherche du *minimum* se fait dans le sens contraire au gradient, mais en tenant compte des gradients calculés dans les étapes précédentes ; ce qui rend les calculs plus précis. La nouvelle direction est obtenue par l'équation suivante :

$$h_{i+1} = g_{i+1} + \gamma_i h_i \quad (\text{II-4})$$

Avec
$$\gamma_i = \frac{g_{i+1} \times g_{i+1}}{g_i \times g_i}$$

h_{i+1} : Vecteur représentant la nouvelle direction

g_{i+1} : Gradient calculé pour le nouveau point

La méthode du gradient conjugué converge mieux que celle de la plus grande pente. Elle présente l'intérêt d'échapper aux comportements oscillatoires autour du minimum tout en accélérant la convergence, mais elle n'est pas adaptée aux structures présentant de mauvais contacts ou pas très bien relaxées. Pour pallier à cela, il est conseillé d'utiliser les deux algorithmes de minimisation l'un après l'autre.

Méthode Newton-Raphson :

Contrairement aux deux méthodes précédentes qui n'utilisent comme information que la pente, la méthode de Newton-Raphson utilise la dérivée seconde de l'énergie. Cette méthode permet donc la prédiction de la courbure locale de l'hypersurface énergétique par le calcul du Hessien (matrice des dérivées secondes). Leur inconvénient réside dans le fait qu'elle nécessite un espace mémoire important et qu'elle ne peut pas être appliquée aux gros systèmes (macromolécules).

Comme le montre la figure suivante, les algorithmes de minimisation ne conduisent pas toujours vers le minimum global mais plutôt vers le minimum le plus proche ; ils dépendent tous de la conformation de départ.

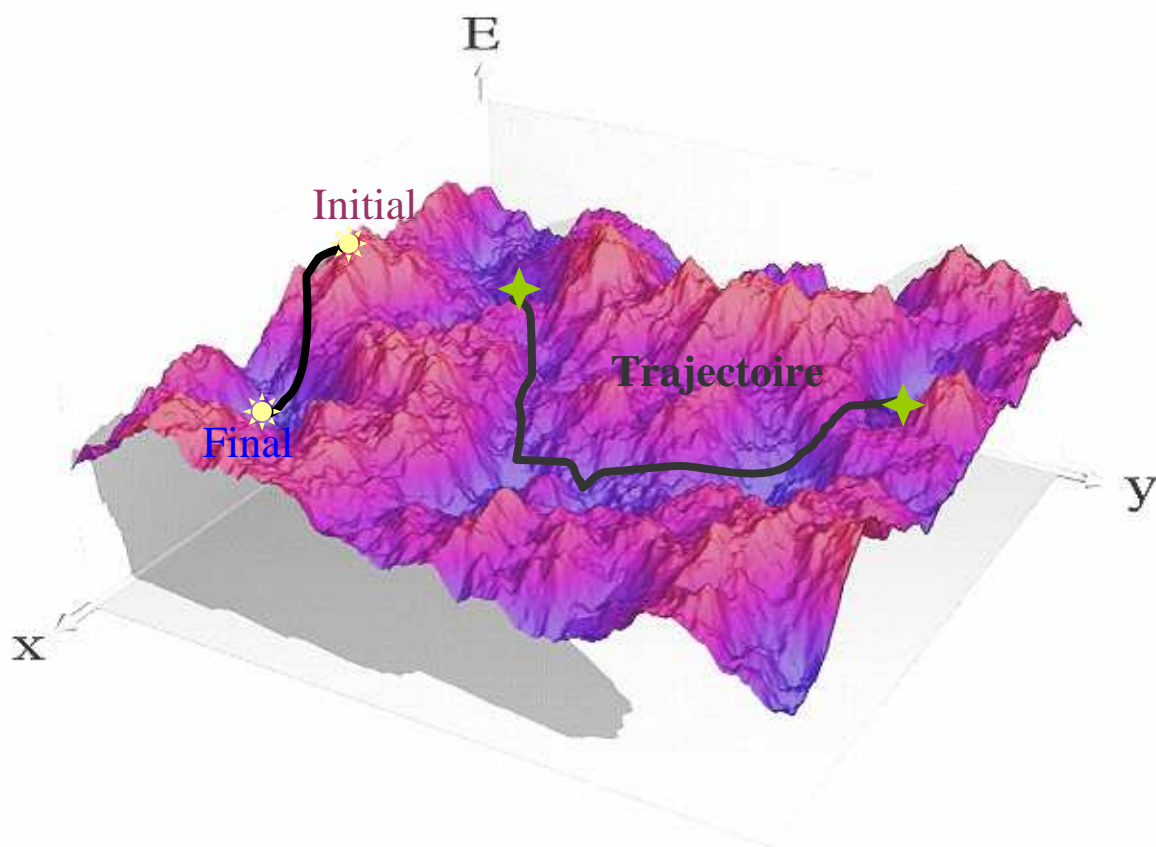


Figure 2 : Représentation d'une portion de l'hypersurface d'énergie potentielle

Afin d'éviter d'être piégé dans le minimum le plus proche, il est fortement conseillé de faire un échantillonnage de l'espace conformationnel de la molécule. Il existe, pour cela,

plusieurs techniques permettant d'explorer l'espace conformationnel d'une molécule. Cependant, cette exploration se fait de manière non exhaustive, à cause du grand nombre de degrés de liberté associé au système étudié.

C2 - Rotation des angles de torsion

Les différences structurales existant entre deux conformères concernent d'abord les angles dièdres, puis les angles de valence et les longueurs de liaison. Par conséquent, une des méthodes efficaces pour l'exploration de l'espace conformationnel de la molécule consiste à faire varier chacun de ses angles dièdres par petits incréments en fixant les angles de valence et les liaisons. Pour cela, le nombre de conformères théoriques est donné par P^i .

Avec **P**: le nombre d'états possibles par torsion pour chaque angle dièdre

i: le nombre d'angles dièdres contenus dans la molécule

Par exemple, pour une molécule possédant 2 angles dièdres et si une torsion de 20° est effectuée, 324 conformères seront alors générés ($P = 360/20 = 18 \rightarrow 18^2 = 324$). Idem pour une molécule ayant 3 angles de torsion, 5832 conformères générés....etc. Il est certain que l'espace conformationnel est parcouru de manière assez exhaustive mais le nombre de conformations monte *en crescendo* avec l'augmentation du nombre d'angles de torsions ; ce qui rend cette technique trop longue et onéreuse pour les systèmes ayant plus de 4 angles dièdres. Pour remédier à ce problème d'autres techniques, dites méthodes statistiques, ont été développées pour l'échantillonnage de l'espace conformationnel de composés. Parmi ces méthodes, nous citons la dynamique moléculaire, le recuit simulé et la méthode de Monté Carlo.

D- Méthodes statistiques

D1 – Dynamique Moléculaire

La Dynamique Moléculaire (DM) est la science de la simulation d'un système de particules. Elle a été appliquée à des systèmes allant de l'atome jusqu'à une galaxie. La dynamique moléculaire produit des trajectoires c'est à dire les coordonnées de tous les atomes du système en fonction du temps ⁽¹⁾.

Comme pour la minimisation d'énergie, la dynamique moléculaire change les degrés de liberté du système pas à pas. Dans ce cadre, les pas représentent les changements des positions atomiques, r_i , en fonction du temps (i.e. vitesses).

La dynamique moléculaire utilise, en général, l'équation de mouvement de Newton pour **simuler** les mouvements atomiques et décrire leur évolution en fonction du temps.

$$F_i = m_i \times a_i \quad (\text{II-5})$$

Avec :

F_i : Force agissant sur l'atome i

m_i : Masse de l'atome i

a_i : Accélération de l'atome i

La force exercée sur un atome peut être calculée à partir de la variation d'énergie entre deux positions.

$$-\frac{dE}{dr_i} = F_i \quad (\text{II-6})$$

L'énergie E est obtenue par mécanique moléculaire

$$-\frac{dE}{dr_i} = m_i \frac{d^2 r_i}{dt^2} \quad (\text{II-7})$$

Connaissant les forces et les masses on peut ensuite déterminer les positions de chaque atome le long d'une série de pas de l'ordre de la femtoseconde (10^{-15} seconde).

La série de positions instantanées obtenue constitue une trajectoire.

En pratique, les accélérations atomiques sont calculées à partir des forces et des masses (II-5).

Les vitesses sont calculées à partir des accélérations par la relation suivante :

$$a_i = \frac{dv_i}{dt} \quad (\text{II-8})$$

Les positions sont calculées à partir des vitesses :

$$v_i = \frac{dr_i}{dt} \quad (\text{II-9})$$

On assigne des vitesses initiales aux atomes en fonction de l'énergie cinétique totale elle-même fonction de la température de simulation. Pour résumer, la dynamique moléculaire est la répétition successive de deux opérations :

- Evaluation de la force qui agit sur chaque atome au temps "t".

- Détermination des coordonnées et des vitesses atomiques au temps " $t + \delta t$ " en fonction des forces subies par chacun d'entre eux. Sachant que δt est le pas d'intégration ⁽¹⁾.

Il faut noter qu'il n'y a pas de solutions analytiques aux équations de mouvement. Pour cela, des intégrations numériques différant par l'algorithme utilisé sont disponibles dans les programmes de dynamique moléculaire. Le plus fréquemment utilisé est l'algorithme de Verlet, détaillé dans un travail antérieur ⁽⁶⁾.

Le résultat d'une simulation de DM consiste en une série de positions atomiques (trajectoire) au cours du temps. La résolution numérique des équations nécessite l'utilisation d'un pas de temps extrêmement petit, de l'ordre de la fs de manière à prendre en compte les vibrations atomiques les plus rapides et en assurant la stabilité des calculs. Ceci explique que le temps maximum de simulation dépasse rarement les 10 à 100 ns pour des systèmes complexes, de plusieurs dizaines de milliers d'atomes (protéines en présence de solvant ou insérées dans des membranes ...).

Les résultats de la simulation peuvent être exploités en analysant les trajectoires permettant d'aider à la compréhension du mécanisme biologique étudié. On peut estimer entre autres : les fluctuations atomiques, les corrélations entre les mouvements de parties distinctes d'une protéine, le potentiel électrostatique vu par un ligand s'approchant d'une protéine, les modes de vibrations d'une macromolécule etc...

Les limitations dans la taille du système et dans le temps de simulation (~10 ns) ne représentent pas, de nos jours, des obstacles incontournables à l'obtention d'informations intéressantes pour la compréhension de la fonction biologique. En effet, le problème de la petite taille peut être résolu partiellement en simulant le système dans des conditions périodiques et, de plus, l'étude des protéines sur des échelles de temps de la nanoseconde, près de leur structure cristallographique, est souvent très riche en informations, concernant les interactions et corrélations entre résidus.

D2 - Monté Carlo

Cette méthode trouve ses racines à la fin de la 2^{ème} guerre mondiale. Elle a été développée par Von Newman Ulam et Metropolis pour l'étude de la diffusion des neutrons dans un matériau fissile. La méthode d'exploration de l'espace conformationnel de Monté Carlo est basée sur la génération aléatoire d'un ensemble de conformations. Le système à

étudier est supposé avoir un grand nombre d'états possibles "i" qui correspondent à des conformations. Ces conformations générées aléatoirement seront gardées si elles ont un poids statistique donné par la relation ci-dessous :

$$P_i = e^{-\frac{E_i}{k_B T}} \quad (\text{II-10})$$

Avec :

E_i : L'énergie d'un état i

k_B : La constante de Boltzmann

T : La température

La méthode de Monté Carlo n'introduit pas la dimension du temps, contrairement à la dynamique moléculaire, et ne permet pas l'obtention d'informations sur les propriétés dynamiques du système.

D3- Recuit Simulé

Cette technique d'échantillonnage repose sur le principe de chauffage puis refroidissement du système étudié. Après avoir subi une dynamique moléculaire à haute température, pour franchir les barrières énergétiques de *minima* locaux, le système est refroidi progressivement puis équilibré à plus basse température. Il est alors piégé dans un puits de potentiel assez profond, une minimisation est alors effectuée pour atteindre le minimum. Dans le but de balayer suffisamment l'espace conformationnel d'une molécule, cette procédure est répétée plusieurs fois, à partir de la structure obtenue précédemment. Le recuit simulé permet, en plus d'explorer l'espace conformationnel d'une molécule, d'étudier ses propriétés dynamiques. Il est bien adapté aux molécules ayant un grand nombre de degrés de liberté.

E- Conclusion

Nous donnons dans le tableau comparatif et récapitulatif suivant les différences entre les deux méthodes de base de la chimie théorique :

Tableau 1 : Comparaison des avantages et des inconvénients des deux approches de base

Mécanique Quantique	Mécanique Moléculaire
<ul style="list-style-type: none"> ✓ Noyaux et électrons distincts ✓ Les interactions électrons-électrons et électrons-noyaux sont explicites ✓ Les interactions sont régies par les charges électroniques et nucléaires (énergie potentielle) et les mouvements électroniques ✓ Les interactions déterminent la distribution spatiale des électrons et des noyaux ainsi que leurs énergies 	<ul style="list-style-type: none"> ✓ Noyaux et électrons modélisés en une seule particule ✓ Les particules sont sphériques (rayons atomiques expérimentaux ou théoriques) et possèdent une charge nette (théorique) ✓ Les interactions sont modélisées par des ressorts et des potentiels classiques ✓ Les interactions doivent être pré-assignées à un jeu spécifique de type d'atomes ✓ Les interactions déterminent la distribution spatiale des particules atomique et leurs énergies

Dans le domaine de la modélisation moléculaire, le choix de la méthode à appliquer dépend du type de calcul à réaliser et de la taille de la molécule. En effet, les méthodes issues de la mécanique quantique sont réservées à l'étude de la réactivité et limitées aux systèmes ne dépassant pas, compte tenu des moyens actuels, une centaine d'atomes. En revanche, les méthodes basées sur la mécanique moléculaire traitent les problèmes essentiellement structuraux et peuvent simuler des systèmes chimiques contenant des milliers d'atomes.

III - METHODES DU DRUG DESIGN

La conception de composés actifs et bénéfiques pour l'homme est favorisée par plusieurs évolutions : informatique, biologie et pharmacochimie. En effet, durant des années déjà, la recherche tente de développer les méthodes du drug design et les premières tentatives, faute de moyens, se sont soldées par des corrélations statistiques entre les descripteurs structuraux et les activités biologiques des systèmes étudiés. L'essor du domaine informatique durant ces vingt dernières années a banalisé l'utilisation des processeurs à grande vitesse et les stations de travail ; ce qui a engendré l'évolution de ce qui est actuellement appelé « conception de médicament assistée par ordinateur » ou encore « Computer Assisted Drug Design » (CADD). Ainsi, le drug design peut être défini comme étant une approche visant à rechercher des molécules thérapeutiques par conception *in silico* sur la base d'informations expérimentales accessibles.

En se basant sur la disponibilité de l'information, deux approches principales sont connues dans le domaine du drug design, l'approche du « ligand-based » et celle du « receptor-based »⁽⁷⁾. La première est appliquée quand des séries de composés manifestent des activités intéressantes sur une cible thérapeutique dont la structure "peut" être inconnue. Elle utilise des outils tels que le QSAR (Quantitative Structure-Activity Relationships) et CoMFA (Comparative Molecular Field Approach). L'approche « receptor-based » est appliquée, quant à elle, lorsqu'un modèle 3D fiable de récepteur est disponible par diffraction aux rayons-X, études RMN ou modélisation par homologie. Les informations relatives au récepteur serviront de base pour la conception d'une tête de série. C'est le problème que tente de résoudre les méthodes de criblage virtuel, docking ou encore le *de novo* design.

Nous présentons donc la méthode, la plus connue, relative à chacune des deux approches utilisées dans le drug design (ligand-based et receptor-based).

A- QSAR (Quantitative Structure-Activity Relationships)

Le QSAR représente une des plus anciennes méthodes dans le drug design puisque ses débuts remontent à la fin du XIXe siècle. En effet, Ehrlich introduit le QSAR comme une étude réalisée sur la base unique d'un ensemble de ligands reconnus affins à une protéine⁽⁷⁾. Une étude QSAR ou de relations structure-activité consiste à déterminer les éléments structuraux responsables de l'activité de ligands vis-à-vis de leur récepteur⁽⁸⁾. Elle a pour objectif de proposer des outils permettant la détermination prévisionnelle de la structure des molécules actives dans une classe pharmacologique donnée. Elle prend en compte l'étude des paramètres critiques pour l'émergence de l'activité et son optimisation : paramètres électroniques, stériques et de lipophilie. Les méthodes prévisionnelles, couplées aux apports de la modélisation moléculaire, visent à permettre une approche rationnelle de la conception du médicament.

En analysant, d'un point de vue quantitatif, les relations structure-activité existantes entre diverses molécules médicamenteuses, le chercheur tente d'une part, d'identifier et d'apprécier les propriétés physico-chimiques des médicaments et, d'autre part, d'évaluer l'impact de ses diverses propriétés sur leur activité biologique. A l'aide des méthodes statistiques, il est possible d'établir une relation mathématique corrélant les influences réciproques des paramètres concernés. Au moyen de ce type d'équation, il est possible d'affirmer le rôle, plus ou moins important, des propriétés prises en considération. Il est également possible d'extrapoler les résultats à des fins prédictives. Les propriétés physico-chimiques dont il est

question sont essentiellement les propriétés structurales, physiques et chimiques des molécules. Les propriétés physico-chimiques les plus couramment étudiées sont : l'hydrophobicité, les effets électroniques, les facteurs stériques, les liaisons hydrogènes, les distances interatomiques ainsi que les conformations moléculaires.

D'autre part, dans le cas où la structure 3D de la protéine d'intérêt est connue, il est important de tenir compte des informations qu'elle apporte en termes de propriétés moléculaires du site actif. Ces informations permettent de mieux définir un pharmacophore très utile au criblage des chimiothèques ou bien d'utiliser une méthode d'assemblage ligand-protéine ("docking").

B- Le Criblage à Haut Débit ou High Throughput Screening (HTS) :

D'importants changements dans le domaine de la biologie se sont produits au cours des dernières décennies. Il faut noter que de vastes quantités d'informations et de données biologiques, issues de la cartographie complète du génome humain et du protéome, ont été générées et continuent toujours de l'être. L'industrie pharmaceutique s'est d'abord associée au domaine de la robotique pour gérer la surabondance de ces données. Comme l'expérience nécessite énormément de moyens matériels, une nouvelle discipline : la bioinformatique, est née de la nécessité d'organiser, de sauvegarder et d'analyser ce déluge d'informations, à moindre coût. La bioinformatique intervient, entre autres, dans le criblage qui consiste à améliorer l'identification des composés prometteurs et ensuite leur optimisation en réduisant les coûts et en augmentant la probabilité de découverte de médicaments.

Actuellement, nous distinguons deux types de criblage dans l'industrie du drug-design, ayant le même principe de base mais très différents par les moyens employés.

B1-Le criblage réel :

Le criblage réel implique l'examen de centaines de milliers de composés pour identifier d'éventuelles molécules affines à une cible biologique. Le problème est d'identifier, parmi toutes ces molécules, celles qui sont pourvues des propriétés biologiques les plus intéressantes.

La technique consiste à mettre en présence la substance à tester et un système biochimique et ensuite de mesurer l'importance de la réaction éventuelle. Les opérations se déroulent sans intervention humaine et évoluent avec les techniques de la robotique. Cependant, ce procédé, appelé le **criblage réel à haut débit**, est limité aussi bien par le nombre de composés

possibles à tester que par le coût des tests qui est estimé à 1 dollar par molécule. Pour remédier au problème du coût d'argent et de temps, les sociétés pharmaceutiques ont de plus en plus recours à l'outil informatique qui est de moins en moins cher et de plus en plus performant.

B2-Le criblage virtuel

Le criblage ou screening virtuel a émergé comme une approche complémentaire et alternative au problème de coût d'argent et de temps relatif au criblage réel à haut débit de larges librairies de molécules. Il peut être défini comme la recherche, dans une chimiothèque, de molécules complémentaires d'une cible biologique⁽⁹⁾.

Généralement, le processus de criblage virtuel est fragmenté en différentes étapes faisant apparaître une complexité croissante⁽¹⁰⁾. Dans la première étape, les filtres non spécifiques de la cible sont appliqués pour éliminer les structures chimiques ayant des propriétés de non-drug-like.

Le recours aux recherches topologiques de ligands connus est habituellement utilisé dans la prospection de composés ayant des caractéristiques différentes. Comme ces méthodes ne nécessitent pas une détermination de la conformation elles sont appropriées à un criblage rapide de bases de données, pouvant contenir plusieurs millions de molécules⁽¹¹⁾. La détermination de conformations rigides et bioactives d'un ou plusieurs ligands, provenant de méthodes qui déterminent la structure des composés ou bien issues des recherches par pharmacophore 3D, représente une option supplémentaire pour le criblage virtuel de chimiothèque. Finalement, lorsque la structure 3D de la cible biologique est connue ou obtenue au moyen d'une modélisation par homologie, les ligands qui ont traversé les premiers filtres peuvent être soumis au **docking et scoring** moléculaire pour fournir des candidats potentiels aux tests expérimentaux.

Les applications du criblage virtuel sont variées, il permet entre autres de reconnaître les ligands de cibles orphelines ainsi que l'identification et l'optimisation de nouveaux ligands⁽¹⁰⁾.

Au moyen du criblage virtuel, il est également possible de discriminer des vraies touches de ligands choisis par hasard, d'enrichir en vraies touches une liste réduite de ligands. Bien entendu, par sa capacité à tester un très grand nombre de ligands sur des cibles biologiques

bien définies, le criblage virtuel est souvent utilisé comme filtre préalable aux tests expérimentaux.

Comme nous l'avons mentionné plus haut, le docking est considéré comme l'un des principaux outils du criblage virtuel. Nous détaillons ci-dessous cette technique, largement utilisée dans le cadre de nos travaux :

α - Docking moléculaire

Le docking moléculaire qui veut dire "chargement moléculaire" ou encore "arrimage moléculaire" est l'ensemble des mécanismes et interactions intervenant lors de la formation de complexes moléculaires.

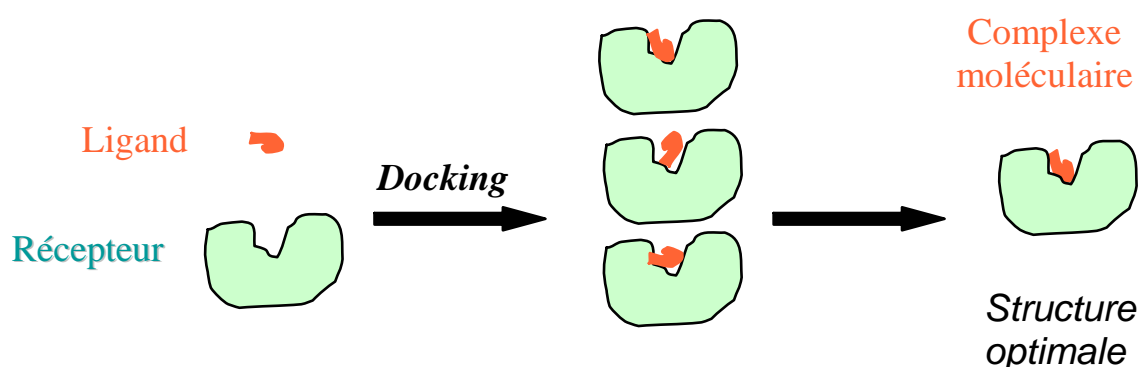


Figure 3 : Illustration schématique du processus de docking.

Les programmes de docking moléculaire, disponibles actuellement, sont basés sur l'hypothèse que les ligands formant des interactions favorables avec le récepteur doivent avoir une affinité de liaison élevée⁽¹²⁾. Ils reposent sur l'exploration systématique de l'espace des configurations pour générer et évaluer un grand nombre de liaisons potentielles. Cependant, les ordinateurs ne possèdent pas l'expérience, l'intuition et l'expertise des scientifiques à analyser les liaisons calculées pour générer des solutions robustes et fiables. Les méthodes de docking moléculaire actuelles nécessitent donc un utilisateur bien averti dans ce domaine et sachant bien exploiter ses résultats et son intuition.

Les principaux défis dans ce domaine sont^(13, 14) :

- La prédiction de l'assemblage des molécules.
- L'identification des facteurs qui déterminent la spécificité de l'interaction.
- L'analyse visuelle des interactions ligand-récepteur permettant une interprétation intuitive et la compréhension du processus de liaison.
- L'identification de nouveaux ligands avec différents modes de liaisons.

Le chercheur utilisant la technique de docking peut être comparé à un serrurier qui essaye successivement les clés d'un trousseau sur une serrure inconnue, la serrure étant la cible moléculaire et le trousseau de clés la chimiothèque contenant les ligands. Les logiciels de docking calculent donc l'énergie d'interaction entre ce ligand potentiel et la macromolécule cible. Les molécules qui présentent des interactions favorables seront ensuite étudiées expérimentalement. Cette technique permet de "tester" quelques dizaines de milliers de molécules par jour sur un ordinateur standard actuel.

Les différents types de docking

Selon les protocoles propres à chaque logiciel, nous pouvons avoir deux types de docking :

1) Le docking rigide : le ligand est flexible afin d'explorer toutes ses conformations possibles, mais la cible est statique ou rigide. Ce type de docking présente, certes, l'avantage de nécessiter un temps de calcul relativement court mais il ne tient pas compte des changements conformationnels, requis par la cible. Nous donnons ci-dessous un exemple d'algorithme de docking rigide, utilisé dans nos travaux de recherche et qui est **MSSH** (Molecular Surface using Spherical Harmonics)⁽¹⁵⁻¹⁷⁾.

