

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE

وزارة التعليم العالي والبحث العلمي

BADJI MOKHTAR - ANNABA UNIVERSITY
UNIVERSITE BADJI MOKHTAR- ANNABA



جامعة باجي مختار - عنابة
Année 2018

Faculté des Sciences
Département de Chimie

THESE

Présentée en vue de l'obtention du diplôme de Doctorat en Sciences

Option : Chimie Organique et Bioorganique

THEME

**MODELISATION DES INTERACTIONS PROTEINE-PETITES
MOLECULES: ETUDE DE LA RELATION STRUCTURE –
FONCTION DANS LE CAS DES LIPASES**

Présentée par : Samah BOUCHAGRA

Devant le jury :

Président	Nour-Eddine AOUF	Professeur	Université d'Annaba
Directrice de thèse	Fatiha BENAMIA	Maître de Conférences/A	Université d'Annaba
Examineurs	Malika BERREDJEM	Professeur	Université d'Annaba
	Touhami LANEZ	Professeur	Université d'EL Oued
	Zine REGAINIA	Professeur	Université de Souk Ahras
	Abdelhak GHEID	Professeur	Université de Souk Ahras
Membre Invité	Zeineddine DJEGHABA	Professeur	Université d'Annaba

Dédicace

A mes parents,

A mon marié et mon fils

A mes sœurs,

A mes frères,

A tous ceux qui sont chers

Remerciements

Ce travail de recherche qui fait l'objet de cette thèse a été réalisé au sein du laboratoire de Chimie Organique Appliquée (LCOA), Groupe « Bioconversion et Synthèse Organique », Université Badji Mokhtar-Annaba.

Tout d'abord, je tiens à exprimer ma plus profonde reconnaissance à Monsieur Zeineddine DJEGHABA, Professeur à l'Université Badji Mokhtar-Annaba, pour m'avoir donné la possibilité de rejoindre son groupe de recherche, pour son soutien et ses conseils judicieux ainsi pour sa grande disponibilité et ses qualités humaines.

Je tiens à remercier très sincèrement ma directrice de thèse Madame Fatiha BENAMIA, Maître de conférences « A » à l'Université Badji Mokhtar-Annaba, pour avoir dirigé mes travaux et pour toute l'aide qu'elle m'a apportée au cours de ce travail. Son intérêt, sa disponibilité, son suivi quotidien de mes travaux et ses précieux conseils m'ont été d'un grand profit. Je la remercie infiniment.

Mes meilleurs remerciements s'adressent également à Monsieur Nour-Eddine AOUF, Professeur à l'Université Badji Mokhtar-Annaba, pour m'avoir fait l'honneur de présider le jury de cette thèse.

J'exprime ma gratitude à Madame Malika BERREDJEM, Professeur à l'Université Badji Mokhtar-Annaba, pour avoir accepté de juger ce travail. Je suis honoré de la compter parmi les membres de ce jury.

Je remercie tout particulièrement Monsieur Touhami LANEZ, Professeur à l'Université Hamma Lakhder-El Oued, pour avoir accepté de faire partie du jury de cette thèse.

Je remercie vivement Monsieur Zine REGAINIA, Professeur à l'Université de Souk Ahras, qui m'a fait l'honneur d'examiner ce travail et de faire partie du jury de cette thèse.

J'adresse mes respectueux remerciements à Monsieur Abdekhak GHEID, Professeur à l'Université de Souk Ahras, pour m'avoir fait l'honneur de participer au jury de cette thèse.

Un grand merci à tous les membres du laboratoire de Chimie Organique Appliquée, pour leur sympathie et leur soutien.

Je remercie également mes chères amies Hafida, Khaoula, Asma, pour leur grand cœur et leur sympathie.

Enfin, il est très difficile pour moi de parler de toi, khalifa mon marie. Les paroles ne suffisent pas pour exprimer mes remerciements. Merci énormément pour ta patience, ton soutien moral, tes conseils précieux.

En conclusion, je dédie ce travail à mon défunt grand père qui m'a légué tant de valeurs. Ce travail est pour toi. Tu resteras à jamais un modèle pour moi. Merci.

Résumé

Le processus de docking est l'un des premières étapes utilisées dans la conception de médicaments. Il consiste à faire interagir un ligand avec un récepteur, généralement de nature protéique. En conséquent, le plus grand avantage des méthodes de docking protéine-ligand est qu'ils peuvent proposer des hypothèses structurales sur la façon dont une petite molécule peut interagir avec sa cible c'est-à-dire la macromolécule. Pour cela nous avons fait appel à cette technique, qui a été assistée par le programme Molegro Virtuel Docker (MVD), pour comprendre le mode d'interaction des enzymes lipolytiques vis-à-vis du (-)-Epigallocatechin-3-gallate (EGCG). Ce dernier est un composé naturel présent dans le thé vert, il présente une activité anti-obésité en diminuant la digestion des lipides alimentaires par le mécanisme d'inhibition des lipases pancréatique et gastrique.

Le présent travail décrit, dans un premier temps, les modes d'interactions possibles de l'Epigallocatechin-3-gallate (EGCG) vis-à-vis de la lipase pancréatique. Pour cela nous avons choisi la LPP et LPH. L'analyse des modes d'interaction de l'EGCG révèle une affinité favorable avec les résidus de la cavité 3 située au voisinage du site actif des deux lipases testées. Les résultats obtenus ont pu élucider les mécanismes d'interaction entre l'EGCG présenté comme un inhibiteur non compétitif expérimentalement, et les deux lipases testées (LPP et LPH). Des liaisons hydrogène importantes ont été formées avec les résidus His 263 (LPP) et His 264 (LPH). Ces dernières font partie de la triade catalytique de ces enzymes.

Une étude complémentaire a été effectuée en testant une lipase gastrique du chien (LGC) qui présente un bon modèle de la lipase gastrique humaine (LGH). Les résultats obtenus montrent une grande affinité avec les résidus de la cavité principale de cette lipase donnant une énergie qui est égale à $-150,27 \text{ kcal mol}^{-1}$. Cette stabilité exprimée par une énergie très faible, résulte de la mise en place des cinq liaisons hydrogène et des plusieurs interactions hydrophobe vis-à-vis des résidus de la cavité principale de la protéine. Une de ces liaisons hydrogène estimée la plus importante, a été formée avec le résidu Ser 153. Cette affinité théorique vers les résidus catalytique prédit un caractère compétitif important d'inhibition pour la lipase gastrique.

Mots clés: Docking Moléculaire, Lipase Pancréatique, Lipases Gastrique, inhibition, (-) - Epigallocatechin-3-gallate (EGCG), Obésité.

Abstract

Molecular docking is one of the first steps used in drug design. This technical modeling is used to study the interaction between ligand receptor, where this last generally is a natural protein. Therefore, the greatest advantage of protein-ligand docking methods is that they can provide structural assumptions about how a small molecule can interact with its target (macromolecule). For this purpose we used this technique, which was assisted by the Molegro Virtual Docker program (MVD), to understand the mode of interaction of lipolytic enzymes toward (-)-Epigallocatechin-3-gallate (EGCG). This flavonoids derivative which is a natural compound present in green tea has an anti-obesity activity by decreasing the digestion of dietary lipids by an inhibition mechanism of pancreatic and gastric lipases.

The present work describes, in a first step, the possible modes of interaction of Epigallocatechin-3-gallate (EGCG) toward a pancreatic lipase. For this study we chose the PPL and PHL. The analysis of the modes of interaction of the EGCG reveals a favorable affinity with the residues of the cavity number 3 located in the active site of the two lipases tested. The obtained results have highlighted the interaction mechanisms between the EGCG (experimentally non-competitive inhibitor) and the two pancreatic lipases tested (LPP and LPH). Significant hydrogen bonds were formed with His 263 (LPP) and His 264 (LPH) residues, which are part of the catalytic triad of these enzymes.

A further study was performed testing a dog gastric lipase (DGL) which is a good model of human gastric lipase (HGL). This great affinity with the residues of the main cavity was explained by the low energy value ($-150.27 \text{ kcal mol}^{-1}$). This stability results from the introduction of the five hydrogen bonds and the several hydrophobic interactions with the residues of the main cavity of the protein. One of these most important estimated hydrogen bonds was formed with Ser 153. This theoretical affinity towards catalytic residues predicts a significant competitive character of inhibition for gastric lipase.

Key-words: Molecular docking, Pancreatic lipase, Gastric lipase, (-) - Epigallocatechin-3-gallate (EGCG), Inhibition, Obesity.

تمثل تقنية الالتحام أحد الخطوات الأولى في تصميم الأدوية حيث تعتمد على تفاعل يحدث بين مركب و مستقبل عادة ما يكون ذو طبيعة بروتينية فالميزة الأساسية لهذه التقنية أنها تسمح باقتراح فرضيات تشرح آلية التفاعل الممكنة لهذا قمنا بتطبيقها باستخدام برنامج (MVD) من أجل فهم طريقة تفاعل الإنزيمات الدهنية مع مركب (EGCG).

هذا الأخير هو مركب طبيعي موجود في الشاي الأخضر لديه نشاط مضاد للسمنة عن طريق خفض هضم الدهون الغذائية وفق آلية تثبيط لليباز البنكرياس والمعدة. من خلال هذا العمل قمنا أولاً بدراسة طبيعة الروابط المتشكلة بين المركب السابق وإنزيم البنكرياس البشري والخنزيري تحليل طبيعة هذه الروابط لكلا الإنزيمين بينت استقرار المركب المدروس في التجويف الثالث المجاور للموقع الفعال, النتائج التي تم الحصول عليها وضحت آلية تفاعل المركب السابق وإنزيمي البنكرياس البشري و الخنزيري والذي اثبت تجريبياً ان له طبيعة تثبيط لا تنافسية, عدة روابط هيدروجينية تشكلت بين المركب المدروس وإنزيمي البنكرياس الخنزيري و البشري من أهمها رابطتين موصولتين بالحمضين الأمينين: His 263 و His 264 الذين يمثلان جزءاً من الثلاثي الفعال, أظهرت دراسة مقارنة أن مركب (EGCG) لديه فعالية تثبيط للليباز البنكرياس البشري أفضل من ليباز البنكرياس الخنزيري.

دراسة أخرى أجريت على الليباز المعدي للكلب الذي يمثل نموذج جيد للليباز المعدي للإنسان نتائج هذه النمذجة الجزئية أظهرت لنا توجه المركب نحو الموقع الفعال لهذا الإنزيم, طبيعة الالتحام الكبيرة الموجودة بين المركب السابق و الأحماض الأمينية الأساسية لهذا الإنزيم تترجم بقيمة طاقه مقدرة ب:- $150,27 \text{ kcal mol}^{-1}$.

هذا الاستقرار ينتج عن تشكل خمسة روابط هيدروجينية والعديد من التفاعلات الهيدروفيبية مع الأحماض الأمينية المشكلة للتجويف الرئيسي واحدة من هذه أهم هذه الروابط الهيدروجينية تم تشكيلها مع الحمض الأميني Ser153.

هذه الألفة النظرية نحو الأحماض الأمينية الأساسية تنبأ بطابع تثبيط تنافسي لمركب (EGCG) بالنسبة للليباز المعدي (LG).

الكلمات المفتاحية: الالتحام , الليباز البنكرياسي , الليباز , المعدي , EGCG , التثبيط , السمنة.

Liste des abréviations

PDB	:	Protein Data Bank
RMN	:	Résonance magnétique nucléaire
RX	:	Rayons X
IC50	:	Concentration d'inhibiteur pour laquelle la vitesse initiale de formation de produit diminue jusqu'à la moitié de sa valeur mesurée en absence d'inhibiteur
Ser	:	Serine
Glu	:	Glutamine
His	:	Histidine
Gly	:	Glycine
Ala	:	Alanine
Trp	:	Tryptophane
Phe	:	Phénylalanine
Cys	:	Cystéine
Leu	:	Leucine
Ile	:	Isoleucine
Tyr	:	Tyrosine
Met	:	Méthionine
MVD	:	Molegro Virtuel Docker
RMSD	:	Root mean square deviation
nRotB	:	Nombre de liaisons rotatives
E_{inter}	:	Energie d'interaction protéine-ligand.
E_{intra}	:	Energie interne du ligand
EGCG	:	Epigallocatechine-3-gallate
LPP	:	Lipase pancréatique du porc (LPP)
LPH	:	Lipase pancréatique humaine
LGH	:	Lipase gastrique humaine
LGC	:	Lipase gastrique du chien
HSL	:	Lipase hormono-sensible humaine
SDF	:	Structure Data File
SN	:	Substitution Nucléophile
RX	:	Rayons X
PMF	:	Potentiel de force moyenne

Liste des abréviations

LPL	: Lipoprotéine lipase
2D	: Deux dimensions
3D	: Trois dimensions
AE	: Algorithme évolutionnaire
AG	: Algorithme génétique
MC	: Monte Carlo
E	: Enzyme
S	: Substrat
I	: Inhibiteur
Å	: Angstrom
CD	: Dichroïsme circulaire
TAG	: Triacylglycérol
PLP	: Potentiel linéaire par paire
TGME	: Ether de mono-octyl tétraéthylène glycol
ΔG	: Énergie d'activation

Liste des tableaux

Partie I: Synthèse Bibliographique

N°	Titre	Page
1	Proportion des catéchines dans le thé vert par comparaison au thé noire.	32
2	Principales propriétés biochimiques et structurales de la LGH et la LPH.	49
3	Avantage et inconvénients des différentes fonctions de score.	70
4	Principaux programmes de docking moléculaire.	74

Partie II: Modélisation des interactions moléculaires « Lipase-EGCG »

N°	Titre	Page
1	Principaux formats d'extensions des fichiers utilisés.	88
2	Cavités détectées pour chaque récepteur : 1ETH, 1LPB, et 1K8Q.	100
3	Paramètre utilisées pour chaque algorithme : MolDock optimizer, MolDock SE (MSE), and Iterated Simplex (Simplex).	106
4	Exemples de charges d'atomes. Modèle de charges prisent en compte par MolDock Score [Grid] pour différents ligands et protéines.	107
5	Energies d'interaction (kcal mol^{-1}) des meilleures poses générées par MolDock Score, Rerank Score et H-bond Score, obtenues pour chaque cavité.	123
6	Les énergies d'interaction (kcal.mol^{-1}) des meilleures poses obtenues pour chaque fonction du score et algorithme.	125
7	Longueurs et orientations spatiales des différentes liaisons hydrogène obtenues.	127
8	Les énergies d'interaction Moldock Score, Rerank Score, et H-Bond Score des meilleures poses obtenues pour chaque cavité.	129
9	Les résidus non similaires obtenus lors de l'alignement dans l'intervalle [200, 270].	133
10	Valeurs des énergies obtenues pour chaque cavité.	135
11	Protocoles de docking moléculaire.	137
12	Valeur d'énergie et la distance des liaisons hydrogène formées.	138

Liste des figures

Partie I: Synthèse Bibliographique Chapitre I

N°	Titre	Page
1	Structure chimique de l'épigallocatechin-3-gallate.	29
2	Image des feuilles de <i>Camellia</i> .	30
3	Structure chimique et classification des Polyphénols de thé.	31
4	Forme mésomères du phénol.	33
5	Oxydation mono-électrique d'un phénol avec les formes mésomères du radical correspondant	34
6	Effet hydrophobe d'un cycle aromatique.	35
7	Hydrolyse régiosélective des esters par les lipases.	40
8	Structure de la lipase de <i>Rhizomucor miehei</i> avec le lid en conformation fermée en bleu (PDB ID : 3GTL) et en conformation ouverte en violet (PDB ID : 4GTL).	41
9	Schéma du repliement α/β . Les flèches représentent les feuilletts β et les rectangles les hélices α . Le rectangle noir représente le volet amphiphile.	42
10	Représentation de la structure secondaire du repliement α/β des hydrolases	43
11	Mécanisme catalytique de l'hydrolyse d'un ester par une enzyme lipolytique.	44
12	Structure de la lipase pancréatique porcine complexée avec une molécule de tétraéthylène glycol mono-octyl éther dans son site actif. PDB: 1ETH.	46
13	Structure cristallographique du complexe LPH/colipase en modèle de ruban ; PDB : 1LPB.	47
14	Le tractus gastro-intestinal humain.	48
15	Structures 3D superposées de la lipase gastrique humaine (LGH) et de la lipase gastrique de chien (LGC).	49

Liste des figures

Partie I: Synthèse Bibliographique Chapitre II

N°	Titre	Page
1	Étapes typiques d'un docking.	62
2	Interaction entre deux molécules de charges différentes.	71
3	Exemples d'une liaison hydrogène.	72
4	Patch hydrophobe.	73
5	Représentation schématique du cycle évolutionnaire d'un algorithme.	78

Partie II: Modélisation des interactions moléculaires « Lipase-EGCG » Chapitre III

N°	Titre	Page
1	Structure 3D du ligand: Epigallocatechine-3- gallate (EGCG).	89
2	Structure tridimensionnelle de la LPP (code 1ETH) avec les résidus.	90
3	Modèle simplifiée de la structure tridimensionnelle de la LPP.	90
4	Structure 3D du complexe lipase pancréatique-colipase.	92
5	Complexe 1LPB (lipase pancréatique humaine-procolipase de porc).	92
6	Modèle simplifiée de la LPH utilisé pour la modélisation.	93
7	Structure de la LGC complexée avec une molécule de β -OG et un inhibiteur alkylphosphonate dans son site actif.	94
8	Représentation simplifiée de la LGC.	94
9	Schéma représentative de la forme bidimensionnelle de la grille autour et à travers la protéine (traits noire), ainsi que ces points d'intersections.	97
10	Paramètre utilisé pour la détection des cavités.	98
11	Illustration des cinq cavités du 1LPB détectées (en vert) par MVD.	99
12	Cavités détectés (en vert) dans les récepteurs 1K8Q.	99
13	Illustration des sept cavités du 1ETH détectées (en vert) par MVD.	99
14	Orientation et hydrophobicité de la cavité principale, A(LPP) ; B(LPH), C(LGC). Zone hydrophobe (colorée en bleu), zone hydrophile (colorée en rouge).	101
15	Représentation du cycle typique d'un algorithme évolutionnaire « AE ».	104

Liste des figures

Partie II: Modélisation des interactions moléculaires « Lipase-EGCG »

Chapitre IV

N°	Titre	Page
1	Structure chimique (2D) de l'Epigallocatechin-3-gallate.	116
2	Structure 3D du l'inhibiteur phosphonate ($C_{15}H_{33}O_2P$). Les liaisons rotatives sont représentées en vert.	117
3	Structure 3D du l'inhibiteur phosphonate ($C_{12}H_{27}O_3P$). Les liaisons rotatives sont représentées en vert.	118
4	Superposition des géométries d'inhibiteur phosphonate ($C_{15}H_{33}O_2P$).	119
5	Comparaison de la conformation expérimentale de l'inhibiteur phosphonate ($C_{12}H_{27}O_3P$). (colorée par type d'atome) et sa conformation optimale simulée par MVD (colorée en vert).	119
6	Structure 3D du l'épigallocatechin-3-gallate (EGCG). Les liaisons rotatives sont représentées en vert.	120
7	Illustration du site catalytique de la LPP avec les résidus principaux (Ser153, His264, et Asp177).	122
8	Orientation des meilleures poses de docking dans les cavités : 2, 3, 4, 7.	124
9	Positionnement du ligand EGCG (coloré en blanc) au sein de la cavité N°3 (vert).	126
10	Interactions de type liaison hydrogène entre EGCG et les résidus du site d'interaction. Les liaisons hydrogène présentées par un trait bleu discontinue.	127
11	Interactions hydrophobes entre EGCG et PPL, produites par le programme LIGPLOT +. Les pointes rouges sur les arcs représentent les interactions hydrophobes avec les résidus (en noir).	128
12	Représentation des interactions EGCG-LPH. Les liaisons hydrogène sont présentées par un trait bleu discontinue.	130
13	Représentation 2D des interactions hydrophobes.	131
14	Superposition de la lipase pancréatique du porc (en orange) et la lipase pancréatique humaine (en bleu).	131
15	Alignement des séquences 1ETH.A.PDB et 1LPB.B.PDB.	132
16	Zoom sur le site catalytique de la LGC avec les résidus principaux (Ser153, His353, et Asp328).	135
17	Illustration de l'espace de recherche qui couvre la structure 3D de la cible centré sur la cavité 1.	136
18	Illustration de l'espace de recherche qui couvre la structure 3D de la cible centré sur le ligand de référence.	137
19	Orientation des trois poses de EGCG obtenues, ver la cavité principale de lipase gastrique.	138
20	Représentation des interactions EGCG-LGC avec les liaisons hydrogène présentées par un trait bleu discontinue.	139
21	Interaction hydrophobe entre la lipase gastrique et l'EGCG.	140

Liste des schémas

N°	Titre	Page
1	Biosynthèse des catéchines dans les feuilles de <i>Camillia sinensis</i> .	33
2	Protocole générale de Docking moléculaire.	64
3	Principe général d'un programme de docking.	96
4	Mode d'action d'un inhibiteur non compétitif.	121

Sommaire

Remerciements.....	3
Résumé.....	5
Abstract.....	6
ملخص.....	7
Liste des abréviations.....	8
Liste des tableaux.....	10
Liste des figures.....	11
Liste des schémas.....	14
Sommaire.....	15
Introduction générale.....	19
Références Bibliographiques.....	23

Partie I : Synthèse Bibliographique

Chapitre I : Structures mises en jeu.....	26
I.1. Introduction	26
I.2. Structures mises en jeu dans le docking moléculaire.....	29
I.2.1. Ligand : EGCG.....	29
I.2.1.1. Source végétale de l'EGCG.....	29
I.2.1.2. Biosynthèse des catéchines de thé vert.....	32
I.2.1.3. Propriétés chimiques de l'EGCG.....	33
I.2.1.4. Effet thérapeutique de l'EGCG.....	37
I.2.2. Récepteur: Lipases.....	39
I.2.2.1. Volet amphiphile et activation interfaciale.....	40
I.2.2.3. Caractéristiques structurales.....	41
I.2.2.4. La Lipase <i>Pancréatique du Porc</i> (LPP).....	45
I.2.2.5. La lipase <i>pancréatique humaine</i> (LPH).....	46
I.2.2.6. La lipase gastrique.....	47
I.3. Conclusion.....	51
Références Bibliographiques.....	52
Chapitre II : Technique de Docking moléculaire.....	61

II.1. Introduction.....	61
II.2. Etapes typiques du Docking moléculaire.....	61
II.2.1. Détermination des structures.....	62
II.2.2. Préparation des structures.....	63
II.2.3. Docking moléculaire.....	64
II.2.4. Prédiction et évaluation.....	66
II.2.4.1. Algorithme de docking.....	66
II.2.4.2. Fonctions de score.....	67
II.2.4.3. Interaction protéine-ligand.....	70
II.2.5. Logiciels utilisés.....	74
II.3. Conclusion.....	79
Références Bibliographiques.....	80

Partie II : Modélisation des interactions moléculaires « Lipase-EGCG »

Chapitre III : Matériels et méthodes.....	88
III.1. Logiciel de simulation moléculaire.....	88
III.2. Structure du ligand.....	89
III.3. Structures des récepteurs.....	89
III.3.1. Lipase <i>pancréatique de porc</i> (LPP).....	90
III.3.2. Structure de la lipase <i>pancréatique humaine</i> (LPH).....	91
III.3.3. Structure de la lipase <i>gastrique du chien</i> (LGC).....	93
III.4. Préparation des récepteurs et du ligand.....	95
III.5. Définition du site actif.....	97
III.6. Algorithme de recherche implémentée dans MVD.....	102
III.6.1. MolDock Optimizer.....	104
III.6.2. L'algorithme MolDock SE (Simplex Evolution).....	104
III.6.3. L'algorithme Simplex Itéré.....	105
III.6.4. Fonction de score.....	106
Références Bibliographiques.....	110
Chapitre IV : Résultats et discussion.....	114
IV.1. Description de l'EGCG.....	115
IV.2. Fiabilité du programme MVD.....	116
IV.2.1. Ecart quadratique moyen (RMSD).....	117

IV.2.2. Superposition des ligands.....	118
IV.3.Docking de l'épigallocatechin-3-gallate (EGCG).....	121
IV.3.1. Cas de la lipase pancréatique de porc (LPP).....	121
IV.3.1.1. Etude des interactions entre EGCG et LPP.....	126
IV.3.2. Cas de la lipase pancréatique humaine (LPH).....	128
IV.3.2.1. Etude des interactions EGCG-LPH.....	129
IV.3.3. Comparaison des interactions entre LPP-EGCG et LPH-EGCG.....	131
IV.3.4. Cas de la lipase gastrique (LG).....	133
Références Bibliographiques.....	141
Conclusion générale.....	144

Introduction générale

Introduction générale

Les protéines, un assemblage tridimensionnel d'acides aminés, sont omniprésentes dans nos organismes afin d'assurer une multitude de fonctions biologiques. Le rôle fondamental des protéines est d'agir comme catalyseurs enzymatiques connues sous le nom de biocatalyseurs, afin d'augmenter la vitesse de la quasi-totalité des réactions chimiques qui se produisent dans notre organisme. Toute vie dépend du bon fonctionnement d'une grande variété de ces biocatalyseurs. Leurs structures biologiques correspondent à des fonctions bien définies ce qui rend la vie d'un organisme fortement reliée au rôle de ces structures.

Parmi ces catalyseurs enzymatiques on cite les lipases. Ces dernières qui de part leur occurrence dans l'ensemble du monde vivant, ont fait l'objet de nombreuses études dans des domaines variés tant biomédical (réactifs, kits de diagnostic, méthodes thérapeutiques) qu'industriel (mise en œuvre de technologies à base d'enzymes dans l'industrie agro-alimentaire) et les détergents¹, en particulier, les lipases digestives des mammifères qui comptent parmi les premières étudiées. L'absorption des acides gras par l'organisme est rendue possible par l'hydrolyse des lipides et des graisses alimentaires par les lipases digestives. Elles sont également des cibles thérapeutiques intéressantes pour le traitement de pathologies comme l'obésité ou pour le développement de nouveaux antibiotiques.

Notre recherche consiste à étudier l'inhibition des enzymes impliquées dans la maladie de l'obésité, avec un inhibiteur naturel. Cette étude a pour but de minimiser la formation des complexes et par la suite retarder sa progression. Afin de rationaliser les propriétés de cet inhibiteur et déterminer le processus réactionnel impliquant ce composé, nous avons essayé d'optimiser et modéliser les complexes formés par des méthodes théoriques utilisant l'outil informatique. Nous avons essayé d'étudier, au point de vue énergétique, les interactions et élucider des mécanismes d'interactions (récepteur -ligand), en vue d'inhiber les enzymes impliquées dans cette maladie.

En se basant sur l'étude théorique par le biais de l'outil informatique, on a essayé d'expliquer et déterminer le mode d'interaction du complexe par la fixation de l'inhibiteur dans l'enzyme, avec une meilleure et forte complémentarité. Ces résultats aideront au développement d'un outil thérapeutique efficace pour lutter contre le développement de la maladie d'obésité. La recherche en biologie ne peut, actuellement, se passer des outils informatiques pour traiter le flot de données produites et optimiser ses avancées. L'un de ces outils est la modélisation moléculaire et plus précisément l'amarrage moléculaire (plus souvent connu sous le terme anglo-saxon "docking").

Introduction générale

Le rôle principal du "docking" moléculaire est de prédire et reproduire des complexes protéine-ligand^{2,3}. Le programme de docking est utilisé pour placer des représentations générées par ordinateur d'une petite molécule appelé ligand dans une structure cible représentée le plus souvent par une protéine. Le ligand peut être aussi placé dans une partie de la cible définie par l'utilisateur. Par exemple, le site actif d'une enzyme dans lequel le ligand se placera dans diverses positions, conformations et orientations.⁴

Les succès réalisés dans ce domaine et l'essor des techniques de chimie combinatoire, ainsi que l'exploitation systématique de la diversité chimique provenant d'autres cibles. Ces succès ont élargi l'utilisation de la méthode de docking moléculaire. Cette méthode a permis aussi d'accéder au criblage des bases de données et l'optimisation des molécules.

Ainsi cette méthode, permettant de cribler des milliers de composés pour une protéine cible, est couramment utilisée en pharmaco-chimie pour l'obtention de nouveaux médicaments⁴. Une telle approche serait difficilement réalisable en biologie traditionnelle où le récepteur est classiquement une protéine ou une protéine oligomérique et le ligand est une petite molécule.⁵

C'est dans ce contexte que nous nous sommes intéressés à l'application de cette alternative prometteuse utilisant la technique de modélisation moléculaire, regroupées sous le nom « Docking moléculaire » pour la prédiction des modes d'interactions possibles et l'étude des interactions moléculaires entre l'Epigallocatechine (EGCG) et trois lipases digestives: la lipase pancréatique du porc (LPP), la lipase pancréatique humaine (LPH), et la lipase gastrique du chien (LGC). Cette dernière présente un bon modèle de la lipase gastrique humaine (LGH).

Pour étudier ces interactions, nous avons choisi un programme de docking moléculaire qui est le Molegro Virtual Docker (MVD). Ce programme a été développé pour aider à la mise au point de molécules à activité thérapeutique. Pour cela nous avons choisi le (-)-Epigallocatechin-3-gallate (EGCG). Ce dernier est un composé naturel présent dans le thé vert. Il a été détecté comme un puissant inhibiteur des lipases pancréatique (LP), selon un mode non compétitif. Ce composé naturel possède des propriétés biologiques et médicinales, y compris les effets antioxydants, anti-cancérogènes, anti-obésité, antibactériens, antiviraux et anti-enzymatiques et anti-métastases⁶. Dans cette étude nous nous sommes intéressés particulièrement à l'effet anti-obésité que possède ce composé.

Introduction générale

L'obésité est devenue un problème majeur de santé publique, en particulier dans les pays industrialisés. De plus en plus répandue et souvent grave, elle prédispose à nombre de maladies⁷, diminue l'espérance de vie et entraîne des dépenses de soins et de prévention croissantes. À ce jour, il n'existe pas de traitements médicamenteux miracles qui remplace les anciens médicaments coupe-faim (les amphétamines, la fenfluramine). Ces derniers ne sont plus utilisés car ils avaient trop d'effets secondaires⁸. Les effets de ces médicaments varient d'un individu à l'autre et ne donneront pas forcément les résultats attendus. De plus, leur prise doit être combinée à une alimentation équilibrée. Tandis que quelques tasses de thé vert chaque jour pourrait ralentir le gain de poids et être une nouvelle option naturelle, dans la lutte contre l'obésité.

Une option tout à fait pertinente puisqu'un polyphénol du thé vert comme l'Epigallocatechin-3-gallate (EGCG) peut diminuer la digestion des lipides alimentaires par le mécanisme d'inhibition des lipases gastrique et pancréatique. Le nombre d'articles relatant l'inhibition enzymatique des lipases digestives, en utilisant l'Epigallocatechin-3-gallate, comme inhibiteur s'est multiplié dernièrement. Une étude cinétique a été également réalisée, suggèrent l'inhibition non compétitive de ce composée⁹, les paramètres thermodynamique indique que l'association entre la lipase pancréatique et l'EGCG est gouvernée par des interactions de type spontanée, tel que : liaison hydrogène, interaction hydrophobe, et interaction électrostatique. Le résultat de dichroïsme circulaire suggère un changement dans la conformation protéique lors de la fixation de l'EGCG¹⁰. Seulement, ces travaux expérimentaux ne fournissent pas d'informations nécessaires qui permettent de répondre à des questions plus fondamentales telles que :

- Quelle est la cavité que représente la région où l'inhibiteur non compétitif (EGCG) va se fixer ?
- Comment le mode d'interaction lipase-EGCG peut déformer la géométrie de site active ?
- Quelles sont les résidus responsables du mécanisme d'inhibition.