L'interaction entre protéines peut être décrite par l'analyse des surfaces de reconnaissance mutuelle complémentaire. Pour modéliser, décrire et analyser les surfaces en interaction (distribution des charges, zones hydrophiles/hydrophobes), il est possible de décrire chaque objet à l'aide d'harmoniques sphériques permettant l'approximation de la surface moléculaire avec une précision voulue. Pour cela, nous avons utilisé l'algorithme de calcul et de comparaison de surface : **MSSH**.

La série des coefficients harmoniques sphériques 1D, peut être comparée à des empreintes digitales de la molécule avec lesquelles sa surface peut être recouverte.

Comme cette surface moléculaire peut être assimilée à une sphère, chaque point de la sphère possède une coordonnée sphérique unique $Y_l^m(\theta, \phi)$ telle que :

$$Y_l^m(\theta, \phi) = \sqrt{\frac{2l+1}{4\pi} \frac{(l-m)!}{(l+m)!}} P_l^m(\cos \theta) e^{im\phi} \quad (\text{III-1})$$

Avec l et m , deux nombres entiers déterminant le nombre et l'arrangement spatial des nœuds dans chaque fonction et $P_l^m(\cos \theta)$ est la fonction associée de Legendre

Ainsi n'importe quelle valeur unique de surface 3D peut être évaluée en codant la distance radiale des points de cette surface, à partir de l'origine, comme une somme de fonctions harmoniques sphériques.

$$r(\theta, \phi) = \sum_{l=0}^L \sum_{m=-l}^l C_{lm} Y_l^m(\theta, \phi) \quad (\text{III-2})$$

Avec:

$r(\theta, \phi)$: Fonctions représentant les distances entre les points de la surface et de l'origine.

C_{lm} : Groupes des coefficients d'expansion des fonctions harmoniques sphériques

L : L'ordre qui détermine le degré de précision de la représentation

2) Le docking flexible : Dans ce cas, le ligand est totalement flexible, alors que la cible l'est partiellement. En fait, seuls les résidus jouant un rôle dans l'interaction avec le ligand sont concernés par cette flexibilité. Ceci représente un véritable challenge pour le docking à cause de la flexibilité des ligands présentant plusieurs rotations libres. Ce type de docking a la qualité de mieux représenter la réalité mais nécessite des temps CPU énormes, d'où le recours au calcul en parallèle.

Nous avons utilisé dans nos travaux, l'algorithme de docking flexible existant dans le logiciel GOLD (Genetic Optimization for Ligand Docking) qui est un produit de la collaboration entre GlaxoSmithKline plc (Université de Sheffield) et CCDC (Cambridge Crystallographic Data Centre)^(18,19). GOLD assure une flexibilité complète du ligand et prend partiellement en compte celle du récepteur. Il génère les conformations en utilisant un algorithme génétique basé sur les principes de l'évolution pour placer le ligand dans le site de liaison préalablement défini.

L'algorithme génétique est un "*algorithme stochastique itératif*" qui opère sur des ensembles de points *codés*, à partir d'une *population initiale*, et qui est bâti à l'aide de trois opérateurs : *sélection, croisement et mutation*. Comme le montre l'organigramme suivant, le premier opérateur fait évoluer la population vers les *optima* du problème, tandis que les deux derniers sont des opérateurs d'exploration de l'espace.

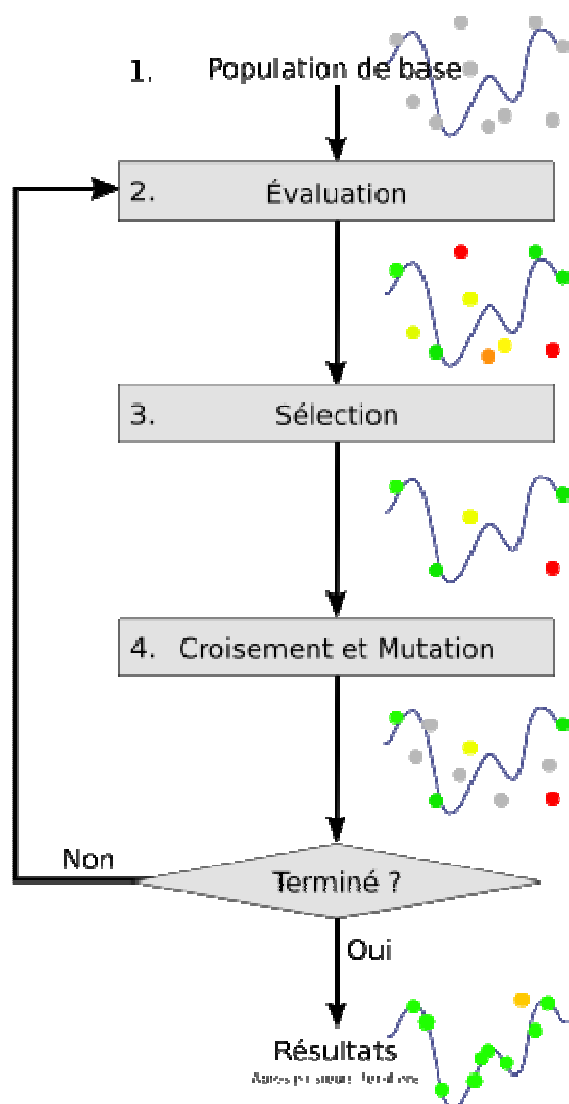


Figure 4 : Organigramme de l'algorithme génétique

L'algorithme génétique est une procédure qui s'inspire des mécanismes de sélection naturelle et des phénomènes génétiques. Le principe de base consiste à simuler le processus d'évolution naturelle dans un environnement hostile. Cet algorithme utilise un vocabulaire similaire à celui de la génétique, cependant, les processus auxquels il fait référence sont beaucoup plus complexes.

On parle ainsi d'individu dans une population et l'individu est composé d'un ou plusieurs chromosomes. Ces derniers sont eux-mêmes constitués de gènes qui contiennent les caractères héréditaires de l'individu.

Pour un problème d'optimisation donné, un individu représente un point de l'espace d'état et, en ce qui nous concerne, il correspond à une conformation. On lui associe la valeur du critère

à optimiser ou plus généralement du critère d'adaptation au problème (fitness). L'algorithme génère ensuite de façon itérative des populations d'individus sur lesquelles on applique des processus de sélection, de croisement et de mutation. La sélection a pour but de favoriser les meilleurs éléments de la population ; on cherche ainsi à privilégier la reproduction des "bons" éléments au détriment des "mauvais". Des opérateurs d'exploration de l'espace sont ensuite utilisés pour "élargir" la population (croisement) et introduire de la nouveauté (mutation) d'une génération sur l'autre.

Tous les logiciels et programmes de docking moléculaire font l'évaluation en utilisant des fonctions de scores ou fitness. Cette dernière peut être définie comme étant l'estimation rapide et simplifiée des énergies d'interaction prises en compte au cours du docking.

Il existe plusieurs types de fonctions de score qui sont globalement plus ou moins équivalentes en termes de prédiction.

La fonction de score de GOLD, utilisée pour décider quels individus (conformations) survivront et produiront des descendants pour la prochaine génération peut être de deux types voire trois : le Goldscore (utilisé par défaut) et le Chemscore⁽¹⁸⁻²⁰⁾. Il offre également la possibilité à l'utilisateur de définir sa propre fonction de score. Dans nos travaux, nous avons utilisé la fonction "Goldscore" qui s'adapte mieux à nos besoins de filtrage. Cette fonction prend en compte, dans l'estimation de l'énergie de liaison, différents termes dus aux interactions de type :

- liaison hydrogène entre le ligand et la protéine (*external H-bond*)
- van der Waals entre le ligand et la protéine (*external vdw*)
- van der Waals interne du ligand (*internal vdw*)
- torsions du ligand (*internal torsion*)

Le principal avantage de GOLD est sa fiabilité à prédire des structures cristallographiques pour des complexes de type ligand/récepteur. Notons un taux de succès de 72% à un test standardisé employant des structures issues de la Protein Data Bank⁽²⁰⁾.

Cependant, certains défauts peuvent être signalés :

- Seuls les complexes présentant au moins une liaison hydrogène sont traités.
- La fonction de score présentée par défaut prend insuffisamment en compte des zones de contact hydrophobe et, le paramétrage d'une fonction de score personnalisée nécessite de nombreux réglages.

– GOLD décrit comme acceptables des "configurations" dans lesquelles le ligand pénètre "assez nettement" la cavité du récepteur ; ce qui conduit à de faux positifs.

- Il favorise les "configurations" qui interdisent la présence de toute molécule d'eau entre la protéine et le ligand, selon la croyance qu'un "bon" ligand est celui qui chasse toute molécule d'eau du récepteur ; ce qui n'est pas en accord avec l'affirmation de Ladbury. En effet, selon cette affirmation, lors de l'étude de systèmes biologiques, négliger les effets des molécules d'eau du solvant diminue les chances que l'on a d'obtenir des résultats satisfaisants.

Nous présentons le tableau ci-dessous issu d'une étude comparative des 8 programmes de docking les plus fréquemment utilisés

Tableau 2 : Forces et faiblesses de quelques logiciels de docking⁽²¹⁾.

Programme	Forces	Faiblesses
DOCK	Sites de liaisons étroits Cavités ouvertes Ligands hydrophobes de petite taille	Ligands flexibles Ligands fortement polaires
FLEXX	Sites de liaisons étroits Ligands hydrophobes de petite taille	Ligands fortement flexibles
FRED	Sites de liaisons larges Ligands flexibles Ligands hydrophobes de petite taille Calcul très rapide	Ligands de petite taille Ligands polaires Ligands enfouis
GLIDE	Ligands flexibles Ligands hydrophobes de petite taille	Classement de ligands fortement polaires Temps CPU important
GOLD	Sites de liaisons de petite taille Ligands hydrophobes de petite taille	Classement de ligands fortement polaires Classement de ligands dans de grandes cavités
SLIDE	Flexibilité des chaînes latérales	Sensibilité aux coordonnées introduites du ligand
SURFLEX	Cavités grandes et ouvertes Sites de liaisons de petite taille Ligands très flexibles	Temps CPU important pour les ligands de grande taille
QXP	Optimisation des modes de liaisons connus	Sensibilité aux coordonnées introduites du ligand

Comme nous l'avons mentionné précédemment, le docking a pour objectif d'estimer la "viabilité" du complexe moléculaire formé. Cette évaluation ou estimation se fait en analysant

de visu, sur un écran avec une bonne performance graphique, les différentes interactions entre le ligand et le récepteur. Pour cela, il est nécessaire au modélisateur d'avoir quelques connaissances sur ce type d'interactions intermoléculaires qui peuvent se manifester sous plusieurs aspects et différant, en particulier, par leur intensité.

◆ Interactions intermoléculaires

Rappelons que la structure tridimensionnelle de la protéine est aussi bien stabilisée par des interactions faibles et non covalentes (effets électrostatiques, van der Waals et hydrophobes) que par des liaisons covalentes (ponts disulfure). Ces interactions que nous détaillons ci-dessous se maintiennent quand la protéine se lie avec une molécule, dite son ligand.

Interactions hydrophobes :

L'effet hydrophobe est induit par les composés apolaires. Ces derniers sont dépourvus de groupements chargés ou d'atomes aptes à former des liaisons hydrogène et, sont prédisposés à s'associer avec des composés également apolaires. Ils ont tendance à fuir le milieu aqueux environnant ; ce qui leur permet, dans le cas des protéines, de se regrouper au coeur de celle-ci et, de réduire leur surface de contact avec le solvant. Il faut noter que ce type d'interaction est fondamental pour le repliement des protéines et limite la surface de contact avec l'eau.

Dans les protéines, les résidus qui s'avèrent hydrophobes sont : I, L, A, V, M, W, et F.

Interactions de van der Waals :

Ce type d'interactions se manifeste entre tous les atomes constituant le système étudié. Elles sont répulsives à courte distance (inférieure à la somme des rayons de van der Waals) et attractives à longue distance.

Interactions électrostatiques :

Ces interactions s'exercent entre atomes ou molécules polarisés. L'énergie de cette interaction est de l'ordre de quelques kJ/mol, mais pour les gros systèmes (macromolécules), elle devient importante à cause du nombre élevé d'atomes impliqués ; ce qui leur confère un rôle important dans la stabilisation des systèmes protéiniques.

Interactions ioniques ou ponts salins:

Les interactions ioniques se manifestent entre les atomes de charge opposée (mais relativement forte) et sont appelées : ponts salins. Ces ponts peuvent concerner les résidus chargés positivement (R, K, H et l'extrémité C-terminale) et ceux chargés négativement (D, E et l'extrémité N-terminale).

Liaisons hydrogène :

Les liaisons hydrogène résultent de l'interaction entre un atome électronégatif, tel que l'atome d'oxygène ou l'azote, et un atome d'hydrogène, porté par un atome électronégatif⁽²²⁾.

Les acides aminés polaires ont la possibilité de former des liaisons hydrogène entre eux ou avec des molécules d'eau ; ce qui leur donne une caractéristique hydrophile.

La liaison hydrogène est utilisée en biologie comme un vecteur des interactions intermoléculaires et elle contribue à la reconnaissance moléculaire.

Pont disulfure ou cystine :

La liaison covalente S-S, appelée généralement pont disulfure, est la seule liaison covalente entre 2 cystéines distantes dans la séquence primaire mais proches dans la structure tertiaire. Chimiquement, cette interaction est le résultat de l'oxydation entre deux groupements thiol (-SH) appartenant à deux cystéines (C) distantes de 2.2 Å au niveau de la structure tertiaire. Les ponts disulfures sont donc des points importants à considérer dans un processus de prédiction du repliement et ne concernent jamais la liaison du ligand au récepteur.

Le pont disulfure impose une forte contrainte topologique à la chaîne polypeptidique due à son énergie de liaison estimée à 60 kcal/mol.

Nous récapitulons ces différentes interactions moléculaires ainsi que les énergies mises en jeu, dans la figure 5 et dans le tableau 3, respectivement :

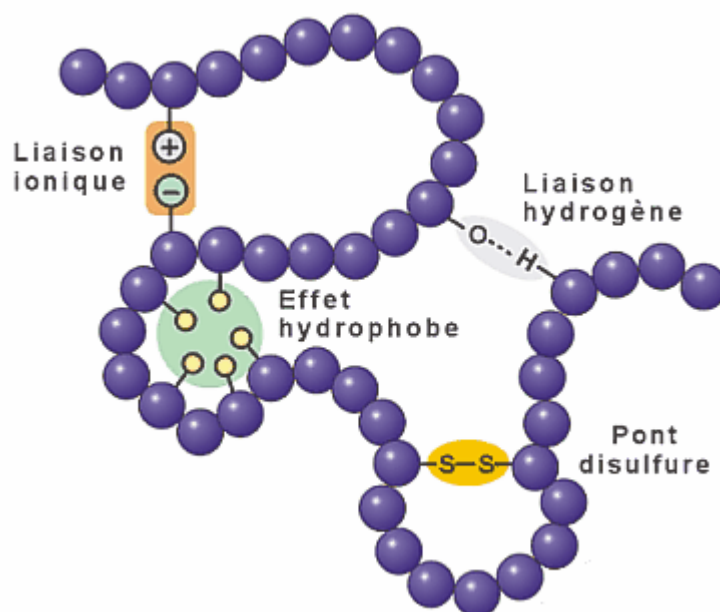


Figure 5 : Représentation des principales interactions dans les protéines.

Tableau 3 : Comparaison des énergies relatives aux différents types d'interactions

Types d'interactions	Energie (kcal/mol)
van der Waals	0,24 - 1
Hydrophobes	0,5 - 2
Hydrogène	1-10
Electrostatique	0,5 - 10
Ionique	40-200
Pont disulfure	60
Covalente	100-500

REFERENCES BIBLIOGRAPHIQUES

- 1- Minoux, H., Thèse de Doctorat, Etude et conception assistées par ordinateur d'antagonistes des intégrines de la famille β_3 . Nancy, UHP, **1998**.
- 2- Akacem, Y., Thèse de Doctorat d'état, Etude Théorique des Mécanismes de Réaction des Composés du Phosphore d'Intérêt Biologique. Alger, USTHB, N° d'ordre : 05/1996-E/CH.
- 3- Cheikh, F., Thèse de Doctorat d'état, L'approche de la Chimie Quantique en Sciences de l'Univers : Processus en Milieux Interstellaire et Atmosphérique. Alger, USTHB, N° d'ordre : 02/2002-E/CH.
- 4- Lokbani - Azzouz, N. S., Thèse de Doctorat d'état, Etude Théorique de Clusters Inorganiques et Organométalliques : Approche Théorique de la Fluxionalité de Squelette. Alger, USTHB, N° d'ordre : 06/2002-E/CH.
- 5- Leach, A. R., Molecular Modelling. Essex, LONGMAN, **1996**.
- 6- Taïri-Kellou, S., Thèse de Doctorat d'état, Conception Assistée par Ordinateur des Complexes ligand/récepteur de la Cholécystokinine, Alger, USTHB, N° d'ordre : 04/2002-E/CH.
- 7- Patrick, G. L., Chimie Pharmaceutique. Bruxelles, de boeck, **2003**.
- 8- Fanelli, F., Menziani, C. et De Benedetti, P. G., Molecular dynamics simulations of m3-Muscarinic receptor activation and QSAR analysis. *Bioorganic & Medicinal Chemistry*, 11, 1465-1477, **1995**.
- 9- Hibert, M., Haiech, J., Des gènes aux médicaments : nouveaux défis, nouvelles stratégies. *médecine/ sciences*, 16, 1332-1339, **2000**.
- 10- Evers, A., Hessler, G., Matter, H. et Klabunde, T., Virtual screening of biogenic amine-binding G-protein coupled receptors: Comparative evaluation of protein- and ligand-based virtual screening protocols. *J. Med. Chem.*, 48, 5448-5465, **2005**.
- 11- Sukjoon, Y., Smellie, A., Hartsough, D. et Filikov, A., Surrogate docking: structure-based virtual screening at high throughput speed. *Journal of Computer-Aided Molecular Design*, DOI 10.1007/s10822-005-9002-6, **2005**.
- 12- Tress, M., De Juan, D., Grana, O., Gomez, M. J., Gomez-Puertas, P., Gonzalez, J. M., Lopez, G. et Valencia A., Scoring docking models with evolutionary information. *PROTEINS : Structure, Function, and Bioinformatics*, 60, 275-280, **2005**.

- 13- Duhovny, D., Nussinov, R. et Wolfson, H. J., Efficient unbound docking of rigid molecules. Springer, 2452, 185-200, **2002**.
- 14- Bindewald, E. et Skolnick, J., A scoring function for docking ligands to low-resolution protein structures. J. Comput. Chem., 26, 374-383, **2005**.
- 15- Cai, W., M. Zhang, and B. Maignet, New approach for representation of molecular surface. Journal of Computational Chemistry, 19, 1805-1815, **1998**.
- 16- Cai, W., X. Shao, and B. Maignet, Protein-ligand recognition using spherical harmonic molecular surfaces: towards a fast and efficient filter for large virtual throughput screening. Journal of Molecular Graphics and Modelling, 20, 313-328, **2002**.
- 17- Cai, W., Xu, J., Shao, X. et Maignet, B., SHEF : an efficient approach for virtual screening using coefficients of spherical harmonic molecular. Submitted to publication, **2006**.
- 18- Jones, G., Willet, P., Glen, R. C., Leach, A. R. et Taylor, R., Development and validation of a genetic algorithm for flexible docking. J. Mol. Biol., 267, 727-748, **1997**.
- 19- Jones, G., Willet, P. et Glen, R. C., Molecular recognition of receptor sites using a genetic algorithm with a description of desolvation. J. Mol. Biol., 245, 43-53, **1995**.
- 20- Jones, G., Willet, P. et Glen, R. C., A genetic algorithm for flexible molecular overlay and pharmacophore elucidation. Journal of Computer-Aided Molecular Design, 9, 532-549, **1995**.
- 21- Kellenberger, E., Rodrigo, J., Muller, P. et Rognan, D., Comparative evaluation of eight docking tools for docking and virtual screening accuracy. PROTEINS: Structure, Function, and Bioinformatics, 57, 225-242, **2004**.
- 22- Gerschel, A., Liaisons intermoléculaires les forces en jeu dans la matière condensée. Paris, EDP Sciences /CNRS Editions, **1995**.

CHAPITRE III

METHODOLOGIE

I –INTRODUCTION

Notre travail de recherches qui consiste à cribler *in silico* une chimiothèque sur la cible périphérique de la cholécystokinine nécessite, d'une part, une collection de molécules et, d'autre part, l'accès à un modèle structural 3D de la cible. La stratégie que nous avons adoptée consiste à soumettre chacun des ligands, de l'ensemble des objets tridimensionnels, provenant de la chimiothèque à des filtres successifs, simulant des contraintes auxquelles doivent obéir des molécules susceptibles d'être de futurs médicaments (fig.1). Le but final des opérations successives est de parvenir à la sortie du dernier filtre à une liste de touches ou molécules, appelées candidats-médicaments.

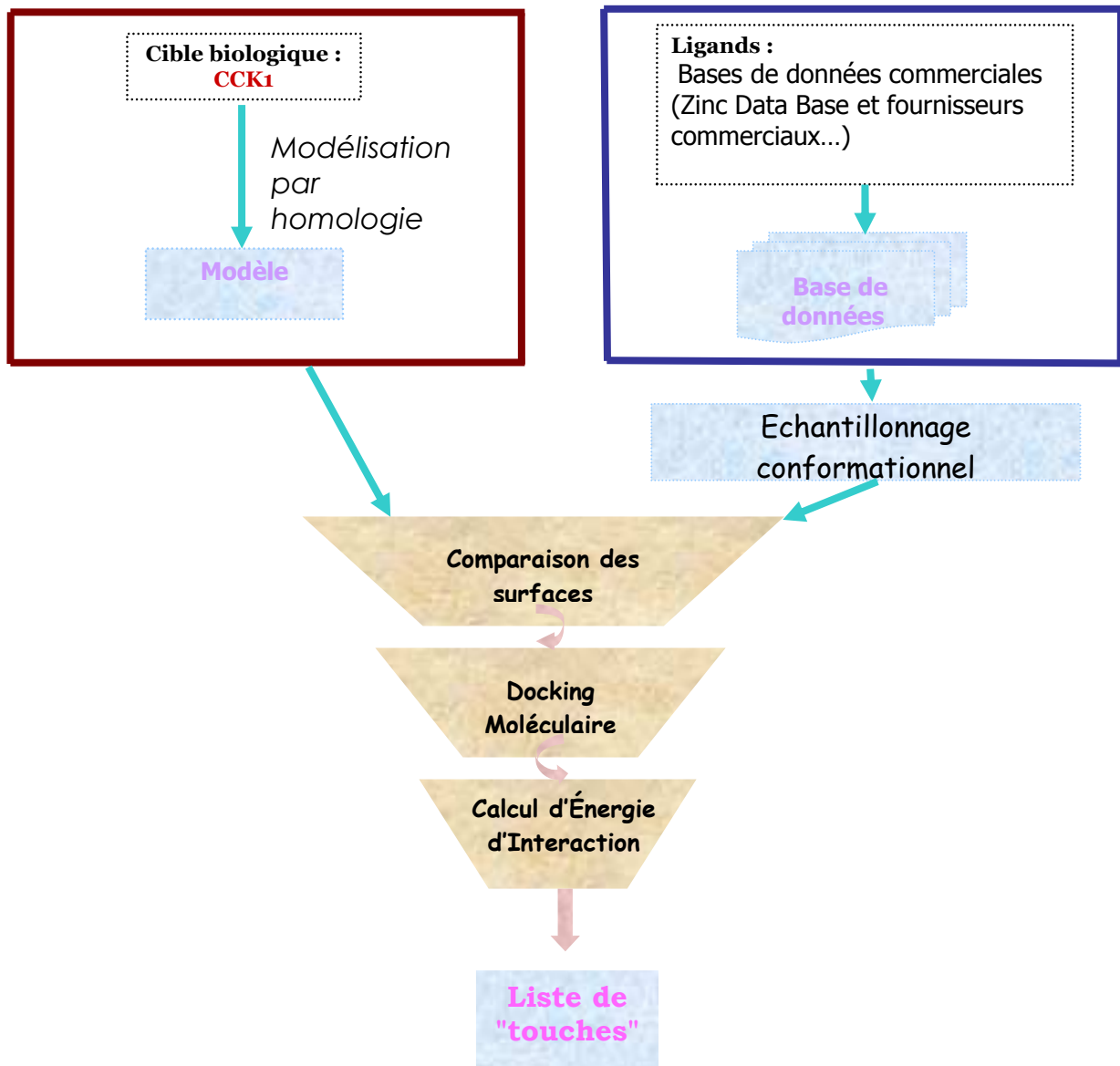


Figure 1 : Illustration du criblage utilisé

II - PRESENTATION DU LOGICIEL UTILISE

Nous avons effectué les travaux à l'aide de la plate-forme du « Virtual Screening Manager for Grids » (VSM-G) développée par l'équipe de recherches eDAM (Dynamique des Assemblages Membranaires) de l'UMR7565 SRSMC (Structure et Réactivité des Systèmes Moléculaires Complexes) à l'Université Henri Poincaré de Nancy. Cet outil logistique, programmé en langage JAVA, offre une grande portabilité ; il est doté d'une interface graphique simple d'utilisation. Analogue à une trousse de bricoleur, les programmes inclus peuvent être utilisés successivement ou de manière complètement indépendante. Au moment où nous avons commencé nos travaux, seule la version sous Windows que nous présentons ci-dessous existait.