Pour répondre à des questions de ce type une analyse des interactions entre les enzymes et l'inhibiteur est primordiale.

Introduction générale

Les détails de ces interactions, au niveau moléculaire, sont donc d'un très grand intérêt et peuvent être étudiés par cristallographie aux rayons « X » ou résonance magnétique nucléaire (RMN). Cependant, la mise en œuvre de ces techniques requiert la cristallisation préalable du complexe enzyme-ligand, ceci est en pratique très difficile car la durée d'existence de ces espèces est très courte¹¹. Par conséquent, il est indispensable de trouver d'autre alternative afin de comprendre les bases moléculaires dictant l'activité inhibitrice de l'Epigallocatechin-3-gallate pour les lipases digestives.

C'est ainsi que nous avons décidé d'aborder cette problématique par une approche théorique utilisant le docking moléculaire afin d'expliquer le mécanisme d'action de EGCG qui agit comme agent hypolipidémiant, vis-à-vis des lipases. Le docking moléculaire *in silico*, vise à prédire la structure d'un complexe moléculaire à partir des molécules isolées, ce qui est considérablement plus facile à mettre en œuvre, moins cher et plus rapide que l'utilisation de l'une des méthodes expérimentales mentionnées ci-dessus. Les logiciels de docking sont donc des outils très utiles en biologie, pharmacie et médecine, car la plupart des principes actifs sont de petites molécules (ligand) qui interagissent avec une cible biologique (récepteur) d'intérêt thérapeutique. Ces résultats aideront au développement d'un outil thérapeutique efficace pour lutter contre le développement de la maladie d'obésité.

A cet effet, la présente thèse se subdivise en deux parties distinctes:

La première partie de cette thèse est une synthèse bibliographique qui comporte deux chapitres. Dans le premier chapitre, nous avons effectué une recherche bibliographique représente un aperçu général sur les molécules mises en jeu dans notre étude de docking. Le second chapitre concerne les principales approches de la modélisation par le docking moléculaire. Ce deuxième chapitre sera consacré à la description de cette technique utilisée en présentant divers aspects du docking moléculaire.

La deuxième partie de cette thèse consiste à décrire l'étude concernant la modélisation des interactions moléculaires « Lipase-EGCG ». Cette partie est divisée en deux chapitres (3 et 4). Le troisième chapitre porte sur une description du matériel utilisé, la démarche expérimentale et les méthodes employées. Le dernier chapitre de ce mémoire sera consacré à exposer les résultats que nous avons obtenus et leur discussion.

Enfin une conclusion générale résume l'ensemble du travail réalisé et présente les perspectives d'études que nous envisagerons pour l'avenir.

Références Bibliographiques

- 1: Soufi W, Merad M, Boukli F, Ghalem S. The Complementarity Effect for Cdc25 Phosphatase Inhibitors. *Advances in Molecular Imaging*, 2011; 17-23.
- 2: Levinthal C, Wodak SJ, Kahn P, Dadivanian AK. Hemoglobin interaction in sickle cell fibers. Theoretical approaches to the molecular contacts. *Proceedings of the National Academy of Sciences*. USA 1975; 72: 1330–34.
- 3: Kuntz ID, Blaney JM, Oatley SJ, Langridge R, Ferrin TE. A geometric approach to macromolecule-ligand interactions. *Journal of Molecular Biology*, 1982; 161: 269–88.
- 4: Romano T. Structure-Based Drug Design: Docking and Scoring. *Current Protein and Peptide Science*, 2007; 8: 312-328.
- 5: Shivaji BB, Rama N, Lakshmi V. G. Nargund, DevarajuK S, Vedamurthy A. B, Shruti SD. Synthesis and biological evaluation of novel pyrazole derivatives as urease inhibitors. *Der Pharma Chemica*, 2011; 3 (5):73-80.
- 6: Nguyen TT, Moon YH, Ryu YB, Kim YM, Nam SH, Kim MS, Kimura A, Kim D. *Enzyme and Microbial Technology*, 2013; 52: 26-3.
- 7: Lunagariya NA, Patel NK, Jagtap SC, Bhutani KK. Inhibitors of pancreatic lipase: state of the art and clinical perspectives. *EXCLI Journal*, 2014; 13: 897-921.
- 8: Liang LF, Wang T, Cai YS, He WF, Sun P, Li YF, Huang Q, Taglialatela-Scafati O, Wang HY, Guo YW. Brominated polyunsaturated lipids from the Chinese sponge *Xestospongia testudinaria* as a new class of pancreatic lipase inhibitors. *European Journal of Medicinal Chemistry*, 2014; 79: 290-297.
- 9: Wang S, Sun Z, Dong S, Liu Y, Liu Y. Interactions between (–)-Epigallocatechin Gallate Analogs and Pancreatic Lipase. *Plos One*, 2014; 9 (11): e111143.
- 10: Wu X, He W, Yao L, Zhang H, Liu Z, Wang W, Ye Y. Characterization of binding interaction of (–)- Epigallocatechin-3-gallate from green Tea and lipase. *Journal of agricultural and food chemistry*, 2013; 61 (37): 8829-8835.
- 11: Eduardo B. Simulation moléculaire appliqué à l'acylation de flavonoïde catalysé par lipase, thèse doctorat l'institut polytechnique Lorrain 2009, 22.

Partie I

Synthèse Bibliographique

Chapitre I

Structures mises en jeu

I.1.Introduction

Les triacylglycérol acyl-hydrolases, ou lipases, sont des enzymes atypiques de par leur mécanisme d'action et leur spécificité de substrats. En fonction du micro-environnement de l'enzyme, elles peuvent agir en tant qu'hydrolases en milieu aqueux ou comme catalyseurs en synthèse organique^{1,2}. En tant qu'hydrolases, elles sont responsables du catabolisme des triglycérides, leurs substrats préférentiels, en acide gras et en glycérol. Chez de nombreux êtres vivants, cette réaction est capitale de par son rôle physiologique majeur dans le métabolisme des graisses et des lipides. De plus, certaines lipases sont capables d'hydrolyser des phospholipides, des esters de cholestérol et même parfois certains esters synthétiques³. Les lipases forment une classe d'enzymes hétérogènes de par leur origine, qu'elles soient animales, végétales ou microbiennes, ce qui augmente encore leurs potentialités. Toutes ces propriétés ont conduit au développement de nombreuses applications aussi bien au point de vue industriel, notamment dans l'industrie agro-alimentaire et dans l'industrie chimique, qu'en médecine humaine.

Notre alimentation apporte quotidiennement différents lipides et nutriments liposolubles⁴. Ces lipides sont rendus bio-disponibles par un processus complexe impliquant une série d'étapes physico-chimiques et enzymatiques tels que ⁵:

- l'émulsification des lipides, c'est-à-dire leur dispersion sous forme de gouttelettes dans le système digestif aqueux,
- l'hydrolyse des lipides à l'interface lipides/eau par des lipases dans l'estomac et dans le duodénum,
- le transport des nutriments lipidiques vers les différentes cellules utilisatrices de l'organisme via le sang sous forme de lipoprotéines.

Les graisses alimentaires représentent une source efficace d'énergie pour l'organisme. En effet, la quantité d'énergie métabolisée à partir des lipides est significativement plus importante que celle métabolisée à partir des carbohydrates ou des protéines. Cependant, la quasi-totalité des lipides ingérés étant assimilés par l'organisme, une surcharge alimentaire en lipides peut provoquer des troubles de santé importants: troubles cardiovasculaires, hyperlipémies et obésité.

Ces troubles sont fréquemment rencontrés dans les pays industrialisés où les populations ont souvent des régimes alimentaires trop riches en graisses saturées. Pour lutter contre ces différentes pathologies, une hygiène alimentaire consistant à limiter l'ingestion de corps gras est nécessaire mais pas toujours suffisante. En effet, si certains corps gras sont faciles à détecter et à éliminer de l'alimentation (beurre, huiles...), d'autres, de par leur intégration dans les aliments (viandes, laitages..) le sont beaucoup moins. Dans les cas où la mise en place d'un régime alimentaire s'avère insuffisante, des traitements pharmacologiques sont alors proposés.

La plupart des traitements actuels visent surtout à lutter contre les hyperlipémies en utilisant, entre autres, des inhibiteurs de la synthèse du cholestérol. Les recherches actuelles se tournent, de plus en plus, vers des produits induisant une inhibition plus générale de l'activité lipolytique⁶. Dans une telle perspective, l'une des orientations qui a été particulièrement étudiée est celle de l'inhibition de la lipase pancréatique, qui est une enzyme-clé de la digestion des triglycérides alimentaires. Pour inhiber la lipase pancréatique et obtenir une activité thérapeutique sur les hyperlipémies et l'obésité, différentes approches ont été proposées. En conséquence, le demandeur s'est donné pour but de pourvoir à un nouvel inhibiteur de la lipase pancréatique, qui réponde mieux aux besoins de la pratique que les inhibiteurs de la lipase de l'art antérieur, notamment en ce qu'il présente une réelle spécificité d'action vis-à-vis de la lipase pancréatique⁷. Le développement des inhibiteurs des lipases commence à prendre un espace important dans la littérature, dont leur intérêt est de bloquer une large gamme de maladies. Récemment, plusieurs molécules d'origine naturelle ou bien synthétisées, ont été testées comme médicaments anti-obésité.⁸

Ces dernières années, l'intérêt porté aux antioxydants naturels, en relation avec leurs propriétés thérapeutiques, a augmenté considérablement. Des recherches scientifiques dans diverses spécialités ont été développées pour l'extraction, l'identification et la quantification de ces composés à partir de plusieurs substances naturelles à savoir, les plantes médicinales et les produits agroalimentaires⁹⁻¹¹. Parmi ces nouveaux composés potentiellement intéressants, les antioxydants, tels que les flavonoïdes, ont été particulièrement étudiés en raison de leur utilisation dans les domaines pharmaceutiques, cosmétiques et alimentaires pour leurs effets bénéfiques pour la santé¹². C'est dans ce contexte que nous nous sommes intéressés aux polyphénols dérivés des flavonoïdes. Les polyphénols constituent le plus grand groupe des substances végétales secondaires. On en trouve, par exemple, dans les raisins et le thé.

A partir de ces données récentes trouvées dans la littérature, qu'a été fondé l'objectif principal de notre travail. Ce dernier porte sur l'étude concernant les interactions de type ligand-enzyme issues de l'inhibition de lipases digestives par l'épigallocatechin-3-gallate (EGCG). Ce composé, qui est un polyphénol que l'on trouve en abondance principalement dans le thé vert, dérive de la famille des flavonoïdes. On le trouve principalement dans la catéchine. EGCG est un puissant antioxydant capable de neutraliser les espèces réactives oxygénées et les radicaux libres lourdement impliqués dans le vieillissement et les maladies chroniques dégénératives.

La recherche a montré que l'EGCG pourrait avoir des effets bénéfiques dans le cas de nombreuses maladies incluant le cancer, l'athérosclérose, le diabète, les maladies neurodégénératives ou l'excès de poids^{13, 14}. Ce composé possède de fortes propriétés protectrices pour les cellules.¹⁵

Le mécanisme et le mode d'action de cette molécule sur les lipases utilisées comme anti-obésité restent sans explication. Ainsi, les méthodes de calculs théoriques et particulièrement le docking moléculaire, peuvent être un support indispensable pour élucider le mode d'action de cette molécule lors de ses interactions avec la protéine cible.

Avec le développement des outils informatiques, la modélisation moléculaire et plus précisément le docking moléculaire (assemblage ou arrimage moléculaire) a rapidement investi le domaine de la recherche en biologie. Celui-ci peut être défini comme la recherche du meilleur appariement entre deux molécules. Le docking moléculaire a pour objectif essentiel de prédire la conformation (position et orientation relative) la plus favorable du ligand au sein de son récepteur, comme il peut servir à l'optimisation de molécules et au criblage de bases de données.¹⁶

En pharmacie, la découverte et la mise au point de nouvelles substances médicamenteuses peuvent passer par le criblage de bases de données avec des millions de composés pour une même protéine cible, ce qui ne serait pas réalisable en biologie classique. En général, la cible est une protéine et le ligand peut être une petite molécule organique¹⁷. La technique de docking moléculaire sera détaillée dans le chapitre 2.

I.2. structures mises en jeu dans le docking moléculaire

I.2.1. Ligand : EGCG

Le gallate d'épigallocatechine (EGCG) un polyphénol extrait du thé vert. C'est un composé constitué d'ester d'épigallocatechine et d'acide gallique. C'est le flavanol le plus abondant du thé, il est actuellement l'objet d'un grand intérêt public et est utilisé dans de nombreuses recherches scientifiques pour ses multiples bienfaits pour la santé^{18,19}. La structure chimique du gallate d'épigallocatechine (EGCG) est à trois cycles aromatiques (A, B et D) qui sont reliés entre eux par un noyau pyranne (C) (fig.1).

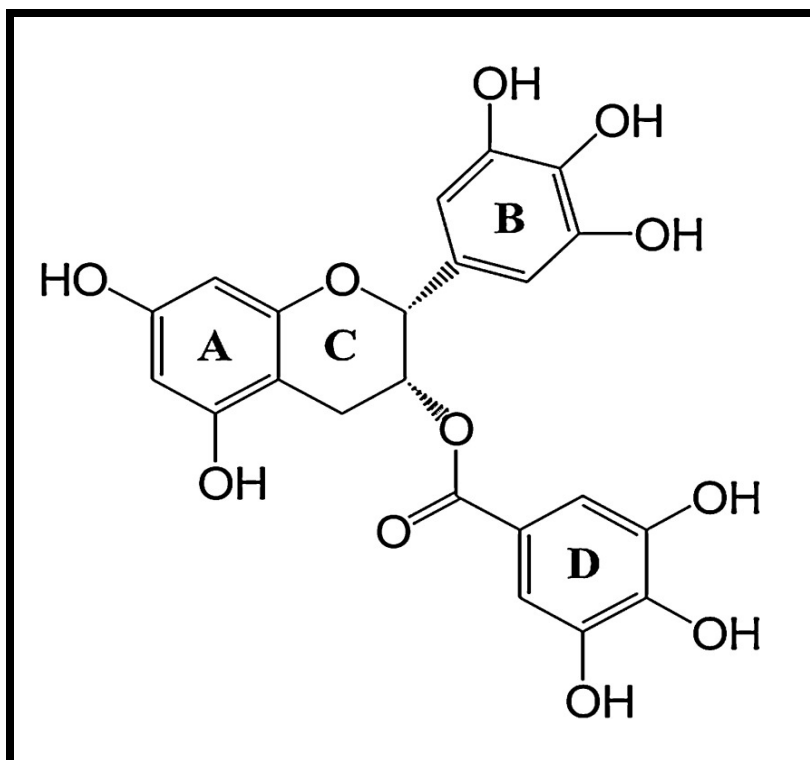


Figure 1. Structure chimique de l'épigallocatechine-3-gallate.²⁰

I.2.1.1. Source végétale de l'EGCG

Les feuilles de *Camellia* par exemple (fig.2) contiennent environ huit catéchines: la catéchine, l'épicatéchine (EC), la gallocatéchine (GC), l'épigallocatechine (EGC), la catéchine gallate (CG), la gallocatéchine gallate (GCG), l'épicatéchine gallate (ECG) et l'épigallocatechine gallate (EGCG)²¹.

Parmi les 36% de polyphénols totaux, l'épigallocatechine gallate (EGCG) est la forme prédominante dans le thé vert avec une teneur dans les feuilles comprise entre 48 et 55%.



Figure 2. Image des feuilles de *Camellia*.²²

De tous les composés antioxydants présents dans le thé vert, les principaux constituants sont les polyphénols, y compris les acides phénoliques et les catéchines²³ (fig.3). Les jeunes pousses contiennent de 200 à 340 mg de catéchine, gallocatéchine et ses dérivés par gramme de feuilles sèches. Dans le thé noir, leur teneur est réduite de moitié environ en raison de leur oxydation en polyphénols plus complexes pendant la fermentation.

Une analyse simple de la concentration de catéchines dans l'eau du thé, par comparaison à la teneur en catéchines dans les feuilles du thé à été menée par Schneider *et al.*,²⁴. Il apparaît dans ce rapport que les catéchines sont nettement plus nombreuses dans une infusion de thé vert que dans le thé noir. L'écart en catéchines EGC et EGCG est frappant.

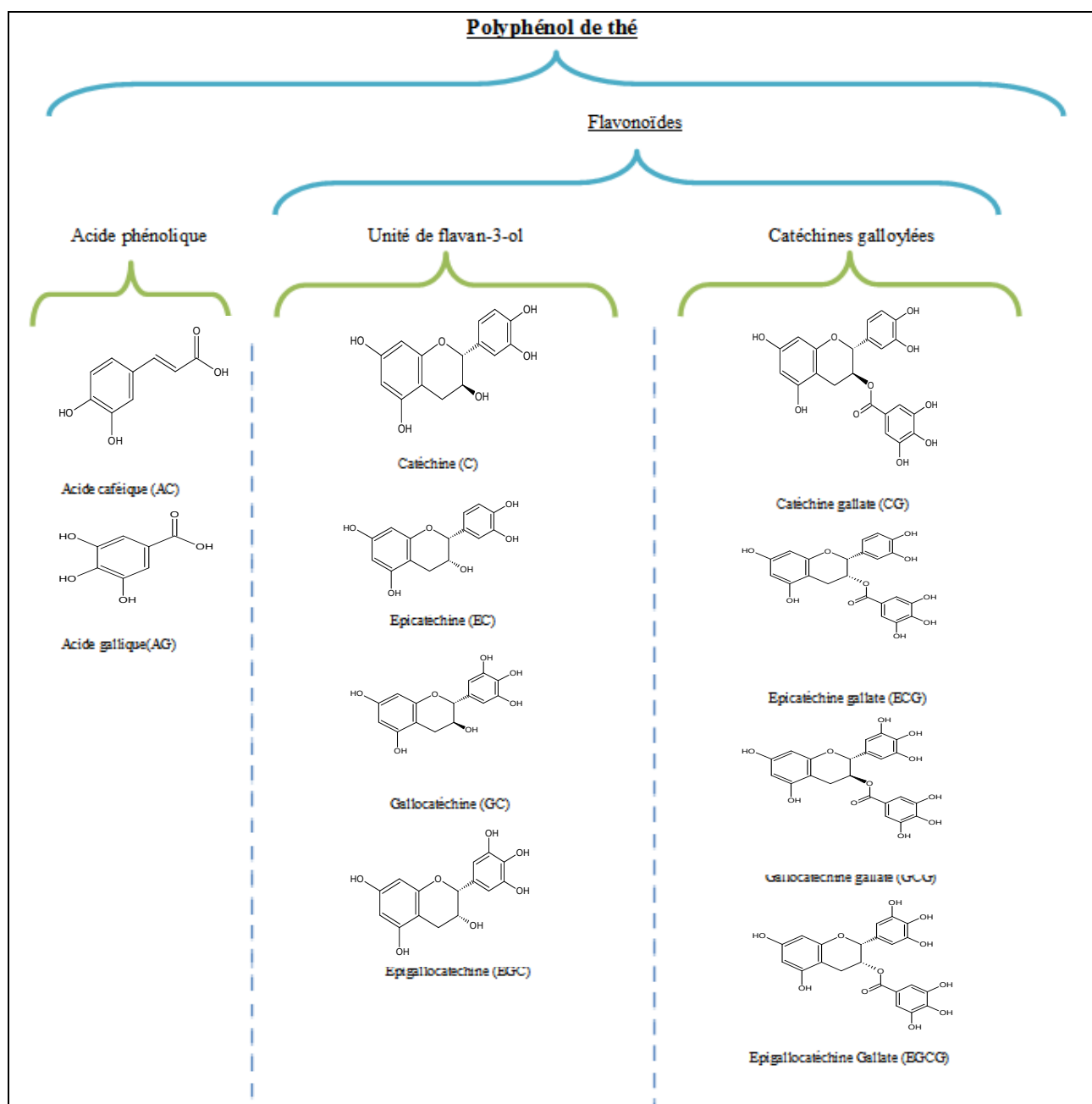


Figure 3. Structure chimique et classification des Polyphénols de thé²³.

Tableau 1. Proportion des catéchines dans le thé vert par comparaison au thé noire.

Catéchines	Proportion de thé d'extrait sec (%)	
	vert	noire
Epicatechine (EC)	1,98	1,21
Epicatechine gallate (ECG)	5,20	3,86
Epigallocatechine (EGC)	8,42	1,1
Epigallocatechine Gallate (EGCG)	20,30	4,63

On note aussi que les catéchines sont présentes aussi dans le chocolat et dans de nombreux fruits. Les principales sources comprennent les abricots, les pommes, les poires, les raisins, les pêches, le cacao et le vin rouge.²⁵

I.2.1.2. Biosynthèse des catéchines de thé vert

La croissance des feuilles de *Camillia sinensis* s'accompagne de changements quantitatifs et qualitatifs du contenu phénolique. Le schéma 1 résume les voies de synthèse des principaux catéchines, d'après les données fournies par la littérature.²⁶

Les catéchines du thé vert sont issues de la condensation de molécule d'acide acétique malonique et d'acide shikimique. L'acide gallique présent dans les feuilles de *Camillia sinensis* est produit par la voie de l'acide shikimique. (-)-Epigallocatechine est produite par hydroxylation de la molécule (-)-Epicatechine alors que la (-)-Epicatechine gallate et l'épigallocatechine gallate sont synthétisées par estérification de la (-)-Epicatechine avec l'acide gallique.

La présence de la lumière favorise l'augmentation de la production des catéchines dans les feuilles du thé. Ce phénomène est relié essentiellement à l'activité de la phénylalanine-ammonia-lyase (PAL) qui est l'enzyme clé dans la biosynthèse des catéchines (noyau B). Quand les feuilles du thé sont couvertes (absence de la lumière), l'activité de cet enzyme décroît rapidement.²⁷

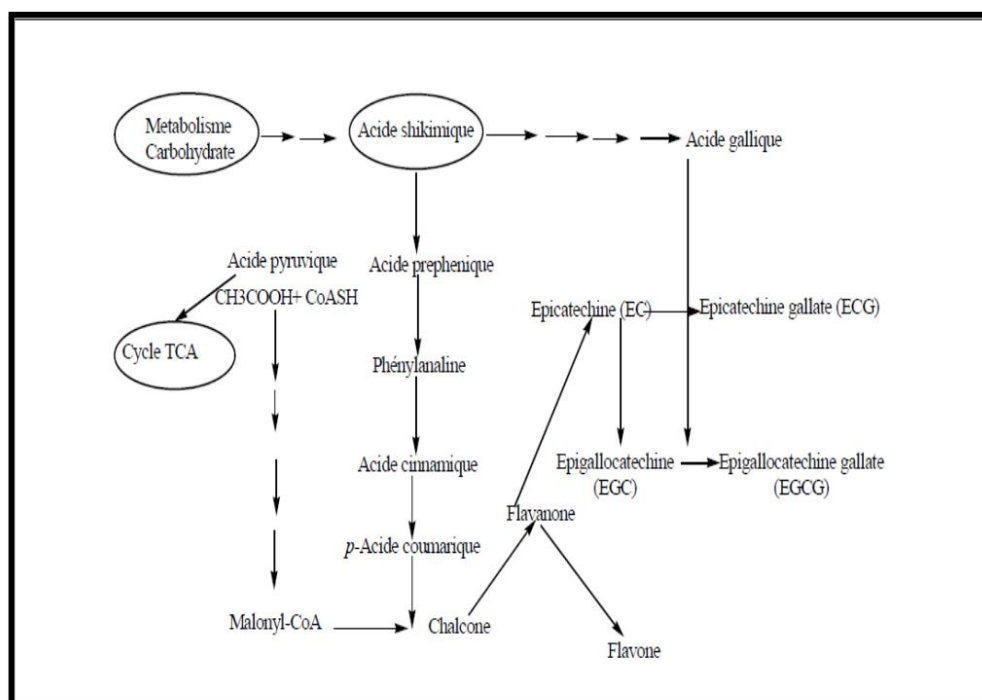


Schéma 1. Biosynthèse des catéchines dans les feuilles de *Camellia sinensis*.²⁶

I.2.1.3. Propriétés chimiques de l'EGCG

Les propriétés chimiques des polyphénols sont essentiellement liées à celles des noyaux phénoliques²⁸, particulièrement des substituant à effet mésomère attracteur d'électrons (-M) et substituant à effet mésomère donneur (+M). La conjugaison d'une des paires libres de l'atome d'oxygène avec le cycle traduit l'effet (+M) du groupement OH. Ce phénomène augmente la délocalisation électronique (8 électrons délocalisés sur 7 atomes) et confère une charge négative partielle aux atomes C2, C4 et C6 qui se traduit par la représentation de 4 formes mésomères (fig. 4)

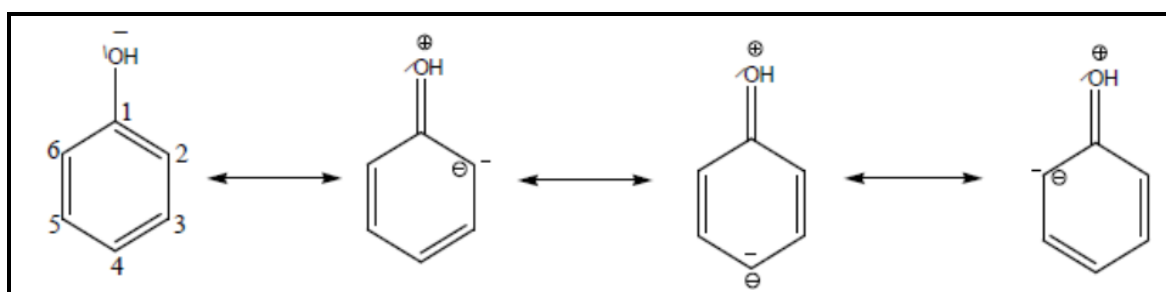


Figure 4. Forme mésomères du phénol.²⁹

De ces caractères de base découlent les différentes propriétés physico-chimiques suivantes :

a) Nucléophilie

La nucléophilie des composés phénoliques est portée par l'atome d'oxygène et les atomes de carbone en *ortho* et *para* du groupement OH (suite à l'effet « +M »). Cette propriété est à l'origine des réactions régiosélectives des positions *ortho* et *para* de substituant électrophiles aromatique (alkylation, acylation, etc..). Les substituant de type 1,3-dihydroxy comme il est le cas du résorcinol et 1,3,5-trihydroxy dans le cas du phloroglucinol, permettent une accumulation de densité électronique sur les sommets C2, C4 et C6 (tous *ortho* ou *para* des groupements OH), accentuant ainsi le caractère nucléophile.²⁷

Le cycle A de la catéchine et des autres flavanols présente deux centres (C6 et C8) fortement nucléophiles. Cette force est due à la position en *ortho* et en *para* de trois groupements OH ou OR ayant chacun un effet +M. Le noyau A est également activé par le groupement carboné saturé en C4. Cette nucléophilie autorise des réactions de substitutions électrophiles aromatiques.

b) Propriétés réductrices

Le potentiel (PI) ou énergie d'ionisation d'un atome ou d'une molécule est l'énergie qu'il faut fournir à un atome neutre pour arracher un électron (le moins lié) à l'état gazeux et former un ion positif. Plus un composé aromatique est substitué par des groupements donneurs d'électrons, plus son PI est faible et plus son caractère réducteur est grand. Il peut alors subir une oxydation mono-électronique qui conduit au radical correspondant. Dans le cas d'un phénol Ar-OH, le radical cation formé est un acide fort qui se déprotonne aussitôt pour conduire à un radical phénoxy ou aryloxy ArO (fig. 5).³⁰

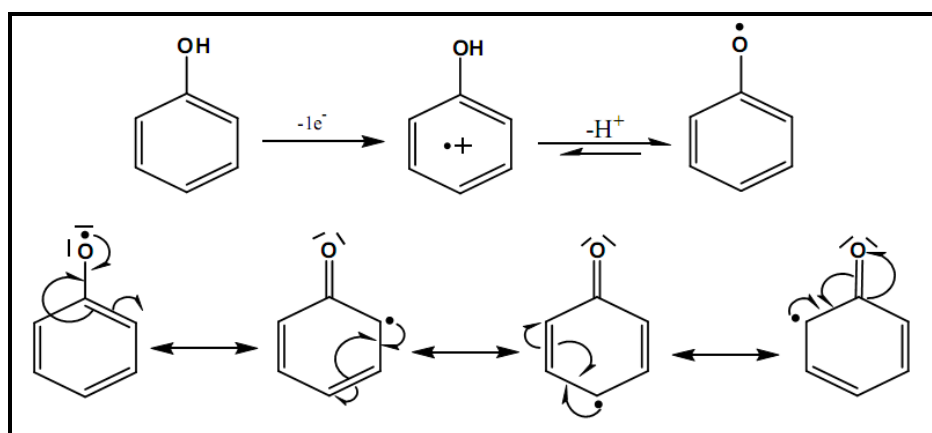


Figure 5. Oxydation mono-électronique d'un phénol avec les formes mésomères du radical correspondant.³¹

Le radical aryloxy ($\text{ArO}^\bullet$) peut être formé directement par transfert d'hydrogène phénolique vers un radical de hautes énergies telles que les radicaux oxy ($\text{RO}^\bullet$) et peroxy ($\text{ROO}^\bullet$) formés par exemple au cours de l'autoxydation des lipides.³¹

Ces réactions de transfert d'atome H et/ou d'électrons avec conversion d'un radical très réactif en radical aryloxy stabilisé par résonance, sont l'un des principaux mécanismes d'action antioxydant des phénols. La capacité du phénol à céder un atome H peut être quantifiée par l'énergie de dissociation homolytique de la liaison OH (*bond dissociation energy, BDE*). Plus la BDE d'un phénol est faible, plus son caractère donneur d'hydrogène est fort.

De nombreuses études *in vitro* mettant en œuvre les flavonoïdes confirment que la présence d'une structure catéchol est un déterminant majeur de l'activité antioxydante en accord avec la faible BDE des groupements OH correspondants.^{32,33}

c) Polarisabilité

La polarisabilité des phénols leur permet de développer de fortes interactions moléculaires de dispersion (composante attractive des interactions de Vander Waals) avec d'autres composés polarisables. Ce phénomène résulte du couplage entre les fluctuations électroniques de deux molécules voisines. Ainsi, en solution aqueuse, l'interaction du noyau benzénique apolaire du phénol avec une autre entité polarisable telle qu'un second cycle aromatique est favorisée par l'effet hydrophobe (fig. 6).²⁷

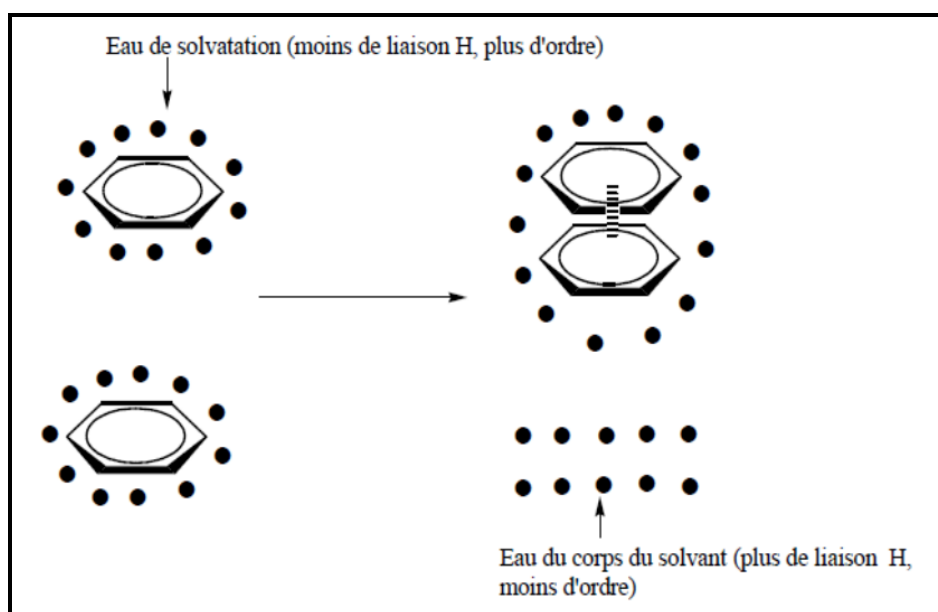


Figure 6. Effet hydrophobe d'un cycle aromatique.

Les molécules d'eau de solvation (fig.6) s'organisent de manière à maintenir entre elle autant de liaison hydrogène que possible, et de ce fait, l'empilement de deux noyaux benzéniques dans l'eau a deux conséquences avantageuses :

- développement d'une forte interaction de dispersion entre les deux noyaux.
- relargage d'une partie des molécules d'eau de solvation dans le corps du solvant. Ce dernier phénomène, appelé « Effet hydrophobe » puisqu'il minimise la surface de contact entre les deux substances en contact (solutés et l'eau), qui se traduit par une augmentation du nombre de liaison hydrogène entre le noyau et la molécule d'eau ($\Delta H < 0$) et une relative désorganisation ($\Delta S > 0$) qui tendent à stabiliser le complexe moléculaire formé entre deux molécules empilées.

La combinaison des interactions de dispersion et de l'effet hydrophobe constitue la principale force motrice pour la complexation moléculaire des phénols dans l'eau.

d) Liaison hydrogène

Les phénols sont des donneurs de liaison hydrogène (liaison H) en raison du caractère acide du proton du groupe OH. Ce sont aussi des accepteurs de liaison H. En fait, seule la paire libre de l'atome d'oxygène qui n'est pas conjuguée avec le cycle et donc est capable d'accepter une liaison H en provenance d'un donneur. Ainsi, un phénol est capable de donner une liaison H et d'en recevoir une seulement. Notons que ces liaisons H se renforcent mutuellement (coopérativité). Par exemple, en donnant une liaison H, le phénol allonge sa liaison OH. Cet état de pré-dissociation accentue la densité électronique sur le centre O et donc favorise son caractère accepteur de liaison H.³¹

e) Acidité

La coupure hétérolytique de la liaison OH (déprotonation) entraîne la formation d'un ion phénate dans lequel la délocalisation électronique de l'atome O vers le cycle aromatique (effet +M) est fortement augmentée (fig.4). Ce phénomène corrélé avec la forte solvation de l'anion phénate par formation de liaison H avec l'eau permettent d'expliquer les propriétés des acides faibles formés par la dissociation des phénols dans l'eau. Les propriétés caractéristiques des phénols (nucléophilie, caractère réducteur, polarisabilité) sont amplifiées lors de la formation des anions phénates correspondants.^{27,31}

Les groupements OH en position *para* et *ortho* des noyaux phénoliques de polyphénols présentent un caractère acide renforcé, ce qui permet une dissociation plus au moins partielle à pH neutre. Cette exaltation de l'acidité est due à la stabilisation de l'ion phénate correspondant par délocalisation de la densité électronique vers le groupement à effet (-M). Elle peut être traduite en termes de formes mésomères.^{27,31}

I.2.1.4. Effet thérapeutique de l'EGCG

Comme il a été défini au début de ce chapitre, La substance antioxydante la plus importante dans le thé vert est l'Epigallocatechine-3-gallate (EGCG). Ce polyphénol puissant de la famille des catéchines a un effet antioxydant très significative sur plusieurs maladies. Comme par exemple, il peut favoriser l'autodestruction des cellules cancéreuses, diminuer le risque de maladie cardiovasculaire et l'obésité¹⁴. On distingue divers effets thérapeutiques de l'EGCG sur l'organisme humain. Parmi ces effets on cite :

a) Effet protecteur contre le cancer

Plusieurs études épidémiologiques suggèrent que les personnes buvant régulièrement du thé vert ont des cancers moins fréquents et moins graves. Il a démontré que l'EGCG agit sur la cyclooxygénase (Cox-2) dans des cellules humaines exerçant un effet protecteur contre le cancer de la prostate. La surexpression de la Cox-2 a été impliquée dans de nombreuses maladies incluant les cancers. Ils ont démontré que l'EGCG inhibe la Cox-2 sans affecter l'expression de la Cox-1 dans les cellules humaines de cancer de prostate, qu'ils soient ou non dépendants des androgènes.³⁴

Dans une autre étude, l'EGCG de thé vert a affecté l'activité et l'expression du PSA (antigène spécifique de la prostate). Celui-ci est capable d'affecter la migration des cellules métastases ou d'autres processus importants du cancer.³⁵

b) Effet positif sur les cellules cérébrales

Le stress oxydant résultant de l'inflammation peut jouer un rôle crucial dans les maladies neurodégénératives. Chez l'animal, après une ischémie unilatérale cérébrale, l'EGCG de thé vert protège des lésions neuronales et de l'œdème cérébral.³⁶

La mort des cellules nerveuses qui se produit dans les maladies de Parkinson ou d'Alzheimer ou dans d'autres maladies neurodégénératives ne résulte pas seulement de lésions oxydatives mais de toute une série de réactions complexes impliquant l'inflammation, le déclin de la

protection neurochimique, l'excès de fer ou une accumulation de protéines dangereuses, comme le bêta amyloïde. L'EGCG pourrait interrompre cette réaction en chaîne et représenter ainsi un agent préventif ou thérapeutique pour les maladies d'Alzheimer ou de Parkinson.³⁷

c) Effet préventif contre l'obésité et le diabète

Des études d'intervention ont montré que la consommation de thé vert pouvait réduire le poids corporel et la masse grasse abdominale. L'EGCG régule les gènes impliqués dans l'oxydation et le stockage des graisses, ainsi que ceux de la signalisation de l'insuline et du métabolisme du glucose.

Des études in vitro ont montré que les catéchines du thé préviennent l'hyperglycémie en augmentant l'activité de l'insuline et en protégeant les cellules β du stress oxydant ont montré que l'EGCG mime l'insuline en diminuant l'expression des gènes contrôlant la néoglucogenèse et donc en diminuant la production du glucose hépatique.^{38, 39}

L'EGCG augmente jusqu'à 40% la thermogénèse. Elle favorise l'oxydation des graisses. Elle augmente l'utilisation des acides gras par les cellules du foie et limite l'activité de l'amylase en diminuant ainsi l'absorption intestinale des glucides.

Des travaux récents ont montré également que l'EGCG diminue l'appétit, le poids corporel, le sucre sanguin et les niveaux d'insuline. Elle inhibe également l'activité de l'amylase. Cette dernière qui est une enzyme digérant l'amidon qui se trouve dans la salive et les intestins. L'amidon étant dégradé plus lentement, l'augmentation du glucose sérique est limitée, réduisant les envies irrésistibles de grignoter entre les repas.⁴⁰

d) Effet protecteur sur le système cardio-vasculaire

Au niveau cardiovasculaire, l'EGCG protège la cellule endothéliale et améliore le flux sanguin chez des patients souffrant d'atteintes coronaires.⁴¹ Elle exerce de nombreux effets vasculaires protecteurs à travers différents mécanismes, incluant des effets antioxydants, anti-inflammatoires, anti-proliférateurs et en abaissant les lipides.

Elle est également capable de réguler le tonus vasculaire. L'EGCG améliore la fonction endothéliale et le flux sanguin chez des patients souffrant de maladies des artères coronaires. Elle exerce de nombreux effets vasculaires protecteurs à travers différents mécanismes, incluant des effets antioxydants, anti-inflammatoires, anti-thrombose, anti-proliférateurs et en abaissant les lipides. Elle est également capable de réguler le tonus vasculaire. L'EGCG active

l'oxyde nitrique synthase endothéliale dans les cellules tapissant les vaisseaux sanguins ou les cellules endothéliales. La libération de l'oxyde nitrique provoque la dilatation des parois des vaisseaux sanguins, ce augmente le diamètre des vaisseaux et améliore le flux sanguin. L'EGCG réduit également l'expression des cytokines cellulaires qui favorisent l'inflammation sous-tendant l'athérosclérose et les maladies cardio-vasculaires. Elle pourrait ainsi inhiber l'inflammation et la prolifération des cellules des muscles lisses dans la paroi des vaisseaux sanguins prévenant le blocage vasculaire.⁴²

I.2.2. Récepteur: Lipases

Les lipases font partie de la classe des hydrolases d'esters carboxyliques (E.C.3.1.1.3.). D'origine bactérienne, fongique, pancréatique, hépatique, gastrique et de propriétés diverses, elles peuvent catalyser l'hydrolyse d'un grand nombre d'esters carboxyliques, mais montrent une forte spécificité envers les glycériques. Les triacylglycérols naturels étant insolubles dans l'eau, les lipases hydrolysent dans les conditions physiologiques les liaisons esters carboxyliques à l'interface de la phase aqueuse dans laquelle l'enzyme est en général initialement soluble.⁴³

Les lipases digestives jouent un rôle fondamental dans le métabolisme lipidique en permettant l'absorption des lipides alimentaires. L'hydrolyse des triglycérides (TG) sous l'action des lipases gastrique et pancréatique, génère des sn-2-monoacylglycérols et des acides gras qui sont absorbés au niveau de l'intestin grêle. Ensuite, ces lipides hydrolysés sont resynthétisés au niveau intestinal et se retrouvent véhiculés par le flux sanguin, sous forme de chylomicrons et de lipoprotéines. La lipoprotéine lipase et la lipase hépatique hydrolysent à leur tour ces lipoprotéines pour procurer de l'énergie aux différents tissus de l'organisme. L'excès des lipides est stocké sous forme de TG dans le tissu adipeux et ils sont mobilisés sous l'action de la lipase hormono-sensible en cas de besoin de l'organisme.⁴⁴

Le rôle physiologique de ces lipases dépend de leur spécificité d'action par rapport aux TG, cette spécificité existe sous trois notions différentes ⁴⁵:

- La régiosélectivité qui représente la préférence à hydrolyser les esters primaires (en position externe sn-1 ou sn-3) ou secondaires (en position interne sn-2) des TG (fig.7);
- La typo-sélectivité ou spécificité par rapport à un type d'acide gras donné ;

- L'énantiosélectivité qui représente la préférence pour un énantiomère d'un produit à hydrolyser plutôt qu'un autre dans le cas d'une molécule chirale.

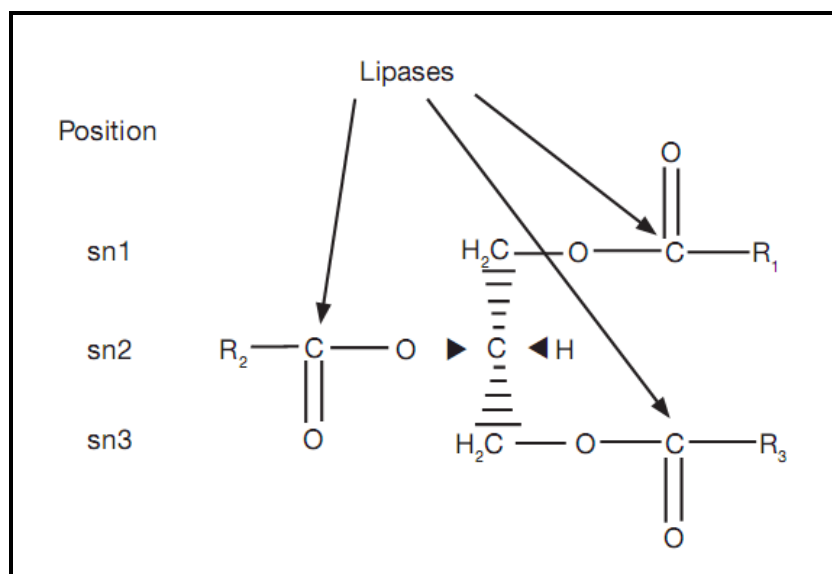


Figure 7. Hydrolyse régiosélective des esters par les lipases.⁴⁶

I.2.2.1. Volet amphiphile et activation interfaciale

La majorité des lipases sont caractérisées par le phénomène d'activation interfaciale⁴⁷. Ce mécanisme d'activation interfaciale est attribué à l'existence d'une boucle amphiphile appelée volet, recouvrant le site actif de l'enzyme en solution. Lors de l'interaction avec une interface lipide/eau, ce volet subirait un réarrangement conformationnel, découvrant ainsi le site actif^{48,49}. Lorsque l'enzyme est au contact d'une interface eau/lipides, ce volet s'écarterait permettant ainsi la fixation du substrat et son hydrolyse.^{50, 51}

La détermination des premières structures de lipases par cristallographie, a révélé la présence de ce volet amphiphile fréquemment rencontré chez les lipases et constitué d'une ou plusieurs hélices α présentant une longueur et une complexité variables selon l'enzyme considérée. La résolution de structure de lipases cristallisées en présence d'inhibiteurs mimant le substrat dans le site actif a permis de mettre en évidence plusieurs conformations intermédiaires de ce volet, aussi appelé « lid » (fig.8)⁵². En conformation qualifiée de « ouverte », la surface hydrophobe exposée est augmentée et le site catalytique devient accessible au substrat.⁵³

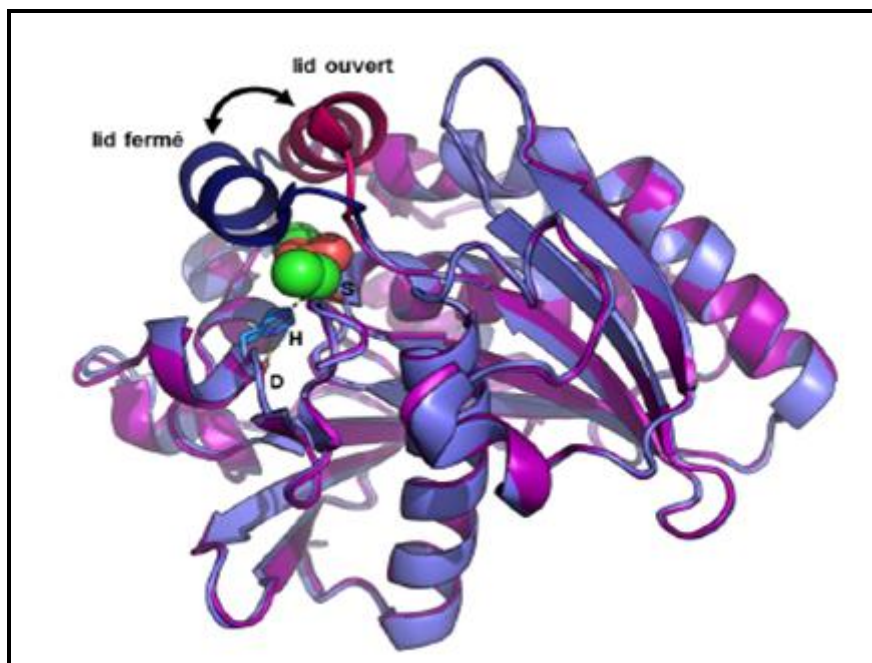


Figure 8. Structure de la lipase de *Rhizomucor miehei* avec le lid en conformation fermée en bleu (PDB ID : 3GTL) et en conformation ouverte en violet (PDB ID : 4GTL).⁵²

Le passage d'une conformation fermée vers une autre ouverte, a fourni une possible explication mécanistique au phénomène d'activation interfaciale mentionné précédemment. Ainsi, en l'absence d'interface hydrophobe, la conformation fermée est prédominante expliquant la faible activité catalytique observée. Lorsque la concentration en substrat est augmentée, et passe au-dessus de la concentration micellaire critique, la formation d'une interface hydrophobe par le substrat insoluble permet une transition de la lipase vers une conformation ouverte ayant une activité catalytique accrue.⁵⁴

I.2.2.2. Caractéristiques structurales

Plus d'une centaine de structures tridimensionnelles de triacylglycérol hydrolases (EC3.1.1.3) sont aujourd'hui disponibles dans la base de données « Protein Data Bank » (PDB). L'analyse de ces structures permet de mettre en évidence certaines caractéristiques communes à ces protéines qui représentent la quasi-totalité des lipases et sont donc considérées comme les lipases « classiques ». Le repliement de deux structures secondaires régulières hélices alpha (α) et feuillets beta (β), de la protéine va constituer la structure tertiaire.

a) Repliement α/β hydrolase

Malgré leurs origines diverses dans le monde vivant, la majorité des lipases connues aujourd'hui possèdent une structure commune, basée sur une succession de feuillets β parallèles et d'un feuillet β antiparallèle reliés entre eux par des hélices α (fig. 9). Ce type de repliement est appelé repliement α/β hydrolase⁵⁵. Il n'est pas spécifique aux lipases et aux enzymes lipolytiques, puisqu'il se retrouve chez d'autres familles de protéines⁵², comme par exemple les protéases et les déshydrogénases⁵⁶. Ces enzymes ont tous en commun un domaine structural central typique formé par huit brins β connectés par six hélices α formant un repliement. Autour de ce domaine central, viennent se greffer diverses structures peptidiques responsables des propriétés catalytiques de l'enzyme ainsi que de sa spécificité de substrat.

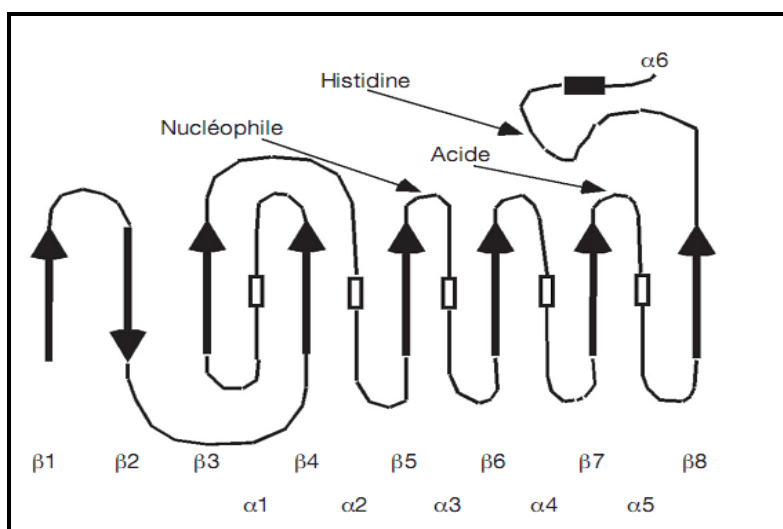


Figure 9. Schéma du repliement α/β . Les flèches représentent les feuillets β et les rectangles les hélices α . Le rectangle noir représente le volet amphiphile.⁵⁷

b) Triade catalytique

Le site catalytique apparaît toujours sous la forme d'une triade hautement conservée : sérine / acide aspartique / histidine (Ser/Asp/His ou S/D/H). C'est un élément très conservé dans la structure des lipases. Pour chaque sous-famille d'hydrolases, l'ordre d'apparition de ces résidus aminés dans la séquence peptidique est fixe: le résidu nucléophile (la sérine), précède les deux autres résidus et se positionne après le feuillet $\beta 5$. Le résidu acide, l'acide aspartique, apparaît très souvent après le feuillet $\beta 7$ et enfin, l'histidine se place toujours après le dernier feuillet⁵⁸. L'acide aspartique est parfois remplacé par un acide glutamique. Les brins (1 à 8) sont représentés par des flèches bleues et les hélices (A à F) par des cylindres

rouges (fig.10). Le positionnement relatif des acides aminés de la triade catalytique est indiqué par des points rouges.

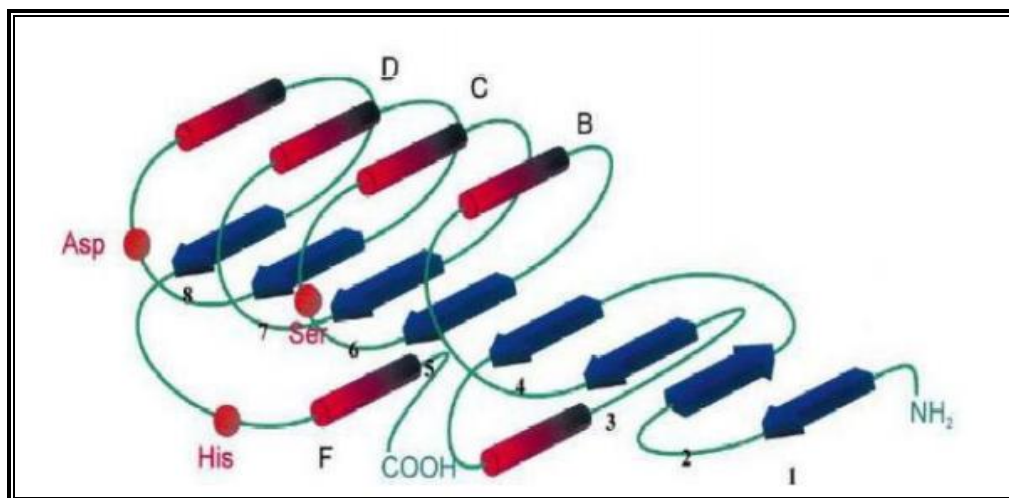


Figure 10. Représentation de la structure secondaire du repliement α/β des hydrolases⁵⁹.

La sérine catalytique est toujours située dans un coude nucléophile situé à l'extrémité C-terminale du cinq brin β et immédiatement suivi d'une hélice α (fig.10). L'histidine de la triade catalytique est située à l'extrémité carboxy-terminale du dernier brin du feuillet constituant le repliement α/β tandis que l'acide carboxylique est généralement situé à l'extrémité du septième brin, sauf pour la lipase pancréatique où il est situé à l'extrémité du sixième brin.^{52, 61}

L'acide glutamique va jouer le rôle de résidu acide et l'histidine, le rôle de résidu basique alors que la sérine sera le résidu nucléophile. Ainsi l'aspartate va former une liaison hydrogène avec le groupe imidazole de l'histidine permettant à celle-ci d'adopter une orientation favorable en direction de l'oxygène de la sérine. De plus, l'interaction avec l'aspartate va modifier le pKa du groupement imidazole de l'histidine. Durant l'étape de catalyse, le groupement imidazole de l'histidine rend l'oxygène de la sérine plus nucléophile et permet l'attaque du carbonyle du substrat. La conservation de l'aspartate / glutamate souligne l'importance de l'interaction avec l'histidine pour que cette dernière puisse jouer son rôle d'accepteur de proton⁶¹.

c) Mécanisme catalytique

Le mécanisme d'action des lipases a été exploité à partir de celui proposé dans le cas de la chymotrypsine⁶². Ce mécanisme réactionnel se divise en deux étapes principales :

➤ Etape d'acylation

L'acylation provient de l'attaque nucléophile d'une des fonctions ester carboxylique du substrat par la sérine de la triade catalytique (fig.11). Le groupement hydroxyle de la sérine activé par le réseau de relais de charge Ser-His-Asp, attaque le carbone de la fonction ester carboxylique du substrat pour former un premier intermédiaire tétraédrique (fig.11, chemin A). Des liaisons hydrogènes entre l'oxygène chargé négativement (oxyanion) de l'intermédiaire tétraédrique et des groupes N-H peptidiques stabilisent la charge négative de l'oxyanion (fig.11, chemin B). Au cours de cette catalyse acido-basique, la liaison ester carboxylique est clivée et le groupement alcool partant (le diacylglycérol lorsque le substrat est un triacylglycérol) emporte un proton de l'ion imidazolium du résidu histidine. La chaîne acyle de la liaison ester du substrat reste liée de façon covalente à l'enzyme sous forme d'un intermédiaire acyl-enzyme (fig.11, chemin C).

➤ La désacylation

La désacylation provient de l'attaque nucléophile d'une molécule d'eau sur cet intermédiaire acyl-enzyme pour donner un deuxième intermédiaire tétraédrique (fig.11, chemin D). Ce dernier se dissocie en libérant un acide carboxylique (acide gras) et en restituant la lipase dans son état initial (fig.11, chemin E).⁶³

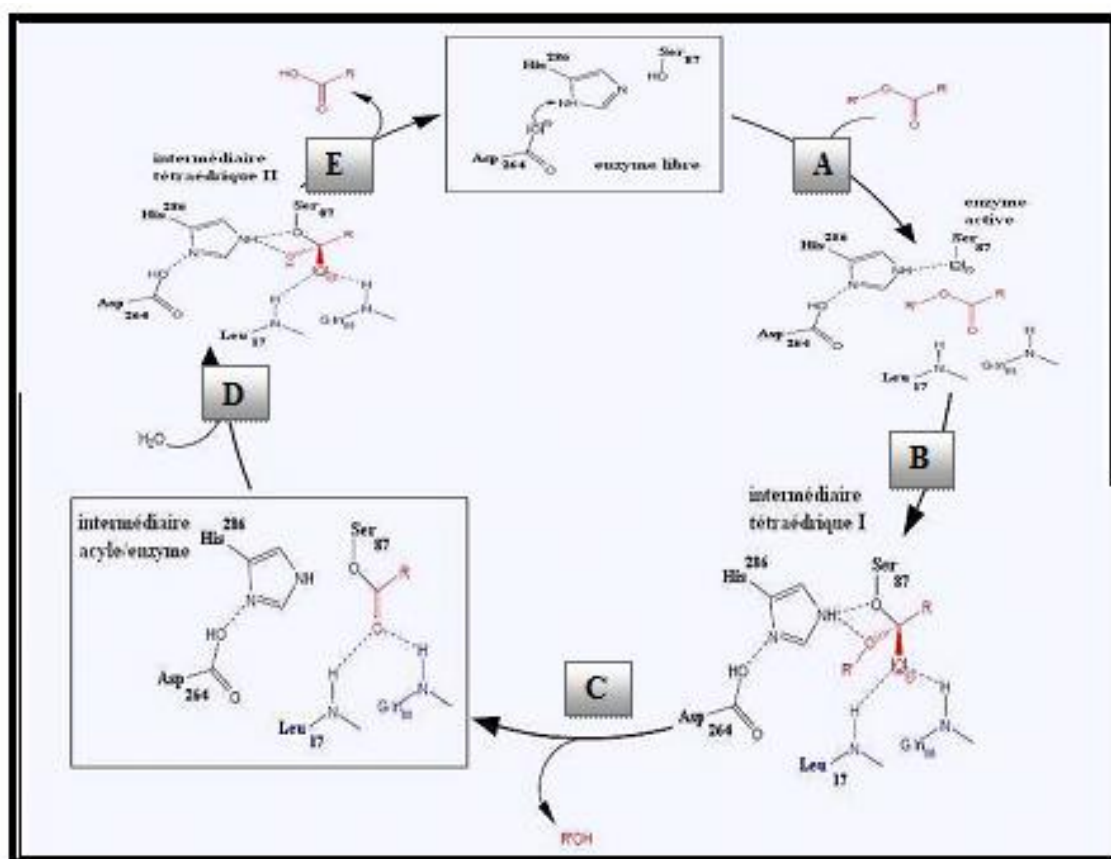


Figure 11. Mécanisme catalytique de l'hydrolyse d'un ester par une enzyme lipolytique.⁶³

I.2.2.3. La Lipase *Pancréatique du Porc* (LPP)

La lipase *pancréatique de porc* possédait la particularité de « s'activer à l'interface » huile-eau⁶⁴. Alors que l'enzyme est faiblement active sur un substrat soluble (triglycérides à courtes chaînes à faible concentration), elle multiplie son activité de 2 à 3 ordres de grandeur lorsque le substrat s'agrège et forme une émulsion au-delà de sa limite de solubilité. Les lipases (triacylglycerol hydrolases, EC 3.1.1.3) ont alors été définies comme une famille particulière d'estérases, actives sur un substrat insoluble, les triglycérides, et possédant la propriété cinétique d'activation interfaciale.⁶⁴

Cet enzyme est une glycoprotéine comportant 449 aminoacides, dont les chaînes glycanes sont liées à l'asparagine (Asn) en position 166 ; elle comporte deux domaines séparés par la liaison Phe-Ala⁶⁵. Chaque domaine porte un site de reconnaissance, à savoir : un site de reconnaissance interfaciale (domaine N-terminal), site de l'hydrolyse proprement dite et un site de reconnaissance pour son partenaire protéique (domaine C-terminal), la colipase. Le domaine N-terminal (résidus 1-335), qui porte le centre actif de l'enzyme, est séparé du domaine C-terminal (résidus 336- 449), par une zone étranglée très résistante à la protéolyse.

Cette organisation en deux domaines répond aux deux fonctions spécifiques énoncées ci-dessus : le domaine N- terminal est responsable de la catalyse, alors que le domaine C-terminal est impliqué dans la reconnaissance de la colipase⁶². La colipase est une petite molécule (10 kDa), très réticulée du fait de la présence de cinq ponts disulfures. Elle porte trois sites de reconnaissance essentiels à sa fonction, à savoir : un site de reconnaissance interfaciale⁴⁹. Ces trois sites sont typologiquement distincts (fig.12).

La LPP (fig.12) est représentée en surface jaune-orange, les résidus hydrophobes (Val, Leu, Ile, Tyr, Trp, Met, Pro, Ala, Phe) sont colorés en blanc, les molécules des détergents sont représentées par des sphères vertes et la colipase est entourée par un liseré rouge. La figure 12 schématise aussi le volet et les boucles $\beta 9$ et $\beta 5$, qui sont importants pour la reconnaissance du substrat.