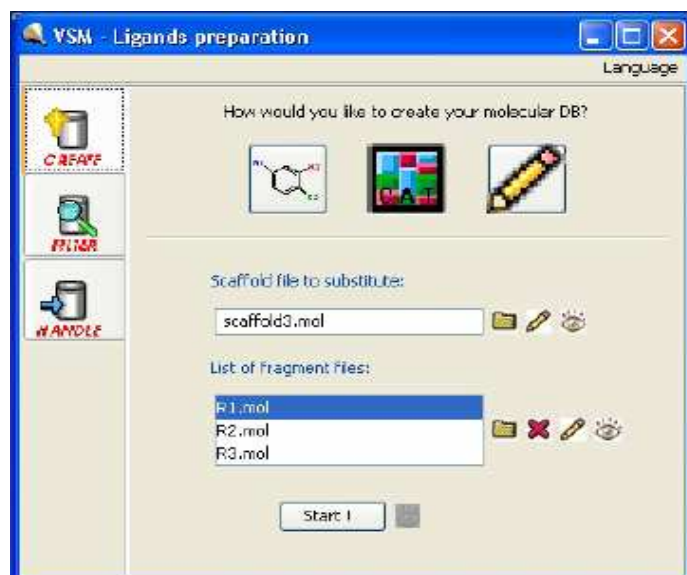
Dès le lancement de VSM-G une fenêtre, présentant trois choix à l'utilisateur, s'ouvre



Les deux premiers choix concernant la préparation des ligands et celle de la cible biologique sont totalement indépendants.

A - Préparation des ligands

Un clic de souris sur la *préparation des ligands* conduit à l'ouverture de la fenêtre ci-dessous



Les 3 onglets dont nous présentons les fonctions ci-après permettent une gestion complète des ligands.

A1- Création de bases de données moléculaires (Create)

Ce module permet de créer la base de données si celle-ci n'est pas créée de manière externe au programme, comme c'est le cas dans nos travaux.

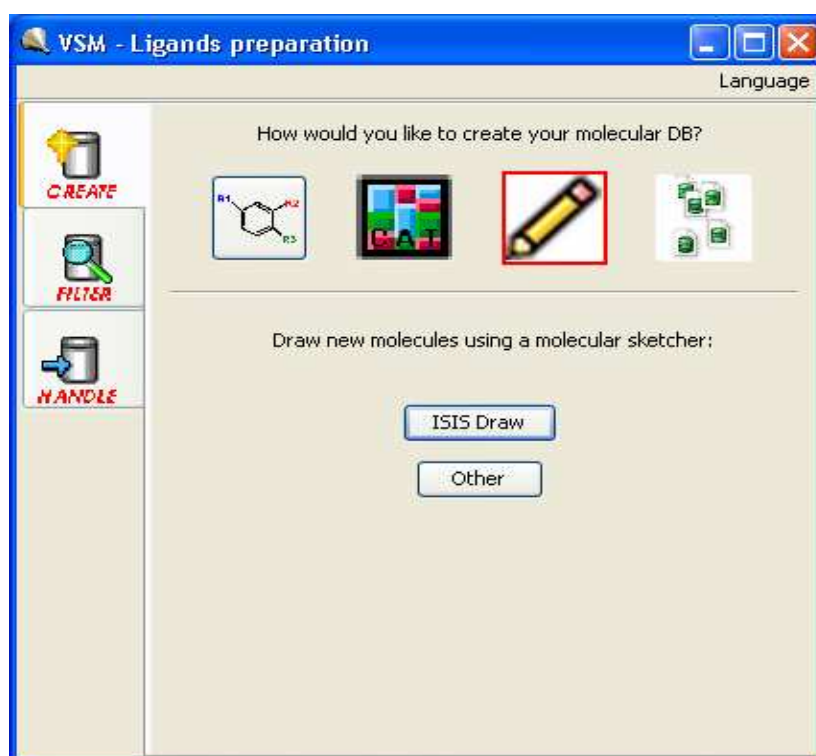
Cette opération peut se faire de 4 façons différentes :

α - Par chimie combinatoire virtuelle (voir fenêtre précédente) en greffant des substituants sur des squelettes de molécules

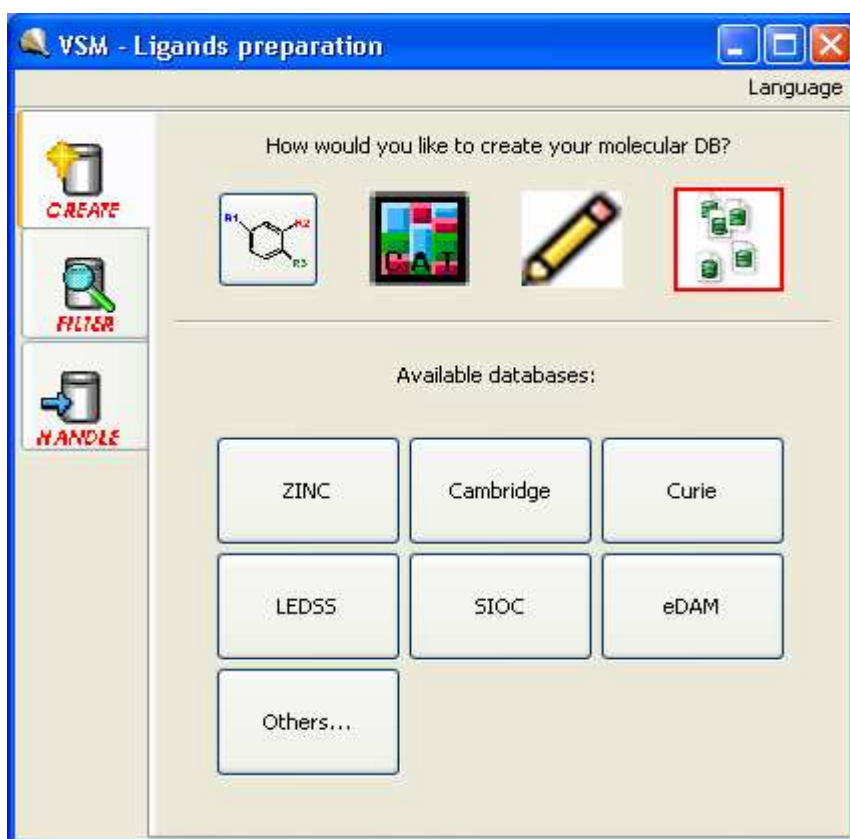
β - Par concaténation de fichiers moléculaires ; ce qui permet de fusionner plusieurs sous bases.



γ- Par saisie de molécules grâce au programme Isis Draw



δ- Par téléchargement à partir de bases de données disponibles sur le web (Zinc Data Base, Curie, Cambridge,) ou commerciales. Il faut noter que quelques bases de données sont proposées par VSM-G.



A2- Filtrage de bases de données (Filter)

Cette option permet d'extraire des molécules spécifiques à partir de bases de données préalablement établies. Ce filtrage peut se faire au moyen d'une :

- recherche par sous-structure
- exclusion des composés toxiques
- recherche par groupement pharmacophore

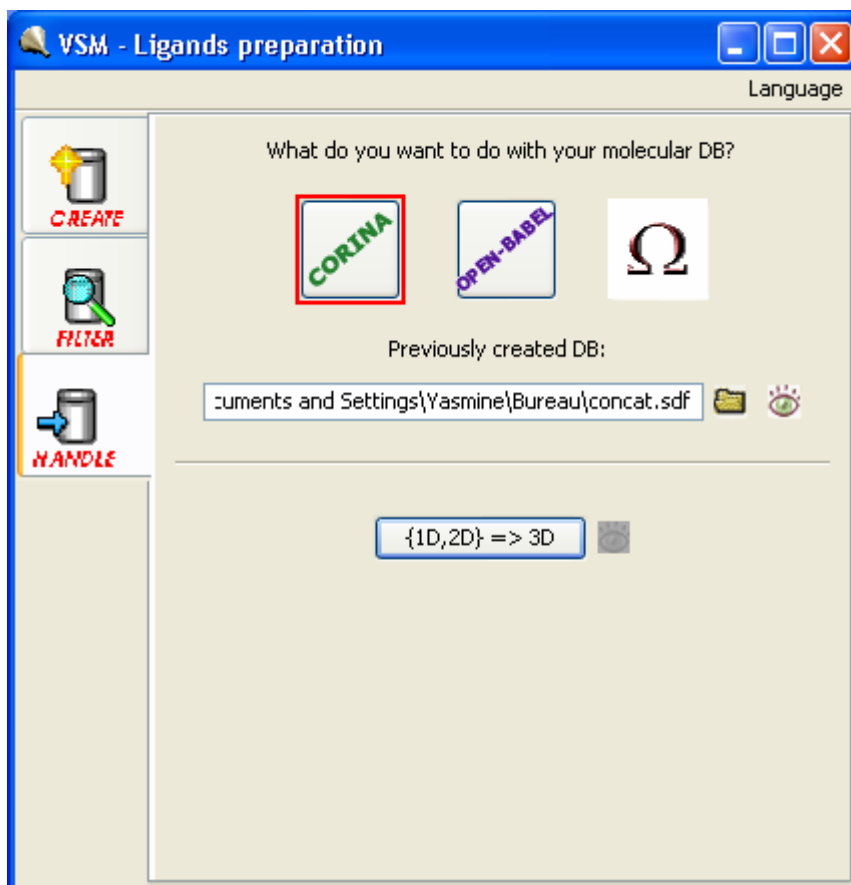


A3- Manipulation de la base de données (Handle)

Cette option du logiciel permet d'uniformiser les formats de fichiers de molécules provenant de chimiothèques distinctes. Ceci peut se faire au moyen de programmes implémentés :

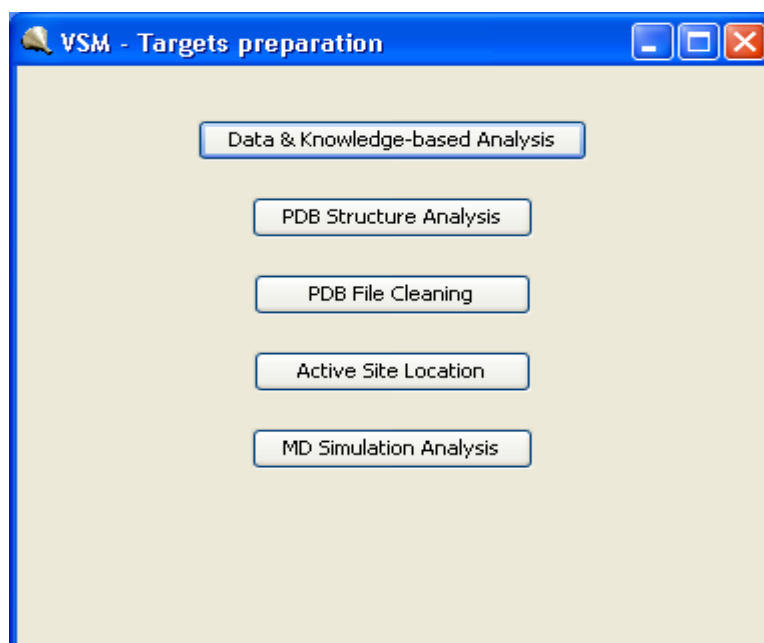
- CORINA⁽¹⁾ permettant les transformations : 1D/2D au 3D
- Open-Babel⁽²⁾ permettant les conversions de formats de fichiers

Le programme OMEGA⁽³⁾ permet d'obtenir un échantillonnage conformationnel de chacune des molécules contenues dans la base de données



B- Préparation de la cible

Dans le cas où la cible biologique est connue VSM-G peut proposer différentes opérations à effectuer sur cette dernière, comme nous le montre la fenêtre ci-dessous.



En cliquant sur « PDB Structure Analysis » VSM-G nous redirige vers un site web contenant un outil logistique d'accès gratuit et appelé Sting Millenium Suite (SMS)⁽⁴⁾. Cet outil est dédié à la visualisation et à l'analyse des structures 3D des protéines permettant principalement de :

- mettre à la bonne norme les fichiers PDB
- corriger la protonation de la protéine
- analyser les points de contact des acides aminés d'une structure PDB
- supprimer des molécules d'eau selon leur positionnement relatif à une interface protéine /ligand ou protéine

Il est important de souligner que nous n'avons pas utilisé ce module dans nos travaux, étant donné que la cible biologique considérée est connue. La préparation du fichier PDB du récepteur, détaillée dans le paragraphe III, s'est effectuée de manière externe à VSM-G.

C- Entonnoir de criblage (Screening Funnel)

L'activation de l'onglet "Screening Funnel", de la fenêtre d'ouverture de VSM-G, conduit à la fenêtre ci-dessous :



Ce module, utilisant le procédé de "l'entonnoir", répond essentiellement au système de criblage virtuel à haut débit par filtres successifs.

Les méthodes de calculs, présentées sous forme de logiciels, sont classées de la plus rapide mais moins efficace à la plus précise. Comme le montre la fenêtre ci-dessus, les trois étapes qui doivent être appliquées successivement sont :

- ☒ le calcul de complémentarité de surfaces : cavité/ligand, au moyen du programme de filtrage MSSH⁽⁶⁾.
- ☒ le docking moléculaire flexible au moyen du logiciel GOLD⁽⁷⁾ qui permet de classer les molécules selon une fonction de score estimant l'énergie d'interaction ligand/récepteur.
- ☒ la dynamique moléculaire, permettant de bien relaxer le complexe, au moyen du logiciel NAMD⁽⁸⁾ puis d'effectuer des calculs d'énergies libres afin d'estimer la stabilité des complexes formés.

Chacune de ces étapes agit comme un filtre moléculaire excluant à chaque niveau un pourcentage de structures moléculaires.

Selon la taille de la chimiothèque, VSM-G présente l'option d'exécuter les calculs relatifs à MSSH, sur un simple PC, ou sur une grille de calcul pouvant atteindre des milliers de noeuds.

III- PRESENTATION DE LA CIBLE BIOLOGIQUE

Rappelons que l'action de la cholécystokinine (CCK) est médiée par deux types de récepteurs membranaires de la famille des RCPGs. Ces deux récepteurs sont notés CCK₁ et CCK₂.

Dans le cadre de ce travail, nous ne disposons que du modèle du récepteur périphérique⁽⁹⁾. Il faut noter que les diverses fonctions biologiques régulées par ce récepteur font de lui une cible de choix dans la recherche thérapeutique de nombreuses pathologies liées à l'alimentation⁽¹⁰⁾.

Le modèle du complexe ligand/CCK₁ avec ligand = dérivé pyridopyrimidique⁽¹¹⁻¹⁴⁾, établi et validé au moyen d'expériences de mutagenèse dirigée⁽¹⁵⁻¹⁷⁾ a servi de point de départ dans nos travaux.

Nous avons ensuite introduit le modèle obtenu sous le logiciel graphique INSIGHT II⁽¹⁸⁾ qui permet d'ioniser les résidus au pH physiologique.

Nous présentons dans les figures 2, 3 et 4 respectivement les modèles du dérivé pyridopyrimidique, du récepteur et du complexe⁽⁹⁾.

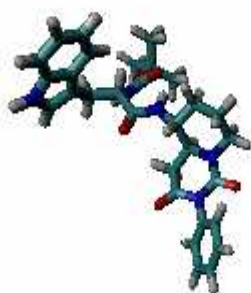


Figure 2 : Dérivé pyridopyrimidique

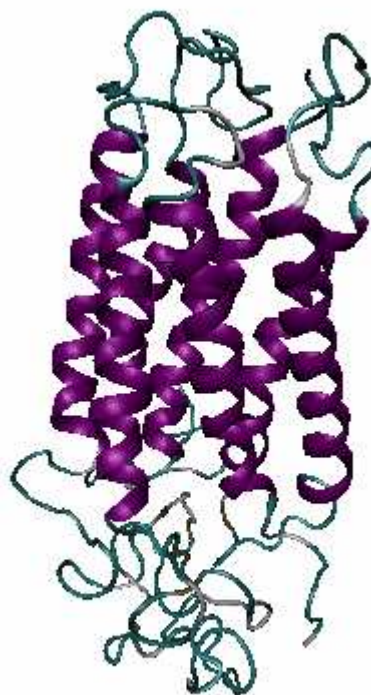


Figure 3 : Modèle du récepteur CCK₁

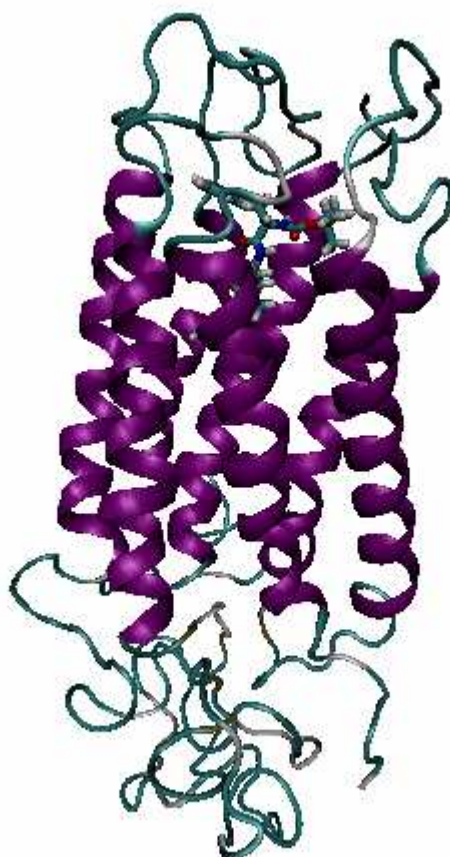


Figure 4 : Complexe Pyridopyrimidine/CCK₁

IV- PRESENTATION DE LA CHIMIOTHEQUE

Afin de gérer toutes les molécules destinées au criblage, l'élaboration d'une chimiothèque devient primordiale. Cette dernière peut être réelle ou virtuelle. La chimiothèque réelle comprend des molécules déjà synthétisées alors que la virtuelle regroupe des molécules conçues virtuellement et que l'on projette de synthétiser.

Nous avons construit notre chimiothèque à partir de la base de données ZINC^(19,20). Cette base de molécules, d'accès gratuit, contient plus de 4,6 millions de molécules provenant de plus d'une douzaine de fournisseurs tels que Asinex, ChemBridge, Sigma-Aldrich,....

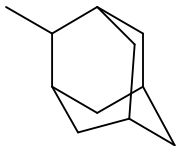
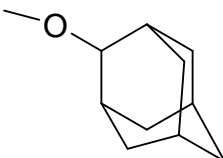
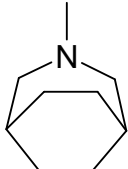

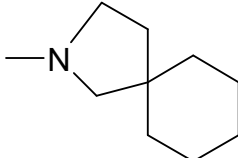
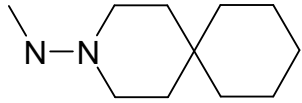
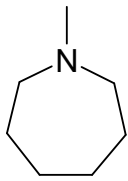
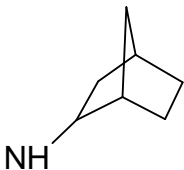
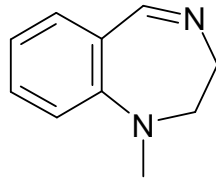
Comme le montre la fenêtre du navigateur ci-dessous, l'interface de dessin moléculaire incorporée à cet outil de recherche, permet une collection de molécules au moyen d'une recherche par sous structure.


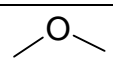
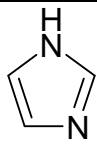
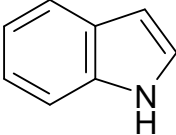
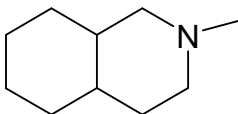
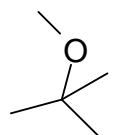
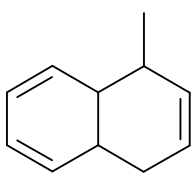
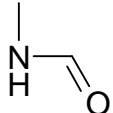
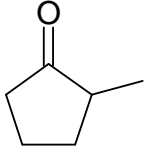
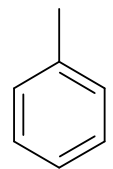
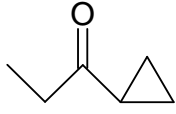
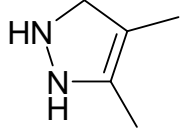
Figure 5 : Fenêtre de ZINC pour la recherche par sous structure

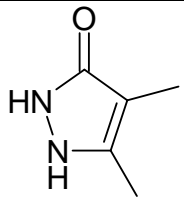
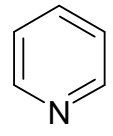
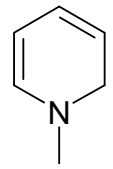
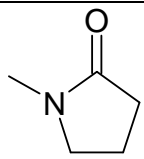
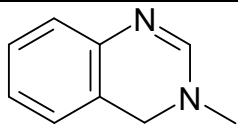
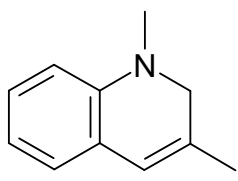
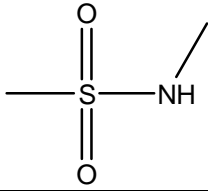
A- Recherche des molécules par sous structure

Dans le but de faire une recherche par sous structure, une consultation exhaustive d'articles récents suivie d'une recherche bibliographique sur des groupements susceptibles de produire des effets antagonistes sur le récepteur CCK₁ nous a conduit à dresser la liste suivante :

Tableau 1 : Sous structures de base de la chimiothèque.

Appellation de la sous structure	Dessin de la sous structure
ADAMANTYL	
ADAMANTYL-O	
ALCOOL	—OH
Azabicyclononane	
Azaspiro44nonane	
Azaspiro45decane	
Azaspiro55undecan3amino	
AZEPANE	
BCHA	
BENZODIAZEPINE	

CARBOXYLATE	
F_O	
IMIDAZOLE	
INDOLE	
ISOQUINOLINE	
METHOXY	
NAPHTYL	
NHCO	
PENTANONE	
PHENYL	
PROPANONE	
PyrazoleH	

PYRAZOLONE	
PYRIDINE	
PYRIDYL	
PYRROLIDINONE	
QUINAZOLINE	
QUINOLINE	
SULFONAMIDE	

Il faut noter que dans la suite de nos travaux, nous avons utilisé le nom de la sous structure pour le nom de la famille de molécules obtenue à partir de la ZINC data base.

L'interface de dessin moléculaire incorporée à cet outil de recherche, permet de dessiner la sous structure. Ensuite, le lancement de la recherche par sous structures ci-dessus, conduit à la fenêtre suivante permettant le téléchargement de toutes les molécules dans le format désiré.

The screenshot displays the ZINC search results page. At the top, it shows the search criteria: 'Beginning: #1 Hits: 100 Display: Standard per page: 100'. Below this are buttons for 'Next Page', 'Download Table', and 'Create Subset'. A 'ZINC' logo is visible on the right. The main content is a table with three rows of results. Each row contains a rank, ZINC ID, chemical structure, supplier information, and properties.

Rank	ZINC Id	Structure	Supplier: Catalog Number	Properties
1	00000018		Asinex: BAS2912638 Asinex: BAS02912638	xLogP: 1.91 H-bond-donors: 1 Mwt: 212.34 apolar & polar desolvation: -1.23 acceptors: 0 rotatable bonds: 4
2	00000023		Ryan Scientific: LT00771761 Sigma-Aldrich: A8676 Ryan Scientific: RJC02271 Sigma-Aldrich: R426113	xLogP: 2.84 H-bond-donors: 2 Mwt: 250.4 apolar & polar desolvation: -1.06 acceptors: 0 rotatable bonds: 8
3	00000031		Sigma-Aldrich: R288985	xLogP: 1.15 H-bond-donors: 2 Mwt: 303.43 apolar & polar desolvation: 1.74 acceptors: 1 rotatable bonds: 5

Figure 6 : Exemple de résultat d'une recherche par sous structure dans la base ZINC

Comme le montre la fenêtre ci-dessus, à chaque molécule sont attribuées certaines propriétés comme le poids moléculaire, son LogP calculé, le nombre de liaisons H donneur et accepteur,...etc. Plusieurs états de protonations selon le milieu, ainsi que de multiples conformations et formes tautomériques sont disponibles. Les molécules se téléchargent gratuitement en plusieurs formats de fichiers tels que SMILES, mol2, SDF et le format de docking flexibase.

Selon la sous structure, la recherche dans la ZINC data base, conduit à une famille de molécules; ce qui a généré **62169** molécules pour l'ensemble des familles. Nous avons

téléchargé la base de molécules en format mol2 concaténé. Notre outil de criblage virtuel utilise ce format (Mol2 3D) pour lire les fichiers d'entrée.

Un des avantages de la base de données ZINC est qu'elle fournit des molécules en format mol2 3D; ce qui évite d'utiliser un convertisseur de structure comme CORINA qui transforme les fichiers 2D en 3D. Un autre aspect positif de cette base virtuelle est qu'elle ne génère pas des molécules n'ayant aucun intérêt à être modélisés tels que les contre ions des sels, des molécules pouvant générer une toxicité et des molécules ne répondant pas à la règle des 5 ou règle de Lipinski⁽²¹⁾.

Notons, qu'avant de commencer le processus de criblage, nous avons écarté les doublons contenus dans notre chimiothèque. Cette opération a réduit la chimiothèque de 62169 à 51947 molécules.

B- Recherche conformationnelle

Afin d'explorer l'espace conformationnel de chaque molécule de notre chimiothèque contenant, à présent, **51947** molécules, nous avons soumis chacune d'elles à une recherche conformationnelle au moyen du logiciel OMEGA. Cet outil permet d'explorer l'espace conformationnel de la molécule en effectuant autant de libres rotations que puisse le permettre la structure du composé considéré.

Nous avons spécifié, au logiciel, de générer pour chaque molécule 20 conformations. Ce nombre est justifié par une estimation globale (générale) du nombre de rotules des molécules qui sont toutes de petite taille. Comme le nombre de conformations par molécule dépend du nombre de rotules de cette dernière, il est important de noter que le nombre de conformations générées ne sera pas égal à 20 pour toutes les molécules.

Cette étape de notre recherche a enrichi notre chimiothèque en objets tridimensionnels. En effet, nous avons obtenu **847220** structures 3D que nous répertorions ci-dessous en fonction de chaque famille.

Tableau 2 : Familles et nombre de structures 3D constituant notre chimiothèque

Nom de la famille	Nombre de molécules	Nombre de conformères
ADAM	157	1483
ADAM-O	31	284
ALCOOL	1987	20927
Azabicyclononane	95	1431
Azaspiro44nonane	80	1134
Azaspiro45decane	16	50
Azaspiro55undecan3amino	6	47
AZEPANE	3505	56207
BCHA	1206	16832
BENZODIAZEPINE	742	9720
CARBOXYLATE	2770	31153
F_O	2192	28804
IMIDAZOLE	151	2387
INDOLE	5927	89389
ISOQUINOLINE	114	730
METHOXY	101	1169
NAPHTYL	2	2
NHCO	10756	146890
PENTANONE	93	746
PHENYL	2461	26717
PROPANONE	279	2175
PyrazoleH	32	393
PYRAZOLONE	37	473
PYRIDINE	1473	16847
PYRIDYL	64	392
PYRROLIDINONE	5215	71537
QUINAZOLINE	2080	24607
QUINOLINE	597	6199
SULFANAMIDE	20000	288495

Pour commencer le processus de criblage, chaque molécule doit avoir son fichier mol2 indépendant des autres. Nous avons donc déconcaténé les fichiers mol2 de chaque famille au moyen du programme "splmol2" interfacé avec VSM-G.

C- Comparaison des surfaces ligand/récepteur avec **MSSH**

C1- Calcul des fingerprints :

Comme nous l'avons dit dans le chapitre III, le logiciel MSSH, de comparaison de surfaces entre le ligand et la cavité du récepteur, commence d'abord par le calcul des "fingerprints" pour les deux partenaires. Rappelons que ce calcul se fait sur la base des fonctions harmoniques sphériques.

Ce calcul conduit à 3 fichiers de sortie en format différent :

- .xyz : donne les coordonnées de la surface moléculaire pour la visualisation
- .coef : donne les coefficients des coordonnées de la structure 3D de la molécule
- fp.out : donne des informations sur l'ensemble du processus de calcul de surface

Nous avons donc soumis chacune des structures 3D de notre chimiothèque à ce calcul.

Le calcul de la surface de la cavité nécessite le fichier PDB du récepteur et également l'introduction des coordonnées du barycentre de la cavité. Ces dernières sont obtenues grâce à VMD.

Bien entendu, le temps de calcul de ces surfaces est fonction des performances du PC utilisé ; il est de l'ordre de 1 à 2 minutes, par structure 3D, sur un Pentium IV (3.20 GHz, 3.19 GHz). Pour traiter nos **847220** objets 3D, organisés en familles, nous avons réparti nos calculs sur des PC structurés en clusters et également sur la grille : Grid5000⁽²²⁾.

C2- Comparaison des fingerprints

Cette comparaison nécessite les fichiers en format .coef de chaque échantillonnage de molécules et également celui de la cavité. Ces calculs, qui prennent relativement peu de temps (~ 2s/molécule sur un Pentium IV), génèrent un tableau classant la complémentarité des surfaces des conformères selon le meilleur RMSD par rapport à celle de la cavité du récepteur.

C3- Filtrage des molécules :

Le premier filtrage de notre chimiothèque consiste à éliminer les objets 3D dont la complémentarité de surfaces avec la cavité est mauvaise. La question qui se pose alors est : à partir de quelle valeur du RMSD allons nous écarter ces objets ? Pour cela, nous nous sommes basés sur un travail analogue⁽²³⁾ s'effectuant sur des molécules dont l'activité antagoniste sur CCK₁ est connue expérimentalement. Il ressort de cette étude que la molécule la plus active présente un RMSD égal à 0,56. Ce résultat nous a conduit à écarter tous les objets dont le RMSD est supérieur à 0,86 tout en gardant une seule conformation pour chaque molécule. Ainsi, le nombre de molécules issues de ce premier filtrage est égal à **12443**. Nous avons conforté ce seuil grâce à une visualisation via VMD de la complémentarité des surfaces ligand/cavité.

D- Préparation des fichiers pour le docking moléculaire GOLD

Il est important de noter que le logiciel GOLD reconnaît les structures telles qu'elles sont dans le milieu physiologique. Pour cela, un programme développé par l'équipe eDAM permet d'établir l'état d'ionisation des ligands. Quant à l'ionisation du récepteur, nous l'avons effectué sous INSIGHT II. Les fichiers de données sont en format mol2 pour le récepteur et les ligands. La technique du docking sous GOLD nécessite également un fichier contenant les coordonnées de la cavité.

Comme ce docking fait appel à la flexibilité du ligand, nous avons opté pour 20 conformations de ce dernier. Les résultats du docking sont donnés sous forme de classement des molécules selon leur meilleur score. Rappelons que ce score estime l'affinité entre le ligand et le récepteur. L'analyse des résultats conduit à retenir les 20% premiers ligands présentant les meilleurs scores⁽²⁴⁾. Ce filtrage par docking semi flexible réduit notre base de données à **363** molécules présentant une affinité potentielle à l'égard du récepteur CCK₁.

E- Calcul des énergies d'interactions ligand/récepteur

Afin d'effectuer un calcul d'énergies d'interactions plus rigoureux, il est préférable d'avoir recours à un logiciel utilisant un champ de force. Nous poursuivons donc le criblage au moyen du programme Discover d'INSIGHT II implanté sur la station de travail Silicon Graphics. Une minimisation est d'abord effectuée sur le complexe afin de relaxer la structure. Pour cela, nous avons employé la méthode du gradient conjugué en effectuant

1000 itérations. Le champ de force employé est CVFF^(25, 26) qui utilise le potentiel de Lennard-Jones pour le terme de vdW. L'énergie d'interaction que nous avons calculée, pour les **363** complexes issus de GOLD, comprend uniquement deux termes : van der Waals et électrostatique.

Sur les **363** complexes issus du précédent filtrage (GOLD) **279** possèdent une énergie d'interaction négative.

Nous n'avons pas pu soumettre les complexes au dernier filtre de VSM-G qui consiste à réaliser une dynamique moléculaire nécessitant beaucoup de temps de simulation, nous avons pour cela opté pour un calcul d'énergie libre de liaison on line et au moyen d'un programme disponible sur le serveur BAPPL⁽²⁷⁾.

F- Calcul des énergies libres de liaisons $\Delta G_{\text{binding}}$

La superposition des classements de scores des trois filtres utilisés, sur les 100 meilleures molécules de chacun de ces filtres, fait ressortir du lot quatre ligands figurant simultanément dans les classements issus de MSSH, GOLD et Discover. Dans le but de conforter nos résultats, nous avons tenu compte du facteur entropique en soumettant les quatre candidats-médicaments à un calcul d'énergie libre de liaison. Le protocole de calcul utilisé par BAPPL est rapide et approprié au criblage virtuel de chimiothèque. Il consiste à prédire les affinités de liaison entre une protéine et son ligand. Cette technique de calcul est basée sur une fonction de score empirique comprenant l'énergie de van der Waals, l'énergie électrostatique et celle des interactions hydrophobes. Il faut noter que cette méthode de calcul a été validée sur un ensemble hétérogène de 161 complexes, ligand/récepteur, donnant de bonnes corrélations⁽²⁸⁾ entre les énergies libres de liaison calculées (prédites) et expérimentales.

REFERENCES BIBLIOGRAPHIQUES

- 1- <http://www.mol-net.de/software/corina>
- 2- <http://openbabel.sourceforge.net>
- 3- <http://www.eyesopen.com/products/applications/omega.html>
- 4- <http://sms.cbi.cnptia.embrapa.br>
- 5- <http://www.ks.uiuc.edu/Research/vmd/>
- 6- <http://www.ustc.edu.cn>
- 7- <http://www.ccdc.cam.ac.uk/>
- 8- <http://www.ks.uiuc.edu/Research/namd/>
- 9- Martin-Martinez, M., Marty, A., Jourdan, M., Escrieut, C., Archer, E., Gonzalez-Muniz, R., Garcia-Lopez, M.T., Maigret, B., Herranz, R. et Fourmy, D. Combination of molecular modeling, site directed mutagenesis and SAR studies to delineate the binding site of pyridopyrimidine antagonists on the human CCK₁ receptor. *J. Med. Chem.*, 48, 4842-4850, **2005**.
- 10- Herranz, R. Cholecystokinin antagonists: Pharmacological and therapeutic potential. *Med. Res. Rev.*, 23, 559-605, **2003**.
- 11- Bartolomé-Nebreda, J.M., Patiño-Molina, R., Martín-Martínez, M., Gómez-Monterrey, I., García-López, M. T., González-Muñiz, R., Cenarruzabeitia, E., Latorre, M., Del Río, J. et Herranz, R. 5-(Tryptophyl)amino-1,3-dioxoperhydroprido[1,2-c]pyrimidine-based potent and selective CCK₁ receptor antagonists: structure-activity relationship studies on the substituent at N2-position. *J. Med. Chem.*, 44, 2219-2228, **2001**.
- 12 - Bartolomé-Nebreda, J.M., García-López, M. T., González-Muñiz, R., Cenarruzabeitia, E., Latorre, M., Del Río, J. et Herranz, R. 5-(Tryptophyl)amino-1,3-dioxoperhydroprido[1,2-c]pyrimidine-based potent and selective CCK₁ receptor antagonists: structure-activity relationship studies on the central 1,3-dioxoperhydroprido[1,2-c]pyrimidine scaffold. *J. Med. Chem.*, 44, 4196-4206, **2001**.
- 13- Bartolomé-Nebreda, J.M., Gómez-Monterrey, I., García-López, M. T., González-Muñiz, R., Martín-Martínez, M., Cenarruzabeitia, E., Latorre, M., Del Río, J. et Herranz, R. 5-(Tryptophyl)amino-1,3-dioxoperhydroprido[1,2-c]pyrimidine-based potent and selective CCK₁ receptor antagonists: structural modifications at the tryptophan domain. *J. Med. Chem.*, 42, 4659-4668, **1999**.
- 14- Martín-Martínez, M., Bartolomé-Nebreda, J.M., Gómez-Monterrey, I., González-Muñiz, R., García-López, M.T., Ballaz, S., Barber, A., Fortuño, A., Del Río, J. et

- Herranz, R. Synthesis and stereochemical structure-activity relationships of 1,3-dioxoperhydropyrido[1,2-c]pyrimidine derivatives: potent and selective cholecystokinin-A receptor antagonists. *J. Med. Chem.*, *40*, 3402-3407, **1997**.
- 15- Gigoux, V., Maigret, B., Escricut, C., Silvente-Poirot, S., Bouisson, M., Fehrentz, J-A., Moroder, L., Gully, D., Martinez, J., Vaysse, N. et Fourmy, D. Arginine 197 of the cholecystokinin-A receptor binding site interacts with the sulfate of the peptide agonist cholécystokinine. *Protein Sci.*, *8*, 2347-2354, **1999**.
- 16- Gigoux, V., Escricut, C., Fehrentz, J-A., Poirot, S., Maigret, B., Moroder, L., Gully, D., Martinez, J., Vaysse, N. et Fourmy, D. Arginine 336 and asparagine 333 of the human cholecystokinin-A receptor binding site interact with the penultimate aspartic acid and the C-terminal amide of cholécystokinine. *J. Biol. Chem.*, *274*, 20547-20464, **1999**.
- 17- Escricut, C., Gigoux, V., Archer, E., Verrier, S., Maigret, B., Behrendt, R., Moroder, L., Bignon, E., Silvente-Poirot, S., Pradayrol, L. et Fourmy, D. The biologically crucial C-terminus of cholecystokinin and the non-peptide agonist SR-146,131 share a common binding site in the human CCK₁ receptor. Evidence for a crucial role of Met-121 in the activation process, *J. Biol. Chem.*, *277*, 7546-7555, **2002**.
- 18- Accelrys/MSI. Road San Diego, CA 92121 USA.
- 19- <http://zinc.docking.org>
- 20- Irwin, J.J. and Shoichet, B.K. ZINC - A Free Database of Commercially Available Compounds for Virtual Screening. *J. Chem. Inf. Model.*, *45*, 177-182, **2005**.
- 21- C. Lipinski et al., *Adv. Drug. Del. Rev.*, *23*, 3-25, **1997**.
- 22- <http://www.grid5000.fr>
- 23- Maouche-Toumi, A., Thèse de doctorat en cours, Alger, USTHB.
- 24- <http://www.ccdc.cam.ac.uk/support/documentation/#gold>
- 25- Hagler, A.T., Dauber, P. et Lifson, S., *J. Am. Chem. Soc.*, *101*, p: 5131, **1978**.
- 26- Hagler A.T., Lifson S., Dauber P., *J. Am. Chem. Soc.*, *101*, p: 5122, **1979**.
- 27- www.scfbio-iitd.res.in/software/drugdesign/bappl.jsp
- 28- Jain, T. et Jayaram, B. An all atom energy based computational protocol for predicting binding affinities of protein-ligand complexes. *FEBS Letters*, *579*, 6659-6666, **2005**.

CHAPITRE IV

RESULTATS, STRATEGIES ET DISCUSSION

I - INTRODUCTION

Dans ce chapitre, nous présenterons les principaux résultats issus de chaque filtre du criblage que nous avons effectué. Nous donnerons également la stratégie que nous avons mise en œuvre pour valider le protocole de criblage adopté dans cette étude.

II – RESULTATS ET ANALYSE

A- MSSH

Rappelons que le premier filtre, de l'outil de criblage que nous avons utilisé : VSM-G, est le MSSH qui consiste à calculer et à comparer les surfaces du ligand et de la cavité du récepteur. Cette comparaison s'effectue sur la base de la valeur du RMSD, notons que plus la valeur du RMSD est petite plus la complémentarité de surfaces est grande. Dans la littérature⁽¹⁻⁴⁾, on dit qu'un ligand s'emboîte mal dans la cavité du récepteur lorsque le RMSD entre les deux surfaces est supérieur à 2Å. Dans notre cas, l'étude par complémentarité de surfaces a donné des valeurs de RMSD comprises entre [**0.618** et **1.203**]. Dans le tableau, ci-dessous, figure le détail des valeurs de RMSD selon les sous structures.

Tableau 1 : Valeurs des RMSD issus du docking MSSH

Sous structure	RMSD	
	Minimum	Maximum
ADAM	0.684	1.016
ADAM-O	0.797	0.999
ALCOOL	0.713	1.189
Azabicyclononane	0.687	1.004
azaspiro44nonane	0.784	1.013
azaspiro45decane	0.783	0.905
azaspiro55undecan3amino	0.867	0.885
AZEPANE	0.618	1.158
BCHA	0.676	1.081
BENZODIADIAZEPINE	0.674	0.979
CARBOXYLATE	0.761	1.180
F_O	0.718	1.192
IMIDAZOLE	0.838	1.084
INDOLE	0.634	1.164
ISOQUINOLINE	0.704	0.993
METHOXY	0.761	1.070
NAPHTYL	0.851	0.878
NHCO	0.743	1.187
PENTANONE	0.782	1.024
PHENYL	0.725	1.191
PROPANONE	0.662	1.143

pyrazoleH	0.798	1.024
PYRAZOLONE	0.792	1.024
PYRIDINE	0.871	1.192
PYRIDYL	0.758	1.174
PYRROLIDINONE	0.729	1.124
QUINAZOLINE	0.725	1.203
QUINOLINE	0.799	1.080
SULFONAMIDE	0.727	1.135

Le domaine de variation des valeurs du RMSD indique que les **847 220** objets 3D soumis à MSSH ont une bonne complémentarité de surface vis-à-vis de la cavité du récepteur utilisée. Ce résultat peut être dû à la petite taille de nos ligands qui semblent s'imbriquer aisément dans la poche. Nous illustrons dans les figures ci-dessous, à titre indicatif, les représentations des complémentarités des surfaces ligand/cavité du récepteur avec le plus petit RMSD puis avec le plus grand RMSD

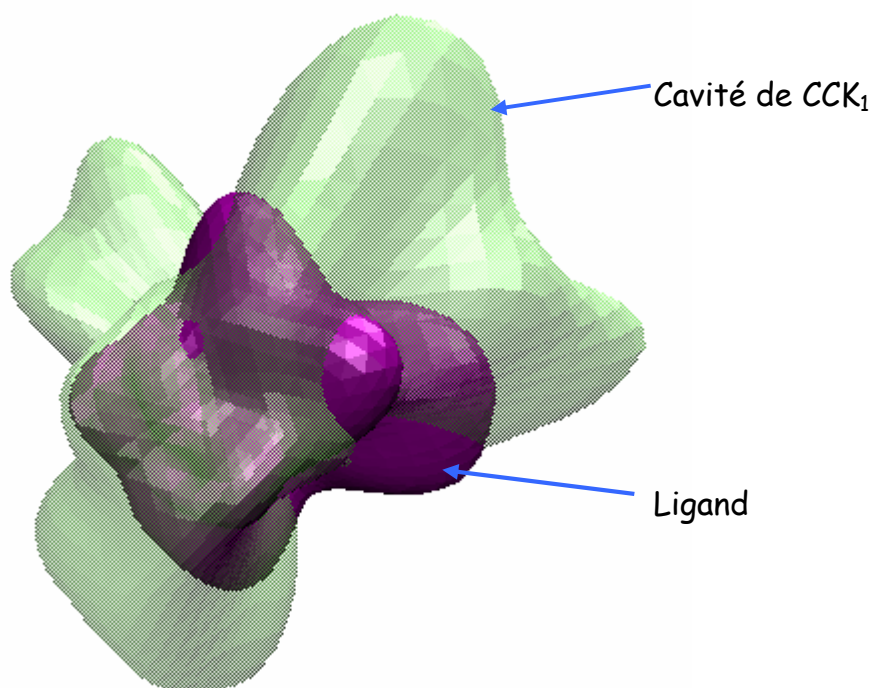


Figure 1 : Complémentarité des surfaces de la cavité du récepteur CCK₁ et du ligand avec
RMSD = 0.618

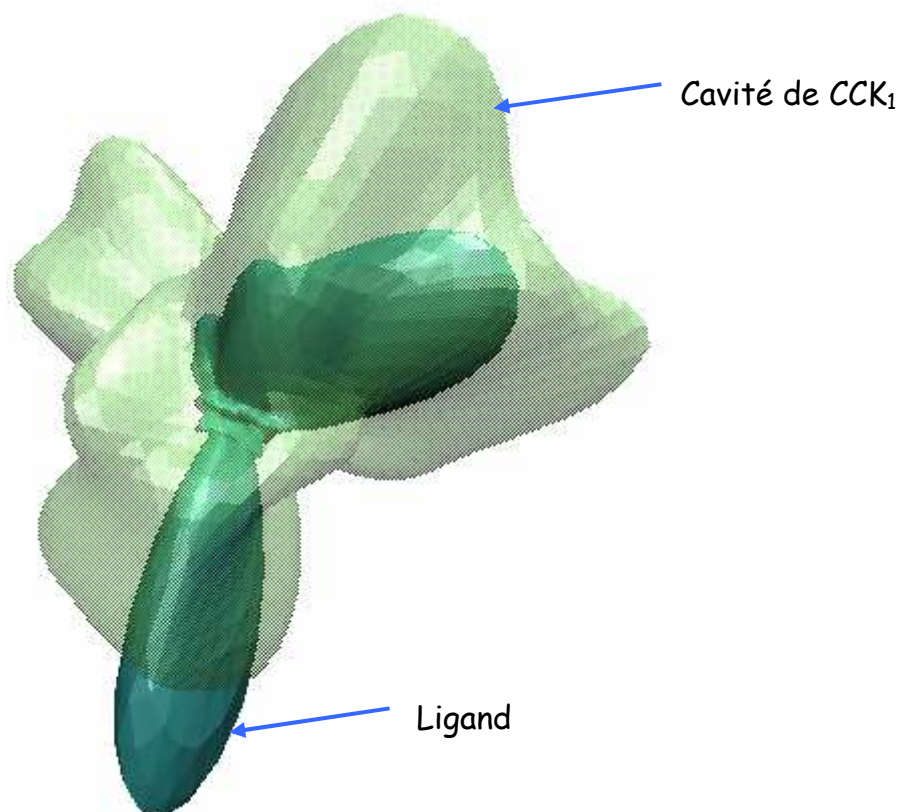


Figure 2 : Complémentarité des surfaces de la cavité du récepteur CCK₁ et d'un ligand avec le plus grand RMSD = 1.204

Dans le but d'analyser nos résultats et de procéder au filtrage, nous nous sommes basés, comme nous l'avons mentionné dans le chapitre précédent, sur un travail analogue⁽⁵⁾ s'effectuant sur des molécules dont l'activité antagoniste sur CCK₁ est connue expérimentalement. Il ressort de cette étude que la molécule la plus active présente un RMSD égal à 0,56. Ce résultat nous a conduit à écarter tous les objets 3D, issus de MSSH, dont le RMSD est supérieur à 0,86 et en gardant que la meilleure conformation pour chaque composé. Ce qui réduit notre chimiothèque à **12 443** molécules.

Afin de poursuivre le processus de criblage, nous avons soumis les **12 443** molécules au filtre suivant qui consiste à effectuer un docking semi-flexible au moyen du logiciel GOLD.

B- GOLD

Nous avons fixé le nombre de docking à 20 pour chaque molécule ; ce choix résulte d'un compromis entre le temps CPU du docking et la flexibilité des molécules constituant notre

chimiothèque. Il faut noter que la moyenne des libres rotations de ces molécules est voisine de 10. Les calculs de docking ont été répartis sur un cluster contenant 8 nœuds (8 PC).

Le docking semi flexible de GOLD a généré des scores de fitness compris entre [90,28 et -149,32]. Nous donnons ci-dessous le détail de ces résultats :

Tableau 2 : Résultats issus du docking par GOLD

Sous structure	Score de fitness GOLD		
	Maximum	Minimum	Nombre de score négatif
ADAM	67.10	21.58	0
ADAM-O	54.93	29.77	0
ALCOOL	71.03	-127.98	13
azabicyclononane	62.72	30.50	0
azaspiro44nonane	70.37	9.61	0
azaspiro45decane	49.44	19.97	0
azaspiro55undecan3amino	48.07	28.92	0
AZEPANE	90.28	-149.32	27
BCHA	79.43	-65.56	11
BENZODIAZEPINE	86.73	35.06	0
CARBOXYLATE	75.03	-128.21	10
F_O	68.04	-128.15	3
IMIDAZOLE	67.78	39.79	0
INDOLE	89.22	-15.03	2
ISOQUINOLINE	56.50	-5.50	1
METHOXY	66.26	24.42	0
NAPHTYL	36.23	31.75	0
NHCO	72.03	-109.69	20
PENTANONE	59.48	19.45	0
PHENYL	69.49	-128.18	22
PROPANONE	74.39	-33.69	4
PyrazoleH	66.37	40.74	0
PYRAZOLONE	61.63	40.54	0
PYRIDINE	72.80	-116.80	3
PYRIDYL	61.84	35.06	0
PYRROLIDINONE	71.14	-14.32	1
QUINAZOLINE	84.60	33.43	0
QUINOLINE	60.74	-7.80	1
SULFONAMIDE	73.61	-128.15	206

Il est important de souligner que 324 molécules possèdent des scores de fitness négatifs, indiquant leur mauvaise adaptabilité à la cavité considérée ; ce qui nous a conduit à leur exclusion de la suite du criblage.

Nous avons poursuivi le criblage en l'effectuant uniquement sur 20% des premières molécules issues du classement de GOLD⁽⁶⁾. Ce critère de sélection conduit à ne retenir que les molécules ayant un score de fitness compris entre **72.2** et **90.28**. A l'issue de ce filtre, notre chimiothèque se réduit à **363** composés.

Comme nous l'avons noté dans le chapitre précédant, le logiciel GOLD utilise l'algorithme génétique pour l'exploration de l'espace conformationnel de la molécule et se base sur des fonctions de scores empiriques pour estimer l'énergie d'interaction entre les résidus du ligand et de la cavité. Ceci nous a conduits à valider les scores de gold par un calcul d'énergie d'interaction au moyen de Discover inclus dans le logiciel de modélisation INSIGHT II⁽⁷⁾.