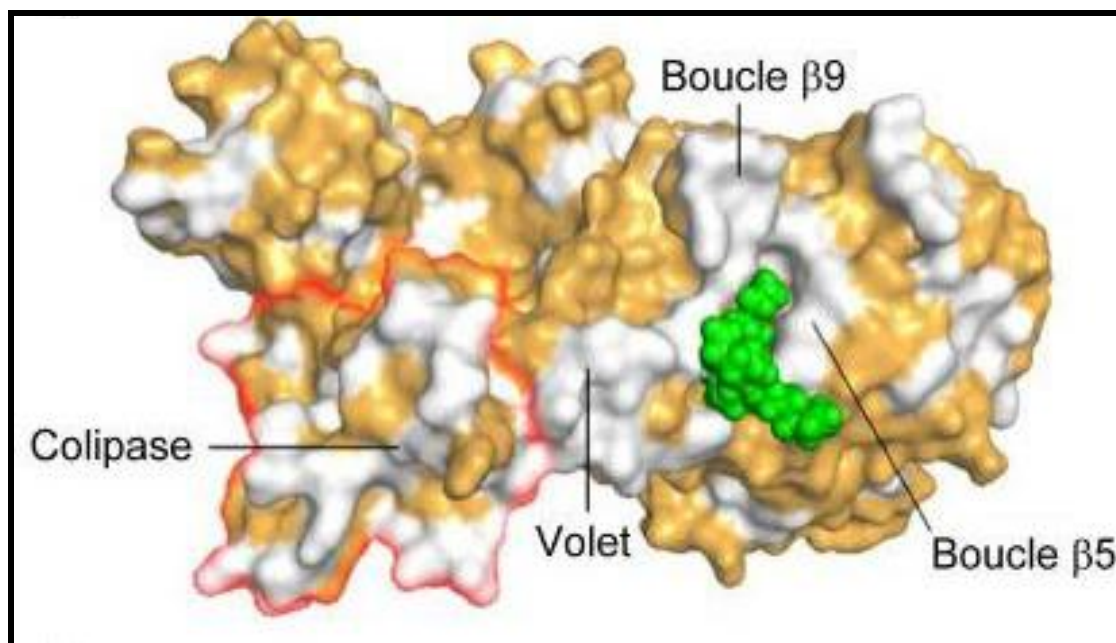


Figure 12. Structure de la lipase pancréatique porcine complexée avec une molécule de tétraéthylène glycol mono-octyle dans son site actif. PDB: 1ETH.⁶⁷

I.2.2.4. La lipase pancréatique humaine (LPH)

Les lipases pancréatiques humaines sont également des glycoprotéines de masses moléculaires d'environ 50 kDa, très conservées chez les mammifères. L'obtention de données cristallographiques a permis une étude précise de leur structure et leur mode de fonctionnement.⁵⁷

Elles sont constituées de deux domaines protéiques. Le premier domaine N-terminale, constitué de 335 acides aminés, possède un repliement α/β hydrolase et contient le site actif. Ce dernier est recouvert par un volet amphiphile mobile de 23 acides aminés. Le domaine C-terminale, constitué de 114 acides aminés, possède une structure de type b-sandwich et interagit avec un cofacteur protéique de 10 kDa, la colipase (fig.13).⁶⁸

La LPH a été obtenue sous sa forme ouverte, complexée avec la colipase porcine, en orange. Le volet est représenté en bleu, la boucle $\beta 9$ en rose, la sérine catalytique en rouge et les ponts disulfures en jaune. Dans l'intestin grêle, la LPH doit agir sur les TAG en présence des sels biliaires qui sont des compétiteurs pour l'adsorption à l'interface huile-eau. Un cofacteur spécifique de 10 kDa également présent dans le suc pancréatique, la colipase, permet de contrecarrer l'effet inhibiteur des sels biliaires en ancrant la LPH à l'interface sous la forme d'un complexe stœchiométrique avec la colipase.

La LPH est une enzyme régiosélective hydrolysant seulement les fonctions esters en positions sn-1 et sn-3 des triacylglycérides (TAG), et générant des monoacylglycérol (2-MAG). Elle ne possède pas d'activité significative sur les phospholipides les esters de cholestérol et les galactolipides.⁶⁹

Les formes ouvertes des lipases pancréatiques exposent d'importantes surfaces hydrophobes au solvant. Ces dernières sont composées de résidus positionnés à proximité du site actif et sur la face intérieure du volet⁴⁹. La colipase, une fois ancrée sur le domaine C-terminale de la lipase pancréatique, expose à son tour une surface hydrophobe à proximité du volet. L'ensemble constitue un véritable plateau lipophile de 50 Å capable d'interagir fortement avec l'interface lipide-eau (fig. 13).

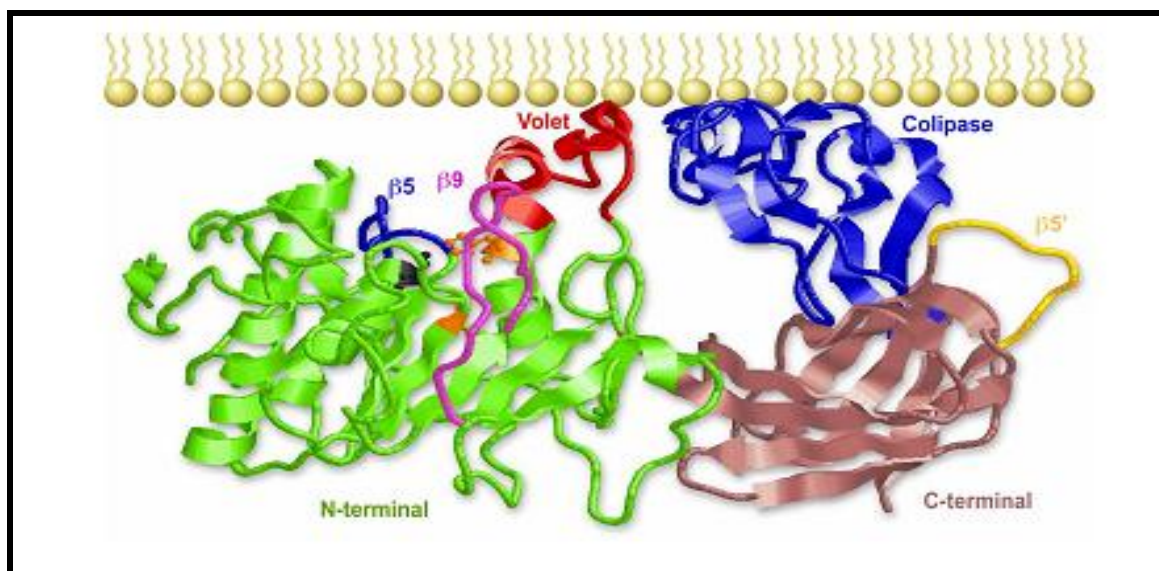


Figure 13. Structure cristallographique du complexe LPH/colipase en modèle de ruban ; PDB : 1LPB.⁵⁷

Dans cette représentation, La colipase est représentée en bleu. Les trois boucles essentielles ($\beta 5$, $\beta 9$ et volet) de la LPH entourant le site actif sont indiquées en bleu, violet et rouge respectivement. La boucle $\beta 5'$ (résidus 405-414) du domaine C-terminal est colorée en jaune.

I.2.2.5. La lipase gastrique

La lipase gastrique, sécrétée par la muqueuse gastrique, hydrolyse les lipides alimentaires dans l'estomac. Elle ne nécessite pas la présence de cofacteurs pour être active ; elle est stable et active à des valeurs de pH proche de 1. Cette lipase possède une certaine résistance à la pepsine ainsi qu'aux protéases gastriques.⁷⁰

La lipase gastrique humaine est une protéine glycosylée de masse moléculaire d'environ 50 kDa qui hydrolyse préférentiellement les positions sn-1 et sn-3 des TAG à chaînes longues contenus dans le bol alimentaire⁷¹. Elle facilite ainsi l'action ultérieure de la lipase pancréatique humaine sécrétée au niveau du duodénum^{72,73} (fig.14). La lipase gastrique conserve son activité jusqu'à l'entrée de l'intestin, permettant ainsi de compenser en partie l'absence de la LPH dans les cas d'insuffisances pancréatiques.^{72,74}

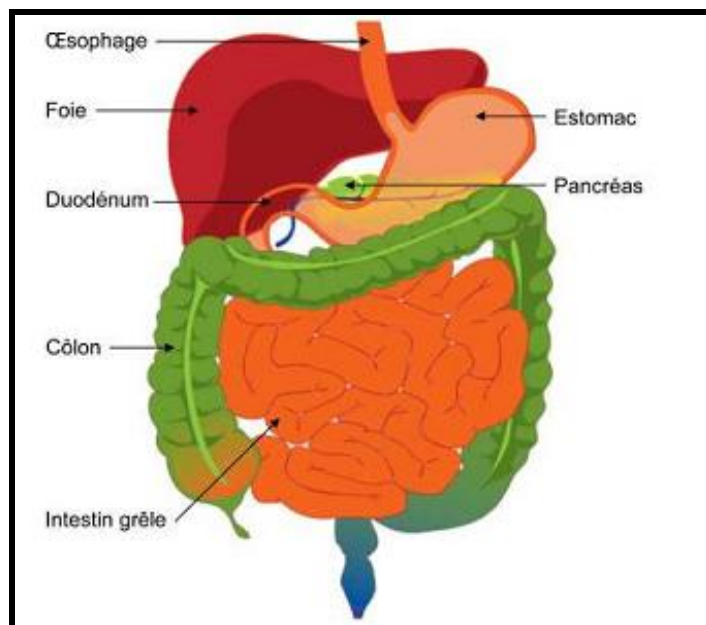


Figure 14. Le tractus gastro-intestinal humain.⁶⁸

La lipase gastrique du chien (LGC) et la lipase gastrique humaine (LGH) ont été exprimées sous formes recombinantes⁷⁶⁻⁷⁸, ce qui a permis l'obtention de leur structure cristallographique⁷⁸⁻⁸⁰. Elles sont composées d'une structure globulaire dont le repliement α/β , au centre, est entouré par six hélices α . Le site actif, composé d'une triade catalytique Ser-Asp-His, est recouvert d'un volet bloquant son accès.

La LGH a été résolue sous sa forme fermée à une résolution de 3 Å, tandis que la LGH a été obtenue sous sa forme ouverte, à une résolution de 2,7Å, avec un inhibiteur alkyl-phosphonate dans son site actif. La structure de la LGH a révélé que les quatre sites potentiels de N-glycosylations étaient occupés⁷⁹. Des études de mutagenèses dirigées ont montré que la glycosylation protégeait la LGH de l'action protéolytique de la pepsine, une protéase sécrétée au niveau de l'estomac.⁸¹ La grande identité de séquence entre la LGH et la LGC a permis de comparer leurs structures respectives, mettant en évidence un réarrangement structural du volet comptant parmi les plus complexes observés chez les lipases (fig.15).^{82, 68}

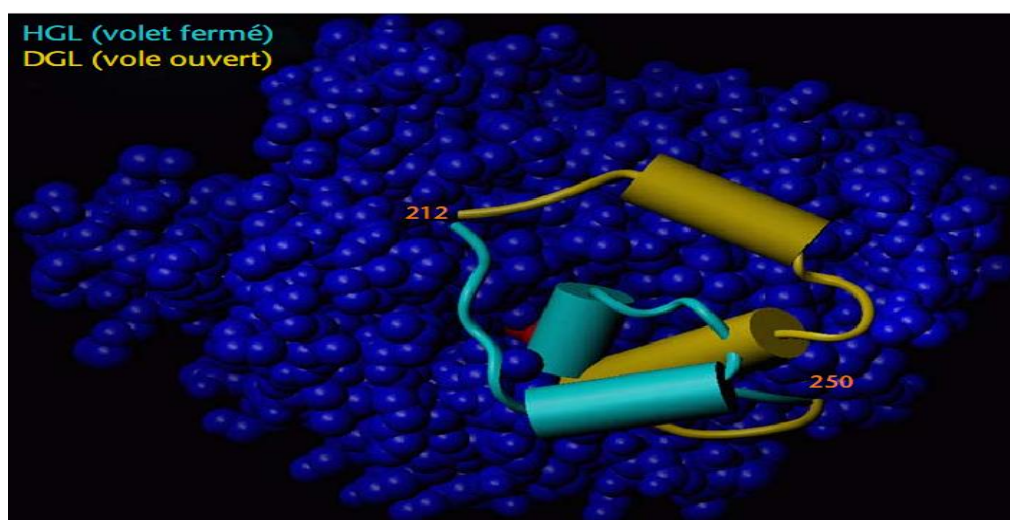


Figure 15. Structures 3D superposées de la lipase gastrique humaine (LGH) et de la lipase gastrique de chien (LGC).

L'activité à pH acide (4-6) de la LGH était liée à une adsorption plus importante de l'enzyme à l'interface lipide-eau. De plus, sa forme ouverte serait plus stable dans cette gamme de pH grâce à la formation d'un pont salin entre deux résidus chargés, l'un appartenant au volet (Glu255) et l'autre à la partie N-terminale de la protéine (Lys4).⁸³

Les principales propriétés biochimiques et structurales des lipases gastrique et pancréatique humaines sont résumées dans le tableau 2.

Tableau 2. Principales propriétés biochimiques et structurales de la LGH et la LPH.

Propriétés	LGH	LPH
MM (kDa)	50 (379 acides aminés)	48 (449 acides aminés)
Glycosylation	La LGH est une glycoprotéine (15 % de chaînes glycanique) avec 4 sites potentiels de glycosylation.	La LPH est une glycoprotéine. La chaîne glycanique est portée par Asn166.
Origine cellulaire	La LGH est sécrétée par les cellules principales du fundus	La LPH est produite par les cellules acineuses du pancréas où elle est stockée dans les granules de zymogènes
Mg de lipase sécrétée /repas liquide (cas d'un individu sain)	21,6 ± 14	253,5 ± 95
Présence de cofacteur	Sans cofacteur	Elle nécessite un cofacteur spécifique, la colipase.

Stabilité vis-à-vis aux pH	La LGH est stable et active à pH acide ($1 < \text{pH} < 5$). Cet enzyme présente un pH optimum d'action de 5,4 en utilisant une émulsion d'Intralipide TM comme substrat.	La LPH est complètement dénaturée et inactivée dans des conditions optimales de pH pour la LGH. In vitro, le pH optimum d'hydrolyse des TG à chaînes courtes ou longues est de 8,0. En présence de sels biliaries, son pH optimum d'action est de 6,5.
Effet des sels biliaries sur l'activité lipasique	Les sels biliaries semblent agir comme activateurs de la LGH lorsque l'on utilise de la TC4 comme substrat. Toutefois, les sels biliaries inhibent l'activité de la LGH en hydrolysant l'émulsion d'intra-lipide.	Les sels biliaries sont de puissants inhibiteurs de l'activité de la LPH. La colipase permet à la LPH d'agir en présence des sels biliaries.
Ponts disulfures et Résidus SH libres	La LGH comporte un pont disulfure et une cystéine libre.	La LPH comporte six ponts disulfures et deux cystéines libres
Structure primaire	La LGH appartient à la famille des lipases acides de mammifères incluant les lipases préduodénales et lysosmales. Ces dernières ne présentent aucune homologie de séquences avec les autres familles de lipases.	La LPH constitue un membre de la superfamille des lipases pancréatiques qui est subdivisée en plusieurs sous-familles selon les homologies des séquences en acides aminés. Cette superfamille comporte aussi des enzymes lipolytiques et des protéines non enzymatiques.
Structure tertiaire	La LGH se présente sous la forme d'un seul domaine globulaire. On distingue le corps central ou « core », qui est constitué de 8 feuillets β et qui est recouvert par un domaine d'extrusion ou « cap ». Ce dernier est formé par 3 hélices α couvrant la sérine catalytique lorsque l'enzyme se trouve dans sa forme fermée.	La structure tridimensionnelle de la LPH montre la présence de deux domaines N- et C terminaux le domaine N-terminal présente un repliement de type α / β et contient le site actif. Alors que le domaine C-terminal possède une structure de type β sandwich et permet la fixation de la colipase.
Synergie entre les deux lipases	Des expériences réalisées in vitro, suggèrent fortement que l'action de la LGH conditionne l'activité ultérieure de la lipase pancréatique, grâce aux acides gras libérés. In vivo, 10 à 30 % des TG alimentaires sont hydrolysés au niveau de l'estomac. La préhydrolyse gastrique génère un substrat directement hydrolysable par la lipase pancréatique. En utilisant une émulsion d'Intralipide TM comme substrat, les acides gras à longues chaînes inhibent la lipolyse qui est catalysée par la LGH. Par contre ces acides gras libérés, grâce à la lipase gastrique, permettent l'activation rapide du complexe lipase-colipase pancréatique (Gargouri et al., 1986 c). Il existe donc une synergie entre la lipase gastrique et la lipase pancréatique humaine.	

I.3. Conclusion

Les lipases appartiennent à la famille des enzymes lipolytiques qui catalysent l'hydrolyse d'un grand nombre d'esters d'acides gras plus particulièrement les triacylglycérols. Elles sont également des cibles thérapeutiques intéressantes pour le traitement de diverses pathologies comme l'obésité. L'une des approches développées pour lutter contre ces maladies est l'inhibition de ces enzymes par des composés naturels et/ou de synthèses.

C'est dans ce contexte que nous nous sommes intéressés à appliquer cette approche théorique de modélisation pour étudier l'effet inhibiteur d'un composé naturel qui est le (-)-Epigallocatechin-3-gallate (EGCG), un polyphenol naturel présent dans le thé vert et qui appartient à la famille des flavonoïdes. Cette molécule a été testée dans ce travail comme inhibiteur des lipases digestives.

Au-delà de l'aspect d'inhibition de la lipase digestive, les recherches concernant l'étude des interactions mises en jeu entre la lipase et l'inhibiteur, plusieurs méthodes de recherche utilisant l'ordinateur ont été créées. Ces méthodes ont été développées pour aider les médecins chimistes à comprendre les interactions, au niveau moléculaire, qui est la base de la plupart des mécanismes biologiques.

Parmi les méthodologies employées, celle de docking moléculaire est de plus en plus utilisée. La présentation de la technique de docking sera développée dans le chapitre 2.

Références Bibliographiques

- 1: Faber K. Biotransformations in Organic Chemistry, 4th Edition, Ed. Springer-Verlag Berlin New York 2000: p 5.
- 2: Gotor-Fernandez V, Brieva R, Goto V R.J. Mol. Catal B : Enzymatic 2006; (40): 111-120.
- 3 : Patrick F, Jacqueline D, Philippe T. Les lipases sont des hydrolases atypiques : principales caractéristiques et applications. *Biotechnologie, Agronomie, Société et Environnement*, 2008 ; 12(2): 119-130.
- 4: Carey MC, Small DM, Bliss CM. Lipid digestion and absorption. *The annual review of physiology*, 1983; 45: 651-677.
- 5: Favé G, Coste TC, Armand M. Physicochemical properties of lipids: new strategies to manage fatty acid bioavailability. *Cellular and Molecular Biology*, 2004; 50: 815-831.
- 6 : Fave G, Peyrot J, Hamosh M, Armand M. Digestion des lipides alimentaires: intérêt de la lipase gastrique humaine. *Cahiers de Nutrition et Diététique*, 2007; 42(4): 183.
- 7: <https://www.google.com/patents/WO1998030588A1>
- 8: Tcheriatchoukine D. L'obésité: découverte récentes relatives aux mécanismes moléculaires à l'origine de nouvelles stratégies thérapeutiques. Thèse de doctorat, Université Henri Poincaré – Nancy 2010.
- 9: Sanchez-Moreno C. Review: methods used to evaluate the free radical scavenging activity in foods and biological systems. *Food Science and Technology International*, 2002; 8 (3): 121-137.
- 10: Marc F, Davin A, Deglène-Benbrahim L, Ferrand C. Méthodes d'évaluation du potentiel antioxydant dans les aliments. *Erudit, M/S: medicine sciences*, 2004; 20(4): 458-463.
- 11: Huang D, Ou B, Prior RL. The chemistry behind antioxidant capacity assays. *Journal of Agricultural and Food Chemistry*, 2005; 53: 1841-1856.
- 12: Chebil L. Acylation des flavonoïdes par les lipases de *Candida antarctica* et de *Pseudomonas cepacia* : études cinétique, structurale et conformationnelle. Thèse de doctorat Institut national polytechnique de lorraine 2006.

13 : Nutra news. Science, Nutrition, Prévention et Santé. Édité par la Fondation pour le libre choix. www.nutranews.org.

14: Chu C, Deng J, Man Y, and Yili Q. Green Tea Extracts Epigallocatechin-3-gallate for Different Treatments. *Biomedical Research International*, 2017; 20(17): 5615-5647.

15: Dr rath. Substances phytobiologiques (Phytobiologicals) 5ème Partie:L'Epigallocatechine gallate (EGCG). Lettre d'information sur la Santé 2010, partie 5. www.phytobiologicals.com

16: Duhovny D, Nussinov R, Wolfson HJ. Efficient unbound docking of rigid molecules 2002.

17: De Oliveira EB. Simulation moléculaire appliquée à l'acétylation de flavonoïdes catalysés par des lipases : influence des structures de la lipase et des flavonoïdes sur la régiosélectivité de la bioconversion. Thèse de doctorat d'université : Procédés biotechnologiques et alimentaires. Nancy Institut National Polytechnique de Lorraine. France 2009: 187.

18: Min KJ, Taeg KK. Anticancer effects and molecular mechanisms of epigallocatechin-3-gallate. Department of Immunology. School of Medicine, Keimyung University Daegu, Korea 2014: 16-24.

19: Santiago SL, Osorio SL, Neri JR, Carvalho RM, Toledano M. effect of the flavoid Epigallocatechin-3-gallate on resin-dentin bond. *The journal of adhesive dentistry*, 2013; 15: 535-40.

20: Kyoung-Jin M, Taeg-Kyu K. Anticancer effects and molecular mechanisms of epigallocatechin-3-gallate. Department of Immunology, School of Medicine, Keimyung University Daegu, Korea 2014: 16-24.

21: Yang XR, Ye CX, Xu JK, Jiang YM. Simultaneous Analysis of Purine Alkaloids and Catechins in Camellia Sinensis, Camellia Ptilophylla and Camellia Assamica Var. Kucha by HPLC. *Food Chemistry*, 2007; 100: 1132–1136.

22: David Keighley (ed.), *The origins of Chinese civilization*, Berkeley, University of California Press, 1983 Li Hui-li « The domestication of plants in China: ecological considerations » p. 21-63.

- 23:** Guang-Jian D, Zhiyu Z, Xiao-Dong W, Chunhao Yu, Tyler C, Chun-Su Y, Chong-Zhi W. Epigallocatechin Gallate (EGCG) is the Most Effective Cancer Chemopreventive Polyphenol in Green Tea. *Nutrients*, 2012; 4:1679-1691.
- 24:** Schneider R, Lüdde T, Töpfer S, Imming P. Tee gegen den Lärm der Welt Pharmazeutische Zeitung Online, Ausgabe 2008.
- 25:** Manach C, Scalbert A, Morand C, Rémésy C, Jiménez L. Polyphenols: food sources and bioavailability. *The American Journal of Clinical Nutrition*, 2004; 79: 727–747.
- 26:** Sweetman S C. Martindale the Complete Drug Reference. Thirty-third Edition, London-Chicago : *Pharmaceutical Press*, 2002: 1681.
- 27:** Nkhili EZ. Polyphenols de l'alimentation: extraction, interaction avec les ions de fer et du cuivre, oxidation et pouvoir antioxydant. Diplôme de doctorat. Université d'Avignon-France et Université Cadi Ayyad-Marrakech 2009.
- 28:** Philip Denton, Chris Rostron. *Pharmaceutics: the science of medicine design*. Oxford university press 2013: 90.
- 29:** Sarni-Manchado P, Cheynier V. Les polyphénols en agroalimentaire. Lavoisier,ed Tec & Doc 2006.
- 30:** Dangles O. The physic-chemical properties of polyphenols. *Agro-food industry Hi-tech* Lavoisier 2006, 29-50.
- 31:** Boubekri C. Etude de l'activité antioxydante des polyphénols extraits de *Solanum melongena* par des techniques électrochimiques. Thèse de doctorat. Université Mohamed Khider – Biskra 2014.
- 32:** Cos P, Ying L, Calomme M, Hu JP, Cimanga K, Van-Poel B, Pieters L, Vlietinck AJ, Van Den BD. Structure-activity relationship and classification of flavonoids as inhibitors of xanthine oxidase and superoxyde scavengers. *Journal of natural products*, 1998; 61: 71-76.
- 33:** Foti M, Piattelli M, Baratta MT, Ruberto G. Flavonoids, coumarins, and cinnamic acids as antioxidants in a micellar system. Structure-activity relationship. *Journal of Agricultural and Food*, 1996; 44: 497-501.
- 34:** Hussain T. Green tea constituent epigallocatechin-3-gallate selectively inhibits Cox-2 without affecting Cox-1 expression in human prostate cancer cells. *International Journal of Cancer*, 2005; 113: 660–669.

- 35:** Pezzato E, Sartor L, Dell'Aica I, Dittadi R, Gion M, Belluco C, Lise M, Garbisa S. Prostate carcinoma and green tea: PSA-triggered basement membrane degradation and MMP-2 activation are inhibited by (-) Epigallocatechin-3-gallate. *International Journal of Cancer*, 2004; 112(5): 787-792.
- 36:** Lee H. Protective effect of green tea polyphenol EGCG against neuronal damage and brain edema after unilateral cerebral ischemia in gerbils. *Journal of Neuroscience Research* , 2004; 77(6): 892-900.
- 37:** Darch E. l'alimentation idéale des jeunes. Edition lanore, Francois-Xavier sorlot, Paris 2010: 143.
- 38:** Bolling BW, Chen CY, Blumberg JB. Tea and health: preventive and therapeutic usefulness in the elderly? *Current Opinion in Clinical Nutrition and Metabolic Care*, 2009; 12(1): 42-48.
- 39:** Waltner-Law ME, Wang XL, Law BK, Hall RK, Nawano M, Granner DK. Epigallocatechin gallate a constituent of green tea, represses hepatic glucose production. *The Journal of Biological Chemistry*, 2002; 277 (38): 34933-34940.
- 40:** Kreydiyyeh SI. Tea extract inhibits intestinal absorption of glucose and sodium in rats. *Comparative Biochemistry and Physiology Part C. Pharmacol. Toxicol. Endocrine*, 1994; 108: 359-365.
- 41:** Widlansky ME. Acute EGCG supplementation reverses endothelial dysfunction in patients with coronary artery disease. *Journal of American College of Nutrition*, 2007; 26(2): 95-102.
- 42:** Khurana S, Krishnan V, Hollingsworth A, Piche M, Tai TC. Polyphenols: Benefits to the Cardiovascular System in Health and in Aging *Nutrients* 2013; 5, 3779-3827.
- 43:** Linden G, Transformation des Produits Alimentaires par les Enzymes, Ed. Techniques Ingénieur, 1998.
- 44 :** Morak M , Schmidinger H , Riesenhuber G , Rechberger GN. , Kollrosier M , Haemmerle G, Zechner R , Kronenberg F , Hermetter A. Adipose triglyceride lipase (ATGL) and hormone-sensitive lipase (HSL) deficiencies affect expression of lipolytic activities in mouse adipose tissues. *Molecular & Cellular Proteomics* , 2012; 11(12): 1777–1789.

- 45: Jensen RG, Galluzzo DR, Bush VJ. Selectivity is an important characteristic of lipases (acylglycerol hydrolases). *Biocatalysis*, 1990; 3: 307-316.
- 46: Muralider RV, Chiromamilla RR, Marchan R, Ramachandran VN, Ward OP, Nigan P. Understanding lipase stereo selectivity. *World journal of microbiology & biotechnology*, 2002; 18: 81-97.
- 47: Schmid RD, Verger R. Lipases: Interfacial enzymes with attractive applications. *Angew. Chem. Int. Ed* 1998; 37: 1608–1633.
- 48: Khan FI, Lan D , Durrani R, Weiqian H, Zexin Z, Yonghua W. The lid domain in lipase: Structural and functional determinant of enzymatic properties. *Frontier in Bioengineering and Biotechnology*, 2017; 5: 16.
- 49: Van Tilbeurgh H, Egloff MP, Martinez C, Rugani N, Verger R, Cambillau C. Interfacial activation of the lipase-procolipase complex by mixed micelles revealed by X-Ray crystallography. *Nature*, 1993; 362(6423): 814-820.
- 50: Brzozowski AM, Derewenda U, Derewenda ZS, Dodson GG, Lawson DM, Turkenburg J, Bjorkling F, Huge-Jensen B, Patkar SA, Thim L. A model for interfacial activation in lipases from the structure of a fungal lipase-inhibitor complex. *Nature*, 1991; 351: 491-494.
- 51: Brzozowski AM, Savage H, Verma CS, Turkenburg JP, Lawson DM, Svendsen A, Patkar SA. Structural origins of the interfacial activation in *Thermomyces (Humicola) lanuginosa* lipase. *Biochemistry*, 2000; 39: 15071-15082.
- 52: Derewenda ZS, Derewenda U, Dodson GO. The crystal and molecular structure of the *Rhizomucor miehei* triacylglyceride lipase at 1.9 Å resolution. *Journal of Molecular Biology*, 1992; 227:818-839.
- 53: Grochulski P, Li Y, Schrag JD, Bouthillier F, Smith P, Harrison D, Rubin B, Cygler M.. Insights into interfacial activation from an open structure of *Candida rugosa* lipase. *Journal of Molecular Biology*, 1993; 268: 12843-12847.
- 54: Guillaume N. Dynamique fonctionnelle des protéines: études d'une lipase et d'une protéine A de la membrane externe de bactérie. thèse de doctorat. Université Toulouse 2015.
- 55: Ollis DL, Cheah E, Cygler M, Dijkstra B, Frolow F, Franken SM, Harel M, Remington SJ, Silman I, Schrag J. The alpha/beta hydrolase fold. *Protein Engineering*, 1992; 5(3): 197-211.

- 56:** Franken SM, Rozeboom HJ, Kalk KH, Dijkstra BW. Crystal-structure of Haloalkane Dehalogenase: An enzyme to detoxify Halogenated. Alkanes. *Embo Journal*, 1991; 10(6): 1297-1302.
- 57:** Egloff MP, Sarda L, Verger R, Cambillau C, Van Tilbeurgh H. Crystallographic study of the structure of colipase and of the interaction with pancreatic lipase. *Protein Science*, 1995; 4: 44-57.
- 58:** Schrag JD, Li YG, Wu S, Cygler M. Ser-His-Glu triad forms the catalytic site of the lipase from *Geotrichum candidum*. *Nature*, 1991; 351: 761-764.
- 59:** Bornscheuer UT. Microbial carboxyl esterases: classification. Properties and application in biocatalysis. *FEMS Microbiology Reviews*, 2002; 26(1):73-81.
- 60:** Cygler M, Schrag JD. Structure and conformational flexibility of *Candida rugosa* lipase. *Biochimica Biophysica Acta*, 1999; 1441: 205-214.
- 61:** Dodson G, Wlodawer A. Catalytic triads and their relatives. *Trends Biochemistry Science*, 1998; 23(9): 347-352.
- 62:** Blow DM, Birktoft JJ, Hartley BS. Role of a buried acid group in the mechanism of action of chymotrypsin. *Nature*, 1969; 221: 334-337.
- 63:** Beer H, Wohlfahrt G, McCarthy JE, Schomburg D, Schmid RD. Analysis of the catalytic mechanism of a fungal lipase using computer-aided design and structural mutants. *Protein*, 1996; 6: 507-517.
- 64:** Sarda L, Desnuelle P. Actions of pancreatic lipase on esters in emulsions. *Biochimica Biophysica Acta*, 1958; 30(3): 513-521.
- 65:** Bousset-Risso EA. Limited proteolysis of porcine pancreatic lipase. Lability of the Phe 335 - Ala 336 bond towards chymotrypsin. *FEBS LETTERS*, 1985; 182(2): 323-326.
- 66:** Abousalham C, Chaillan B, Kerfelec E, Foglizzo C, Chapus. Uncoupling of catalysis and colipase binding in pancreatic lipase by limited proteolysis. *Protein Engineering Design and Selection*, 1992, 5(1): 105-111.
- 67:** Hermoso J, Pignol D, Kerfelec B, Crenon I, Chapus C, Fontecilla-Camps JC. Lipase activation by nonionic detergents: The crystal structure of the porcine lipase-colipase-

tetraethylene glycol monoethyl ether complex". *Journal of Biological Chemistry*, 1996; 271(30): 18007-18016.