C- Discover

Cette étape a pour objectif de classer les complexes selon l'énergie d'interaction calculée dans un champ de forces.

Discover est un module du logiciel INSIGHT II permettant, entre autres, de relaxer les complexes et de calculer les énergies d'interactions entre le ligand et le récepteur. Pour cela, nous avons minimisé les structures issues de GOLD, au moyen du gradient conjugué et en effectuant 1000 itérations dans le champ de forces CVFF. Les résultats sont récapitulés dans l'annexe 2.

Les résultats issus de ce calcul d'énergie d'interaction présentent pour certains complexes des énergies positives, exprimant ainsi leur instabilité. De plus, au cours de l'analyse des résultats, un cas particulier a attiré notre attention qui est celui d'une molécule présentant une énergie d'interaction positive alors qu'elle possède le meilleur score dans le docking par GOLD. Ce résultat semble être dû à la nature de la fonction de score qu'utilise GOLD par défaut. En effet, pour chaque groupement d'un système donné, la fonction de score (dans notre cas la Goldscore) définit des paramètres dont la somme constitue le score de fitness. Il faut noter que lorsqu'un groupement n'est pas paramétré, la Goldscore lui attribue une valeur nulle engendrant ainsi un très haut score et un "faux positif". Le groupement en question, non défini dans GOLD, est illustré dans la structure de cette molécule (figure 3) qui concerne également certains membres de la famille AZEPANE.

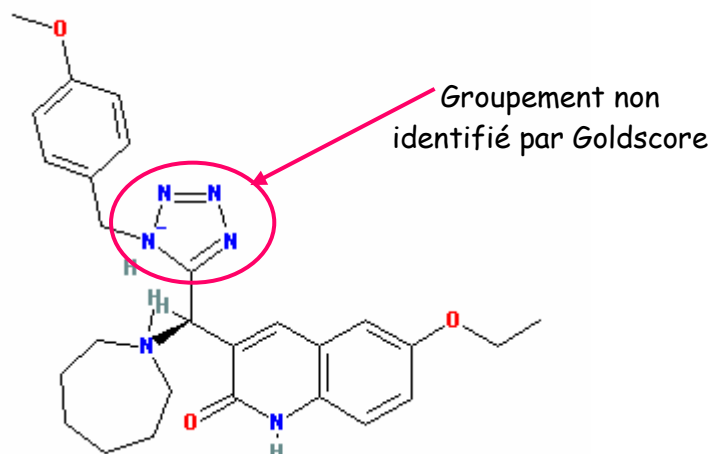
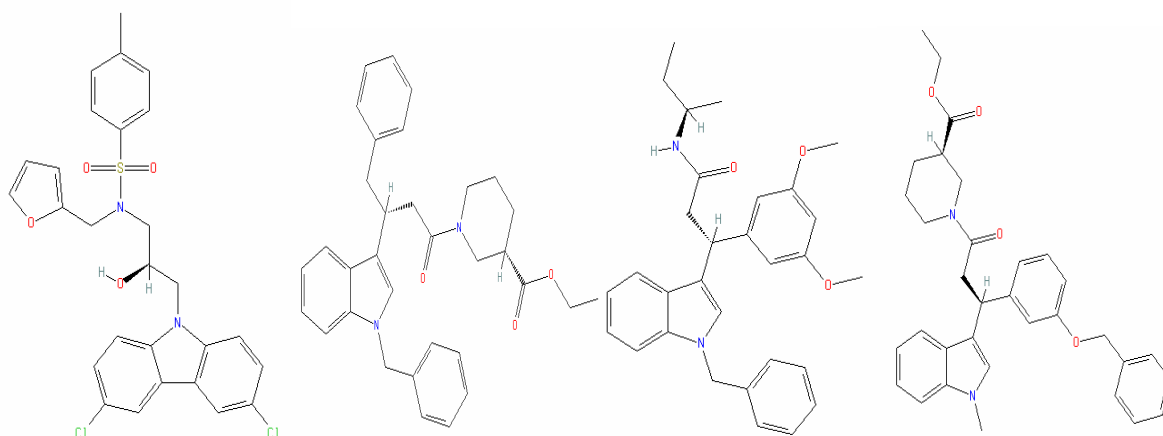


Figure 3 : Structure 2D de la molécule possédant le meilleur score de GOLD

III- DISCUSSION

La poursuite de l'analyse des résultats montre que, dans l'ensemble, les scores issus des 2 calculs concordent. Afin de déterminer notre liste de touches, nous avons superposé les 3 classements de scores des différents filtres utilisés : MSSH, GOLD et Discover. L'analyse des 100 meilleurs scores issus de chaque filtre fait ressortir 4 molécules communes. Ces composés constituant désormais notre liste de "touches" ou des candidats-médicaments possèdent de bons classements par rapport aux 3 programmes de calcul. Nous donnons ci-dessous la structure bi et tridimensionnelle des molécules formant cette liste de touches :



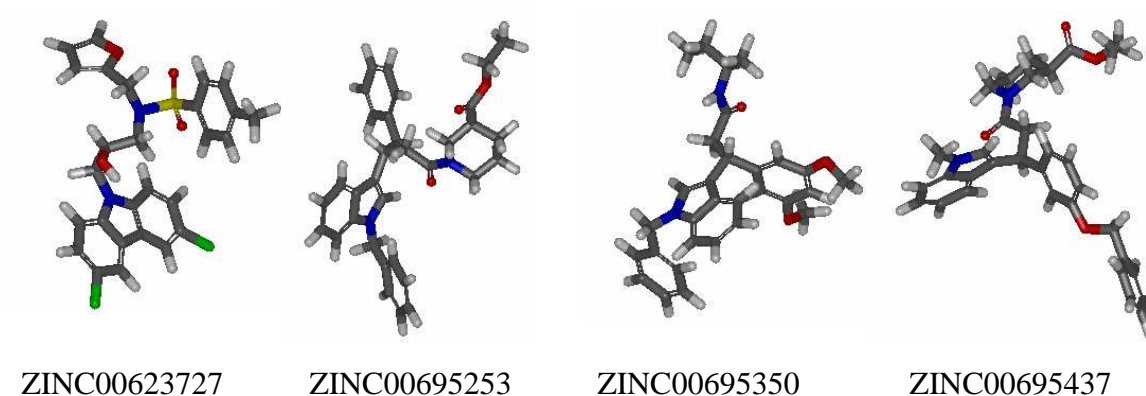


Figure 4: Structure 2D et 3D des ligands de la liste de touches et leur code Zinc respectif

Les quatre candidats-médicaments issus du criblage de notre chimiothèque appartiennent à la famille des "indoles". Notons que l'indole est un des groupements fonctionnels associés à une importante activité biologique du récepteur CCK_1 ⁽⁸⁻⁹⁾.

Nous récapitulons ci-dessous certaines caractéristiques, issues de la base de données Zinc, relatives aux quatre molécules constituant la liste de touches.

Tableau 3 : Caractéristiques de la liste de touches

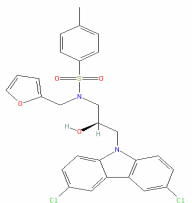
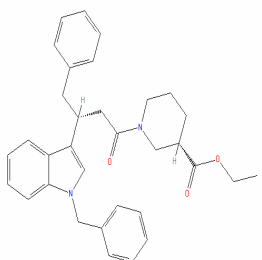
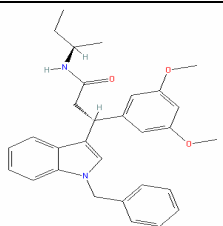
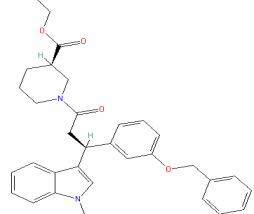
Code Zinc et formule moléculaire	Nom IUPAC	Fournisseurs	Caractéristiques			
			PM g/mol	HBD	HBA	RL
ZINC00623727 C ₂₇ H ₂₄ Cl ₂ N ₂ O ₄ S	N-(3-(3,6-dichloro-9H-pyrazol-9-yl)-2-hydroxypropyl)-N-(furan-2-ylmethyl)-4-methylbenzenesulfonamide	°Asinex °ChemDiv °IBScreen °Pharmeks °PubChem °Sigma-Aldrich	543.46	1	6	8
ZINC00695253 C ₂₆ H ₃₆ N ₂ O ₃	Ethyl 1-(3-(1-benzyl-1H-indol-2-yl)-4-phenylbutanoyl)piperidine-3-carboxylate	ComGenex	508.66	0	5	10
ZINC00695350 C ₂₈ H ₃₄ N ₂ O ₃	3-(1-benzyl-1H-indol-3-yl)-N-sec-butyl-3-(3,5-dimethoxyphenyl)propanamide	ComGenex	470.61	1	5	10
ZINC00695437 C ₃₂ H ₃₅ N ₂ O ₄	Ethyl 1-(3-(3benzyloxy)phenyl)-3-(1-methyl-1H-indol-3-yl)propanoyl)piperidine-3-carboxylate	ComGenex	524.66	0	6	10

PM : poids moléculaire, HBD : nombre de liaison hydrogène donneur, HBA : nombre de liaison hydrogène accepteur, RL : nombre de rotations libres.

Dans le souci de valider les résultats du dernier filtre utilisé, Discover, nous avons effectué un calcul d'énergie libre d'interaction au moyen du serveur BAPPL⁽¹⁰⁾, disponible sur le web, sur les 4 molécules susceptibles d'être des antagonistes CCK₁.

Nous présentons les résultats de l'ensemble des différents programmes de calcul employés, dans le tableau suivant :

Tableau 4 : Récapitulatif des résultats obtenus pour les 4 candidats-médicaments

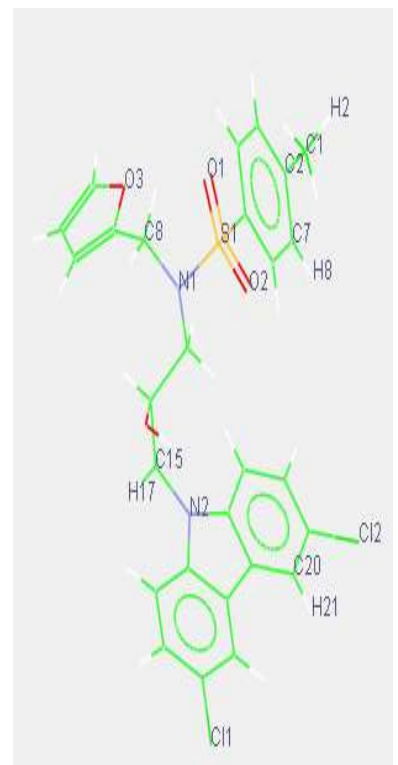
Molécule	Structure 2D	Score MSSH (Å)	Score GOLD	Energie d'interaction (kcal/mol)	Energie libre de liaison (kcal/mol)
1		0.675	80.28	-113.89	0.3
2		0.643	80.69	-125.56	-0.13
3		0.671	80.76	-110.64	-0.17
4		0.684	80.44	-127.67	-0.17

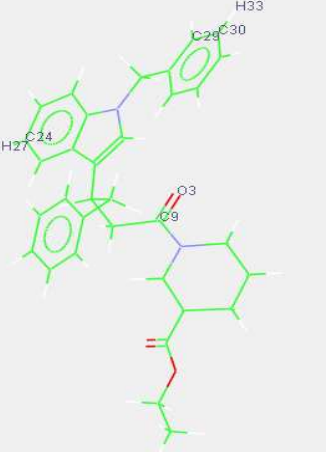
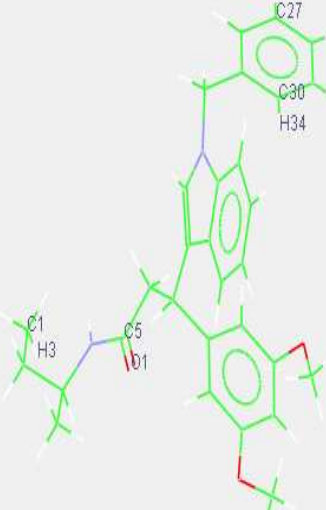
Comme expliqué dans le chapitre III, la prédiction des affinités de liaison entre le ligand et le récepteur, en tenant compte du facteur entropique, au moyen de BAPPL, passe par une fonction de score qui résulte de la somme de 3 composantes énergétiques (vdW, électrostatique et hydrophobe). Les résultats prédisent une bonne interaction entre le ligand et le récepteur, les valeurs des énergies libres étant négatives. De plus, nous avons comparé nos valeurs avec une étude similaire faite sur des antagonistes dotés d'activités expérimentales⁽⁵⁾ qui montre que des composés actifs possèdent des valeurs d'énergies libres (ΔG_{bind}) et d'interaction (E_i) voisines des valeurs que nous avons obtenues. En effet, une des molécules actives, de cette étude⁽⁵⁾, présente les valeurs d'énergies suivantes : $\Delta G_{\text{bind}} = -0.19$ kcal/mol et $E_i = -140$ kcal/mol. Cependant, le cas de la molécule, notée 1 dans le tableau 4, présente une énergie libre de liaison positive alors que son énergie d'interaction indique une bonne affinité

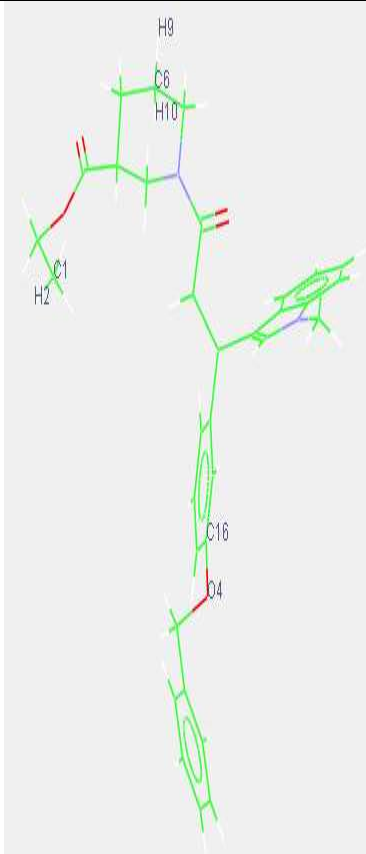
sur le récepteur CCK₁. Suite à une comparaison structurale, nous imputons ce désaccord à une mauvaise définition des paramètres liés au chlore.

Nous présentons ci-dessous l'analyse des modes de liaisons issus des dockings effectués. Au moyen du logiciel SILVER⁽¹¹⁾ nous avons visualisé les interactions mises en jeu et nous les récapitulons dans le tableau suivant :

Tableau 5 : Interactions entre le récepteur CCK₁ et les ligands constituant la liste de touches
(e) : électrostatique ; (h) : hydrophobe

Structure moléculaire	Liaisons hydrogènes			Autres interactions voisines		
	Résidu CCK ₁	Atome du ligand	Distance (Å)	Résidu CCK ₁	Atome du ligand	Distance (Å)
	• Arg336 (H)	SO ₂	1,799	• Arg336 (H)	H (C15)	1,893
	• Arg336 (NH)	SO ₂	2,305	• Ala343 (C=O)	Carbone 8	2,660
	• Arg336 (NH)	N	2,544	• Thr118 (C)	SO ₂	2,528
				• Asn194 (H)	H (C1)	2,670
				• Asn194 (H)	H (C2)	2,710
				• Asn194 (H)	H (C7)	2,494
				• Thr117 (H)	H (C7)	1,230
				• Met121 (H)	H (C20)	1,835
				• Gly122 (N)	Cl 2	2,450 (e)
				• Gly122 (C)	Cl 2	2,805

	<ul style="list-style-type: none"> • Arg336 (NH) • Arg336 (NH₂) 	C=O C=O	2,199 1,968	<ul style="list-style-type: none"> • Gln207 (H) • Ser180 (O) • Arg345 (H) • Thr118 (CO₂) 	H (C30) Carbone 29 H (C24) C=O	1,850 2,686 1,672 2,668 (e)
				<ul style="list-style-type: none"> • Ser180 (C=O) • Val113 (C=O) • Thr117 (O) • Thr117 (H) • Arg336 (H) 	Carbone 27 Carbone 1 C=O H (C1) H (C30)	2,492 2,548 2,236 (e) 1,337 1,380

				<ul style="list-style-type: none"> • Asn333 (O) Carbone 16 2,490 • Asn333 (O) Oxygène 4 1,994 (e) • Leu50 (H) H (C1) 1,830 • Leu50 (H) H (C2) 1,883 • Val113 (H) H (C6) 1,760 • Val113 (O) Carbone 6 2,624 • Thr117 (H) H (C6) 1,722 • Phe330 Phenyl 2,295 (h) • Ile352 Phenyl 2,300 (h) • Arg336 Indole 2,300 (h)
--	--	--	--	--

Dans nos résultats nous avons retrouvé des acides aminés, appartenant au récepteur CCK₁, importants dans la formation de liaisons antagonistes/CCK₁ et cités dans la littérature⁽¹²⁻²⁵⁾. Le résidu arginine 336, en particulier, est apparu dans toutes les interactions formées par les 4 molécules de la liste de touches. Il faut également noter l'importance de l'interaction hydrophobe formée par le résidu Phe330 et le groupement phényle du quatrième ligand. Cette interaction en forme de "T" est la même que celle formée par l'antagoniste pyridopyrimidique et le récepteur CCK₁. Notre résultat, en accord avec les travaux de la littérature⁽¹²⁾, montre que cette interaction semble conférer la stabilité nécessaire au ligand pour se maintenir au fond de la cavité.

A titre comparatif, nous donnons dans le tableau suivant les résidus, du récepteur CCK₁, mis en évidence dans des interactions pertinentes, dans la littérature et dans nos travaux. Les résidus, intervenant dans au moins 2 études différentes, sont indiqués avec la même couleur.

Tableau 6 : Comparaison de nos résultats avec ceux de la littérature

Résidus pertinents de la littérature		Résidus pertinents de nos travaux	Molécule
Travaux de l'équipe de Martin-Martinez ⁽¹²⁻¹⁶⁾	Autres travaux ⁽¹⁷⁻²⁵⁾		
Arg197, Arg336, Asn333, Ile329, Ile352, Phe330, Gly122, Pro177, Thr118, Val125, Leu214.	Met195, Arg197, Ser348, Ile329, Asn333, Arg336, Ile352, Val125, Leu214, Ile329, Phe330, Thr117, Lys115, Leu50, Ile51, Leu53, Cys94, Leu356, Trp39, Gln40, Met121, Phe218, Trp326, Tyr360.	Arg336, Thr118, Met121, Thr117, Ala343, Gly122, Asn194.	1
		Arg336, Thr118, Gln207, Ser180, Arg345.	2
		Arg336, Thr117, Val113, Ser180.	3
		Arg336, Asn333, Ile352, Phe330, Leu50, Val113, Thr117.	4

Nous représentons dans la figure suivante le site de liaison, molécule4/CCK₁. Les résidus mis en évidence, uniquement, par notre étude sont colorés en violet. Les autres résidus, colorés en jaune et en rouge, ont été mis en évidence antérieurement à nos travaux et signalons que certains sont communs à notre étude (Tableau 6).

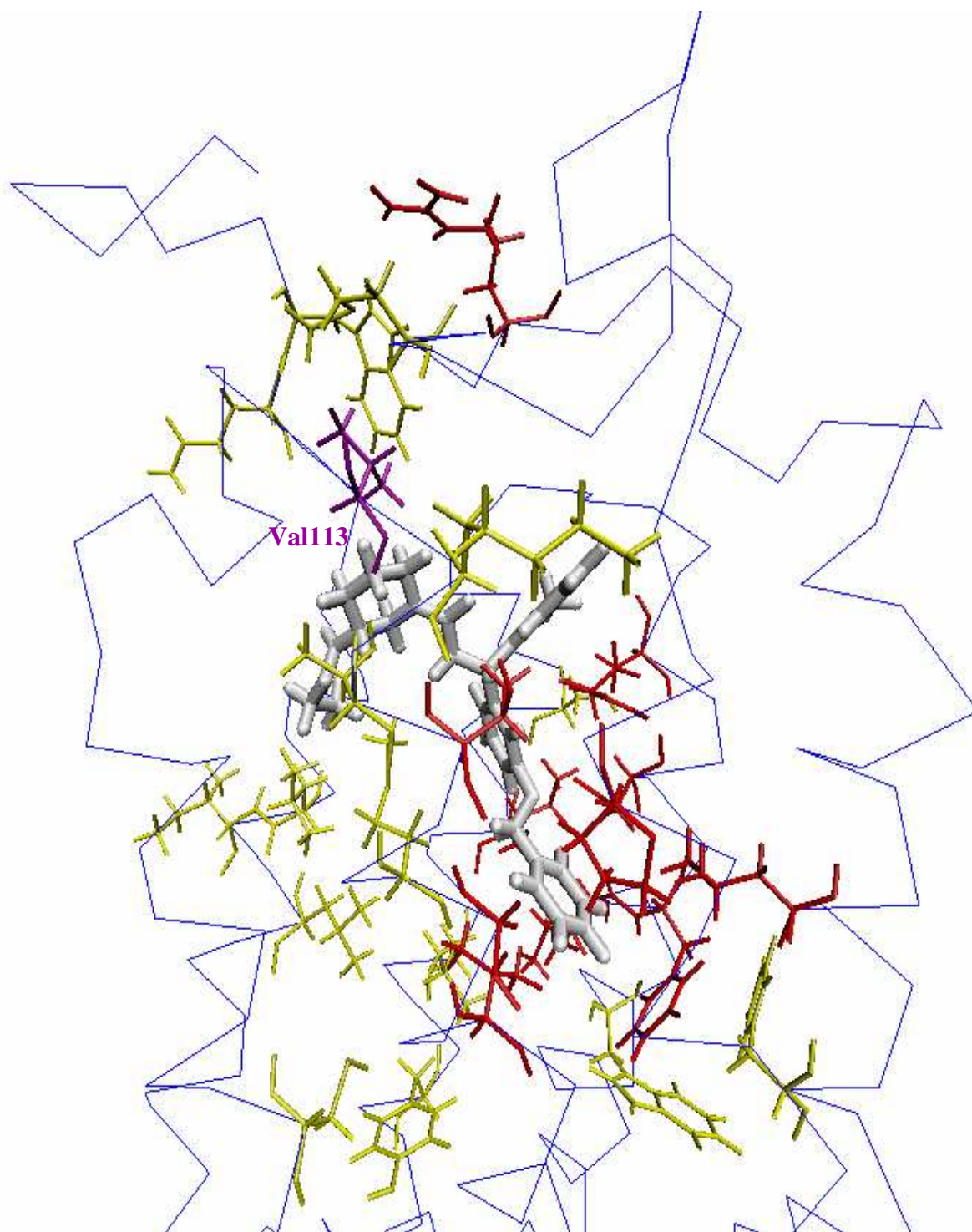


Figure 5 : Représentation des résidus entourant la molécule 4 (en gris) dockée dans le récepteur CCK₁. La trace du récepteur est représentée en bleu.

III – STRATEGIES ET DISCUSSION

Dans la seconde partie de nos travaux, nous avons voulu valider le protocole de filtrage utilisé, moyennant un ensemble de programmes. La stratégie consiste à comparer les résultats de deux filtres successifs en estimant le nombre de molécules communes à leur classement respectif. Le but de cette corrélation est de valider la position de chaque filtre, dans l'entonnoir. Le programme, écrit en Fortran, calcule le taux de molécules communes apparaissant simultanément dans les classements de scores, issus des différentes techniques de filtrage employées dans notre criblage virtuel.