68: Delorme V. De Mycobacterium tuberculosis à la protéomique chimique: Utilisation et greffage d'inhibiteurs de lipases et carboxylestérases. Thèse de doctorat. Université Paul-Cézanne. Marseille 2012.

69 : Sias B, Ferrato F, Grandval P, Lafont D, Boullanger P, De Caro. Human pancreatic lipase-related protein 2 is a galactolipase. *Biochemistry*, 2004. 43(31), 10138-10148.

70: Fickers P, Destain J, Thonart P. Les lipases sont des hydrolases atypiques : principales caractéristiques et applications. *Biotechnology Agronomy Society and Environment Environ*, 2008; 12(2): 119-130.

71: Gargouri Y, Moreau H, Verger R. Gastric lipases: biochemical and physiological studies. *Biochimica Biophysica Acta*, 1989; 1006(3): 255-271.

72 : Gargouri Y, Pieroni G, Rivière C, Lowe PA, Saunière JF, Sarda L, Verger R. Importance of human gastric lipase for intestinal lipolysis: an in vitro study. *Biochimica Biophysica Acta*, 1986; 879(3): 419-423.

73: N'Goma, J. C. B, Amara S, Dridi K, Jannin V, Carrière F. . Understanding the lipid-digestion processes in the GI tract before designing lipid-based drug-delivery systems. *Therapeutic Delivery*, 2012; 3(1): 105-124.

74: Hamosh M. Lingual and gastric lipases: Their role in fat digestion Boston. *CRC Press*, 1990; 6(6):421-8.

75: Joliff G, Vaganay S, Legay C, Benicourt C. Secretion of an active recombinant dog gastric lipase from baculovirus-infected insect cells. *Biotechnology Letters*, 1998; 20(7): 697-702.

76: Gruber V, Berna PP, Arnaud, T, Bournat, P, Clement C, Mison D, Olganier B, Philippe L, Theisen M, Baudino S, Benicourt C, Cudrey C, Bloes C, Duchateau N, Dufour S, Gueguen C, Jacquet S, Ollivo C, Poncetta C, Zorn N, Ludevid D, Van Dorsselaer A, Verger R, Doherty A, Merot B, Danzin C. Large-scale production of a therapeutic protein in transgenic tobacco plants: effect of subcellular targeting on quality of a recombinant dog gastric lipase. *Molecular Breeding*, 2001; 7(4): 329-340.

- 77: Canaan S, Dupuis L, Rivière M, Faessel K, Romette J-L, Verger R, Wicker-Planquart C. Purification and interfacial behavior of recombinant human gastric lipase produced from insect cells in a bioreactor. *Protein Expression and Purification*, 1998; 14(1): 23-30.
- 78: Roussel A, Canaan S, Egloff MP, Rivière M, Dupuis L, Verger R, Cambillau C.). "Crystal structure of human gastric lipase and model of lysosomal acid lipase, two lipolytic enzymes of medical interest. *Journal of Biological Chemistry*, 1999; 274(24): 16995-17002.
- 79: Roussel A, Miled N, Berti-Dupuis L, Rivière M, Spinelli S, Berna P, Gruber V, Verger R, Cambillau C. Crystal structure of the open form of dog gastric lipase in complex with a phosphonate inhibitor. *Journal of Biological Chemistry*, 2002; 277(3): 2266-2274.
- 80: Miled N, Roussel A, Bussetta C, Berti-Dupuis L, Rivière M, Buono G, Verger R, Cambillau C, Canaan S. Inhibition of dog and human gastric lipases by enantiomeric phosphonate inhibitors: a structure-activity study. *Biochemistry*, 2003; 42(40): 11587-11593.
- 81: -Planquart C, Canaan S, Rivière M, Dupuis L. Site-directed removal of N-glycosylation sites in human gastric lipase. *European Journal of Biochemistry*, 1999; 262(3): 644-651.
- 82: Aloulou A, Rodriguez JA, Fernandez S, Van Oosterhout D, Puccinelli D; Carrière F. Exploring the specific features of interfacial enzymology based on lipase studies. *Biochimica and Biophysica Acta-Molecular and Cell Biology of Lipids*, 2006; 1761(9): 995-1013.
- 83: Chahinian H, Snabe T, Attias C, Fojan P, Petersen SB, Carrière F. "How gastric lipase, an interfacial enzyme with a Ser-His-Asp catalytic triad, acts optimally at acidic pH. *Biochemistry*, 2006; 45(3): 993-1001.
- 84: Nacim ZOUARI, Lipase digestive d'un animal primitif, le scorpion : purification, caractérisation biochimique et localisation cellulaire, thèse doctorat l'Ecole Nationale d'Ingénieurs de Sfax 2006.

Chapitre II

Technique de Docking moléculaire

II.1. Introduction

La chimoinformatique, et plus spécifiquement la modélisation moléculaire, sont des techniques permettant de comparer les propriétés physico-chimiques associées aux molécules chimiques, et d'analyser les interactions supramoléculaires responsables d'un phénotype biologique, points clés pour la conception de nouveaux ligands. Plusieurs modèles mathématiques sont disponibles pour définir les conformations associées à chaque structure, des méthodes de mécanique quantique aux méthodes de mécanique moléculaire en considérant notamment la complexité de la molécule et les propriétés que l'on souhaite modéliser.¹

Le docking moléculaire connu aussi sous le nom « amarrage moléculaire pour les puristes francophones », fait partie des méthodes de modélisation. Cette technique utilise les mêmes principes que les autres méthodes de modélisation, sauf qu'elle a la particularité de combiner deux molécules ou plus en même temps. Cette propriété fait du docking l'une des plus importantes méthodes de modélisation moléculaire. Pour cela différentes méthodes et algorithmes ont été mis au point pour l'élaboration de cette technique.²

Le rôle principal de cette technique est étudié puis prédire les interactions probables entre des ligands (substrat, activateur ou inhibiteur) et les acides aminés composant la structure de récepteur (protéine). Le docking moléculaire se déroule en deux étapes distinctes:

- La première étape consiste à positionner le ligand dans le site choisi de la protéine.
- La seconde étape de cette méthode permet l'évaluation des interactions énergétiques potentielles entre le ligand et la protéine. Ces deux étapes diffèrent chacune du programme de docking utilisé.³

II.2. Etapes typiques du Docking moléculaire

Le processus de docking moléculaire consiste à faire interagir un ligand avec le récepteur (généralement de nature protéique). La technique de docking comprend quatre étapes principales (fig.1):

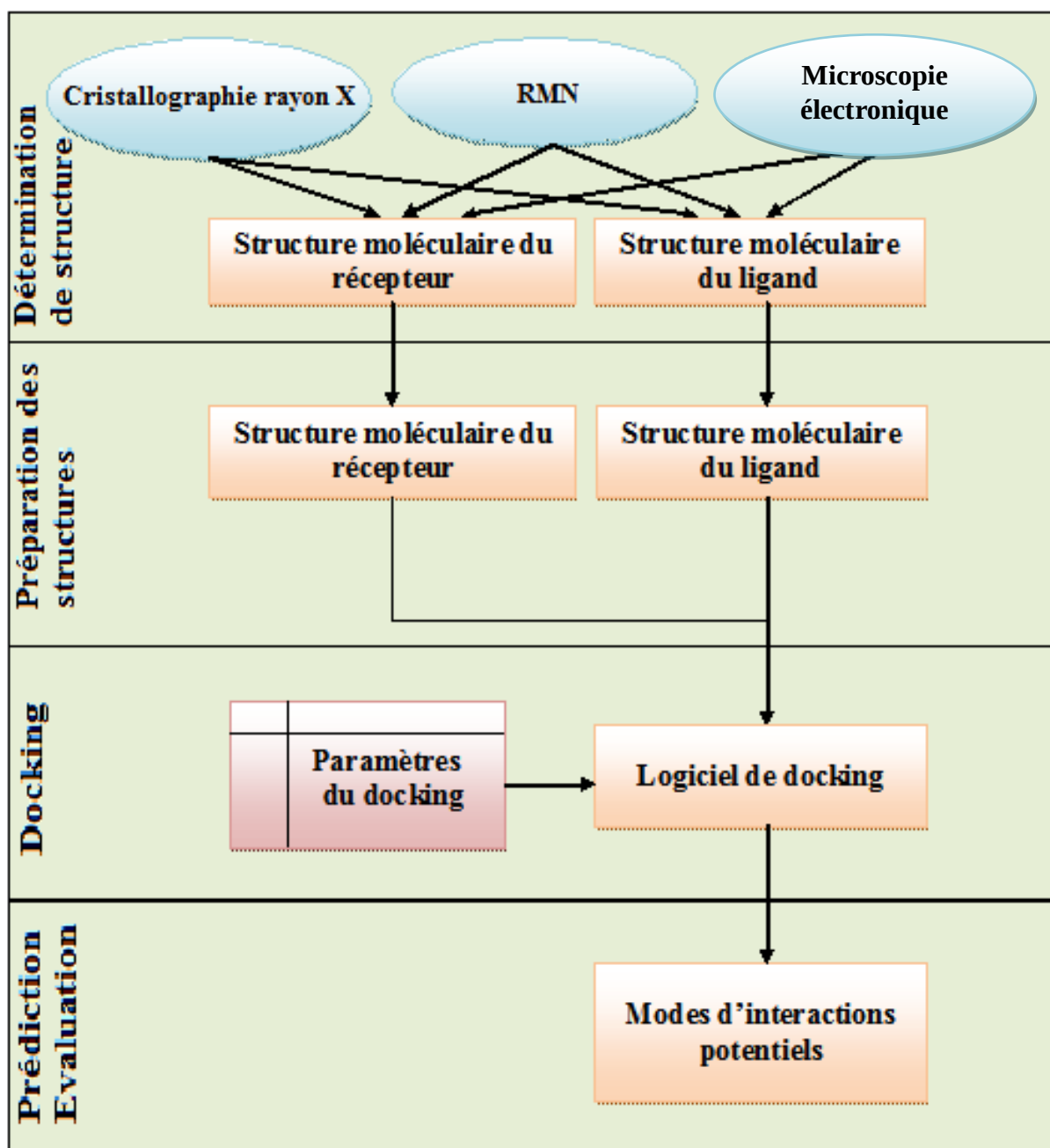


Figure 1. Étapes typiques d'un docking.⁴

II.2.1. Détermination des structures

Cette étape préliminaire du docking moléculaire est réalisée pour déterminer les structures moléculaires mises en jeu (le récepteur et le ligand). Ces structures sont obtenues en faisant appel à trois méthodes d'analyse (la résonance magnétique nucléaire (RMN), la microscopie électronique et la cristallographie par rayons X). Ces méthodes expérimentales (fig.1) permettent aujourd'hui de déterminer la structure des protéines.⁵

La majorité des structures protéiques sont disponibles via la «Protein Data Bank »⁶, Un grand nombre de structures d'une même molécule protéinique avec ou sans ligand permet d'avoir une information pertinente. Dans le cas où la méthode expérimentale n'est pas possible ou n'a pas été encore réalisée, la solution à envisager en premier est une approche par modélisation comparative.⁷

II.2.2. Préparation des structures

La préparation du récepteur en vue du docking ne peut être réalisée qu'une fois les structures mises en jeu sont identifiées. Pour cela il faut veiller à résoudre les problèmes de clashes stériques, ainsi que ceux des états de protonation.⁴

Il est indispensable de suivre les étapes ci-dessous pendant la préparation du récepteur qui sera utilisé pour le traitement du docking. L'hierarchie de ces étapes se présente comme suit :

- 1- Suppression de toute molécule d'eau située à la surface de la protéine dans le cas de la simulation en l'absence de l'eau.
- 2- Vérification des parties manquantes et qui sont à proximité du site actif connu, dans la chaîne polypeptidique étudiée.
- 3- Vérification de l'existence des chaînes latérales qui sont à modéliser. Si une chaîne latérale manque à proximité du site actif, il faut la positionner au mieux, afin d'augmenter les possibilités d'interactions entre le ligand et la protéine.
- 4- Vérification du clash stérique après l'ajout des atomes d'hydrogènes manquants sur la protéine et le ligand.
- 5- Vérification de la protonation des histidines en contact avec le ligand car il arrive souvent que le tautomère du groupement imidazole dans la molécule de l'histidine soit mal choisi lors du positionnement des atomes d'hydrogène. Si un tel problème se pose il faudra donc corriger ces erreurs à la main. La correction de ce mauvais positionnement est observée par l'existence d'un contact direct entre le groupement hydroxyle ou amide du ligand et la fonction amide du groupement imidazole de l'histidine. Cet indice révèle que le H de l'histidine est au mauvais endroit et donc à retirer puis le positionner sur l'atome d'azote en sp² de l'autre côté du groupement imidazole.

- 6- Une fois ces étapes sont validé on pourra compléter par les groupements amides manquants sur les résidus comme les glutamines ou l'asparagine.

II.2.3. Docking moléculaire

Dans le domaine de la modélisation moléculaire, le docking est une méthode qui calcule l'orientation préférée d'une molécule vers une seconde lorsqu'elles sont liées pour former un complexe stable⁸. Le docking moléculaire se fait généralement en trois étapes (schéma 1).

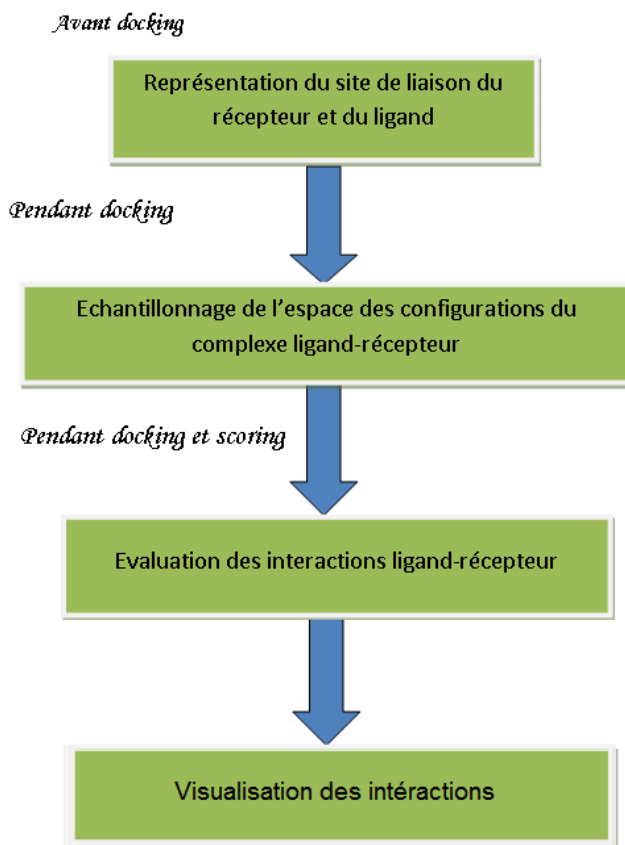


Schéma 2. Protocole générale de Docking moléculaire.

L'utilisation de la technique de «docking» apparaît comme une alternative très intéressante quand on n'a pas une réelle connaissance sur le site actif. Le plus important problème pour l'étape de docking est de parcourir le mieux possible l'espace conformationnel c'est-à-dire l'Echantillonnage de l'espace des configurations du complexe ligand-récepteur.

La complexité de ce problème est fonction de plusieurs paramètres tels que : le nombre de degrés de liberté obtenus pour la translation et la rotation vis-à-vis des conformations de départ possibles du ligand. Dans notre cas nous avons choisi le « Molegro Virtual Docker (MDL) » comme logiciel pour la modélisation.

Ce dernier il présente une plate-forme intégrée pour prédire - protéine ligand interactions qui permet de gérer tous les aspects du processus d'accueil (la préparation des molécules, la détermination des potentiels sites des liaisons de la protéine cible, et la prévision des modes de liaison des ligands).⁹

Contrairement au docking rigide qui ne fait intervenir que six degrés de libertés de rotation et de translation, l'introduction de la flexibilité augmente nettement le nombre de degrés de liberté, l'espace de recherche et donc le coût de calcul. On peut distinguer trois niveaux de docking :

- ✓ le docking rigide est bien sûr le plus simple et demeure encore souvent employé pour l'amarrage protéine-protéine.
- ✓ le docking semi-flexible est "asymétrique" et généralement utilisé pour l'amarrage protéine-ligand, le ligand étant considéré comme flexible, la protéine gardée rigide.
- ✓ le docking flexible, traite la flexibilité des deux molécules, mais la flexibilité permise est limitée, simplifiée par des modèles.

Afin d'éviter des calculs, que les machines ne peuvent résoudre ou seulement dans des temps bien trop importants, plusieurs approximations sont possibles. Les algorithmes de recherche de la flexibilité du ligand peuvent se classer en trois principes, nommés combinatoire, stochastique, et déterministe.⁴

a) Approche combinatoire

Cette approche est basée sur des grilles de valeurs pour chaque degré de liberté, et chacune de ces grilles est explorée de manière combinatoire au cours de la recherche. Dans un premier temps le ligand est découpé en parties rigides et flexibles. Entre les points où des rotations sont possibles, une ou plusieurs «ancres» rigides sont définies, ensuite une première partie rigide est mise en interaction avec le récepteur puis les parties flexibles sont ajoutées de manière successive avec une exploration des angles de torsion.¹⁰

Cette méthode a été incorporée dans plusieurs programmes dont Dock¹¹ et FlexX¹². Le programme FlexX positionne l'ancre à partir d'une modélisation des interactions chimiques et effectue une sélection automatique des fragments de base.^{13,14}

b) Approche stochastique

L'approche stochastique consiste à effectuer des changements aléatoires dans la structure tridimensionnelle du ligand. Habituellement il s'agit de modifier un degré de liberté à chaque fois. L'un des points faibles de cette méthode est l'incertitude de convergence. Pour l'éviter, il faut multiplier les calculs, indépendamment les uns des autres. Un des principaux algorithmes stochastiques est la méthode de Monte Carlo.¹⁵

Pour la méthode de Monte Carlo le ligand est considéré dans son ensemble et les changements s'effectuent aussi bien sur les translations, les rotations que sur les torsions. A chaque mouvement, la molécule est minimisée et son énergie est calculée. La conformation obtenue par cette transformation est testée avec un critère de sélection basé sur l'énergie. Si ce critère est validé, il sera sauvegardé et le programme générera ensuite la prochaine conformation. Les itérations se poursuivront jusqu'à ce que la quantité prédéfinie de conformations soit collectée. Le principal avantage de la méthode de Monte Carlo est que le changement peut être assez important pour permettre au ligand de franchir les barrières énergétiques sur la surface d'énergie potentielle. Un point qui n'est pas facilement atteint par les méthodes de simulation basées sur la dynamique moléculaire¹⁶.

c) Approche déterministe

Le problème des systèmes déterministes est qu'ils peuvent facilement rester piégés dans un minimum local car leurs capacités à surmonter des barrières énergétiques sont faibles. Il s'agit de l'approche la plus simple et la plus directe. L'exemple le plus répandu est la simulation de dynamique moléculaire. Cette méthode est très rarement employée en docking du fait des moyens qu'elle demande et de son biais pour les minima locaux.¹⁷

II.2.4. Prédiction et évaluation

Une fois les paramètres du docking moléculaire sont établis, le programme passe à l'étape de prédiction et d'évaluation. Celle-ci permet la mise au point des modes d'interactions potentiels. Cette étape est réalisée comme suit:

II.2.4.1. Algorithme de docking

En principe, un algorithme de docking doit être capable de générer les modes de liaison attendus pour des ligands. Pour cela, il est nécessaire que l'algorithme d'une recherche conformationnelle puisse explorer l'espace conformationnel possible, de façon efficace et exhaustive.

L'aptitude d'un algorithme à trouver l'emplacement correct du ligand par rapport à son récepteur est habituellement déterminée au moyen de la déviation quadratique moyenne ou RMSD (Root-Mean-Square Deviation) du modèle conçu par le logiciel avec les coordonnées cartésiennes (x_{pose} , y_{pose} , z_{pose}) et ceci vis-à-vis les coordonnées cartésiennes de la structure du cristal ($x_{cristal}$, $y_{cristal}$, $z_{cristal}$). L'écart RMSD est donné par l'équation suivante:

$$RMSD = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_{cristal i} - x_{pose i})^2 + (y_{cristal i} - y_{pose i})^2 + (z_{cristal i} - z_{pose i})^2} \quad Eq.1$$

La valeur admise est une différence maximale de deux angströms au-delà de laquelle la prédiction est considérée comme non adéquate^{18,19}. En général, les erreurs de Docking sont dues à un échantillonnage insuffisant ou à une fonction de score inadéquate.

II.2.4.2. Fonctions de score

La fonction de score est une donnée numérique utile pour quantifier le degré avec lequel un ligand se complexe à un récepteur. C'est globalement une approximation de l'énergie libre résultant du passage de la forme libre de la protéine et du ligand à l'association sous forme de complexe. Le principe thermodynamique est le suivant (Eq. 2):²⁰

$$\Delta G = \Delta G_{complexe} - \Delta G_{ligand} - \Delta G_{protéine} \quad Eq.2$$

L'établissement d'une bonne fonction de score est un important problème du docking. Il arrive souvent que la solution évaluée comme étant la plus probable ne soit pas la forme native attendue. Ceci peut être dû au fait que le complexe natif n'est pas forcément celui qui présente la plus grande surface d'accès, ou encore le plus grand nombre de liaisons hydrogène disponibles. Pour cela il existe différents types de fonctions de score selon les critères sur lesquels celles-ci sont basées²¹. L'affinité chimique peut être calculée par l'énergie libre de Gibbs ΔG_l pour une température T ^{22, 23}:

$$\Delta G_l = -RT \ln K_{eq} \quad Eq.3$$

Où,

R est la constante des gaz parfaits et K_{eq} est la constante d'équilibre. Le calcul de l'énergie libre étant cependant lourd, son application devient impossible lorsqu'il s'agit d'évaluer un

grand nombre de complexes. L'estimation de l'affinité repose généralement sur un calcul d'énergie contenant un certain nombre d'approximations. Pour cela il existe différentes fonctions de score. Ces dernières ajustées dans les programmes de docking ont fait l'objet de nombreuses publications²⁴⁻²⁶. Ces fonctions de score sont généralement classées en quatre familles:

- les fonctions de score basées sur un champ de force,
- les fonctions de score empiriques,
- les fonctions de score consensus,
- les fonctions de score de type knowledge-based.

a) Fonctions de score basées sur un champ de force

Les fonctions de score basées sur un champ de force calculent, par mécanique moléculaire, l'énergie d'interaction du complexe et l'énergie interne du ligand. G-Score³⁰ par exemple, fait appel au champ de force de Tripos, et AutoDock²⁸ à celui d'AMBER. Les interactions entre récepteur et ligand comprennent souvent des termes de Vander Waals et électrostatiques. L'énergie interne du ligand est généralement écrite de manière similaire. Les principales limitations des fonctions de score basées sur des champs de forces proviennent du fait qu'elles ont été écrites pour des modèles en phase gazeuse et ne contiennent donc pas d'effet de solvation ni de terme d'entropie. Des extensions incluant un terme d'entropie pour le ligand (dans G-Score) et les liaisons hydrogène protéine-ligand (dans Gold et Autodock) ont été ajoutées récemment.²⁹

b) Fonctions de score empiriques

Les fonctions de score empiriques sont utilisées pour interpréter l'énergie d'interaction d'un complexe récepteur-ligand à partir d'une équation de sommation d'interactions chimiques localisées.³⁰

Ce type de fonction est basé sur la régression multiple pour ajuster les coefficients de fonctions selon la physique du système. Ajustement à partir d'un jeu de données de complexes récepteurs-ligands avec des affinités mesurées³¹.

Les fonctions de score empiriques contiennent usuellement des termes décrivant les interactions ioniques, interactions hydrophobiques, les ponts ou liaisons hydrogène et les interactions engendrées par le changement d'entropie (pénalité d'entropie). Ceci dit,

ces fonctions de score somment ces différents termes en les pondérant à des termes décrivant les différents types d'interactions moléculaires³². La plupart des logiciels de docking utilise ce type de fonction à cause de son efficacité de point de vue rapidité et précision. Cependant, leur principal inconvénient est leur forte dépendance aux données de paramètre calibration.

c) Les fonctions de score type knowledge-based

Ces fonctions proviennent de l'analyse des structures tridimensionnelles de complexes ligand-protéine déterminés expérimentalement.

Des règles définissant la géométrie préférentielle des interactions sont déduites de ces structures grâce à des moyens statistiques. Cette alternative aux fonctions empiriques est plus tolérante quant aux interactions présentes au sein du complexe et leurs expressions sont moins strictes.³³⁻³⁷

Ces fonctions statistiques, dépendent de leurs groupes d'apprentissage (expérience), elles ne peuvent modéliser que les interactions qui existent dans leurs bases de données expérimentales. Par conséquent, avec ce type de fonctions, il est difficile de modéliser les interactions qui ne dévient pas trop de leur modèle. L'un des potentiels les plus simples est le PLP (Piecwise Linear Potential). Il est basé sur quatre types d'atomes et prend en compte les interactions stériques et les ponts hydrogènes. Les distances interatomiques sont prises en compte en dessous de 5 Å. De même l'état de référence est défini comme étant un complexe entre les deux entités chimiques qui n'interagissent pas.³⁸

Les fonctions de score type knowledge-based ont été utilisées avec succès dans des études d'amarrage de différentes cibles protéiques et ont montré une certaine amélioration dans la prédiction des modes de liaison corrects et le classement des complexes protéine-ligand par rapport aux fonctions de notation empiriques et celles basées sur le champ de force.³⁹

d) Fonctions de score consensus

L'idée principale de ces fonctions est combiner les informations obtenues à partir des différents scores obtenus et qui sont insuffisant des autres⁴⁰. Plusieurs études ont montré que les énergies libres des complexes sont performées par ces fonctions, ainsi les interactions protéine-ligand mieux que les fonctions individuelles^{41,42}. Un exemple de fonction de score consensus est X-CSCORE⁴³ qui combine un PMF⁴⁴⁻⁴⁶, et ChemScore⁴⁷. Le tableau suivant récapitule les avantages et les désavantage des différentes fonctions de score.

Tableau 3. Avantage et inconvénients des différentes fonctions de score.

Fonction de score	Avantages	Inconvénients
Basé sur les champs de force	- Les termes des FFs sont bien étudiés et ont une base physique. - Transférable, rapides quand elles sont utilisés sur une grille pré-calculée.	- Ne tient compte que d'une partie de l'énergie potentielle. - Elles sont parfois augmentée par des termes de solvation et d'entropie. - L'électrostatique est souvent surestimée.
Empirique	-Estimation rapide et directe de l'énergie de liaison.	- Il existe peu de complexes avec à la fois des structures et des énergies de liaison connues précisément. - Forte dépendance sur l'orientation des atomes d'hydrogène. - Pas de véritables pénalités pour les mauvaises structures.
Knowledge-based	-Similaire aux méthodes empiriques, mais plus générales. Existence de données d'inter-distances plus que d'énergies de liaison expérimentales.	- Les PMF sont "pairwise", alors que la probabilité de trouver les atomes A and B à une distance r n'est pas "pairwise" et dépend aussi des autres atomes.

II.2.4. 3. Interaction protéine-ligand

Le fonctionnement des systèmes biologiques est basé sur de remarquables mécanismes d'assemblage et de reconnaissance impliquant des interactions moléculaires non-covalentes de faible énergie. Les messages biologiques qui sont intégrés dans leur structure ne peuvent s'exprimer que par l'intermédiaire d'interactions spécifiques qui s'établissent, de façon souvent réversible, avec d'autres partenaires de choix.

Le problème qui se pose pendant l'ancrage du ligand au sein de la protéine (docking) est la prédiction de la conformation et l'orientation du ligand relative au site actif de la protéine cible. À chaque protéine cible de structure connue le docking se révèle être la clé dans le design de nouveaux médicaments⁴⁸⁻⁵⁰. De nombreux type d'interaction non covalente, ont été mis en évidence dans les complexes protéine- ligand.

Parmi les interactions qui peuvent exister entre le ligand, qui est considéré comme une petite molécule organique, et la protéine on cite celles de type polaire (liaisons hydrogène, liaisons ioniques) et celles de type hydrophobe qui résultent d'un contact entre des groupements hydrophobe.

a) Interaction ionique

Les interactions entre molécules portant des charges électriques sont régies par la loi de Coulomb. L'énergie potentielle d'interaction de deux charges électriques Q_1 et Q_2 séparées par une distance d est obtenue en calculant le travail nécessaire à la séparation de ces deux charges à une distance infinie (fig.2). L'énergie potentielle (V) de deux molécules de charges différentes est exprimée par l'équation 4.

$$V = \frac{Q_1 Q_2}{4\pi\epsilon\epsilon_0 d} \quad \text{Eq.4}$$

Avec :

ϵ_0 : permittivité du vide.

ϵ : permittivité du milieu dans lequel évoluent les charges.

d : distance entre les deux charge.

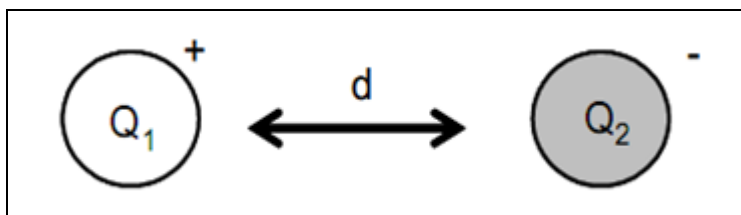


Figure 2. Interaction entre deux molécules de charges différentes.

En milieu aqueux, les acides aminés chargés sont majoritairement exposés à la surface de la protéine et sont donc très solvatés, ce qui réduit considérablement l'interaction entre les charges.

b) Interaction hydrogène

La liaison hydrogène est une liaison chimique non covalente de type dipôle-dipôle entre deux molécules ou entre deux groupements d'une molécule. Elle consiste essentiellement dans l'interaction entre deux molécules dont l'une possède un atome donneur d'électrons (O, N, F) et l'autre possède un atome H accepteur d'électrons (OH, NH₂). Ce type de liaison résulte d'un transfert partiel d'un électron célibataire sur le groupement H. Les liaisons hydrogène peuvent être intramoléculaires ou intermoléculaires⁵¹. Cette liaison est illustrée ci-dessous (fig.3) par des exemples.