A- Comparaison MSSH/GOLD

Dans le but de conforter les résultats obtenus avec MSSH, nous effectuons un docking GOLD sur l'ensemble des molécules de la chimiothèque (**51 947**). Le classement des RMSD issus de MSSH est comparé au classement des scores de fitness issus de GOLD, sur l'ensemble des molécules. Les résultats sont présentés sous forme de tableau et histogramme suivants:

Tableau 7 : Comparaison entre les classements MSSH et GOLD

Nombre de molécules	Taux de molécules communes aux 2 classements de score(%)
100	7
200	11
300	20
400	26
500	27,8
600	30,67
700	32,14
800	31,75
900	32,67
1000	33,2
1500	30,67
2000	29
3000	27,8
4000	27,92
5000	29,94
6000	31,72
7000	33,3
8000	36,13
9000	38,13
10000	40,11
20000	60,03
30000	79,65
40000	99,73

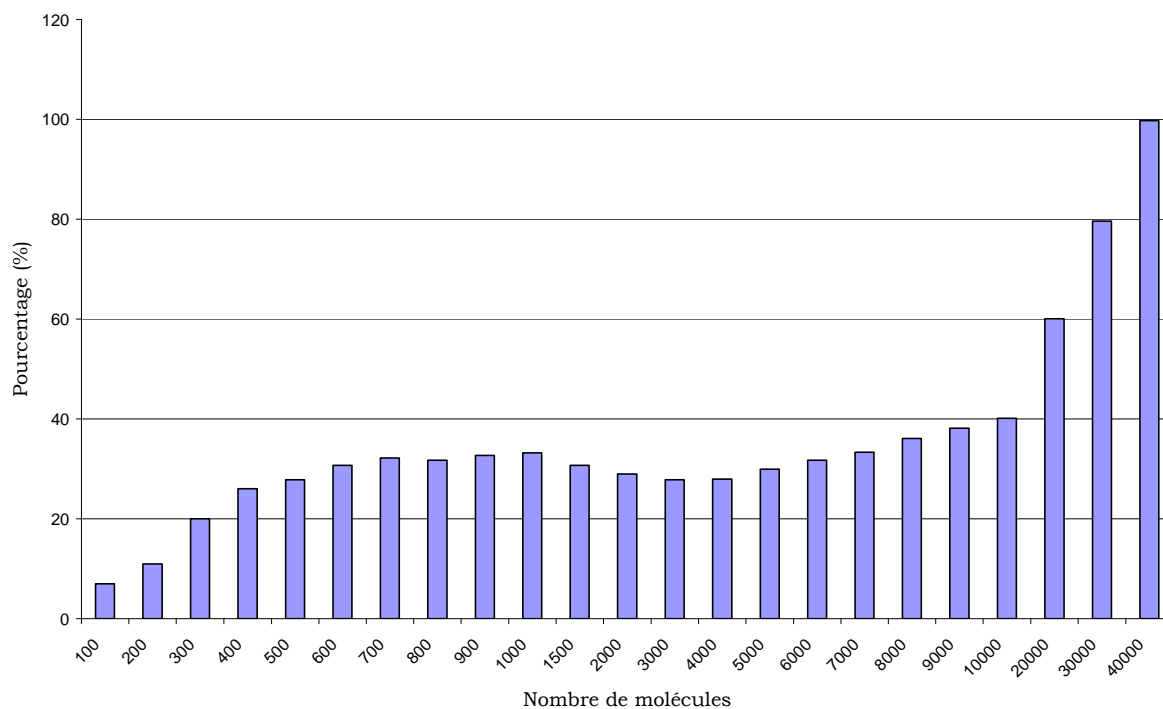


Figure 6 : Taux de molécules communes aux scores **MSSH/GOLD**

L'analyse des résultats des deux classements met en évidence un palier autour des 1000 premières molécules communes aux scores MSSH et GOLD. En d'autres termes, pour ces 1000 meilleurs scores des deux docking, MSSH et GOLD, correspond un taux de 33,2% de molécules en commun. Il est important de noter que pour les 10 000 meilleurs scores, ce taux devient égal à 40,11%. Ce qui veut dire que pour environ 20 % de la chimiothèque (10 000 premières molécules), nous obtenons un taux de molécules communes aux deux scores, égal à environ 40 % de la chimiothèque (taux de 40 %).

L'analyse de ce résultat montre que GOLD peut être utilisé à la suite du filtre MSSH ; ce qui permet d'épargner partiellement les calculs prohibitifs que nécessite GOLD.

Il faut noter que la question de temps que prendrait un docking rigide par rapport à un docking semi-flexible a son importance surtout lorsque le criblage concerne des chimiothèques contenant des millions de molécules. En effet, le temps moyen CPU consommé par GOLD pour une molécule de notre chimiothèque est de 21 mn alors que celui consommé par MSSH est de 2,43 s.

Nous avons, également, cherché à connaître le positionnement des 363 molécules, issues du filtre GOLD, dans le classement comparatif MSSH/GOLD. Les résultats ont montré que dans les 1000 premiers scores nous retrouvons 50% de ces 363 molécules. Il est bon de noter

que le taux de présence devient égal à 75% lorsque l'on pousse l'analyse aux 10 000 premiers scores obtenus par MSSH et GOLD.

B- Comparaison GOLD/Discover

La recherche d'une éventuelle corrélation entre les résultats de GOLD et de Discover s'est faite sur le classement des 363 complexes. Les résultats issus de la comparaison entre les scores de fitness de GOLD et les énergies d'interaction de Discover sont donnés dans le tableau et l'histogramme suivants :

Tableau 8 : Comparaison entre les classements GOLD et Discover

Nombre de molécules	Taux de molécules communes aux 2 classements de score (%)
10	0
20	10
30	6.7
40	5
50	12
100	32
200	60
300	79

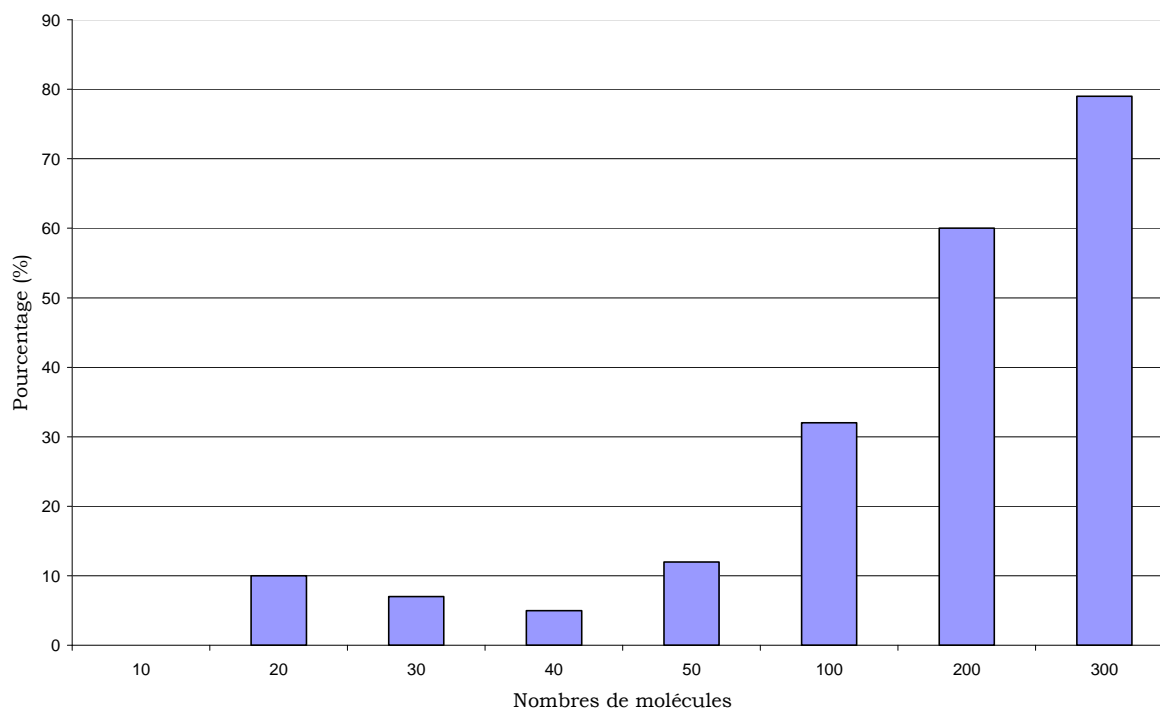


Figure 7 : Taux de molécules communes aux scores GOLD/Discover

À la moitié des 363 molécules issues de GOLD, correspond un taux de molécules communes, aux 2 filtres GOLD et Discover, égal à 60 % montrant ainsi la fiabilité des scores de GOLD à prédire un bon classement (dans la limite des capacités de sa fonction de score).

Le détail des résultats de ces études comparatives est présenté dans l'annexe 3.

IV- CONCLUSION

Les quatre candidats-antagonistes issus de notre criblage, présentent des groupements fonctionnels, révélés importants dans la littérature quant à leur activité sur le récepteur CCK₁. Nous comptons les soumettre ultérieurement à des tests expérimentaux ; ce qui conforterait notre étude théorique.

Malgré la divergence de leurs scores pris séparément, les différentes techniques de criblage utilisées, semblent avoir un bon pouvoir prédictif. En effet, leur application successive, de la plus rapide à la plus précise a permis de filtrer notre chimiothèque de 850 000 objets tridimensionnels à 4 molécules susceptibles de présenter une activité antagoniste sur le récepteur CCK₁.

La "validation" du processus de criblage employé nous a donné des résultats satisfaisants, confirmant ainsi la position "stratégique" de chaque filtre dans l'entonnoir.

Notons également que le processus de validation du protocole, que nous avons mis en oeuvre indépendamment de VSM-G, met en exergue la nécessité d'utiliser des filtres complémentaires et croissants en précision ; ce qui permet de conclure que cette technique de criblage virtuel semble adaptée à la recherche de candidats antagonistes CCK₁.

Cependant, au cours de notre criblage nous n'excluons pas le risque d'écarter des molécules "réellement" affines au récepteur CCK₁. Ces composés exclus sont appelés les "faux négatifs". Il est vrai qu'à chaque étape du filtrage, nous avons essayé de minimiser cet état de fait, mais le risque d'éliminer des "hits" potentiels reste toujours présent.

REFERENCES BIBLIOGRAPHIQUES

- 1- Bursulaya, B. D., Trotoev, M., Abagyan, R. et Brooks III, C. L., Comparative study of several algorithms for flexible ligand docking. *Journal of Computer-Aided Molecular Design*, 17, 755-763, **2003**.
- 2- Bindewald, E. et Skolnick, J., A scoring function for docking ligands to low-resolution protein structures. *J. Comput. Chem.*, 26, 374-383, **2005**.
- 3- Kellenberger, E., Rodrigo, J., Muller, P. et Rognan, D., Comparative evaluation of eight docking tools for docking and virtual screening accuracy. *PROTEINS: Structure, Function, and Bioinformatics*, 57, 225-242, **2004**.
- 4- Erickson, J. A., Jalaie, M., Robertson, D. H., Lewis, R. A. et Vieth M., Lessons in molecular recognition: the effects of ligands and protein flexibility on molecular docking accuracy. *J. Med. Chem.*, 47, 45-55, **2004**.
- 5- Maouche-Toumi, A., Thèse de doctorat en cours, Alger, USTHB.
- 6- <http://www.ccdc.cam.ac.uk/support/documentation/#gold>
- 7- Accelrys/MSI. Road San Diego, CA 92121 USA.
- 8- Varnavas, A., Lassiani, L., Valenta, V., Berti, F., Tontini, A., Mennuni, L., Makovek, F., Anthranilic acid based CCK₁ antagonists : the 2-indole moiety may represent a « needle » according to the recent homonymous concept. *European Journal of Medicinal Chemistry*, 39, 85-97, **2004**.
- 9- Varnavas, A., Lassiani, L., Valenta, V., Mennuni, L., Makovek, F., Hadjipavlou-Litina, D., Anthranilic acid based CCK₁ receptor antagonists: preliminary investigation on their second “touch point”. *European Journal of Medicinal Chemistry*, 40, 563-581, **2005**.
- 10- www.scfbio-iitd.res.in/software/drugdesign/bappl.jsp
- 11- <http://www.ccdc.cam.ac.uk/>
- 12- Martin-Martinez, M., Marty, A., Jourdan, M., Escriet, C., Archer, E., Gonzalez-Muniz, R., Garcia-Lopez, M.T., Maigret, B., Herranz, R. et Fourmy, D. Combination of molecular modeling, site directed mutagenesis and SAR studies to delineate the binding site of pyridopyrimidine antagonists on the human CCK₁ receptor. *J. Med. Chem.* 48, 4842-4850, **2005**.
- 13- Bartolomé-Nebreda, J.M., Patiño-Molina, R., Martín-Martínez, M., Gómez-Monterrey, I., García-López, M. T., González-Muñiz, R., Cenarruzabeitia, E., Latorre, M., Del Río, J. et Herranz, R. 5-(Tryptophyl)amino-1,3-

- dioxoperhydropyrido[1,2-c]pyrimidine-based potent and selective CCK₁ receptor antagonists: structure-activity relationship studies on the substituent at N2-position. *J. Med. Chem.* 44, 2219-2228, **2001**.
- 14- Bartolomé-Nebreda, J.M., García-López, M. T., González-Muñiz, R., Cenarruzabeitia, E., Latorre, M., Del Río, J. et Herranz, R. 5-(Tryptophyl)amino-1,3-dioxoperhydropyrido[1,2- c]pyrimidine-based potent and selective CCK₁ receptor antagonists: structure-activity relationship studies on the central 1,3-dioxoperhydropyrido[1,2-c]pyrimidine scaffold. *J. Med. Chem.* 44, 4196-4206, **2001**.
- 15- Bartolomé-Nebreda, J.M., Gómez-Monterrey, I., García-López, M. T., González-Muñiz, R., Martín-Martínez, M., Cenarruzabeitia, E., Latorre, M., Del Río, J. et Herranz, R. 5-(Tryptophyl)amino-1,3-dioxoperhydropyrido[1,2-c]pyrimidine-based potent and selective CCK₁ receptor antagonists: structural modifications at the tryptophan domain. *J. Med. Chem.* 42, 4659-4668, **1999**.
- 16- Martín-Martínez, M., Bartolomé-Nebreda, J.M., Gómez-Monterrey, I., González-Muñiz, R.,García-López, M.T., Ballaz, S., Barber, A., Fortuño, A., Del Río, J. et Herranz, R. Synthesis and stereochemical structure-activity relationships of 1,3-dioxoperhydropyrido[1,2-c]pyrimidine derivatives: potent and selective cholecystokinin-A receptor antagonists. *J. Med. Chem.* 40,3402-3407, **1997**.
- 17- Gouldson ,P., Legoux, P., Carillon, C., Delpech, B., Le Fur, G.,Ferrara, P., Shire, D., Contrasting role of Leu³⁵⁶ in the human CCK₁ receptor for antagonist SR 27897 and agonist SR 146131 binding. *European Journal of Pharmacology*, 383, 339-346, **1999**.
- 18- Gouldson, P., Legoux, P., Carillon, C., Delpech, B., Le Fur, G.,Ferrara, P., Shire, D., The agonist SR 146131 and the antagonist SR 27897 occupy different sites on the human CCK₁ receptor. *European Journal of Pharmacology*. 400, 185-194, **2000**.
- 19- Gigoux, V., Maigret, B., Escricut, C., Silvente-Poirot, S., Bouisson, M., Fehrentz, J-A., Moroder, L., Gully, D., Martinez, J., Vaysse, N. et Fourmy, D. Arginine 197 of the cholecystokinin-A receptor binding site interacts with the sulfate of the peptide agonist cholecystokinin. *Protein Sci.* 8, 2347-2354, **1999**.
- 20- Gigoux, V., Escricut, C., Fehrentz, J-A., Poirot, S., Maigret, B., Moroder, L., Gully, D., Martinez, J., Vaysse, N. et Fourmy, D. Arginine 336 and asparagine 333 of the human cholecystokinin-A receptor binding site interact with the penultimate aspartic acid and the C-terminal amide of cholecystokinin. *J. Biol. Chem.* 274, 20547-20464, **1999**.

- 21- Escrieut, C., Gigoux, V., Archer, E., Verrier, S., Maigret, B., Behrendt, R., Moroder, L., Bignon, E., Silvente-Poirot, S., Pradayrol, L. et Fourmy, D. The biologically crucial C-terminus of cholecystokinin and the non-peptide agonist SR-146,131 share a common binding site in the human CCK₁ receptor. Evidence for a crucial role of Met-121 in the activation process. *J. Biol. Chem.* 277, 7546-7555, **2002**.
- 22- Kennedy, K., Gigoux, V., Escrieut, C., Maigret, B., Martinez, J., Moroder, L., Fréhel, D., Gully, D., Vaysse, N. et Fourmy, D. Identification of two amino acids of the human cholecystokinin-A receptor that interact with the N-terminal moiety of cholécystokinine. *The Journal of Biological Chemistry*, 5, 2920-2926, **1997**.
- 23- Gigoux, V., Escrieut, C., Silvente-Poirot, S., Maigret, B., Gouilleux, L., Fehrentz, J-A., Gully, D., Moroder, L., Vaysse, N. et Fourmy, D. Met-195 of the cholecystokinin-A receptor interacts with the sulfated tyrosine of cholecystokinin and is crucial for receptor transition to high affinity state. *The Journal of Biological Chemistry*, 23, 14380-14386, **1998**.
- 24- Fourmy, D., Escrieut, C., Archer, E., Galès, C., Gigoux, V., Maigret, B., Moroder, L., Silvente-Poirot, S., Martinez, J., Fehrentz, J. A. et Pradayrol, L., Structure of Cholecystokinin receptor binding sites and mechanism of activation/inactivation by agonists/antagonists. *Pharmacology & Toxicology.*, 91, 313-320, **2002**.
- 25- Archer-Lahlou, E., Tikhonova, I., Escrieut, C., Dufresne, M., Seva, C., Pradayrol, L., Moroder, L., Maigret, B. et Fourmy, D., Modeled structure of a G-protein-coupled receptor: the cholecystokinin-1 receptor. *J. Med. Chem.*, 48, 180-191, **2005**.

CONCLUSION GENERALE ET PERSPECTIVES

Dans cette ère post-génomique et protéomique où l'on connaît de grandes avancées en chimie organique combinatoire et techniques informatiques, le criblage moléculaire virtuel à haut débit s'avère comme une méthode très utile. En effet, cette méthode de travail est incontournable pour qui veut éviter de procéder à de trop nombreux essais, coûteux, par la technique du criblage à haut débit *in vitro*; ceci tout en explorant une diversité moléculaire plus importante. En outre, le criblage virtuel permet d'accélérer et de rationaliser la recherche de molécules innovantes, qui sont des têtes de série pour la recherche de molécules actives sur des cibles biologiques importantes en thérapeutique.

Dans le premier chapitre nous avons présenté l'hormone endogène, cible moléculaire de nos recherches à savoir la cholécystokinine, ainsi que ses récepteurs membranaires. Cette mise à jour des connaissances nous a permis de faire le point sur les sites de liaisons pertinents issus d'études effectuées sur le récepteur périphérique, CCK₁, et quelques uns de ses ligands (agonistes et antagonistes).

Le deuxième chapitre est consacré à la présentation des méthodes issues du drug design. Nous avons donné le détail des techniques usuelles de la modélisation moléculaire et les techniques utilisées en particulier dans le cadre du criblage virtuel par les différents types de docking moléculaire.

Le troisième chapitre concerne la méthodologie que nous avons adoptée pour mener notre recherche de molécules susceptibles d'être des candidats-antagonistes du récepteur CCK₁. Nous avons collecté nos molécules à partir d'une base de données, disponible sur le web et en effectuant une recherche par sous structure 'pharmacophorique'. Le modèle du récepteur utilisé dans nos travaux est le modèle conçu et validé expérimentalement par l'équipe de Madame Martin-Martinez. Ce modèle structural 3D présente une cavité propre aux ligands utilisés pour son élaboration, à savoir les dérivés pyridopyrimidiques. Nous avons criblé notre chimiothèque à l'aide de l'outil logistique Virtual Screening Manager for Grids (VSM-G). Ce dernier comporte une succession de programmes jouant le rôle de filtres moléculaires. En effet, ces techniques de criblage sont classées de la plus rapide, mais moins efficace, à la plus précise.

Dans le quatrième chapitre nous avons présenté, en premier lieu, les résultats issus du criblage de notre chimiothèque. La liste de touches est formée de quatre ligands qui ont reproduit certains aspects décrits dans la littérature concernant les interactions mises en jeu entre le récepteur CCK₁ et les antagonistes non-peptidiques. En second lieu, nous avons, d'abord, exposé la stratégie employée pour valider les techniques de criblages utilisées puis on a présenté les résultats des corrélations entre les filtres successifs.

En conclusion, la liste de touches résultante de notre criblage comportant quatre candidats-antagonistes, présente des groupements fonctionnels révélés importants dans la littérature quant à leur activité sur le récepteur CCK₁. Ces quatre candidats-médicaments seront ultérieurement proposés aux équipes de biologistes pour des tests expérimentaux ; ce qui complétera notre étude théorique.

La validation du processus de criblage employé nous a donné des résultats satisfaisants, validant la position de chaque filtre dans l'entonnoir. Malgré la divergence de leurs scores pris séparément, les différentes techniques de criblage utilisées, semblent avoir un bon pouvoir prédictif. En effet, leur application successive, de la plus rapide à la plus précise a permis de filtrer notre chimiothèque de 850 000 objets tridimensionnels à 4 molécules susceptibles de présenter une activité antagoniste sur le récepteur CCK₁. Notons également que le processus de validation du protocole, que nous avons mis en oeuvre indépendamment de VSM-G, met en exergue la nécessité d'utiliser des filtres complémentaires et croissants en précision ; ce qui permet de conclure que cette technique de criblage virtuel semble adaptée à la recherche de candidats antagonistes CCK₁.

Cependant, au cours de notre criblage nous n'excluons pas le risque d'écarter des molécules "réellement" affines au récepteur CCK₁. Ces composés exclus sont appelés les "faux négatifs". Il est vrai qu'à chaque étape du filtrage, nous avons essayé de minimiser cet état de fait, mais le risque d'éliminer des "hits" potentiels reste toujours réel. Donc, en perspective de nos travaux et dans le but d'accélérer le criblage virtuel au moyen de VSM-G, nous projetons d'introduire en amont du traitement de la chimiothèque, un pré-filtrage par rapport à un modèle statistique QSAR, basé sur des critères topologiques, permettant ainsi d'améliorer la répartition en vrais positifs et vrais négatifs.

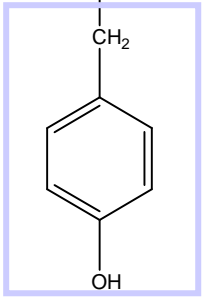
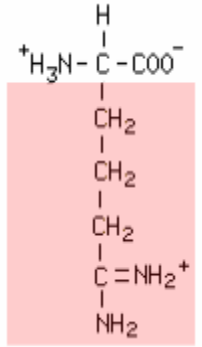
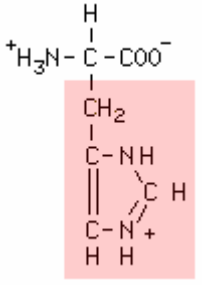
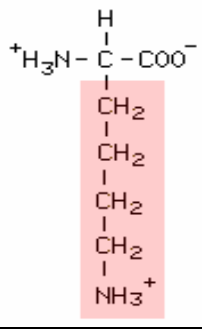
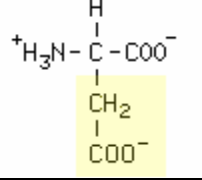
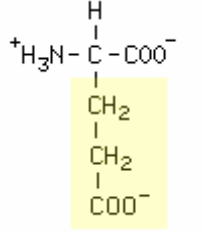
ANNEXES

ANNEXE 1

Liste des 20 acides aminés naturels avec quelques propriétés.

	FORMULE	NOM	CODES	PROPRIETES STRUCTURALES
Acides Aminés hydrophobes (non polaires)	$\begin{array}{c} \text{H} \\ \\ ^+\text{H}_3\text{N}-\text{C}-\text{COO}^- \\ \\ \text{H} \end{array}$	Glycine	Gly G	- sans radical
	$\begin{array}{c} \text{H} \\ \\ ^+\text{H}_3\text{N}-\text{C}-\text{COO}^- \\ \\ \text{CH}_3 \end{array}$	Alanine	Ala A	- Petite taille
	$\begin{array}{c} \text{H} \\ \\ ^+\text{H}_3\text{N}-\text{C}-\text{COO}^- \\ \\ \text{H}-\text{C}-\text{CH}_3 \\ \\ \text{CH}_2 \\ \\ \text{CH}_3 \end{array}$	Isoleucine	Ile I	- Aliphatique
	$\begin{array}{c} \text{H} \\ \\ ^+\text{H}_3\text{N}-\text{C}-\text{COO}^- \\ \\ \text{CH}_2 \\ \\ \text{CH} \\ / \quad \backslash \\ \text{H}_3\text{C} \quad \text{CH}_3 \end{array}$	Leucine	Leu L	- Aliphatique
	$\begin{array}{c} \text{H} \\ \\ ^+\text{H}_3\text{N}-\text{C}-\text{COO}^- \\ \\ \text{CH}_2 \\ \\ \text{CH}_2 \\ \\ \text{S} \\ \\ \text{CH}_3 \end{array}$	Méthionine	Met M	- Aliphatique - Soufré
	$\begin{array}{c} \text{H} \\ \\ ^+\text{H}_3\text{N}-\text{C}-\text{COO}^- \\ \\ \text{CH}_2 \\ \\ \text{C}_6\text{H}_5 \end{array}$	Phénylalanine	Phe F	- Aromatique
	$\begin{array}{c} \text{H} \\ \\ ^+\text{H}_2\text{N}-\text{C}-\text{COO}^- \\ \\ \text{H}_2\text{C} \quad \text{CH}_2 \\ \quad \quad \\ \quad \quad \text{CH}_2 \end{array}$	Proline	Pro P	- Rigide - Contient un groupement secondaire

	$ \begin{array}{c} \text{H} \\ \\ ^+\text{H}_3\text{N}-\text{C}-\text{COO}^- \\ \\ \text{CH}_2 \\ \\ \text{C}=\text{CH} \\ \\ \text{NH} \\ \\ \text{C}_6\text{H}_4 \end{array} $	Tryptophane	Trp	W	-Aromatique -Donneur de liaisons H*
	$ \begin{array}{c} \text{H} \\ \\ ^+\text{H}_3\text{N}-\text{C}-\text{COO}^- \\ \\ \text{CH} \\ \\ \text{H}_3\text{C} \quad \text{CH}_3 \end{array} $	Valine	Val	V	- Aliphatique
Acides Aminés non chargés hydrophiles (polaires)	$ \begin{array}{c} ^+\text{H}_3\text{N}-\text{CH}-\text{COO}^- \\ \\ \text{CH}_2 \\ \\ \text{C}=\text{C} \\ \\ \text{NH}_2 \end{array} $	Asparagine	Asn	N	- Donneur et accepteur de liaisons H* - Contient la fonction amide
	$ \begin{array}{c} ^+\text{H}_3\text{N}-\text{CH}-\text{COO}^- \\ \\ \text{CH}_2 \\ \\ \text{SH} \end{array} $	Cystéine	Cys	C	- Soufré - Responsable de la formation des ponts disulfure
	$ \begin{array}{c} ^+\text{H}_3\text{N}-\text{CH}-\text{COO}^- \\ \\ \text{CH}_2 \\ \\ \text{CH}_2 \\ \\ \text{C}=\text{O} \\ \\ \text{NH}_2 \end{array} $	Glutamine	Gln	Q	- Donneur et accepteur de liaisons H* - Contient la fonction amide
	$ \begin{array}{c} ^+\text{H}_3\text{N}-\text{CH}-\text{COO}^- \\ \\ \text{CH}_2 \\ \\ \text{OH} \end{array} $	Sérine	Ser	S	- Donneur et accepteur de liaisons H* - Hydroxylé
	$ \begin{array}{c} ^+\text{H}_3\text{N}-\text{CH}-\text{COO}^- \\ \\ \text{CH}-\text{OH} \\ \\ \text{CH}_3 \end{array} $	Thréonine	Thr	T	- Donneur et accepteur de liaisons H* - Hydroxylé

	$^+H_3N-CH-COO^-$ 	Tyrosine	Tyr	Y	<ul style="list-style-type: none"> - Aromatique - Donneur et accepteur de liaisons H* - Contient le groupement phénol
Acides Aminés positivement chargés (polaires)	$^+H_3N-C-COO^-$ 	Arginine	Arg	R	<ul style="list-style-type: none"> - Basique - Donneur de liaisons H*
	$^+H_3N-C-COO^-$ 	Histidine	His	H	<ul style="list-style-type: none"> - Basique - Aromatique - Donneur et accepteur de liaisons H*
	$^+H_3N-C-COO^-$ 	Lysine	Lys	K	<ul style="list-style-type: none"> - Basique - étendue et flexible
Acides Aminés négativement chargés (polaires)	$^+H_3N-C-COO^-$ 	Aspartate	Asp	D	<ul style="list-style-type: none"> - Acide - Accepteur de liaisons H*
	$^+H_3N-C-COO^-$ 	Glutamate	Glu	E	<ul style="list-style-type: none"> - Acide - Accepteur de liaisons H*

* liaisons H = liaisons hydrogène.