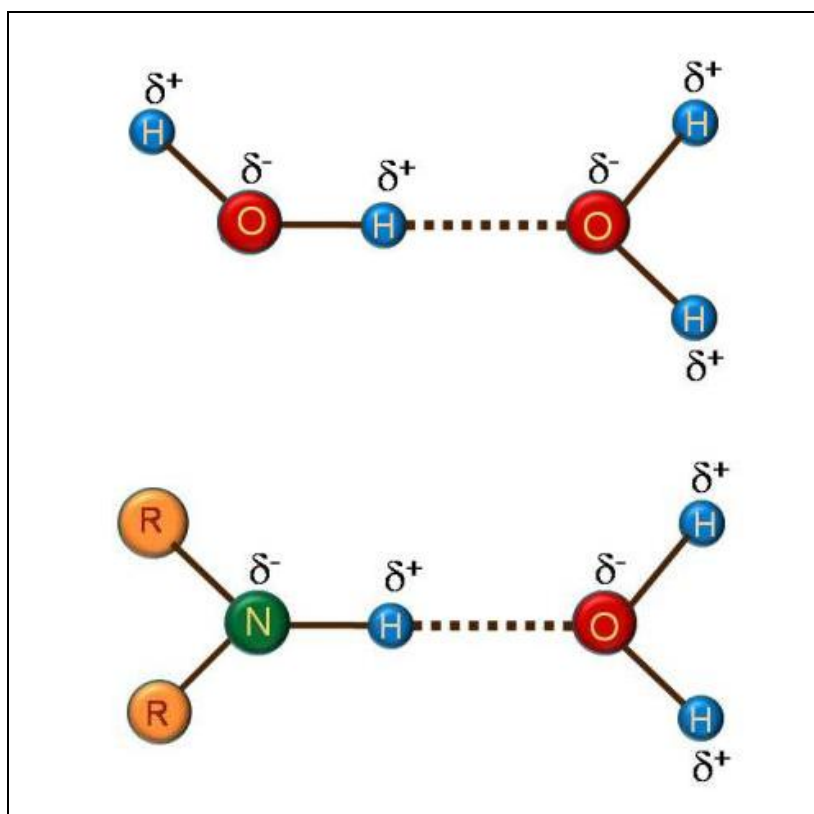


Figure 3. Exemples d'une liaison hydrogène.

Les liaisons hydrogène jouent un rôle est très important en en chimie et en biochimie. Ce sont des liaisons qui donnent à l'eau ses propriétés particulières, comme la capacité des molécules H_2O à s'associer en grandes structures. En outre, elles jouent le rôle de stabilisateur de la structure secondaire des macromolécules biologiques. Les liaisons hydrogènes existent dans les protéines (structure IIIaire et IVaire), les acides nucléiques (double hélice stabilisée par de liaisons H), les interactions moléculaires (reconnaissance de petites molécules par des récepteurs ...). Les liaisons hydrogènes sont plus fortes que les liaisons de Van der Waals ; leurs énergies sont estimées entre 3 et 9 kcal/mol.

c) Les interaction de Van der Waals

Ces interactions ont été étudiées par J.D. van der Waals, physicien hollandais, prix Nobel de physique 1910. Elles reposent sur les interactions entre les dipôles constitués par les molécules. On distingue trois types de ces interactions :

- Interaction dipôle permanent – dipôle permanent ou effet d'orientation de Keesom :
Ce type d'interaction se développe entre deux molécules polaires.

- Interaction dipôle permanent – dipôle instantané ou effet d'induction de Debye : Ce type d'interaction se développe entre une molécule polaire et une molécule quelconque (polaire ou apolaire).
- Interaction dipôle instantané – dipôle instantané ou dispersion de London : Ce type d'interaction se développe entre deux molécules quelconques (polaires ou apolaires).

Ces interactions sont très faibles mais dans le cas des macromolécules, leur nombre élevé va produire au totale une force importante.⁵²

d) Les interactions hydrophobes

Les molécules non polaires et peu polarisables ont tendance à se regrouper, ce qui crée une force de liaison hydrophobe. Il s'agit d'interactions entre molécules ou groupements qui ont très peu d'affinité pour le solvant dans lequel elles sont dissoutes (eau). Les groupements vont se positionner de manière à présenter la plus faible surface de contact avec l'eau. Les groupements vont donc s'attirer mutuellement par des forces de type dispersion (London).

Plusieurs acides aminés hydrophobes se trouvent en contact à l'interface par complémentarité hydrophobe. Ces derniers entraînent une interaction stabilisante pour beaucoup de complexes. Les acides aminés non polaires tendent à créer entre elles des liaisons de type interactions "patch" hydrophobe⁵³ (fig.4).

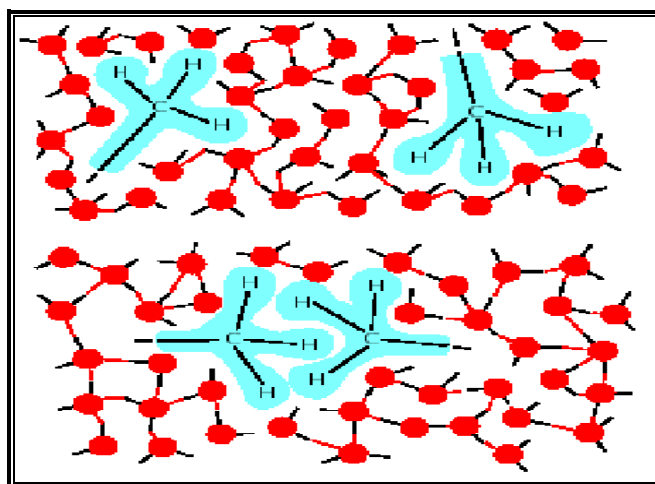


Figure 4. Patch hydrophobe.

Il existe divers type d'interactions hydrophobiques. Parmi ces dernières on cite celles qui se forment entre le centre d'un cycle aromatique et les atomes d'un autre cycle, d'un amide ou

des groupements méthyle. La géométrie de ce type d'interaction est légèrement moins restrictive, ce qui la sépare des interactions purement hydrophobiques. Des études sur l'interaction protéine-ligand ont montré que ce type d'interaction est le plus fréquent¹⁷. Un autre type d'interactions hydrophobe qui se forment lors des contacts hydrophobiques non spécifiques entre les atomes de carbone aliphatique ou aromatique. Ces interactions sont généralement sphériques avec un rayon de 4Å° et elles couvrent la plus grande partie du ligand^{12,54}. Pour le placement du ligand dans le site actif de la protéine, les interactions hydrogène sont utilisées en premier, ensuite celles hydrophobes.

II.2.5. Logiciels utilisés

Le docking moléculaire a pour objectif essentielle de prédire la conformation (position et orientation relative) la plus favorable du ligand au sein de son récepteur. Initié au début des années 1980, ce domaine s'est développé pour devenir, de nos jours, un outil incontournable dans la recherche de produits biologiquement actifs.⁵⁵

A l'heure actuelle, plus de 30 programmes de docking moléculaires (commerciaux ou non) sont disponibles. Les plus fréquemment cités sont respectivement: AutoDock, GOLD, FlexX, DOCK et Molegro virtuel Docker (MVD). Les principaux programmes de docking moléculaire sont résumé ci-dessous (tableau 4).

Tableau 4. Principaux programmes de docking moléculaire.

Nom Editeur	Site Internet
AutoDock Scripps	http://www.scripps.edu/mb/olson/doc/autodock/
FlexX BioSolveIT	http://www.biosolveit.de/FlexX/
Fred OpenEyes	http://www.eyesopen.com/products/applications/fred.html
Dock UCSF	http://dock.compbio.ucsf.edu/
Glide Schrödinger	http://www.schrodinger.com/Products/glide.html
Gold CCDC	http://www.ccdc.cam.ac.uk/products/life_sciences/gold/
ICM Molsoft	http://www.molsoft.com/products.html
LigandFit Accelrys	http://www.accelrys.com/cerius2/c2ligandfit.html
Surflex Biopharmics	http://www.biopharmics.com/products.html

a) GOLD

Ce programme commercial de docking est développé par “Cambridge Crystallographic Data Center” (<http://www.ccdc.cam.ac.uk/>). Il est basé sur trois parties majeures: un algorithme génétique pour explorer les possibilités de modes de liaison⁵⁶, un mécanisme pour placer le ligand dans le site actif et une fonction de score pour classer les différents modes de liaison. La population constituant la première solution est obtenue aléatoirement.

Ces solutions se comportent ensuite comme des “chromosomes” et peuvent subir aléatoirement des événements génétiques de type mutation et de type “cross-over”. Les événements de mutation correspondent à une modification aléatoire de la conformation d’un chromosome, alors que les événements de type “cross-over” sont l’échange de matériel équivalent entre deux chromosomes différents. Les nouvelles conformations sont ensuite évaluées avec la fonction de score et remplacent les conformations initiales moins bonnes. Les chromosomes qui subissent les événements génétiques sont sélectionnées aléatoirement dans la population d’origine, mais les conformations ayant les meilleurs scores sont favorisés et apparaissent plus souvent. Cet algorithme est de type stochastique, car il permet d’explorer de façon aléatoire les différentes conformations du ligand, et de tester leurs interactions avec le site actif tandis que la protéine, reste rigide, fixée dans sa conformation cristallographique.³ Ses principaux avantages sont sa fiabilité à prédire des structures cristallographiques pour des complexes de type protéine-ligand et l’emploi d’un algorithme génétique efficace. De plus, le logiciel dispose d’une interface graphique simple d’emploi, et aisément scriptable et s’adapte particulièrement bien aux environnements de calcul parallèle.⁵⁷

b) Glide

Ce programme développé par la société Schrödinger (www.schrodinger.com/Glide/) utilise un algorithme de docking ayant pour principe une élimination rapide de nombreuses conformations du ligand dont les interactions avec la protéine ne sont pas favorables⁵⁸. Le nombre de conformations du ligand diminue de façon significative et forme un entonnoir jusqu’aux conformations optimales. La première étape de l’algorithme de docking « Glide » est la recherche exhaustive des conformations du ligand.

La génération des conformations passe par le choix d’un “cœur” et des “groupes rotamères” qui ne contiennent pas de liaisons rotatives. Le “cœur” relie alors l’ensemble des groupes rotamères par des liaisons rotatives. Les groupes fonctionnels avec des carbones ou des azotes finissant par des hydrogènes ($-CH_3$, $-NH_2$, $-NH^{+3}$) ne sont pas considérés comme pouvant

tourner. Ces conformations sont ensuite criblées par une méthode heuristique qui élimine les conformations qui ne sont pas compatibles avec le processus de docking. Ce premier filtrage est nécessaire pour éliminer les conformations à haute énergie interne et il est effectué à partir d'une version modifiée du champ de force « OPLS-AA ». L'énergie limite acceptable est déterminée à partir de l'énergie interne la plus faible définie pour l'ensemble des conformations. Les paramètres de ce criblage ont été optimisés sur les conformations des ligands présents dans la base de données des protéines (PDB), qui sont cristallisés dans le site actif de leur protéine. Les conformations restantes sont alors soumises au docking dans le site actif de la protéine. Leur nombre varie de moins de 1000 pour des ligands rigides à plusieurs dizaines de milliers pour des ligands flexibles.³

La première étape du docking effectuée par le programme « *Glide* » est la génération d'une grille de 2 Å d'espacement du site actif, et la répartition régulière sur celle-ci des points d'ancrage. Ces points d'ancrage sont définis en fonction de la distance entre le centre du ligand et sa surface. Le centre du ligand est défini comme étant le milieu du segment reliant les deux atomes les plus éloignés du ligand dans le cas d'un docking rigide et du "cœur" dans le cas où le ligand est flexible. Ce segment est également appelé diamètre du ligand. Chaque conformation est ensuite placée sur chaque point d'ancrage (ou dans un espace cubique autour de celui-ci) et seules les positions du ligand ayant un diamètre inférieur à la distance entre le point d'ancrage et la surface de la protéine sont conservées.

La deuxième étape de docking par *Glide* est l'élimination des conformations ayant des conflits stériques trop importants avec la surface de la protéine. Les conformations restantes effectuent ensuite des rotations autour du diamètre du ligand. Les interactions hydrogènes protéine/ligand et métal/ligand qui apparaissent sont évaluées, et si elles dépassent un certain seuil, l'ensemble des interactions protéine/ligand sont évaluées, sinon la conformation est abandonnée.

La troisième étape de cette procédure de docking consiste à réaliser le déplacement rigide de plus ou moins 1 Å des meilleures positions, selon les directions cartésiennes. *Glide* introduit ensuite à la dernière étape une dose de dynamique moléculaire, en minimisant les meilleures positions (correspondant relativement à 400 positions en général) dans un champ de force de type « OPLS-AA »⁵⁹. Enfin les angles de torsions de trois à six meilleures positions sont étudiés grâce à la méthode Monte-Carlo.⁶⁰

c) AutoDock

Ce programme de docking libre est développé au Scripps Research Institute par l'équipe d'Arthur J. Il est basé sur un algorithme génétique, comme « GOLD »⁶¹. AutoDock et GOLD, reposent sur une fonction objective basée sur un champ de force combiné avec un AE, ces derniers permettant facilement d'implémenter l'heuristique d'échantillonnage d'un côté et la fonction objective de l'autre. La combinaison de ces éléments semble donc remporter un succès certain.⁴ AutoDock permet d'utiliser trois méthodes de recherche différentes³:

- (i) une méthode de recherche utilisant un algorithme génétique classique, où les chromosomes sont réellement des chaînes de caractères. Ces chromosomes sont composés dans l'ordre des trois coordonnées cartésiennes représentant la translation du ligand, suivies de quatre variables définissant un quaternion. Ce dernier qui est un ensemble mathématique qui définit la rotation du ligand dans l'espace donnant à la fin une valeur réelle pour chaque angle de torsion représentant chaque liaison du ligand.
- (ii) une deuxième méthode de recherche locale, basée sur un recuit simulé de type Monte Carlo, peut être ajoutée pour minimiser l'énergie;
- (iii) une méthode mixte de recherche globale et locale, basé sur les principes génétiques énoncés par Jean Baptiste Lamarck au début du XXe siècle. Selon Larmarck, les caractéristiques acquises (phénotype) par un individu au cours de sa vie peuvent se transmettre à sa descendance (génotype), ce qui revient à dire que le phénotype d'un individu peut modifier son génotype. Dans le cadre du docking moléculaire, cette théorie de Larmarck peut être utilisée.

d) AutoDock Vina

AutoDock Vina est un programme de docking dérivé d'AutoDock comme son nom l'indique, qui a également été développé par le groupe de Arthur J. Olson au "Scripps Research Institute" (<http://autodock.scripps.edu/>)⁶². AutoDock Vina utilise une méthode de gradient d'optimisation sophistiquée dans sa procédure d'optimisation locale. Le gradient donne en fait l'algorithme d'optimisation d'un « sens de l'orientation » à partir d'une seule évaluation et un programme de simulation d'amarrage de ligand flexible sur un récepteur rigide.

L'algorithme de positionnement des ligands AutoDock de Vina est une application globale d'une recherche de type itérative locale ("Iterated Local Search global optimizer")^{63,64}. Ce type d'algorithme est utilisé pour sortir des minima locaux qui ne permettent plus d'améliorer la conformation. Pour ce faire, l'algorithme est relancé à partir d'une structure géométrique légèrement différente. Ces différents points de départ sont conservés en mémoire pour ne pas effectuer le même chemin plusieurs fois, et ainsi gagner du temps. Vina utilise l'algorithme d'optimisation locale de Broyden-Fletcher-Goldfarb-Shanno (BFGS).⁶⁵

e) Mologro virtuel Docker (MVD)

Mologro Virtuel Docker (MVD) est un logiciel performant qui permet la prévision des interactions Enzyme-ligand. Il est conçu pour réaliser, automatiquement toutes les étapes du processus. MVD offre un amarrage de haute qualité, il utilise une technique d'optimisation basée sur la rentabilité et la productivité. MolDock Score est basée sur un nouvel algorithme heuristique de recherche qui combine l'évolution différentielle avec un algorithme de prévision de cavité.

Les algorithmes évolutionnaires sont des méthodes d'optimisation génériques et itératives. Elles se reposent sur un processus adaptatif similaire à celui de l'évolution naturelle. Ces méthodes sont capables de gérer plusieurs solutions simultanément, regroupées dans une population. (fig.5).

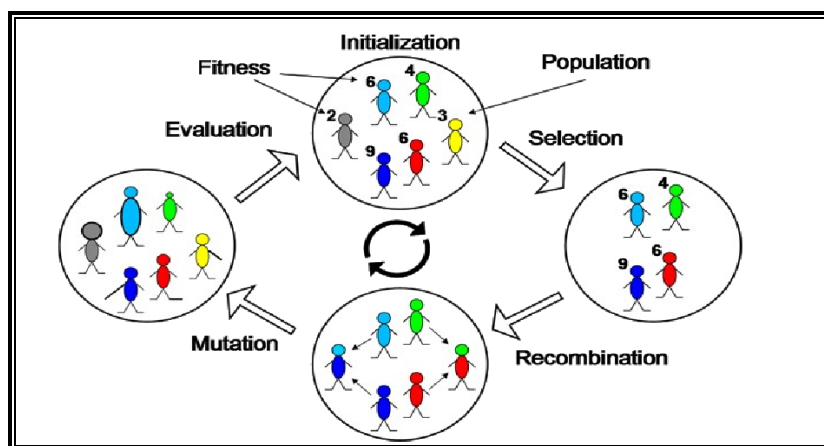


Figure 5. Représentation schématique du cycle évolutionnaire d'un algorithme.⁶⁶

Cette population est exposée à une pression de sélection, implémentée dans une fonction objective (fonction de score) qui elle-même décrit le problème à optimiser, où plusieurs degrés de liberté interviennent, qui à leur tour permettent de définir une solution appelé « génome ».

La sélection des solutions les plus performantes d'après la fonction de score est appelée « parents ». Celle-ci est contre balancée par la génération de nouvelles solutions « enfants ». Celle-ci est réalisée pour maintenir une diversité de population. Les enfants, donc sont générés par modification du génome parental par des opérateurs classés, suivant le nombre de parents sur lesquels ils s'appliquent. Une fois les enfants créés, ils prennent la place des plus mauvaises solutions de la population et donc le processus itératif commence. Cette dernière population (nouvelle population) est de nouveau exposée à la pression de sélection implémentée par la fonction de score. L'évolution continue par une nouvelle itération ou bien génération⁶⁷. Au cours du processus itératif les solutions les mieux adaptées s'émergent, jusqu'à convergence de la population ou après un nombre déterminé de génération. Dans le docking moléculaire, l'algorithme évolutionnaire cherche à décrire les interactions entre le ligand et la cible protéique où les degrés de liberté (ddl) correspondent aux positions, orientations et conformation du ligand et du récepteur

II.3. Conclusion

Le processus de docking est l'un des premières étapes dans la conception de médicaments. Il consiste à faire interagir un ligand avec un récepteur, généralement de nature protéique. En conséquent, le plus grand avantage des méthodes de docking protéine-ligand est qu'ils peuvent proposer des hypothèses structurales sur la façon dont une petite molécule peut interagir avec sa cible macromolécule. Des études ont montré que certains algorithmes de docking sont plus fiables que d'autres pour reproduire le mode de fixation expérimentale de ligand. La contrepartie de ces techniques est généralement une hausse des temps de calcul et des ressources. Le nombre de programme de docking actuellement disponibles est élevé et n'a cessé d'augmenter au cours des dernières décennies. Dans ce travail nous avons utilisé le programme Molegro Virtuel Docker (MVD).

C'est dans ce contexte que nous nous sommes intéressées à la recherche et l'étude des interactions Enzyme-Ligand : cas des inhibiteurs de lipase dans le but de caractériser les mécanismes d'interaction mis en jeu entre l'épigallocatechine gallate (EGCG) qui un polyphénol extrait du thé vert et trois lipases digestives : la lipase pancréatique du porc (LPP), la lipase pancréatique humaine (LPH), et la lipase gastrique (LG). Cette étude sera développée dans les chapitres 3 et 4.

Références Bibliographiques

- 1: <https://www.techniques-ingenieur.fr/base-documentaire/mesures-analyses-th1/etudes-de-structure-et-caracterisation-42386210/modelisation-moleculaire-et-conception-de-nouveaux-ligands-d-interets-biologiques-pha1015/>
- 2: <https://bioinfo-fr.net/la-modelisation-moleculaire>.
- 3: François Martz, Développement d'une nouvelle méthode de docking basée sur les mécanismes enzymatiques et guidée par des groupes prosthétiques, thèse de doctorat de l'université paris 2007.
- 4: Aurélien. G, Conception d'un logiciel de docking et applications dans la recherche de nouvelles molécules actives. Thèse de Doctorat, Université Joseph Fourier - Faculté de pharmacie de Grenoble 2007; 17.
- 5: Howard GC, Brown WE. Modern protein chemistry. *CRC press*, 2001; 272.
- 6: Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. *The Protein Data Bank.Nucleic Acids Res*, 2000; 28: 235-242.
- 7: Al-Lazikani B, Jung J, Xiang Z, Honig B. Protein structure prediction. *Current Opinion in Structural Biology*, 2001; 5: 51-56.
- 8: Lengauer T, Rarey M (Jun). Computational methods for biomolecular docking. *Current Opinion in Structural Biology*, 1996; 6(3): 402-406.
- 9: Thomsen R, Christensen MH. MolDock: a new technique for high-accuracy molecular docking. *Journal medicinal chemistry*, 2006; 49: 3315-332.
- 10: Merzoug A, Chikhi A, Bensegueni A, Boucherit H, Hioual KS, Mokrani H, Teniou S. Conception in silico de nouveaux antipneumocoques par l'inhibition du peptide déformylase. *Les technologies de laboratoire*, 2012; 7: 27.
- 11: Kuntz ID. Structure-based strategies for drug design and discovery. *Science*, 1992; 257: 1078-1082.
- 12: Rarey M, Kramer B, Lengauer T, Klebe G. A Fast Flexible Docking Method Using an Incremental Construction Algorithm. *Journal of Molecular Biology*, 1996; 26: 470-489.

- 13:** Rarey M, Kramer B, Lengauer T. Multiple automatic base selection: protein-ligand docking based on incremental construction without manual intervention. *Journal of Computer-Aided Molecular Design*, 1997; 11: 369-384.
- 14:** Rarey M, Wefing S, Lengauer T. Placement of medium-sized molecular fragments into active sites of proteins. *Journal of Computer-Aided Molecular Design*, 1996; 10: 41–54
- 15:** Holland JH. *Adaptation in Natural and Artificial Systems*, Thèse de doctorat, University of Michigan Press. Ann Arbor, MI 1975.
- 16:** Xuan YM, Hong XZ, Mihaly M, Meng C, Molecular Docking: A powerful approach for structure-based drug discovery. *Current Computer-Aided Drug*, 2011; 7(2): 146–157.
- 17 :** Chikhi A. Calculs et modélisations des interactions peptide deformylase – substances anti-bacteriennes à l’aide de techniques de “docking” (arrimage) moléculaire. Thèse de doctorat, Université Mentouri Constantine 2007.
- 18:** Vieth M, Hirst JD, Kolinski A, Brooks CL. Assessing energy functions for flexible docking. *Journal of Computational Chemistry*, 1998; 19: 1612-1622.
- 19:** Gabb J, Jackson RM, Sternberg MJ. Modelling protein docking using shape complementarity: electrostatics and biochemical information. *Journal of Molecular Biology*, 1997; 272(50): 106–120.
- 20:** Kollman PA, Massova I, Reyes C, Kuhn B, Huo S, Chong L, Lee M, Lee T, Duan Y, Wang W, Domini O, Cieplak P, Srinivasan J, Case DA, Cheatham TE. Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. *Accounts of Chemical Research*, 2000; 33: 889-897.
- 21:** BRUT M, Nouvelle approche méthodologique pour la prise en compte de la flexibilité dans les interactions entre molécules biologiques : Les Modes Statiques Université Toulouse III - Paul Sabatier, 2009.
- 22:** Kollman PA. Free energy calculations: applications to chemical and biochemical phenomena. *Chemical Reviews* 1993; 93: 2395-2417.
- 23:** Simonson T, Archontis G, Karplus M. Free energy simulations come of age: protein-ligand recognition. *Accounts of Chemical Research* 2002; 35: 430-437.

- 24:** Gohlke H, Klebe G. Approaches to the description and prediction of the binding affinity of small-molecule ligands to macromolecular receptors. *Angewandte Chemie International Edition*, 2002; 41 (15): 2644-2676.
- 25:** Ferrara P, Gohlke H, Price DJ, Klebe G. Assessing scoring functions for protein-ligand. *Interactions Journal of Medicinal Chemistry*, 2004; 47: 3032-3047.
- 26:** Wang R, Lu Y, Fang X, Wang S. An extensive test of 14 scoring functions using the PDB bind refined set of 800 protein-ligand complexes. *Journal of Chemical Information and Computer Sciences*, 2004; 44 (6): 2114-2125.
- 27:** Kramer B, Rarey M, Lengauer T, Evaluation of the FlexX incremental construction algorithm for protein-ligand docking. *Proteins*, 1999; 37: 228-241.
- 28:** Morris GM, Goodsell DS, Halliday RS, Huey R, Hart WE. Automated docking using a Lamarckian genetic emphasis on quantitative analysis of protein-ligand algorithm and an empirical free energy function. *Journal of Computational Chemistry*, 1998; 19: 1639-1662.
- 29:** Verdonk ML, Cole JC, Hartshorn MJ, Murray CW. Taylor. Improved protein–ligand docking using GOLD. *Proteins: Structure, Function, and Bioinformatics*, 2003; 52(4): 609–623.
- 30:** Ignarro LJ, Fukuto JM , Griscavage JM , Rogers NE , Byrns RE. Oxidation of nitric oxide in aqueous solution to nitrite but not nitrate: Comparison with enzymatically formed nitric oxide from L-arginine. *Proceedings of the National Academy of Sciences USA*, 1993; 90: 8103-8107.
- 31:** Lafond M , Docking et scoring. *Schrödinger*, 2015.
- 32:** Holloway MK, A priori Prediction of Activity for HIV-1 Protease Inhibitors Employing Energy Minimization in the Active Site. *Journal of Medicinal Chemistry*, 1995; 38: 305-317.
- 33:** Boehm HJ, The development of a simple empirical scoring function to estimate the binding constant for a protein-ligand complex of known three dimensional structure. *Journal of Computer-Aided Molecular Design*, 1994; 8: 243-256.
- 34:** Boehm HJ. Prediction of binding constants of protein ligands: a fast method for the prioritization of hits obtained from de novo design or 3D database search programs. *Journal of Computer-Aided Molecular Design*, 1998; 12: 309-323.

- 35:** Head RD, Smyte ML, Oprea TI, Waller CL, Green SM, Marshall GR. Validate: a new method for the receptor-Based prediction of binding affinities of Novel ligands. *Journal of the American Chemical Society*, 1996; 118: 3959-3969.
- 36:** Muegge I, Martin YA. A General and fast scoring function for protein-ligand interactions: a simplified potential approach. *Journal of Medicinal Chemistry*, 1999; 42: 791-804.
- 37:** Muegge I. Aknowledge-based scoring function for protein-ligand interaction: Probing the reference state. *Perspectives in Drug Discovery and Design*, 2000, 20: 99-114.
- 38:** Perez C, Pastor M, Ortiz AR, Gago FJ. Comparative binding energy analysis of HIV-1 protease inhibitors: incorporation of solvent effects and validation as a powerful tool in receptor-based drug design. *Journal of Medicinal Chemistry*, 1998; 41: 836-852.
- 39:** Schneider G, Böhm HJ. Virtual screening and fast automated docking methods. *Combinatorial Chemistry*, 2002;7: 64-70.
- 40:** Charifson PS, Corkery JJ, Murcko MA, Walters WP. Consensus scoring: A method for obtaining improved hit rates from docking databases of three dimensional structures into proteins. *Journal of Medicinal Chemistry*, 1999; 42: 5100-5109.
- 41:** Bissantz C, Folkers G, Rognan D. Protein-based virtual screening of chemical databases. Evaluation of different docking scoring combinations. *Journal of Medicinal Chemistry*, 2000; 43: 4759-4767.
- 42:** Terp GE, Johansen BN, Christensen IT, Jørgensen FS. A new concept for multidimensional selection of ligand conformations (MultiSelect) and multidimensional scoring (MultiScore) of protein-ligand binding affinities. *Journal of Medicinal Chemistry*, 2001; 44: 2333-2343.
- 43:** Wang RX, Lai LH, Wang SM, Further development and validation of empirical scoring functions for structure-based binding affinity prediction. *Journal of Computer-Aided Molecular Design*, 2002; 16: 11-26.
- 44:** Muegge I. PMF. Scoring revisited. *Journal of Medicinal Chemistry*, 2006; 49: 5895-5902.
- 45:** Muegge I, Martin YC. A general and fast scoring function for protein-ligand interactions: a simplified potential approach. *Journal of Medicinal Chemistry*, 1999; 42: 791-804.

- 46: Muegge I. Effect of ligand volume correction on PMF scoring. *Journal of Computational Chemistry*, 2001; 22: 418-425.
- 47: Eldridge MD, Murray CW, Auton TR, Paolini GV, Mee RP. Empirical scoring functions: The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes. *Journal of Computer-Aided Molecular Design*, 1997; 11: 425-445.
- 48: Leach AR, Kuntz ID. Conformational analysis of flexible ligands in macromolecular receptor sites. *Journal of Computational Chemistry*, 1992; 13: 730-748.
- 49: Böhm HJ. LUDI: rule-based automatic design of new substituent for enzyme inhibitor leads. *Journal of Computer-Aided Molecular Design* 1992; 6(6): 593-606.
- 50: Böhm HJ. The development of a simple empirical scoring function to estimate the binding constant for a protein-ligand complex of known three-dimensional structure. *Journal of Computer-Aided Molecular Design*, 1994; 8 (3): 243-256.
- 51: Chimie Physique 6ème édition de Peter W. *Atkins*, 1997.
- 52 : Gerschel A. Liaisons intermoléculaires : Les forces en jeu dans la matière condensée. Paris, *EDP Sciences*, 2012 ; 2^e édition, 288p.
- 53: Israelachvili JN. Intermolecular and surface forces. Academic Press, 2^e édition, Londres 1997.
- 54: Rarey M, Kramer B, Lengauer T. Docking of hydrophobic ligands with interaction-based matching algorithms. *Bioinformatics*, 1999, 15 (3): 243-250.
- 55: Thiriot E. Modélisation de la reconnaissance et de la catalyse enzymatiques : développements méthodologiques et détermination du mécanisme des Méthionine Sulfoxyde Réductases de classe A. Thèse de doctorat d'université: Chimie informatique et théorique. Nancy: Université de Nancy. France 2009: 291.
- 56: Verdonk ML, Cole JC, Hartshorn MJ, Murray CW. Taylor. Improved protein–ligand docking using GOLD. *Proteins : Structure, Function, and Bioinformatics*, 2003; 52(4): 609-623.
- 57: Boucherit HN, Boucherit H, Chikhi A, Bensegueni A, Merzoug A, Hioual KS, Mokrani. L'amarrage moléculaire : une nouvelle approche pour lutter contre le développement de

teberculose. *Revue de Microbiologie Industrielle, Sanitaire, et Environnementale*, 2013; 7(2): 133-149.

58: Friesner RA, Banks JL, Murphy R B, Halgren TA, Klicic JJ, Mainz DT, Repasky MP, Knoll EH, Shelley M, Perry JK. Glide : a new approach for rapid, accurate docking and scoring. Method and assessment of docking accuracy. *Journal of Medicinal Chemistry*, 2004; 47(7): 1739–1749.

59: Jorgensen WL, Maxwell DS, Tirado-Rives J. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *Journal of the American Chemical Society*, 1996; 118(45): 11225–11236.

60: Metropolis N, Ulam S. The Monte-Carlo method. *Journal of the American Statistical Association*, 1949; 44(247): 335–341.

61: Morris GM, Goodsell DS, Halliday RS, Huey R, Hart WE, Belew RK, Olson AJ. Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *Journal of Computational Chemistry*, 1998; 19(14): 1639–1662.

62: Trott O, Olson AJ. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of Computational Chemistry*, 2010; 31(2): 455–461.

63: Baxter J. Local optima avoidance in depot location. *Journal of the Operational Research Society*, 1981; 815–819.

64: Lourenço HR, Martin OC, Stützle T. Iterated local search: Framework and applications, in Handbook of Metaheuristics. *Springer* 2010: 363–397.

65: Wright S. Nocedal J. Numerical optimization. Springer New York 1999.

66: MolegroVirtual Docker, User manual 2011, 273.

67: Sousa SF, Fernandes PA, Ramos MJ. Protein-Ligand Docking. *Current Status and Future Challenges. Proteins*, 2006; 65: 15-26.

Partie II

***Modélisation des interactions
moléculaires « Lipase-EGCG »***

Chapitre III

Matériels et méthodes

III.1. Logiciel de simulation moléculaire

Toutes les simulations du docking moléculaires effectués dans cette étude ont été réalisés par Molegro Virtual Docker (MVD 2012.5.5.0, Molegro ApS) comme logiciel d'amarrage virtuel. Ce programme possède une interface utilisateur conviviale et il a la possibilité de réaliser un paramétrage fin à tous les niveaux du processus tels que : le docking, le scoring, la flexibilité des résidus et des ligands qui est modulable¹, limitant ainsi son aspect de « boîte noire ».

Ce programme possède une meilleure performance pour étudier et prédire le mode d'interaction ligand-récepteur en comparaison avec d'autres logiciels fréquemment utilisés tels que Autodock, Superflex, Flex and Gold.^{2,3}

MVD permet une meilleure visualisation des interactions hydrophobes qui résultent des simulations du docking. Les résultats obtenus par MDV, sont stockés et traités par la suite avec le logiciel Ligplot⁴. Ce logiciel est un programme de visualisation moléculaire à deux dimensions (2D) pour l'affichage animé et l'analyse des systèmes biomoléculaires de grandes tailles à l'aide d'une interface graphique et de la conception de scripts.⁵

Les différents formats d'extensions avec les quels sont importés le ligand et le récepteur (protéine cible) sont résumés dans le tableau 1.

Tableau 1. Principaux formats d'extensions des fichiers utilisés.

Format	Utilisation	Référence et documentation
pdb	Macromolécules, ligands, complexes	Protein Data Bank http://www.pdb.org
MOL	Ligand (1 molécule, 2D-3D)	MDL Molfile (MDL Informations Systems, Symyx, Accelrys)
SDF	ligands (n molécules, 2D-3D)	SYBYL (Tripos) http://www.tripos.com/tripos_resources/fileroot/pdfs/mol2_format.pdf
MOL2	ligands (1... n molécules, 2D-3D)	SYBYL (Tripos) http://www.tripos.com/tripos_resources/fileroot/pdfs/mol2_formatpdf

III.2. Structure du ligand

La structure 3D du ligand utilisé dans ce travail est importé sous forme PDB de la banque de donnée PubChem (<http://pubchem.ncbi.nlm.nih.gov>) sous le code CID 65064 (fig.1). Comme l'algorithme de docking de MVD explore l'espace conformationnel du ligand de manière randomisée via une approche de Monte carlo, la conformation de départ du ligand peut être prise au hasard. Ainsi, la structure de ligand a été directement utilisée pour la simulation de docking.

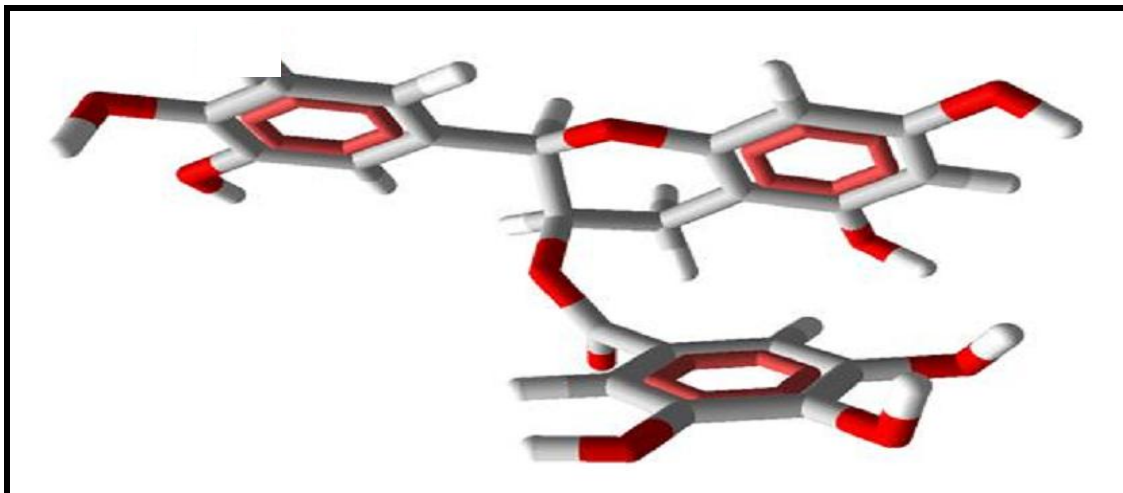


Figure 1. Structure 3D du ligand: Epégallocatéchine-3- gallate (EGCG).

III.3. Structures des récepteurs

Toutes les structures tridimensionnelles des lipases testées dans cette étude ont été tirées de la base de données Bookhaven Protein Data Bank (<http://www.pdb.org>). PDB présente une collection mondiale de données sur la structure tridimensionnelle (3D) de macromolécules biologiques: protéines essentiellement, et acides nucléiques. Ces structures sont essentiellement obtenues soit par cristallographie (rayons X), soit par RMN.

Les enzymes incluses dans cette étude sont : la lipase pancréatique humaine (LPH), la lipase pancréatique de porc (LPP), et la lipase gastrique de chien (LGC), qui est un bon modèle de la lipase gastrique humaine (LGH). Les structures tridimensionnelles des lipases étudiées sont identifiées respectivement par les codes suivant : 1ETH, 1LPB, 3K8Q.

III.3.1. Lipase *pancréatique de porc* (LPP)

La structure tridimensionnelle de la LPP disponible dans PDB est identifiée par le code 1ETH. Il s'agit d'une structure correctement définie avec une bonne résolution ($A=2,8\text{\AA}$)⁶. Cette structure est équipée d'un volet ouvert recouvrant le site actif. Ceci est favorable pour une étude d'ancrage de ligand au sein des cavités de la macromolécule.

Notre étude est basée sur la simplification d'enzyme comme une étape clé et essentielle pour accélérer et simplifier les calculs. Pour cela on élimine une des deux chaînes (fig.2) liées avec les molécules de co-cristallisation, les molécules d'eau, et le cofacteur. Une fois ces chaînes sont éliminées le modèle des enzymes est simplifié d'où on retient seulement un monomère de protéine avec une séquence d'une longueur de 875 acides aminés (fig.3).

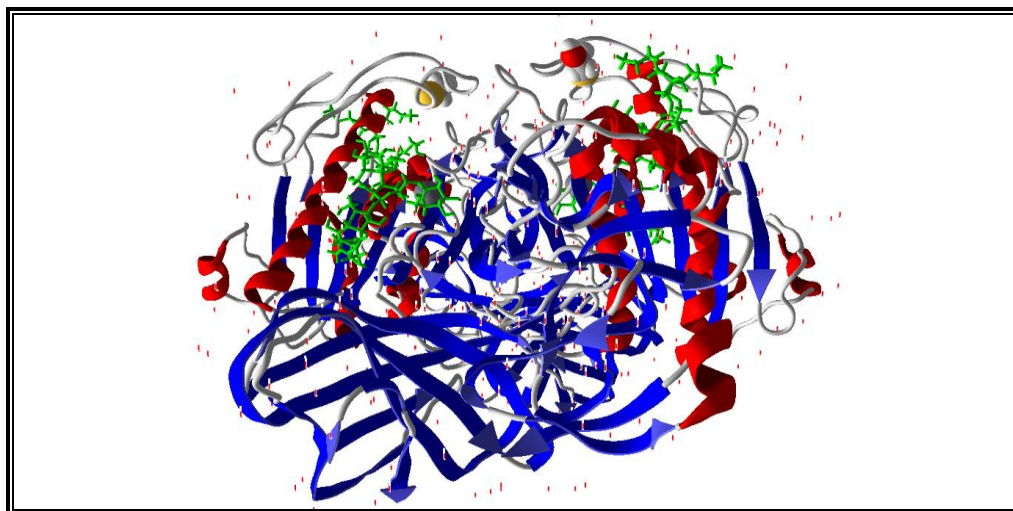


Figure 2. Structure tridimensionnelle de la LPP (code 1ETH) avec les résidus.

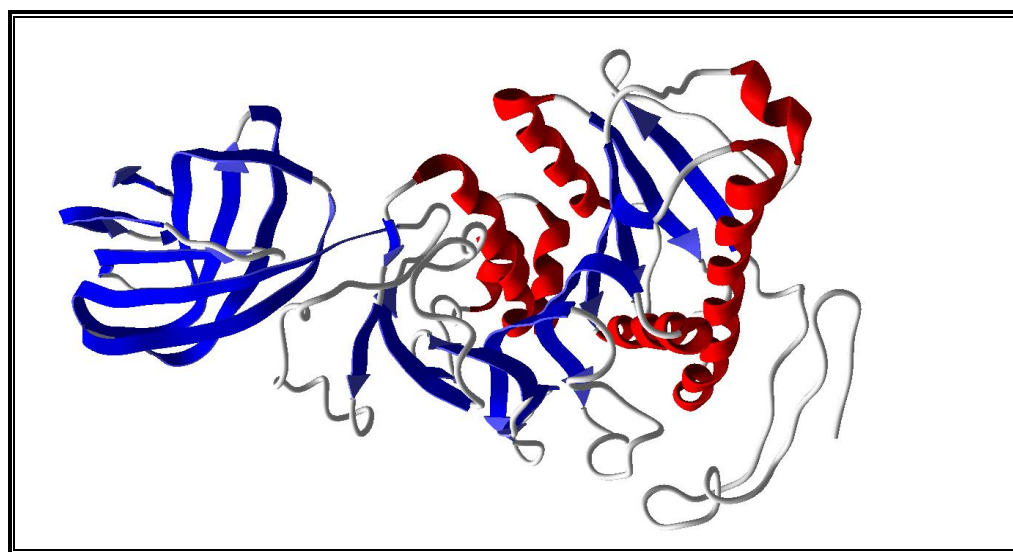


Figure 3. Modèle simplifiée de la structure tridimensionnelle de la LPP.

III.3.2. Structure de la lipase *pancréatique humaine* (LPH)

La structure tridimensionnelle choisie dans cette étude pour LPH est identifiée par le code 1LPB. Cette structure a été choisie suite à la bonne résolution (2,46 Å) et la présence d'un inhibiteur phosphonate (C11 alkylphosphonate) cristallisé lié de manière covalente à la sérine catalytique⁴. Cette liaison avec un des résidus catalytiques du site actif, permet d'avoir des informations d'origine expérimental sur le mode d'ancrage du ligand 1LPB.

Une première structure du complexe « lipase pancréatique humaine-colipase » obtenue par Herman van Tilbeurgh⁷, puis une seconde structure du même complexe obtenue en présence de phospholipides et de sels biliaires, dans laquelle la lipase pancréatique fut observée pour la première fois dans sa conformation « ouverte » et fonctionnelle⁸. Les structures obtenues par Herman van Tilbeurgh apportèrent des informations essentielles sur les relations structure-fonction du complexe lipase-colipase (fig.4).

La figure 4A représente la lipase *pancréatique humaine* (LPH) seule avec le volet contrôlant l'accès au site actif dans sa conformation fermée. La figure 4B expose la relation structure complexe lipase pancréatique humaine-procolipase de porc, avec le volet dans sa conformation ouverte et un inhibiteur (C11 alkylphosphonate) lié de façon covalente à la sérine du site actif. La colipase interagit avec le domaine C-terminal de la lipase et avec le volet dans sa conformation ouverte⁷⁻⁹

Dans cette section (fig.4B) le domaine C-terminal de la LPH (structure β -sandwich) possède une homologie structurale avec les domaines C2 (ou domaine de liaison des lipides dépendant du calcium) que l'on retrouve dans un grand nombre de protéines interagissant avec les lipides et impliquées dans la signalisation cellulaire. Ce domaine possède une boucle hydrophobe, la boucle $\beta 5'$, qui joue un rôle important dans l'adsorption de la lipase à l'interface huile-eau.^{7,8}

Dans la figure 5 on voit de façon plus explicite la relation structure-fonction du complexe [LPH-procolipase de porc] avec le volet dans sa conformation ouverte où un inhibiteur type C11 alkylphosphonate est lié de façon covalente à la sérine du site actif de la macromolécule.

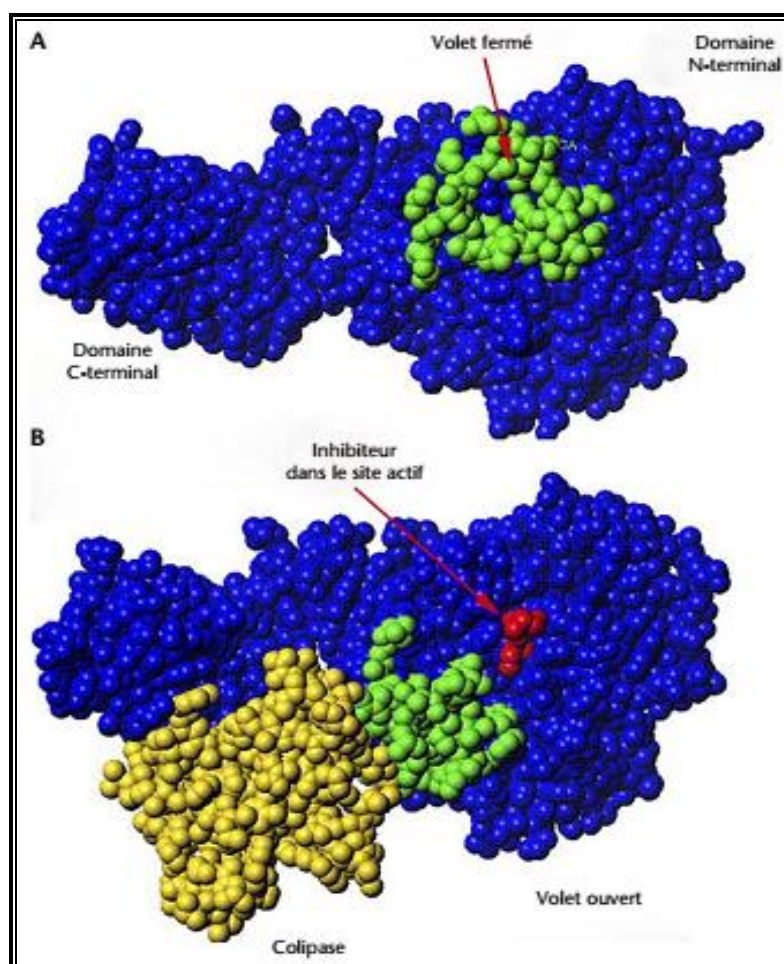


Figure 4. Structure 3D du complexe lipase pancréatique-colipase.¹²

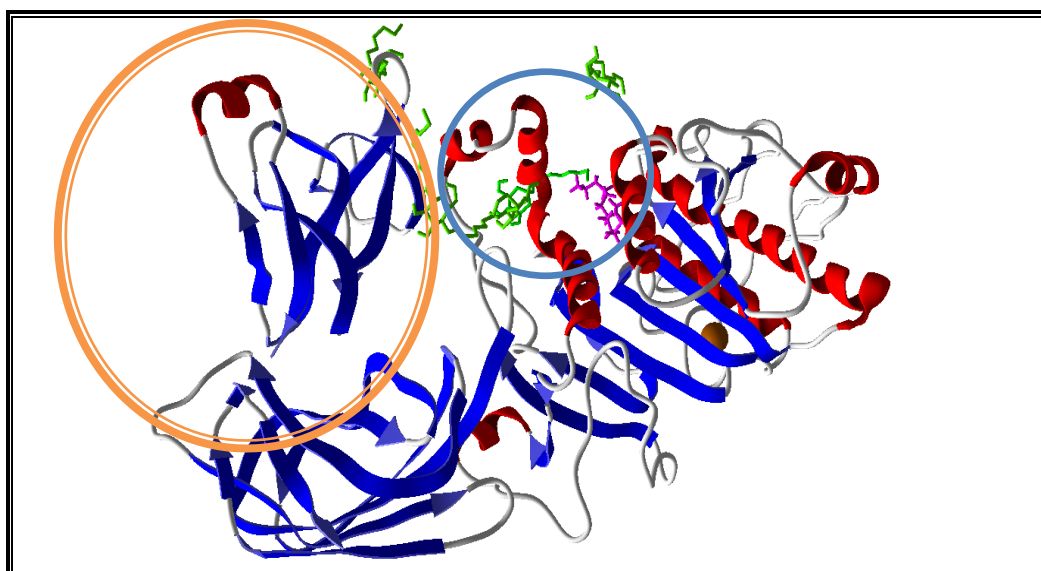


Figure 5. Complexe 1LPB (lipase pancréatique humaine-procolipase de porc).⁹

La LPH (fig.5) a été obtenue sous sa forme ouverte, complexée avec la proco-lipase de porc, entourée en orange. Le volet est représenté en bleu, et un cation de Ca^{+} complexé par les carboxylates des chaînes latérales des résidus Asp192 et Asp195. Pour passer à l'étape de modélisation la structure de LPH a été modifiée et le système ainsi obtenu est schématisé par la figure 6.

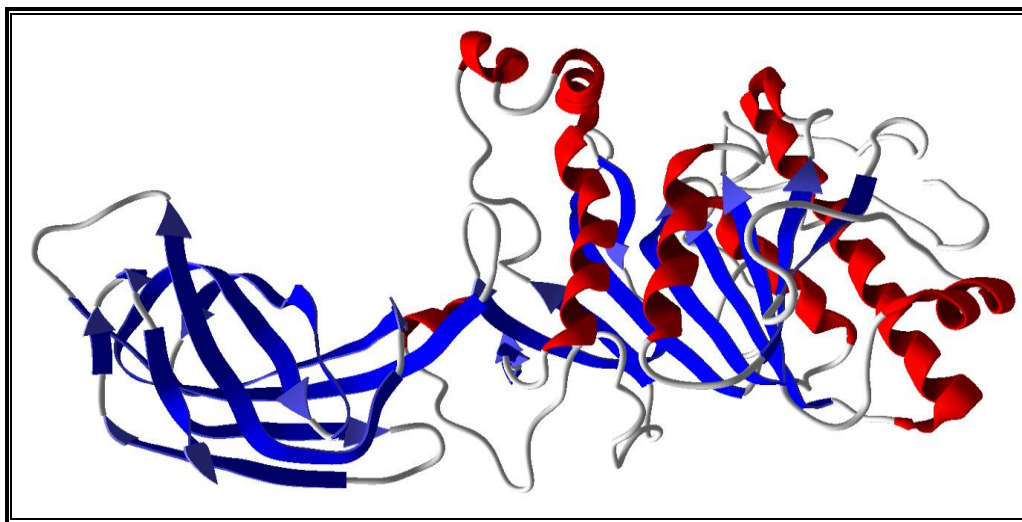


Figure 6. Modèle simplifiée de la LPH utilisé pour la modélisation.

Cette structure a été modifiée comme suit :

- Le cation Ca^{+} a été effacé.
- Les molécules d'eau absorbée ont été supprimées.
- La colipase est exclue de l'espace de travail quand le volet est dans sa forme ouverte.
- L'inhibiteur et les molécules co-cristallisé on été effacés de la chaîne.

III.3.3. Structure de la lipase *gastrique du chien* (LGC)

Dans notre étude nous avons testé la lipase gastrique pour voir l'effet inhibiteur de notre ligand. Pour cela nous avons choisi celle du chien (LGC) à cause de sa très grande ressemblance avec celle de l'humain (LGH).

La structure tridimensionnelle au format PDB de la LGC est disponible dans la banque des données des protéines (PDB) sous le code « 1K8Q », avec une résolution de 2,0Å. Le fichier PDB téléchargé est constitué d'un dimère dont chaque chaîne est composée de 377 acides aminés, 191 molécules d'eau absorbées, un inhibiteur phosphonate ($\text{C}_{15}\text{H}_{33}\text{O}_2\text{P}$) liée de manière covalente à la sérine catalytique, une molécule de β -OG, et un dimère de N-acétyl glucosamine (fig.7). L'inhibiteur alkylphosphonate est coloré en move, la molécule de β -OG

en jaune, les molécules d'eau sont représentées par des points rouges discontinus et les autres résidus en vert.

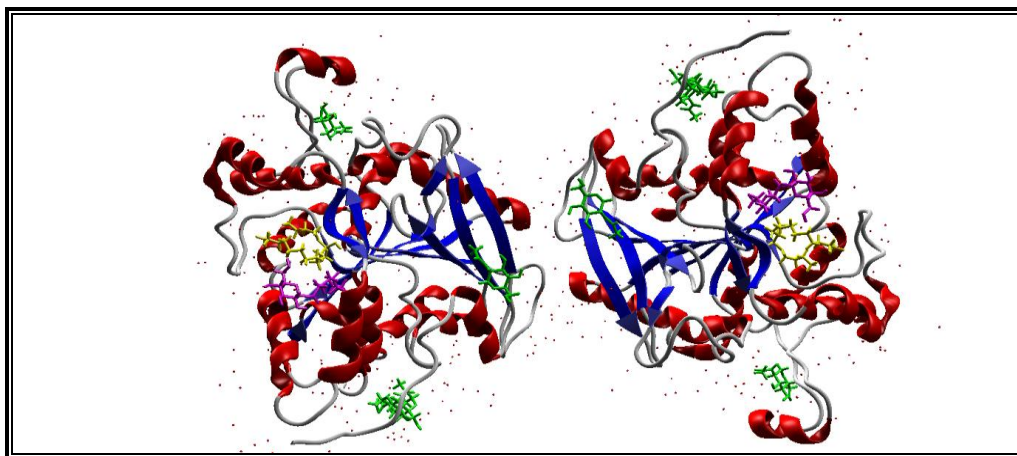


Figure 7. Structure de la LGC complexée avec une molécule de β -OG et un inhibiteur alkylphosphonate dans son site actif.¹³

Cette structure est modifiée pour la simulation comme suit :

- Seule la première chaîne a été conservée ; la deuxième a été effacée.
- Les molécules d'eau ont été enlevées.
- L'inhibiteur phosphonate ($C_{15}H_{33}O_2P$), la molécule de β -OG, et les dimères acétyl glucosamines ont été effacées de la chaîne restante.

Le système modifié ainsi obtenu est représenté par la figure 8.

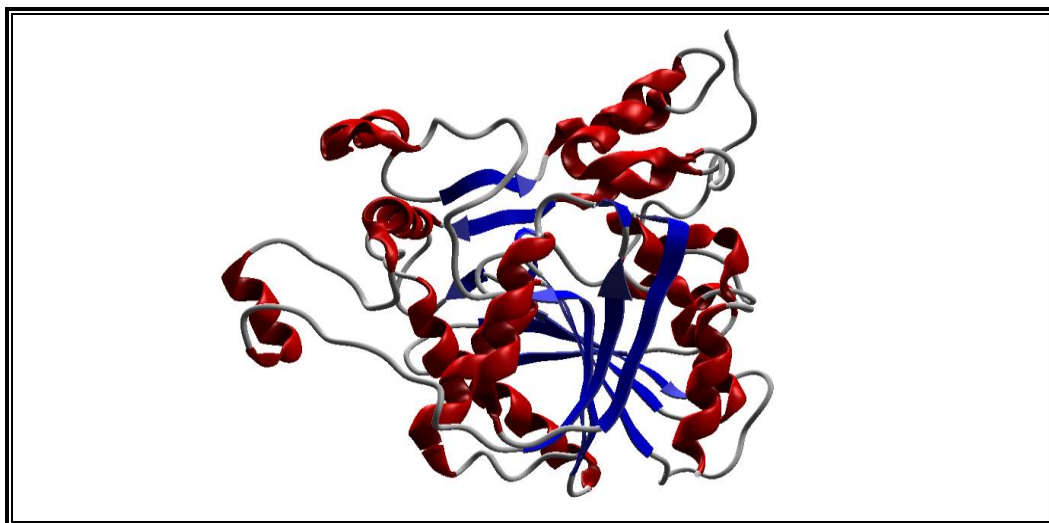


Figure 8. Représentation simplifiée de la LGC.

III.4. Préparation des récepteurs et du ligand

Généralement, les techniques d'amarrage moléculaire permettent de prédire, par des méthodes de calcul, les structures du complexe ligand-protéine en partant d'un ensemble variable de conformations et d'orientations du ligand (flexible ou non flexible) et parfois du récepteur (chaînes latérales, résidus et « backbone » flexible ou non). Le processus global se déroule en différentes étapes présentées ici comme consécutives, mais ce n'est pas toujours le cas (selon le logiciel utilisé).

- 1) La première étape consiste à importer et à préparer le ligand et le récepteur dans l'espace de recherche. Les fonctions de préparation générées automatiquement et qui sont intégrées dans le programme MVD, permettent de préparer le récepteur et le ligand dans l'espace de travail «work space». Dans cet 'espace de recherche le récepteur est considéré comme un corps rigide et le ligand sera flexible. Les principales préparations effectuées par le logiciel MVD sont :
 - ajout des hydrogènes explicites,
 - assignement des charges,
 - correction des types des liaisons (simple ou double),
 - détection des torsions et les liaisons rotatives.
- 2) La seconde étape permet de détecter les cavités présentes dans le récepteur en utilisant l'algorithme intégré pour la détection de cavité. Dans cette étape l'utilisateur peut identifier le site actif, soit à l'aide des données bibliographiques ou par la présence d'un ligand co-cristallisé au niveau de la poche principale et la visualisation ainsi des résidus.
- 3) La troisième étape dite d'amarrage moléculaire (docking wizard) permet de produire les complexes protéines-ligands et ceci à partir des différents algorithmes présents dans le logiciel. Ces algorithmes permettent à la fois de générer des conformations de complexes et aussi de les faire évoluer en utilisant une fonction de score. Cette dernière qui doit être toujours minimisée.
- 4) La quatrième phase de cette étape de préparation, consiste à sélectionner/trier et regrouper les scores. Cette étape est connue sous le nom «clustering». Il s'agit de faire ressortir des conformations représentatives, appelée : poses. Cette étape peut être

- 5) directement issue de la seconde étape ou être complétée par des calculs intermédiaires. Par exemple une minimisation du complexe ou un calcul de dynamique moléculaire.
- 6) Enfin, intervient une étape finale. Cette dernière permet l'analyse de la conformation des poses (appelé «runs») et des valeurs de scores. Cette étape implique souvent une inspection visuelle de chaque pose et des calculs de descripteurs, également des mesures concernant l'écart quadratique des erreurs (RMSD) ou encore des analyses statistiques.

Ce processus de ces étapes est schématisé ci-dessous (schéma 1):

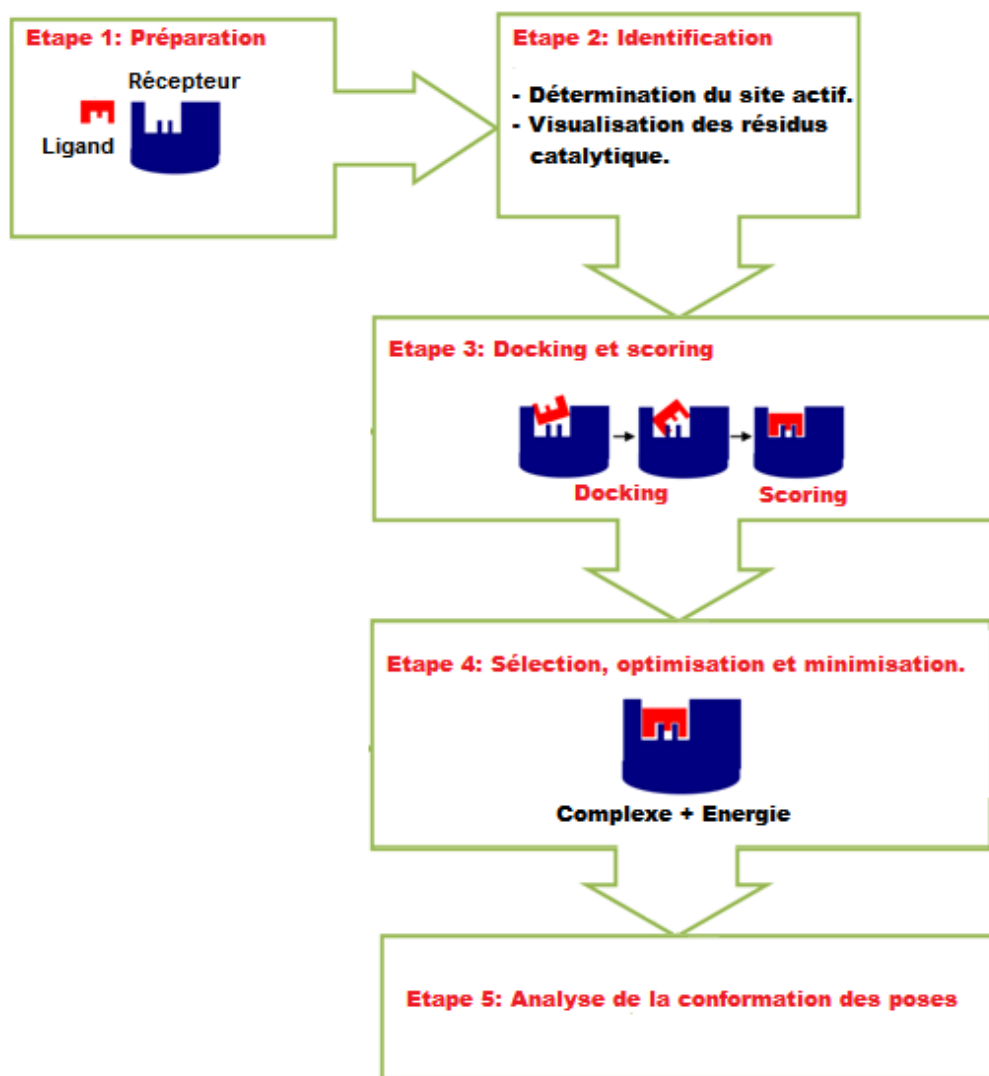


Schéma 3. Principe général d'un programme de docking.