➤ Agoniste :

Un **agoniste** (du latin tardif *agnista*, 'contester', du grec *agnists*, 'contestataire', de *agn*, 'contester') est une molécule (le peptide endogène ou le médicament) interagissant avec un récepteur membranaire et **activant** celui-ci. L'agoniste mime en général le messenger se liant habituellement avec le récepteur en question. Il peut être plus ou moins sélectif pour un type de récepteur.

➤ Agoniste inverse :

En pharmacologie, un **agoniste inverse** est un agent qui interagit avec le même récepteur qu'un agoniste de ce récepteur mais ne produit pas l'effet pharmacologique de l'agoniste.

➤ Antagoniste :

Un **antagoniste** est une molécule interagissant avec un récepteur membranaire et **bloquant** sa fonction.

➤ Règles de Lipinski :

Lipinski et al. ont établi la règle des 5 stipulant que pour devenir un médicament, une molécule doit posséder 3 des 4 propriétés suivantes:

- ◆ Poids moléculaire ≤ 500
- ◆ $\log P_{\text{calc}} \leq 5$
- ◆ Nombre de donneurs de liaisons H ≤ 5
- ◆ Nombre d'accepteurs de liaisons H ≤ 10
- ◆ Nombre de rotules ≤ 5

➤ Cibles orphelines :

Cibles pour lesquelles aucun ligand n'a été identifié.

ANNEXE 2

Tableau regroupant les résultats, concernant les scores de GOLD et les énergies d'interactions calculées par Discover, pour les 363 complexes.

	<i>Code Zinc</i>	<i>Score GOLD</i>	<i>E_i(kcal/mol)</i>	<i>Famille</i>
1	ZINC01039093	74.39	-84.27	Propanone
2	ZINC00821012	72.34	-93.09	
3	ZINC03058581	86.73	-98.74	Benzodiazepine
4	ZINC03055753	82.42	-104.23	
5	ZINC01657815	72.64	-94.41	
6	ZINC02072804	72.71	-118.49	
7	ZINC03055914	81.33	-98.03	
8	ZINC03055937	80.65	-109.21	
9	ZINC03055938	79.33	-114.41	
10	ZINC03059969	79.10	-111.90	
11	ZINC03055755	79.29	-103.75	
12	ZINC03029178	78.53	-110.89	
13	ZINC03055934	77.63	-113.48	
14	ZINC03058582	77.53	180.98	
15	ZINC03051422	76.68	-127.31	
16	ZINC03059834	76.98	165.31	
17	ZINC03055752	75.07	-118.06	
18	ZINC03055946	75.22	-96.10	
19	ZINC03029182	74.96	-106.60	
20	ZINC03029189	74.77	-111.75	
21	ZINC03055713	74.65	-97.62	
22	ZINC03055806	74.15	-109.22	
23	ZINC03055917	74.04	-100	
24	ZINC03055951	74.56	-96.31	
25	ZINC03029166	73.13	-98.36	
26	ZINC03055936	73.61	-108.40	

	<i>Code Zinc</i>	<i>Score GOLD</i>	<i>E_i (kcal/mol)</i>	<i>Famille</i>
27	ZINC03029179	72.51	-111.89	
28	ZINC03029191	72.51	-115.48	
29	ZINC03055737	72.21	-110.22	
30	ZINC03055742	72.59	-107.76	
31	ZINC02881744	79.43	-91.92	BCHA
32	ZINC00969715	77.76	-81.02	
33	ZINC02636876	77.35	-96.75	
34	ZINC00655040	76.68	-82.98	
35	ZINC00979363	74.48	-80.33	
36	ZINC01461745	74.43	207.64	
37	ZINC00646006	74.41	-77.07	
38	ZINC00866604	74.39	-82.40	
39	ZINC00656683	74.23	-86.54	
40	ZINC01112301	74.21	-80.06	
41	ZINC00714120	74.11	-90.98	
42	ZINC00835572	74.04	-98.29	
43	ZINC01462333	73.08	-85.57	
44	ZINC02882084	73.08	-75.84	
45	ZINC00645985	73.65	-88.26	
46	ZINC00656128	73.05	-71.68	
47	ZINC00656515	73.17	-84.31	
48	ZINC00780548	73.65	-92.64	
49	ZINC00870040	73.66	-81.20	
50	ZINC00979184	73.00	-79.41	
51	ZINC01171274	72.32	-82.34	
52	ZINC02899532	72.64	-75.14	
53	ZINC00634663	72.87	-75.18	
54	ZINC00043147	72.80	-73.16	Pyridine

	<i>Code Zinc</i>	<i>Score GOLD</i>	<i>E_i (kcal/mol)</i>	<i>Famille</i>
55	ZINC00051736	72.61	-73.27	
56	ZINC00028437	72.51	-95.21	NHCO
57	ZINC00036885	72.52	-92.16	
58	ZINC00911861	90.28	176.08	Azepane
59	ZINC00911379	88.43	183.52	
60	ZINC00908931	87.19	186.44	
61	ZINC00907720	86.36	182.74	
62	ZINC00908934	83.15	190.33	
63	ZINC00893887	82.88	183.94	
64	ZINC00893488	82.64	181.01	
65	ZINC00911864	82.61	182.58	
66	ZINC00908644	82.60	184.13	
67	ZINC00893030	82.01	180.08	
68	ZINC01005807	81.48	-94.14	
69	ZINC00911385	81.47	186.88	
70	ZINC00910898	81.91	183.61	
71	ZINC00890336	81.32	184.46	
72	ZINC00909486	80.03	178.21	
73	ZINC00912293	80.12	179.02	
74	ZINC00907724	80.81	186.36	
75	ZINC00889033	80.79	182.99	
76	ZINC00909861	79.60	184.48	
77	ZINC00908643	79.80	181.78	
78	ZINC00890061	79.16	184.69	
79	ZINC00909862	78.68	196.27	
80	ZINC00912298	78.08	178.95	

	<i>Code Zinc</i>	<i>Score GOLD</i>	<i>E_i (kcal/mol)</i>	<i>Famille</i>
81	ZINC00912304	78.91	182.19	
82	ZINC00781853	77.28	-71.00	
83	ZINC00930618	77.85	-75.63	
84	ZINC00894277	76.76	199.86	
85	ZINC00907722	76.07	183.42	
86	ZINC00909488	76.93	197.64	
87	ZINC00889315	75.99	207.10	
88	ZINC00754633	75.11	-87.18	
89	ZINC00715656	75.57	-95.32	
90	ZINC00908641	75.97	173.64	
91	ZINC00909484	75.67	174.44	
92	ZINC00922318	75.04	-70.06	
93	ZINC00660280	74.65	-95.34	
94	ZINC00706255	74.68	181.34	
95	ZINC00796429	74.00	-89.57	
96	ZINC00879924	74.73	-88.60	
97	ZINC00892932	74.09	210.60	
98	ZINC00893396	74.73	210.28	
99	ZINC00908639	74.15	191.35	
100	ZINC00909485	74.60	197.98	
101	ZINC00912290	74.39	184.67	
102	ZINC01016329	74.10	-87.93	
103	ZINC00654612	73.54	-86.42	
104	ZINC00660251	73.55	-84.12	
105	ZINC00665621	73.15	-94.22	
106	ZINC00790016	73.05	-100.47	

	<i>Code Zinc</i>	<i>Score GOLD</i>	<i>E_i (kcal/mol)</i>	<i>Famille</i>	
107	ZINC00823177	73.00	207.33		
108	ZINC00893486	73.60	183.05		
109	ZINC00893603	73.76	194		
110	ZINC00893606	73.21	196.63		
111	ZINC00893800	73.47	213.69		
112	ZINC00894002	73.42	199.12		
113	ZINC00907718	73.07	180.89		
114	ZINC00910095	73.69	172.95		
115	ZINC00919681	73.60	-97.72		
116	ZINC00931301	73.84	-76.69		
117	ZINC00983659	73.74	-81.52		
118	ZINC00889141	72.78	199.10		
119	ZINC00890129	72.34	193.37		
120	ZINC00911383	72.28	207.62		
121	ZINC00654334	72.54	-86.94		
122	ZINC00034079	75.03	-391.03		Carboxylate
123	ZINC00046932	72.44	-76.93		
124	ZINC00696742	89.22	162.71		Indole
125	ZINC00695380	88.37	166.30		
126	ZINC00696743	88.36	173.99		
127	ZINC00696647	87.96	163.97		
128	ZINC00696715	87.64	174.91		
129	ZINC00696725	86.07	-109.33		
130	ZINC00696785	86.27	159.85		
131	ZINC00706600	86.03	-115.45		
132	ZINC00705479	85.93	-104.27		

	<i>Code Zinc</i>	<i>Score GOLD</i>	<i>E_i (kcal/mol)</i>	<i>Famille</i>
133	ZINC00695267	85.74	161.24	
134	ZINC00706035	85.12	-127.02	
135	ZINC00634887	84.56	-124.88	
136	ZINC00695353	84.15	160.04	
137	ZINC00696717	84.81	-112.08	
138	ZINC00696633	83.90	153.65	
139	ZINC00696724	83.92	-111.38	
140	ZINC00696726	83.89	-112.91	
141	ZINC00696746	83.27	-106.51	
142	ZINC00695269	82.04	155.31	
143	ZINC00693223	82.86	-102.5	
144	ZINC00638627	82.79	-400.34	
145	ZINC00695461	82.02	147.62	
146	ZINC00696716	82.32	149.11	
147	ZINC00696719	82.99	-114.73	
148	ZINC00696734	82.52	162.58	
149	ZINC00705700	82.51	-99.17	
150	ZINC00633586	81.49	-107.75	
151	ZINC00681091	81.29	-119.31	
152	ZINC00695251	81.35	173.64	
153	ZINC00696637	81.77	-109.86	
154	ZINC00696638	81.35	-104.99	
155	ZINC00696642	81.80	-101.87	
156	ZINC00696644	81.97	-110.51	
157	ZINC00696713	81.13	-111.73	
158	ZINC00696720	81.70	-113.63	
159	ZINC00696732	81.77	-109.35	

	<i>Code Zinc</i>	<i>Score GOLD</i>	<i>E_i (kcal/mol)</i>	<i>Famille</i>
160	ZINC00696745	81.38	-117.64	
161	ZINC00696748	81.65	-102.59	
162	ZINC00707679	81.59	-116.63	
163	ZINC00716539	81.86	-116.05	
164	ZINC00720896	81.23	-84.81	
165	ZINC00723711	81.46	-121.02	
166	ZINC00623727	80.28	-113.89	
167	ZINC00660609	80.53	-86.69	
168	ZINC00688007	80.08	-102.79	
169	ZINC00695249	80.69	176.15	
170	ZINC00695253	80.69	-125.55	
171	ZINC00695258	80.17	-120.42	
172	ZINC00695350	80.76	-110.64	
173	ZINC00695437	80.44	-127.67	
174	ZINC00696635	80.68	-106.95	
175	ZINC00696640	80.24	-97.05	
176	ZINC00696714	80.26	-103.63	
177	ZINC00696718	80.31	-103.34	
178	ZINC00696733	80.62	-107.65	
179	ZINC00623782	79.07	77.81	
180	ZINC00629606	79.05	456.77	
181	ZINC00663600	79.64	-112.61	
182	ZINC00693217	79.82	-113.55	
183	ZINC00695346	79.37	-109.18	
184	ZINC00695431	79.19	150.45	
185	ZINC00696636	79.06	-97.92	

	<i>Code Zinc</i>	<i>Score GOLD</i>	<i>E_i (kcal/mol)</i>	<i>Famille</i>
186	ZINC00696639	79.73	-108.68	
187	ZINC00696645	79.04	-89.29	
188	ZINC00696646	79.60	153.77	
189	ZINC00696712	79.23	-119.73	
190	ZINC00696727	79.09	-111.62	
191	ZINC00696731	79.46	-107.85	
192	ZINC00696749	79.97	-114.61	
193	ZINC00696786	79.35	-114.36	
194	ZINC00706602	79.99	-106.82	
195	ZINC00706603	79.86	-99.05	
196	ZINC00709255	79.22	-113.93	
197	ZINC00709657	78.25	-102.93	
198	ZINC00707617	78.14	-108.5	
199	ZINC00696787	78.19	-103.95	
200	ZINC00696747	78.92	-112.38	
201	ZINC00696729	78.84	-93.25	
202	ZINC00696728	78.77	-107.14	
203	ZINC00696650	78.17	-90.06	
204	ZINC00696643	78.20	-105.93	
205	ZINC00695453	78.69	-122.65	
206	ZINC00695450	78.65	-119.30	
207	ZINC00695382	78.50	165.53	
208	ZINC00695355	78.14	156.28	
209	ZINC00695344	78.47	-119.12	
210	ZINC00693292	78.18	-107.02	
211	ZINC00693288	78.89	-126.91	

	<i>Code Zinc</i>	<i>Score GOLD</i>	<i>E_i (kcal/mol)</i>	<i>Famille</i>
212	ZINC00681682	78.31	-101.5	
213	ZINC00681092	78.03	-118.85	
214	ZINC00679275	78.35	-92.43	
215	ZINC00660934	78.89	-103.35	
216	ZINC00640715	78.44	-120.00	
217	ZINC00633584	78.10	-94.70	
218	ZINC00629628	78.08	-108.90	
219	ZINC00623728	78.44	-120.06	
220	ZINC00723710	77.07	-119.19	
221	ZINC00718126	77.66	-95.12	
222	ZINC00707677	77.00	-110.52	
223	ZINC00696791	77.49	-95.77	
224	ZINC00696751	77.05	-88.92	
225	ZINC00696704	77.13	-107.56	
226	ZINC00695456	77.28	-122.43	
227	ZINC00695449	77.47	-117.85	
228	ZINC00695444	77.97	-112.66	
229	ZINC00695377	77.09	-90.93	
230	ZINC00695337	77.38	-118.84	
231	ZINC00695334	77.57	139.61	
232	ZINC00693293	77.86	-105.62	
233	ZINC00670036	77.55	-107.48	
234	ZINC00659424	77.56	-92.42	
235	ZINC00629624	77.26	-100.64	
236	ZINC00624295	77.31	-115.60	
237	ZINC00624293	76.83	-99.68	

	<i>Code Zinc</i>	<i>Score GOLD</i>	<i>E_i (kcal/mol)</i>	<i>Famille</i>
238	ZINC00626553	76.61	470.87	
239	ZINC00633604	76.14	-103.53	
240	ZINC00640691	76.53	-102.72	
241	ZINC00643922	76.55	-116.68	
242	ZINC00666050	76.04	-98.81	
243	ZINC00680543	76.61	-90.81	
244	ZINC00693291	76.14	-124,07	
245	ZINC00693294	76.34	-111.48	
246	ZINC00693806	76.11	-110.49	
247	ZINC00695261	76.12	-123.59	
248	ZINC00695264	76.19	-97.19	
249	ZINC00695351	76.17	-111.78	
250	ZINC00695357	76.31	-114.19	
251	ZINC00695384	76.51	-109.27	
252	ZINC00695433	76.17	130.08	
253	ZINC00695465	76.91	-116.42	
254	ZINC00696721	76.86	-115.68	
255	ZINC00696737	76.11	-100.78	
256	ZINC00696750	76.71	-105.85	
257	ZINC00705699	76.25	-114.32	
258	ZINC00709693	76.46	-109.27	
259	ZINC00720017	76.06	-94.34	
260	ZINC00629615	75.34	-115.82	
261	ZINC00629626	75.08	-104.11	
262	ZINC00643586	75.11	-112.00	
263	ZINC00680558	75.62	-97.64	

	<i>Code Zinc</i>	<i>Score GOLD</i>	<i>E_i (kcal/mol)</i>	<i>Famille</i>
264	ZINC00659434	75.94	-96.13	
265	ZINC00693220	75.44	-122.41	
266	ZINC00695271	75.42	-89.78	
267	ZINC00695336	75.82	142.34	
268	ZINC00695348	75.37	-108.47	
269	ZINC00695359	75.14	-108.47	
270	ZINC00695375	75.58	-117.55	
271	ZINC00695487	75.11	-112.59	
272	ZINC00696730	75.95	-105.22	
273	ZINC00696740	75.22	-106.24	
274	ZINC00706604	75.40	-80.30	
275	ZINC00713058	75.03	-108.64	
276	ZINC00720787	74.61	-106.51	
277	ZINC00726202	74.02	-96.16	
278	ZINC00712345	74.44	176.44	
279	ZINC00710740	74.91	-84.65	
280	ZINC00707540	74.30	-106.32	
281	ZINC00707253	74.65	-107.72	
282	ZINC00706248	74.00	-103.64	
283	ZINC00696792	74.60	-109.83	
284	ZINC00696752	74.79	156.90	
285	ZINC00696739	74.70	-108.99	
286	ZINC00696735	74.87	146.75	
287	ZINC00696705	74.70	-100.09	
288	ZINC00696689	74.66	172.81	
289	ZINC00696688	74.90	166.93	

	<i>Code Zinc</i>	<i>Score GOLD</i>	<i>E_i (kcal/mol)</i>	<i>Famille</i>
290	ZINC00696649	74.76	-100.92	
291	ZINC00696641	74.96	-90.37	
292	ZINC00695484	74.02	-110.89	
293	ZINC00695473	74.52	-119.26	
294	ZINC00695462	74.93	-124.80	
295	ZINC00695447	74.62	-117.91	
296	ZINC00695398	74.81	-112.72	
297	ZINC00695395	74.29	-98.96	
298	ZINC00695393	74.61	-93.44	
299	ZINC00695373	74.50	-103.05	
300	ZINC00695342	74.47	-106.53	
301	ZINC00693263	74.21	-112.19	
302	ZINC00693227	74.67	-113.9	
303	ZINC00678622	74.86	-82.69	
304	ZINC00659954	74.07	168.03	
305	ZINC00640572	74.57	-112.03	
306	ZINC00637814	74.31	-106.96	
307	ZINC00629641	74.36	-103.86	
308	ZINC00629610	74.49	-108.11	
309	ZINC00561439	74.55	-87.52	
310	ZINC00627313	74.82	-92.63	
311	ZINC00608088	73.64	-100.02	
312	ZINC00622467	73.93	-127.32	
313	ZINC00623733	73.44	-103.98	
314	ZINC00624031	73.93	-93.58	
315	ZINC00634571	73.80	-103.42	

	<i>Code Zinc</i>	<i>Score GOLD</i>	<i>E_i (kcal/mol)</i>	<i>Famille</i>
316	ZINC00634883	73.31	-104.38	
317	ZINC00643588	73.13	-91.28	
318	ZINC00644373	73.29	-117.27	
319	ZINC00660666	73.12	-95.63	
320	ZINC00664528	73.44	-111.05	
321	ZINC00695370	73.84	-101.72	
322	ZINC00695464	73.96	-113.02	
323	ZINC00695470	73.82	-111.07	
324	ZINC00696709	73.98	178.95	
325	ZINC00706428	73.74	-103.49	
326	ZINC00716703	73.98	-94.24	
327	ZINC00720573	73.34	-106.61	
328	ZINC00730097	73.01	-115.77	
329	ZINC00209328	72.77	-77.11	
330	ZINC00624249	72.76	-106.63	
331	ZINC00629635	72.59	-104.37	
332	ZINC00640642	72.38	-99.73	
333	ZINC00643589	72.90	-96.56	
334	ZINC00643590	72.74	-108.12	
335	ZINC00693260	72.98	-111.67	
336	ZINC00695367	72.34	-93.56	
337	ZINC00695386	72.76	-102.93	
338	ZINC00695440	72.80	-126.12	
339	ZINC00695475	72.92	-115.08	
340	ZINC00696703	72.84	-105.75	
341	ZINC00696738	72.51	-109.33	

	<i>Code Zinc</i>	<i>Score GOLD</i>	<i>E_i (kcal/mol)</i>	<i>Famille</i>
342	ZINC00696790	72.94	-91.92	
343	ZINC00707257	72.27	-116.33	
344	ZINC00707673	72.55	-106.00	
345	ZINC00718004	72.99	-115.95	
346	ZINC00719926	72.85	-103.36	
347	ZINC00720900	72.66	-97.20	
348	ZINC00665961	84.60	-110.68	
349	ZINC00704892	81.32	-107.78	
350	ZINC00633493	80.96	-114.81	
351	ZINC00644256	75.34	447.89	
352	ZINC00651175	74.02	-106.66	
353	ZINC00669359	74.85	-106.46	
354	ZINC00629619	73.69	-109.95	
355	ZINC00633526	73.59	-95.24	
356	ZINC00704302	73.58	-102.06	
357	ZINC00633235	73.13	-109.22	
358	ZINC00669357	73.13	-99.56	
359	ZINC00624856	73.02	-93.56	
360	ZINC00622996	72.33	-101.5	
361	ZINC00676192	72.93	-94.47	
362	ZINC00268338	73.61	-64.33	Sulfonamide
363	ZINC00308628	72.61	-87.53	

ANNEXE 3

Nous donnons ci-dessous les résultats issus du programme de comparaison entre les classements de scores GOLD et d'énergie d'interaction par Discover. Le pourcentage est également indiqué

EI no : numéro de classement de la molécule par rapport à son énergie d'interaction

GOLD no : numéro de classement de la molécule par rapport à son score GOLD

pourcentage pour les		10 premiers:	0.00
EI no.	5 GOLD no.	17 molécule	ZINC00706035
EI no.	8 GOLD no.	20 molécule	ZINC00634887
pourcentage pour les		20 premiers:	10.00
EI no.	5 GOLD no.	17 molécule	ZINC00706035
EI no.	8 GOLD no.	20 molécule	ZINC00634887
pourcentage pour les		30 premiers:	6.67
EI no.	5 GOLD no.	17 molécule	ZINC00706035
EI no.	8 GOLD no.	20 molécule	ZINC00634887
pourcentage pour les		40 premiers:	5.00
EI no.	5 GOLD no.	17 molécule	ZINC00706035
EI no.	8 GOLD no.	20 molécule	ZINC00634887
EI no.	35 GOLD no.	50 molécule	ZINC00707679
EI no.	38 GOLD no.	44 molécule	ZINC00716539
EI no.	45 GOLD no.	14 molécule	ZINC00706600
EI no.	48 GOLD no.	27 molécule	ZINC00696719
pourcentage pour les		50 premiers:	12.00
EI no.	2 GOLD no.	77 molécule	ZINC00695437
EI no.	5 GOLD no.	17 molécule	ZINC00706035
EI no.	7 GOLD no.	71 molécule	ZINC00695253
EI no.	8 GOLD no.	20 molécule	ZINC00634887
EI no.	16 GOLD no.	54 molécule	ZINC00723711
EI no.	17 GOLD no.	82 molécule	ZINC00695258
EI no.	21 GOLD no.	62 molécule	ZINC00681091
EI no.	32 GOLD no.	55 molécule	ZINC00696745
EI no.	35 GOLD no.	50 molécule	ZINC00707679
EI no.	38 GOLD no.	44 molécule	ZINC00716539
EI no.	45 GOLD no.	14 molécule	ZINC00706600
EI no.	47 GOLD no.	66 molécule	ZINC00633493
EI no.	48 GOLD no.	27 molécule	ZINC00696719
EI no.	49 GOLD no.	87 molécule	ZINC00696749
EI no.	50 GOLD no.	100 molécule	ZINC03055938
EI no.	51 GOLD no.	99 molécule	ZINC00696786
EI no.	56 GOLD no.	79 molécule	ZINC00623727
EI no.	57 GOLD no.	48 molécule	ZINC00696720
EI no.	58 GOLD no.	89 molécule	ZINC00693217
EI no.	61 GOLD no.	24 molécule	ZINC00696726
EI no.	64 GOLD no.	92 molécule	ZINC00663600
EI no.	68 GOLD no.	18 molécule	ZINC00696717
EI no.	75 GOLD no.	65 molécule	ZINC00696713
EI no.	79 GOLD no.	22 molécule	ZINC00696724
EI no.	84 GOLD no.	19 molécule	ZINC00665961
EI no.	85 GOLD no.	69 molécule	ZINC00695350
EI no.	87 GOLD no.	42 molécule	ZINC00696644
EI no.	91 GOLD no.	46 molécule	ZINC00696637
EI no.	93 GOLD no.	47 molécule	ZINC00696732
EI no.	95 GOLD no.	13 molécule	ZINC00696725
EI no.	100 GOLD no.	73 molécule	ZINC03055937
EI no.	100 GOLD no.	74 molécule	ZINC03055937
pourcentage pour les		100 premiers:	32.00

EI no.	2	GOLD no.	77	molecule	ZINC00695437
EI no.	4	GOLD no.	180	molecule	ZINC03051422
EI no.	4	GOLD no.	181	molecule	ZINC03051422
EI no.	5	GOLD no.	17	molecule	ZINC00706035
EI no.	6	GOLD no.	118	molecule	ZINC00693288
EI no.	7	GOLD no.	71	molecule	ZINC00695253
EI no.	8	GOLD no.	20	molecule	ZINC00634887
EI no.	10	GOLD no.	196	molecule	ZINC00693291
EI no.	11	GOLD no.	197	molecule	ZINC00695261
EI no.	13	GOLD no.	121	molecule	ZINC00695453
EI no.	14	GOLD no.	164	molecule	ZINC00695456
EI no.	16	GOLD no.	54	molecule	ZINC00723711
EI no.	17	GOLD no.	82	molecule	ZINC00695258
EI no.	18	GOLD no.	128	molecule	ZINC00623728
EI no.	19	GOLD no.	129	molecule	ZINC00640715
EI no.	20	GOLD no.	104	molecule	ZINC00696712
EI no.	21	GOLD no.	62	molecule	ZINC00681091
EI no.	22	GOLD no.	123	molecule	ZINC00695450
EI no.	24	GOLD no.	168	molecule	ZINC00723710
EI no.	25	GOLD no.	127	molecule	ZINC00695344
EI no.	26	GOLD no.	144	molecule	ZINC00681092
EI no.	27	GOLD no.	160	molecule	ZINC00695337
EI no.	31	GOLD no.	159	molecule	ZINC00695449
EI no.	32	GOLD no.	55	molecule	ZINC00696745
EI no.	34	GOLD no.	185	molecule	ZINC00643922
EI no.	35	GOLD no.	50	molecule	ZINC00707679
EI no.	36	GOLD no.	174	molecule	ZINC00695465
EI no.	38	GOLD no.	44	molecule	ZINC00716539
EI no.	42	GOLD no.	176	molecule	ZINC00696721
EI no.	43	GOLD no.	162	molecule	ZINC00624295
EI no.	45	GOLD no.	14	molecule	ZINC00706600
EI no.	47	GOLD no.	66	molecule	ZINC00633493
EI no.	48	GOLD no.	27	molecule	ZINC00696719
EI no.	49	GOLD no.	87	molecule	ZINC00696749
EI no.	50	GOLD no.	100	molecule	ZINC03055938
EI no.	50	GOLD no.	101	molecule	ZINC03055938
EI no.	51	GOLD no.	99	molecule	ZINC00696786
EI no.	52	GOLD no.	191	molecule	ZINC00705699
EI no.	53	GOLD no.	190	molecule	ZINC00695357
EI no.	54	GOLD no.	105	molecule	ZINC00709255
EI no.	56	GOLD no.	79	molecule	ZINC00623727
EI no.	57	GOLD no.	48	molecule	ZINC00696720
EI no.	58	GOLD no.	89	molecule	ZINC00693217
EI no.	59	GOLD no.	150	molecule	ZINC03055934
EI no.	59	GOLD no.	151	molecule	ZINC03055934
EI no.	61	GOLD no.	24	molecule	ZINC00696726
EI no.	63	GOLD no.	145	molecule	ZINC00695444
EI no.	64	GOLD no.	92	molecule	ZINC00663600
EI no.	66	GOLD no.	115	molecule	ZINC00696747
EI no.	68	GOLD no.	18	molecule	ZINC00696717
EI no.	71	GOLD no.	108	molecule	ZINC03059969
EI no.	71	GOLD no.	109	molecule	ZINC03059969
EI no.	73	GOLD no.	193	molecule	ZINC00695351
EI no.	75	GOLD no.	65	molecule	ZINC00696713
EI no.	77	GOLD no.	110	molecule	ZINC00696727
EI no.	78	GOLD no.	189	molecule	ZINC00693294
EI no.	79	GOLD no.	22	molecule	ZINC00696724
EI no.	82	GOLD no.	124	molecule	ZINC03029178
EI no.	82	GOLD no.	125	molecule	ZINC03029178
EI no.	84	GOLD no.	19	molecule	ZINC00665961
EI no.	85	GOLD no.	69	molecule	ZINC00695350

EI no.	86	GOLD no.	170	molecule	ZINC00707677
EI no.	87	GOLD no.	42	molecule	ZINC00696644
EI no.	88	GOLD no.	198	molecule	ZINC00693806
EI no.	91	GOLD no.	46	molecule	ZINC00696637
EI no.	93	GOLD no.	47	molecule	ZINC00696732
EI no.	95	GOLD no.	13	molecule	ZINC00696725
EI no.	96	GOLD no.	187	molecule	ZINC00695384
EI no.	97	GOLD no.	188	molecule	ZINC00709693
EI no.	100	GOLD no.	73	molecule	ZINC03055937
EI no.	100	GOLD no.	74	molecule	ZINC03055937
EI no.	101	GOLD no.	98	molecule	ZINC00695346
EI no.	103	GOLD no.	142	molecule	ZINC00629628
EI no.	104	GOLD no.	91	molecule	ZINC00696639
EI no.	106	GOLD no.	138	molecule	ZINC00707617
EI no.	112	GOLD no.	96	molecule	ZINC00696731
EI no.	113	GOLD no.	61	molecule	ZINC00704892
EI no.	115	GOLD no.	51	molecule	ZINC00633586
EI no.	117	GOLD no.	75	molecule	ZINC00696733
EI no.	118	GOLD no.	166	molecule	ZINC00696704
EI no.	119	GOLD no.	154	molecule	ZINC00670036
EI no.	120	GOLD no.	120	molecule	ZINC00696728
EI no.	121	GOLD no.	135	molecule	ZINC00693292
EI no.	123	GOLD no.	72	molecule	ZINC00696635
EI no.	124	GOLD no.	86	molecule	ZINC00706602
EI no.	131	GOLD no.	25	molecule	ZINC00696746
EI no.	136	GOLD no.	133	molecule	ZINC00696643
EI no.	137	GOLD no.	179	molecule	ZINC00696750
EI no.	139	GOLD no.	146	molecule	ZINC00693293
EI no.	141	GOLD no.	57	molecule	ZINC00696638
EI no.	142	GOLD no.	15	molecule	ZINC00705479
EI no.	145	GOLD no.	36	molecule	ZINC03055753
EI no.	145	GOLD no.	37	molecule	ZINC03055753
EI no.	148	GOLD no.	134	molecule	ZINC00696787
EI no.	150	GOLD no.	102	molecule	ZINC03055755
EI no.	150	GOLD no.	103	molecule	ZINC03055755
EI no.	152	GOLD no.	80	molecule	ZINC00696714
EI no.	153	GOLD no.	195	molecule	ZINC00633604
EI no.	157	GOLD no.	117	molecule	ZINC00660934
EI no.	158	GOLD no.	78	molecule	ZINC00696718
EI no.	161	GOLD no.	132	molecule	ZINC00709657
EI no.	162	GOLD no.	84	molecule	ZINC00688007
EI no.	163	GOLD no.	186	molecule	ZINC00640691
EI no.	164	GOLD no.	49	molecule	ZINC00696748
EI no.	165	GOLD no.	29	molecule	ZINC00693223
EI no.	167	GOLD no.	45	molecule	ZINC00696642
EI no.	169	GOLD no.	131	molecule	ZINC00681682
EI no.	172	GOLD no.	199	molecule	ZINC00696737
EI no.	173	GOLD no.	165	molecule	ZINC00629624
EI no.	181	GOLD no.	35	molecule	ZINC00705700
EI no.	182	GOLD no.	88	molecule	ZINC00706603
EI no.	185	GOLD no.	9	molecule	ZINC03058581
EI no.	185	GOLD no.	10	molecule	ZINC03058581
EI no.	188	GOLD no.	58	molecule	ZINC03055914
EI no.	188	GOLD no.	59	molecule	ZINC03055914
EI no.	189	GOLD no.	112	molecule	ZINC00696636
EI no.	191	GOLD no.	81	molecule	ZINC00696640
EI no.	195	GOLD no.	192	molecule	ZINC00695264
EI no.	196	GOLD no.	161	molecule	ZINC02636876
pourcentage pour les 200 premiers: 59.50					
EI no.	1	GOLD no.	210	molecule	ZINC00695375
EI no.	2	GOLD no.	77	molecule	ZINC00695437

EI no.	4	GOLD no.	180	molecule	ZINC03051422
EI no.	4	GOLD no.	181	molecule	ZINC03051422
EI no.	5	GOLD no.	17	molecule	ZINC00706035
EI no.	6	GOLD no.	118	molecule	ZINC00693288
EI no.	7	GOLD no.	71	molecule	ZINC00695253
EI no.	8	GOLD no.	20	molecule	ZINC00634887
EI no.	9	GOLD no.	235	molecule	ZINC00695462
EI no.	10	GOLD no.	196	molecule	ZINC00693291
EI no.	11	GOLD no.	197	molecule	ZINC00695261
EI no.	13	GOLD no.	121	molecule	ZINC00695453
EI no.	14	GOLD no.	164	molecule	ZINC00695456
EI no.	15	GOLD no.	212	molecule	ZINC00693220
EI no.	16	GOLD no.	54	molecule	ZINC00723711
EI no.	17	GOLD no.	82	molecule	ZINC00695258
EI no.	18	GOLD no.	128	molecule	ZINC00623728
EI no.	19	GOLD no.	129	molecule	ZINC00640715
EI no.	20	GOLD no.	104	molecule	ZINC00696712
EI no.	21	GOLD no.	62	molecule	ZINC00681091
EI no.	22	GOLD no.	123	molecule	ZINC00695450
EI no.	23	GOLD no.	267	molecule	ZINC00695473
EI no.	24	GOLD no.	168	molecule	ZINC00723710
EI no.	25	GOLD no.	127	molecule	ZINC00695344
EI no.	26	GOLD no.	144	molecule	ZINC00681092
EI no.	27	GOLD no.	160	molecule	ZINC00695337
EI no.	29	GOLD no.	226	molecule	ZINC03055752
EI no.	29	GOLD no.	227	molecule	ZINC03055752
EI no.	30	GOLD no.	258	molecule	ZINC00695447
EI no.	31	GOLD no.	159	molecule	ZINC00695449
EI no.	32	GOLD no.	55	molecule	ZINC00696745
EI no.	34	GOLD no.	185	molecule	ZINC00643922
EI no.	35	GOLD no.	50	molecule	ZINC00707679
EI no.	36	GOLD no.	174	molecule	ZINC00695465
EI no.	38	GOLD no.	44	molecule	ZINC00716539
EI no.	40	GOLD no.	216	molecule	ZINC00629615
EI no.	42	GOLD no.	176	molecule	ZINC00696721
EI no.	43	GOLD no.	162	molecule	ZINC00624295
EI no.	45	GOLD no.	14	molecule	ZINC00706600
EI no.	47	GOLD no.	66	molecule	ZINC00633493
EI no.	48	GOLD no.	27	molecule	ZINC00696719
EI no.	49	GOLD no.	87	molecule	ZINC00696749
EI no.	50	GOLD no.	100	molecule	ZINC03055938
EI no.	50	GOLD no.	101	molecule	ZINC03055938
EI no.	51	GOLD no.	99	molecule	ZINC00696786
EI no.	52	GOLD no.	191	molecule	ZINC00705699
EI no.	53	GOLD no.	190	molecule	ZINC00695357
EI no.	54	GOLD no.	105	molecule	ZINC00709255
EI no.	55	GOLD no.	252	molecule	ZINC00693227
EI no.	56	GOLD no.	79	molecule	ZINC00623727
EI no.	57	GOLD no.	48	molecule	ZINC00696720
EI no.	58	GOLD no.	89	molecule	ZINC00693217
EI no.	59	GOLD no.	150	molecule	ZINC03055934
EI no.	59	GOLD no.	151	molecule	ZINC03055934
EI no.	61	GOLD no.	24	molecule	ZINC00696726
EI no.	62	GOLD no.	242	molecule	ZINC00695398
EI no.	63	GOLD no.	145	molecule	ZINC00695444
EI no.	64	GOLD no.	92	molecule	ZINC00663600
EI no.	65	GOLD no.	224	molecule	ZINC00695487
EI no.	66	GOLD no.	115	molecule	ZINC00696747
EI no.	67	GOLD no.	285	molecule	ZINC00693263
EI no.	68	GOLD no.	18	molecule	ZINC00696717
EI no.	69	GOLD no.	263	molecule	ZINC00640572

EI no.	70	GOLD no.	223	molecule	ZINC00643586
EI no.	71	GOLD no.	108	molecule	ZINC03059969
EI no.	71	GOLD no.	109	molecule	ZINC03059969
EI no.	73	GOLD no.	193	molecule	ZINC00695351
EI no.	74	GOLD no.	244	molecule	ZINC03029189
EI no.	74	GOLD no.	245	molecule	ZINC03029189
EI no.	75	GOLD no.	65	molecule	ZINC00696713
EI no.	77	GOLD no.	110	molecule	ZINC00696727
EI no.	78	GOLD no.	189	molecule	ZINC00693294
EI no.	79	GOLD no.	22	molecule	ZINC00696724
EI no.	82	GOLD no.	124	molecule	ZINC03029178
EI no.	82	GOLD no.	125	molecule	ZINC03029178
EI no.	83	GOLD no.	296	molecule	ZINC00695484
EI no.	84	GOLD no.	19	molecule	ZINC00665961
EI no.	85	GOLD no.	69	molecule	ZINC00695350
EI no.	86	GOLD no.	170	molecule	ZINC00707677
EI no.	87	GOLD no.	42	molecule	ZINC00696644
EI no.	88	GOLD no.	198	molecule	ZINC00693806
EI no.	91	GOLD no.	46	molecule	ZINC00696637
EI no.	92	GOLD no.	262	molecule	ZINC00696792
EI no.	93	GOLD no.	47	molecule	ZINC00696732
EI no.	95	GOLD no.	13	molecule	ZINC00696725
EI no.	96	GOLD no.	187	molecule	ZINC00695384
EI no.	97	GOLD no.	188	molecule	ZINC00709693
EI no.	99	GOLD no.	286	molecule	ZINC03055806
EI no.	99	GOLD no.	287	molecule	ZINC03055806
EI no.	100	GOLD no.	73	molecule	ZINC03055937
EI no.	100	GOLD no.	74	molecule	ZINC03055937
EI no.	101	GOLD no.	98	molecule	ZINC00695346
EI no.	102	GOLD no.	250	molecule	ZINC00696739
EI no.	103	GOLD no.	142	molecule	ZINC00629628
EI no.	104	GOLD no.	91	molecule	ZINC00696639
EI no.	105	GOLD no.	230	molecule	ZINC00713058
EI no.	106	GOLD no.	138	molecule	ZINC00707617
EI no.	107	GOLD no.	215	molecule	ZINC00695348
EI no.	108	GOLD no.	221	molecule	ZINC00695359
EI no.	111	GOLD no.	269	molecule	ZINC00629610
EI no.	112	GOLD no.	96	molecule	ZINC00696731
EI no.	113	GOLD no.	61	molecule	ZINC00704892
EI no.	115	GOLD no.	51	molecule	ZINC00633586
EI no.	116	GOLD no.	257	molecule	ZINC00707253
EI no.	117	GOLD no.	75	molecule	ZINC00696733
EI no.	118	GOLD no.	166	molecule	ZINC00696704
EI no.	119	GOLD no.	154	molecule	ZINC00670036
EI no.	120	GOLD no.	120	molecule	ZINC00696728
EI no.	121	GOLD no.	135	molecule	ZINC00693292
EI no.	122	GOLD no.	280	molecule	ZINC00637814
EI no.	123	GOLD no.	72	molecule	ZINC00696635
EI no.	124	GOLD no.	86	molecule	ZINC00706602
EI no.	125	GOLD no.	298	molecule	ZINC00651175
EI no.	128	GOLD no.	232	molecule	ZINC03029182
EI no.	128	GOLD no.	233	molecule	ZINC03029182
EI no.	129	GOLD no.	271	molecule	ZINC00695342
EI no.	130	GOLD no.	260	molecule	ZINC00720787
EI no.	131	GOLD no.	25	molecule	ZINC00696746
EI no.	132	GOLD no.	240	molecule	ZINC00669359
EI no.	133	GOLD no.	281	molecule	ZINC00707540
EI no.	134	GOLD no.	220	molecule	ZINC00696740
EI no.	136	GOLD no.	133	molecule	ZINC00696643
EI no.	137	GOLD no.	179	molecule	ZINC00696750
EI no.	139	GOLD no.	146	molecule	ZINC00693293

EI no.	140	GOLD no.	205	molecule	ZINC00696730
EI no.	141	GOLD no.	57	molecule	ZINC00696638
EI no.	142	GOLD no.	15	molecule	ZINC00705479
EI no.	145	GOLD no.	36	molecule	ZINC03055753
EI no.	145	GOLD no.	37	molecule	ZINC03055753
EI no.	146	GOLD no.	225	molecule	ZINC00629626
EI no.	148	GOLD no.	134	molecule	ZINC00696787
EI no.	149	GOLD no.	279	molecule	ZINC00629641
EI no.	150	GOLD no.	102	molecule	ZINC03055755
EI no.	150	GOLD no.	103	molecule	ZINC03055755
EI no.	151	GOLD no.	300	molecule	ZINC00706248
EI no.	152	GOLD no.	80	molecule	ZINC00696714
EI no.	153	GOLD no.	195	molecule	ZINC00633604
EI no.	157	GOLD no.	117	molecule	ZINC00660934
EI no.	158	GOLD no.	78	molecule	ZINC00696718
EI no.	159	GOLD no.	268	molecule	ZINC00695373
EI no.	161	GOLD no.	132	molecule	ZINC00709657
EI no.	162	GOLD no.	84	molecule	ZINC00688007
EI no.	163	GOLD no.	186	molecule	ZINC00640691
EI no.	164	GOLD no.	49	molecule	ZINC00696748
EI no.	165	GOLD no.	29	molecule	ZINC00693223
EI no.	167	GOLD no.	45	molecule	ZINC00696642
EI no.	169	GOLD no.	131	molecule	ZINC00681682
EI no.	171	GOLD no.	246	molecule	ZINC00696649
EI no.	172	GOLD no.	199	molecule	ZINC00696737
EI no.	173	GOLD no.	165	molecule	ZINC00629624
EI no.	175	GOLD no.	249	molecule	ZINC00696705
EI no.	177	GOLD no.	293	molecule	ZINC03055917
EI no.	177	GOLD no.	294	molecule	ZINC03055917
EI no.	181	GOLD no.	35	molecule	ZINC00705700
EI no.	182	GOLD no.	88	molecule	ZINC00706603
EI no.	183	GOLD no.	282	molecule	ZINC00695395
EI no.	184	GOLD no.	202	molecule	ZINC00666050
EI no.	185	GOLD no.	9	molecule	ZINC03058581
EI no.	185	GOLD no.	10	molecule	ZINC03058581
EI no.	187	GOLD no.	295	molecule	ZINC00835572
EI no.	188	GOLD no.	58	molecule	ZINC03055914
EI no.	188	GOLD no.	59	molecule	ZINC03055914
EI no.	189	GOLD no.	112	molecule	ZINC00696636
EI no.	191	GOLD no.	81	molecule	ZINC00696640
EI no.	193	GOLD no.	254	molecule	ZINC03055713
EI no.	193	GOLD no.	255	molecule	ZINC03055713
EI no.	195	GOLD no.	192	molecule	ZINC00695264
EI no.	196	GOLD no.	161	molecule	ZINC02636876
EI no.	198	GOLD no.	264	molecule	ZINC03055951
EI no.	198	GOLD no.	265	molecule	ZINC03055951
EI no.	199	GOLD no.	297	molecule	ZINC00726202
EI no.	200	GOLD no.	206	molecule	ZINC00659434
EI no.	201	GOLD no.	218	molecule	ZINC03055946
EI no.	201	GOLD no.	219	molecule	ZINC03055946
EI no.	202	GOLD no.	158	molecule	ZINC00696791
EI no.	204	GOLD no.	256	molecule	ZINC00660280
EI no.	205	GOLD no.	211	molecule	ZINC00715656
EI no.	208	GOLD no.	149	molecule	ZINC00718126
EI no.	209	GOLD no.	140	molecule	ZINC00633584
EI no.	212	GOLD no.	201	molecule	ZINC00720017
EI no.	215	GOLD no.	52	molecule	ZINC01005807
EI no.	219	GOLD no.	259	molecule	ZINC00695393
EI no.	220	GOLD no.	119	molecule	ZINC00696729
EI no.	223	GOLD no.	241	molecule	ZINC00627313
EI no.	224	GOLD no.	130	molecule	ZINC00679275

EI no.	225	GOLD no.	153	molecule	ZINC00659424
EI no.	227	GOLD no.	97	molecule	ZINC02881744
EI no.	230	GOLD no.	289	molecule	ZINC00714120
EI no.	231	GOLD no.	167	molecule	ZINC00695377
EI no.	232	GOLD no.	184	molecule	ZINC00680543
EI no.	233	GOLD no.	234	molecule	ZINC00696641
EI no.	234	GOLD no.	136	molecule	ZINC00696650
EI no.	235	GOLD no.	213	molecule	ZINC00695271
EI no.	236	GOLD no.	299	molecule	ZINC00796429
EI no.	237	GOLD no.	114	molecule	ZINC00696645
EI no.	238	GOLD no.	169	molecule	ZINC00696751
EI no.	239	GOLD no.	247	molecule	ZINC00879924
EI no.	241	GOLD no.	290	molecule	ZINC01016329
EI no.	243	GOLD no.	266	molecule	ZINC00561439
EI no.	244	GOLD no.	222	molecule	ZINC00754633
EI no.	246	GOLD no.	76	molecule	ZINC00660609
EI no.	247	GOLD no.	283	molecule	ZINC00656683
EI no.	250	GOLD no.	64	molecule	ZINC00720896
EI no.	251	GOLD no.	236	molecule	ZINC00710740
EI no.	253	GOLD no.	275	molecule	ZINC01039093
EI no.	253	GOLD no.	276	molecule	ZINC01039093
EI no.	255	GOLD no.	182	molecule	ZINC00655040
EI no.	256	GOLD no.	239	molecule	ZINC00678622
EI no.	257	GOLD no.	277	molecule	ZINC00866604
EI no.	261	GOLD no.	148	molecule	ZINC00969715
EI no.	263	GOLD no.	214	molecule	ZINC00706604
EI no.	264	GOLD no.	284	molecule	ZINC01112301
EI no.	267	GOLD no.	274	molecule	ZINC00646006
EI no.	271	GOLD no.	147	molecule	ZINC00930618
EI no.	277	GOLD no.	163	molecule	ZINC00781853
EI no.	278	GOLD no.	228	molecule	ZINC00922318
EI no.	280	GOLD no.	111	molecule	ZINC00623782
EI no.	281	GOLD no.	194	molecule	ZINC00695433
EI no.	282	GOLD no.	152	molecule	ZINC00695334
EI no.	283	GOLD no.	207	molecule	ZINC00695336
EI no.	284	GOLD no.	238	molecule	ZINC00696735
EI no.	285	GOLD no.	40	molecule	ZINC00695461
EI no.	286	GOLD no.	38	molecule	ZINC00696716
EI no.	287	GOLD no.	106	molecule	ZINC00695431
EI no.	288	GOLD no.	23	molecule	ZINC00696633
EI no.	289	GOLD no.	94	molecule	ZINC00696646
EI no.	290	GOLD no.	39	molecule	ZINC00695269
EI no.	291	GOLD no.	137	molecule	ZINC00695355
EI no.	292	GOLD no.	243	molecule	ZINC00696752
EI no.	293	GOLD no.	12	molecule	ZINC00696785
EI no.	294	GOLD no.	21	molecule	ZINC00695353
EI no.	295	GOLD no.	16	molecule	ZINC00695267
EI no.	296	GOLD no.	34	molecule	ZINC00696734
EI no.	297	GOLD no.	2	molecule	ZINC00696742
EI no.	298	GOLD no.	6	molecule	ZINC00696647
EI no.	299	GOLD no.	126	molecule	ZINC00695382
EI no.	300	GOLD no.	4	molecule	ZINC00695380

pourcentage pour les 300 premiers: 79.00