III.5. Définition du site actif

Molegro virtual Docker (MVD) comprend un outil pour la détection de cavités dans la protéine (fig. 9). Cet outil se base sur un système de grille fonctionne comme suit :

- D'abord une grille tridimensionnel d'espacement fixée à $0,8\text{\AA}$, ayant les mêmes dimensions maximal de la protéine selon les axes X Y Z est construite suivant la structure de la molécule. Pour chaque point de la grille si la distance entre la protéine et l'atome lourd le plus proche (hydrogène) est inférieur ou égale $1,2\text{\AA}$, le point est placé comme occupé dans le cas contraire le point sera classé comme libre.¹⁴
- Ensuite les points libres sont effacés en partant de l'extérieur de la grille vers la surface de la protéine jusqu'à ce le nuage de point restant aient une largeur maximale prédéfini par l'utilisateur. Chaque nuage de point non effacé représente une cavité pour reproduire au mieux les dimensions du site actif dans la simulation.¹⁵

Les enzymes ayant une topologie de surface assez irrégulière, qui permet de détecter plusieurs cavités. Cette procédure automatique, permet au modélisateur, en se basant sur ces connaissances de l'enzyme étudiées, de sélectionner parmi les cavités détectées celles qui constituent le vrai site de liaison pour y faire le docking du ligand dans les étapes suivants.¹⁵

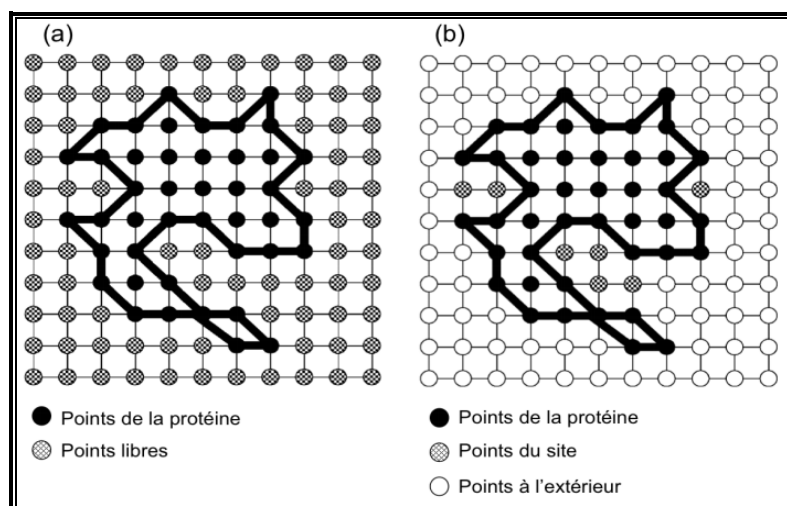


Figure 9. Schéma représentative de la forme bidimensionnelle de la grille autour et à travers la protéine (traits noire), ainsi que ces points d'intersections.

La figure 9 montre que les points sont classées occupées, sont représentées par des cercles noires et celles libres par des cercles gris. Les points libres sont éliminés (cercle blanc) de

manière à obtenir des cavités dont l'ouverture maximale est définie par l'utilisateur, les points libres restantes définissent les cavités qui seront prises pour le traitement de docking.¹⁶

Cette étape consiste à déterminer les régions connectées. Deux points de la grille sont connectés s'ils sont voisins. Les régions dont le volume est inférieur à 10 Å sont considérées comme non pertinentes (fig.10), les cavités trouvées sont ensuite classées en fonction de leur volume.

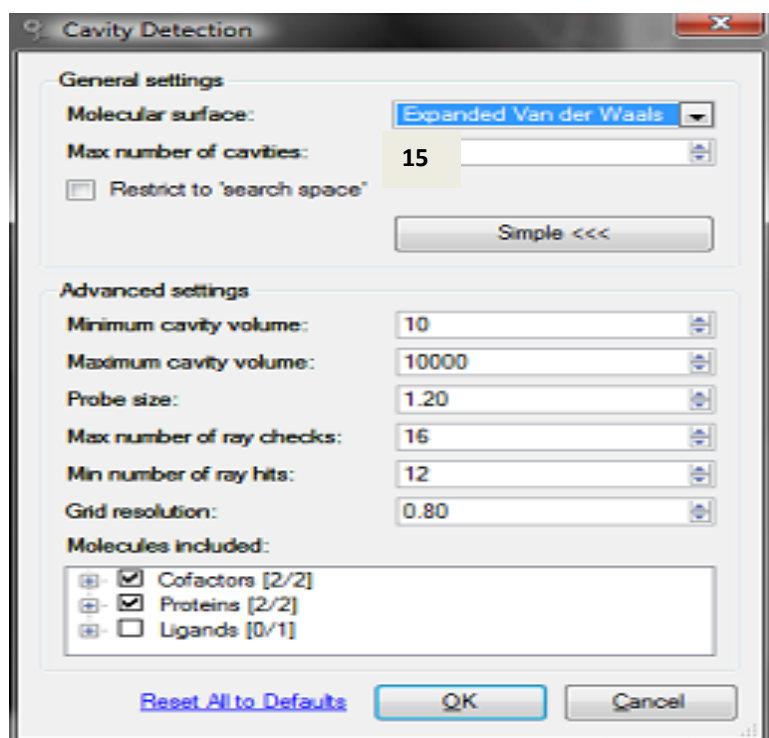


Figure 10. Paramètre utilisé pour la détection des cavités.

Dans notre travail, nous avons détecté par le biais de cet algorithme, cinq cavités pour la LPH, LGC, et sept cavités pour la LPP, qui sont illustrés respectivement par les figures 11, 12, 13. Les cavités détectées pour les trois récepteurs sont classées selon leur volume dans le tableau 2.

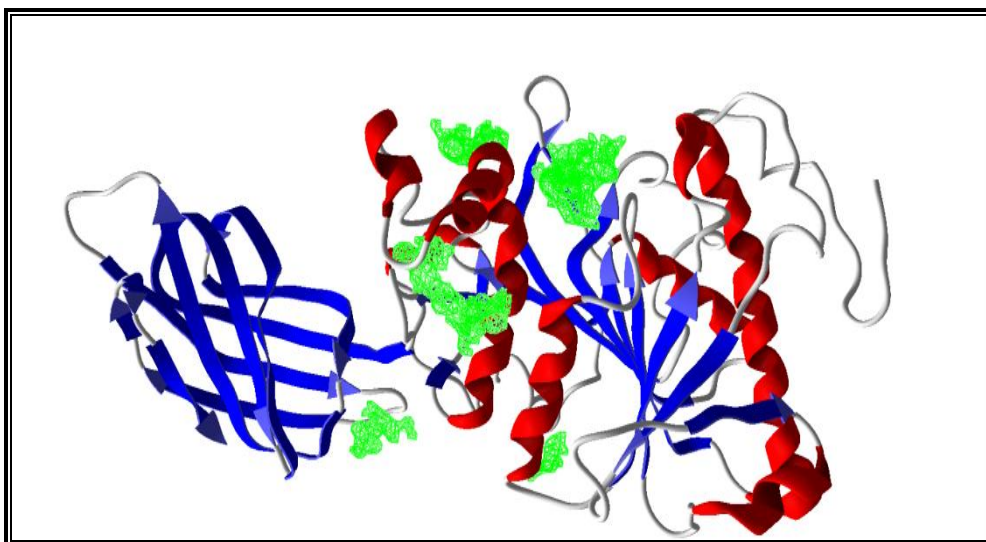


Figure 11. Illustration des cinq cavités du 1LPB détectées (en vert) par MVD.

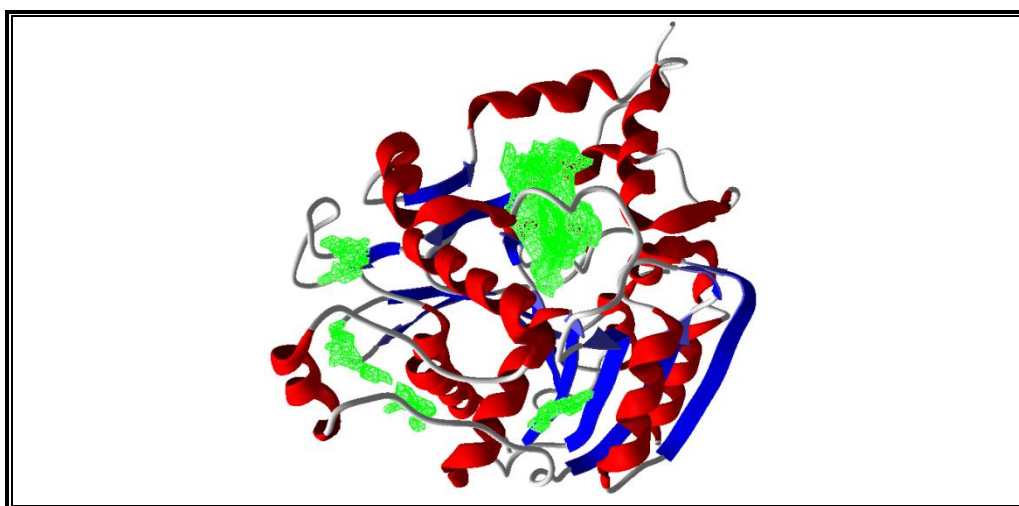


Figure 12. Cavités détectés (en vert) dans les récepteurs 1K8Q.

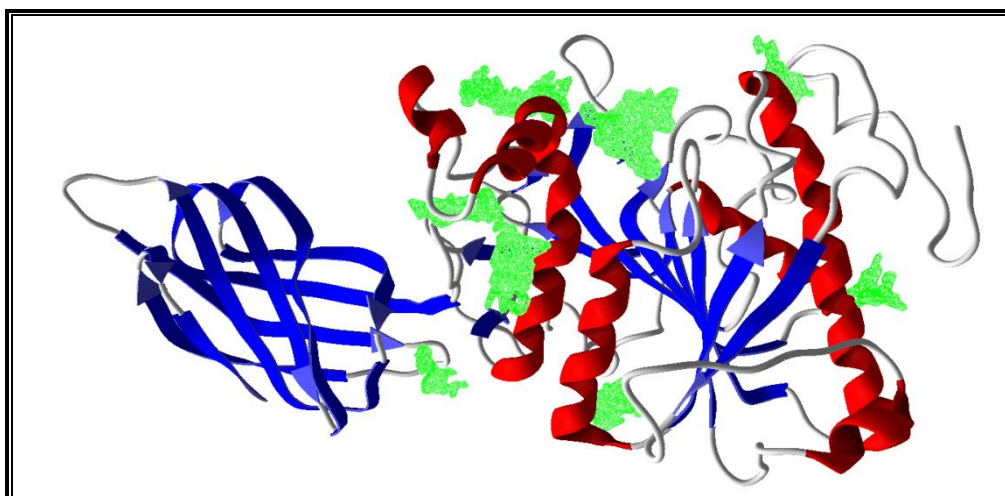


Figure 13. Illustration des sept cavités du 1ETH détectées (en vert) par MVD.

Tableau 2. Cavités détectées pour chaque récepteur : 1ETH, 1LPB, et 1K8Q.

Récepteur	N° de la cavité	Volume de la cavité (Å ³)
1ETH	1	101,52
	2	74,52
	3	54,64
	4	20,52
	5	14,47
	6	14,46
	7	12,09
1LPB	1	121,8
	2	92,67
	3	52,22
	4	28,16
	5	14,84
1K8Q	1	340,48
	2	31,74
	3	18,43
	4	15,36
	5	14,33

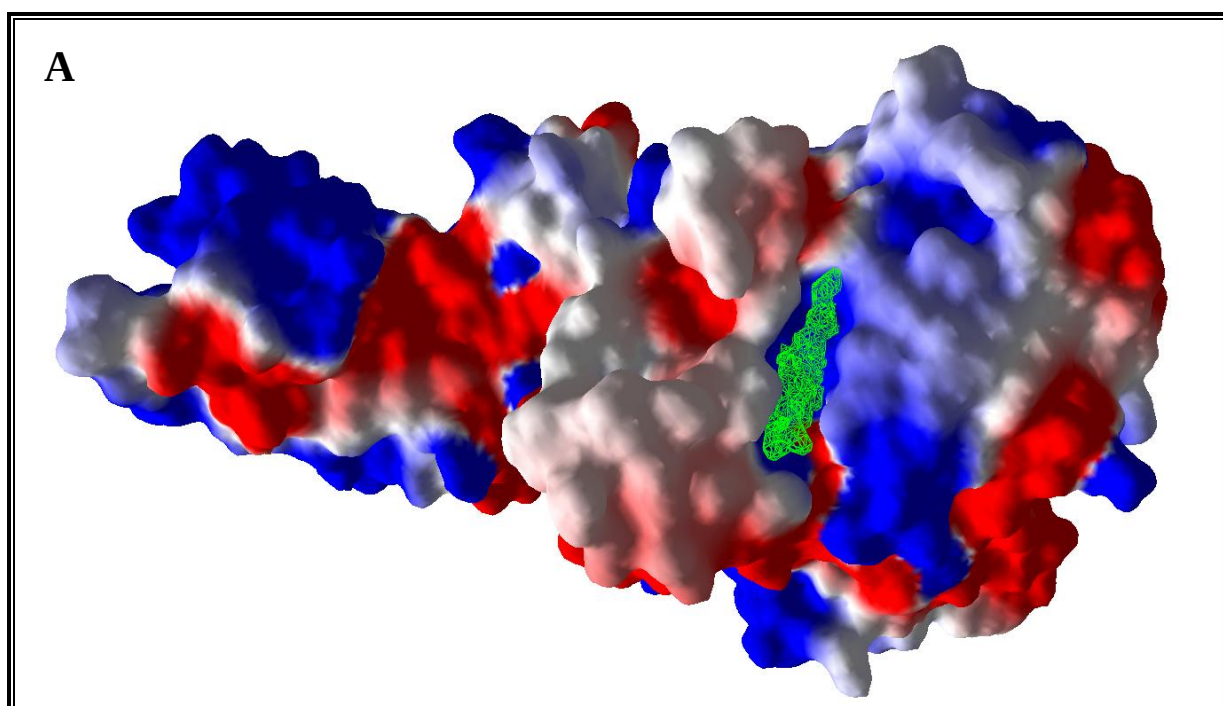
Le site actif des lipases a été défini en se basant sur la description, de leurs cavités, à partir des données cristallographiques trouvées dans la littérature (fig.14).

- Pour la LGC, la cavité catalytique se compose d'une gorge profonde, avec les dimensions Approximative de 20 Å à 7 Å de largeur et 20 Å de profondeur, la triade catalytique constituée, par les résidus Ser 153, His 353, Asp 324 est localisé au fond de la cavité ; les résidus de trou oxyanionique sont Q154 and L67, la cavité de LGC est essentiellement hydrophobe, tapissé par les résidus aliphatique Leu 195, Leu 326, Leu 68, Leu 67, Ile192, Ile 243, Val 272, Val 279, Val 185, Val 182, seul une petite région autour des résidus catalytique possède un caractère hydrophile dû à la présence des chaine latérale des résidus Thr 190, Thr188, Gln145.

- Pour la LPP, la triade catalytique localisée dans la partie centrale qui se trouve au fond de la cavité est constituée par les résidus Ser 153, His 264 et Asp176. La partie interne de la cavité possède deux régions : une région hydrophobe constituée par les résidus Phe78, Tyr115, Ala179, Pro181, Phe216 et une petite région faiblement hydrophile, délimité par les résidus : Leu 256, Val 270, Ile 79, Leu 154.
- Pour la LPH, la cavité de site actif est large et allongée, constitué par les résidus Ser 152, His 263 et Asp176, la majorité de la cavité présente une région hydrophobe constituée par les résidus Leu 213, Ile 209, Leu 153, Pro 180. La région hydrophile est plus importante dans la LPH que dans la LPP.

Les données cristallographiques trouvées dans la littérature montrent que la cavité 1 présente la cavité principale. Cette confirmation a été validée par la visualisation des résidus catalytique autour de cette cavité, ainsi que la présence d'inhibiteur compétitif co-cristallisé dans cette poche pour les deux lipases LPH, LGC.

A partir de ces résultats nous avons donc décidé d'adopter cette cavité dans le processus de docking concernant notre étude.



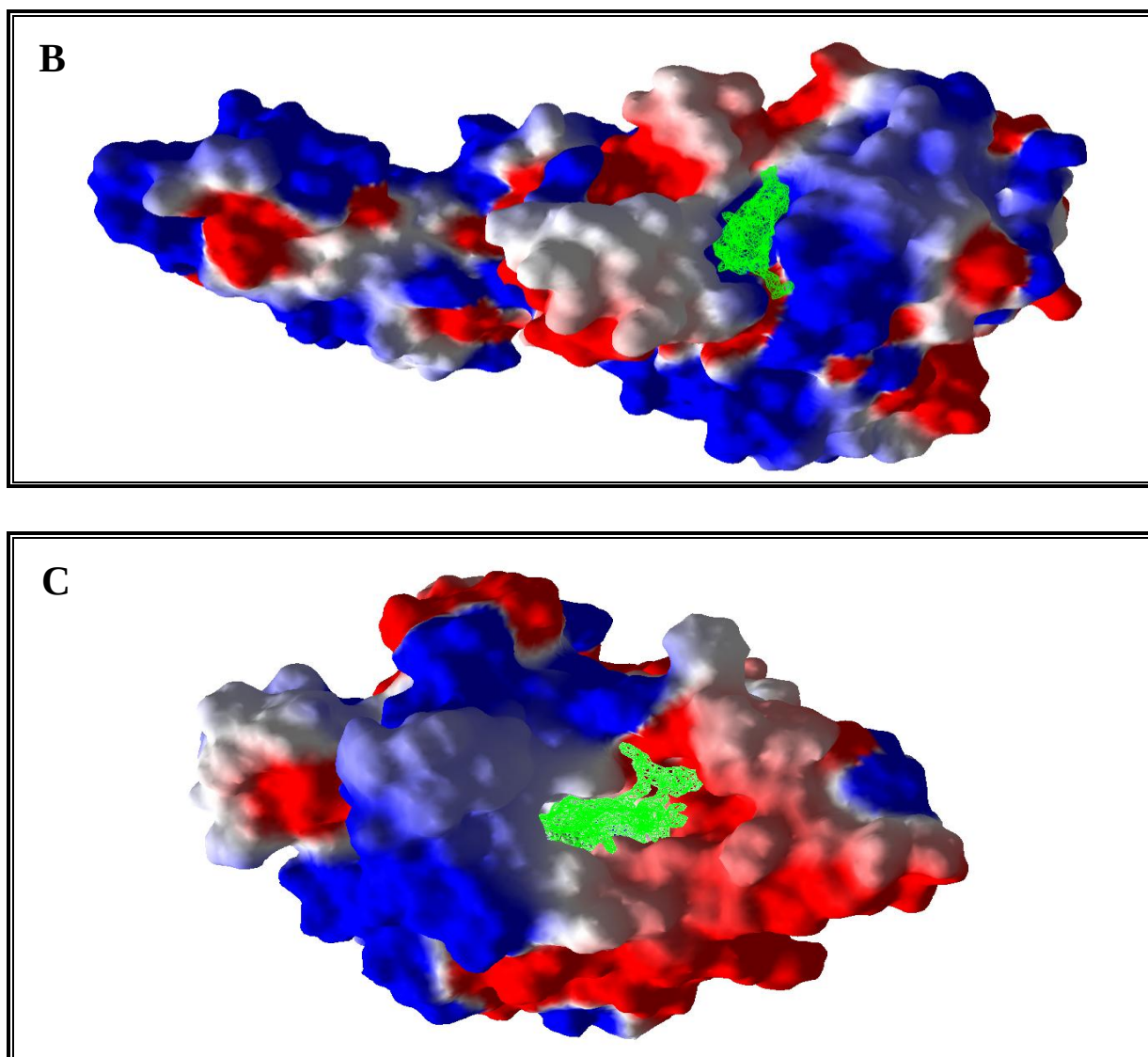


Figure 14. Orientation et hydrophobicité de la cavité principale, A(LPP) ; B(LPH), C(LGC). Zone hydrophobe (colorée en bleu), zone hydrophile (colorée en rouge).

III.6. Algorithme de recherche implémentée dans MVD

Le docking moléculaire permet de positionner le ligand au sein du site du récepteur. Ce dernier correspond aux différentes enzymes (LPP, LPH, LGC) testées dans notre travail. Le complexe ligand-récepteur formé, adoptera la conformation la plus stable, engendrant le niveau énergétique le plus faible. La recherche de la conformation la plus favorable entre de ligand (flexible) et le récepteur (rigide). En effet, dans de tels cas, la structure cristallographique peut être considérée comme plus représentative de l'état de la protéine dans son milieu conventionnel, ce qui augmente les chances de simuler correctement la complexation des ligands¹⁷.

Les algorithmes évolutionnaires sont des méthodes d'optimisation génériques et itératives, reposant sur un processus adaptatif similaire à celui de l'évolution naturelle. Ils sont classés, avec les réseaux de neurones et les systèmes à logique floue. Contrairement à la plupart des techniques d'optimisation, ces algorithmes sont capables de gérer simultanément plusieurs solutions et de les regrouper dans un ensemble appelé « population ».¹⁸

Cette population est exposée à une pression de sélection, intégrée dans une fonction de scoring qui décrit le problème à optimiser et dans laquelle plusieurs degrés de liberté interviennent. L'ensemble des degrés de liberté permettant de définir une solution est appelé «génome». La sélection des solutions les plus performantes d'après la fonction de score, appelées «parents» sont compensée par la génération de nouvelles solutions appelées « enfants ». Cette sélection vient dans le but de maintenir la diversité de la population.

Les enfants sont générés par modification du génome parental par des opérateurs, classés suivant le nombre de parents sur lesquels ils s'appliquent. Les opérateurs unaires sont appelés «mutations», ceux binaires appelés « recombinaisons ». Quand les enfants créés, prennent la place des plus mauvaises solutions de la population, celle-ci sera de nouveau exposée à la pression de sélection implémentée par la fonction de scores, et l'évolution continue par une nouvelle itération (génération). Au fur et à mesure que les générations et les solutions sont de mieux en mieux adaptées au problème décrit, la fonction objective émerge. Ces étapes se poursuivent jusqu'à la convergence de la population, donnant à la fin un nombre déterminé de générations.¹⁸

Cette dernière appliquée au problème du docking, cherche à décrire les interactions entre le ligand et le récepteur, donnant ainsi les degrés de liberté correspondant aux positions, orientations et conformations du ligand et du récepteur. De telles méthodes sont générales car elles peuvent s'appliquer à n'importe quel type de problème sans adaptation algorithmique particulière. Le cycle typique d'un algorithme évolutionnaire « AE » est schématisé ci-dessous.

Dans cette étude trois algorithmes évolutionnaire. Ces algorithmes sont: MolDock optimizer, MolDock SE (MSE), et Iterated Simplex (Simplex).

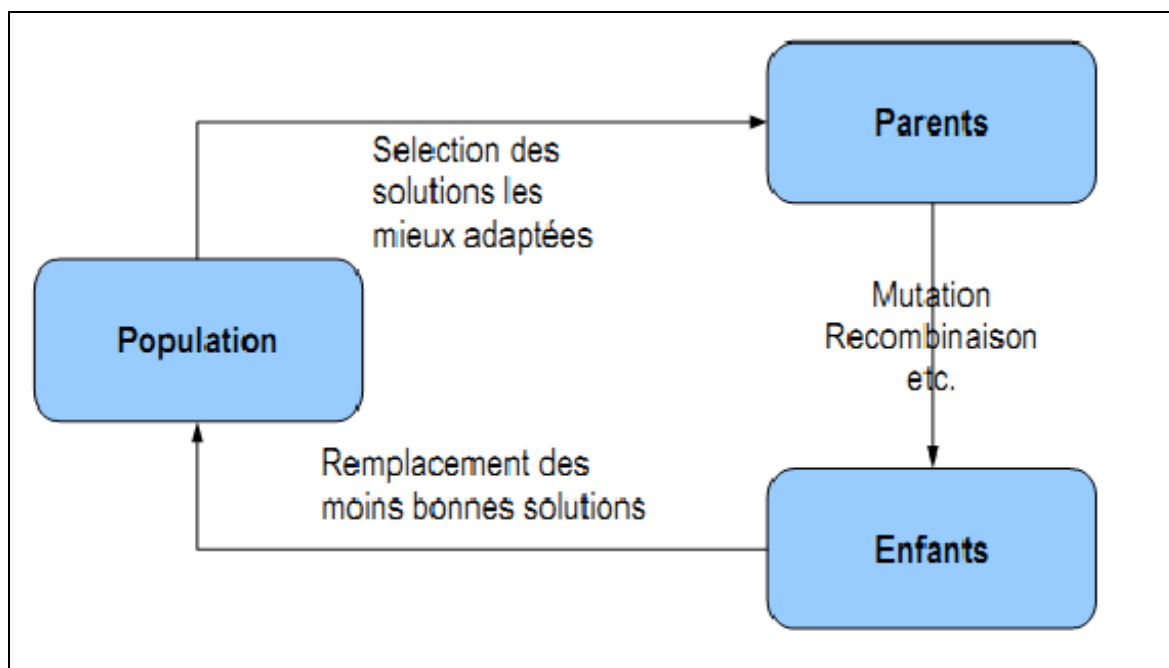


Figure 15. Représentation du cycle typique d'un algorithme évolutionnaire « AE ».

III.6.1. MolDock Optimizer

MolDock Optimizer¹⁹ ou optimiseur du docking moléculaire, est un algorithme intégré dans MVD, basé sur un algorithme évolutionnaire^{20,21}. Il appartient à la famille des algorithmes évolutifs (EA). Ces derniers sont utilisés comme technique d'optimisation itérative, inspirée de la théorie darwinienne. L'algorithme d'évolution différentielle guidé, est utilisé pour MolDock Optimizer. Cet algorithme est basé sur une variante EA appelée évolution différentielle (DE). L'algorithme DE a été introduit par Storn et Price en 1995.²²

Comparé à des techniques EA plus connues (par exemple, algorithmes génétiques, programmation évolutive et stratégies d'évolution), DE utilise une approche différente pour sélectionner et modifier les solutions candidates (individus). L'idée novatrice principale en DE est de créer une progéniture à partir d'une différence pondérée des solutions parentes.

III.6.2. L'algorithme MolDock SE (Simplex Evolution)

L'algorithme Simplex Evolution (SE) est connu pour mieux fonctionner sur certains complexes où l'algorithme standard MolDock échoue¹⁹. Il s'agit d'une heuristique de recherche alternative, qui peut être utilisée avec les fonctions de score MolDock ou MolDock Grid. L'implémentation de MolDock SE sera modifiée pour être adaptée au docking (en

incluant l'étape de génération de pose, et la façon dont les simplexes initiaux sont créés), quand d'autres algorithmes basés sur la recherche simplex parallèle existent. L'algorithme d'évolution simplex effectue une recherche combinée locale / globale sur les poses générées par le générateur de pose. La recherche locale est effectuée à l'aide de l'algorithme de recherche locale Nelder-Mead, mais contrairement au schéma original de Nelder-Mead, l'algorithme a été étendu pour tenir compte de la position des autres individus de la population.²³

III.6.3. L'algorithme Simplex Itéré

Cet algorithme est généralement plus robuste (résultats de docking reproduisant w.r.t avec des scores similaires) que les SE et MolDock Optimizer. Cet algorithme fonctionne comme suit:

- D'abord, une population initiale de poses est créée (le nombre initial de poses est déterminé par le paramètre de taille de population).
- Ensuite, le processus suivant sera exécuté jusqu'à ce que des itérations maximales se produisent: Chaque individu de la population sera affiné en utilisant l'algorithme de recherche locale Simplex (également appelé Nelder-Mead). L'algorithme Simplex fonctionnera pour des pas maximums ou jusqu'à la différence fractionnaire entre les meilleurs et les plus mauvais sommets dans le Simplex (w.r.t). Dans ce cas la fonction de score d'ancrage utilisée est inférieure à une valeur de tolérance donnée. Lorsque tous les individus ont été affinés, le meilleur individu (nommé: meilleure solution d'itération) sera encore affiné en utilisant le même algorithme Simplex mais avec une valeur de tolérance inférieure. Lorsque les itérations maximales ont eu lieu, l'algorithme se termine et renvoie le meilleur solution (s) trouvée (s).¹⁹

Pour ce qui nous concerne, la procédure de docking de ligand au sein de la protéine a été effectuée pour chaque algorithme avec les paramètres suivant (tableau 3).

Tableau 3. Paramètres utilisés pour chaque algorithme : MolDock optimizer, MolDock SE (MSE), and Iterated Simplex (Simplex).

<i>Algorithme de recherche</i>	<i>MolDock optimizer</i>	<i>MolDock SE</i>	<i>Simplex</i>
Taille de population	50	50	20
Interactions maximales	2000	1500	100
Facteur d'échelle	0,5	-	-
Taux de croisement	0,9	-	-
Seuil d'énergie	-	100	-

D'autres paramètres de docking sont utilisés. Le nombre de tours (runs) est fixé à 30, quand le ligand se place dans la cavité détectée et le casier d'optimisation des interactions hydrogène est déjà cochée. Il a été démontré que cette option réduit le nombre des interactions hydrogènes peu probable. L'écart minimal pour la discrimination des poses de ligand est égal à 2Å°. Cette valeur évite que trop de poses semblables soient retenues dans la liste de sauvegarde, ce qui mènerait à la rétention des complexes redondants. Le «scoring» de poses obtenues pour chaque ligand (après 30 tours de recherche) mène à un classement des meilleures poses basé sur leurs énergies d'interaction.

III.6.4. Fonction de score

Lors de la simulation du docking plusieurs poses candidates peuvent être obtenues pour chaque ligand. L'option du scoring est de sélectionner les meilleures poses avec les énergies les plus basses. Deux fonctions de score sont testées dans cette étude, MolDock Score et PLANTS Score.

a) Moldock Score

Moldock Score utilisée par MVD est un dérivé, qui amélioré à partir des fonctions de score PLP (piecewise linear potential). Ces fonctions qui ont été initialement proposées par Gehlhaar et al.^{24,25} et développées plus tard par Yang et al.²⁶ avec l'insertion de nouveaux termes de liaisons hydrogène et de charges.

La fonction de score, EScore, est définie par les termes d'énergie suivants (Eq.5):

$$E_{\text{scor}} = E_{\text{inter}} + E_{\text{intra}} \quad \text{Eq.5}$$

Avec :

E_{inter} : énergie d'interaction ligand-protéine.

E_{intra} : énergie d'interaction intramoléculaire du ligand.

L'énergie d'interaction ligand-protéine est donnée par l'équation suivante:

$$E_{\text{inter}} = \sum_{i \in \text{ligand}} \sum_{j \in \text{protein}} \left[E_{\text{PLP}}(r_{ij}) + 3320 \frac{q_i q_j}{4r_{ij}^2} \right] \quad \text{Eq.6}$$

Cette équation prend en compte tous les atomes lourds du ligand (i) et les atomes lourds de la protéine (j) ainsi que les atomes des cofacteurs et des molécules d'eau dans le cas échéant.

Avec :

E_{PLP} : le potentiel linéaire PLP

La valeur numérique 3320 est donnée pour fixer l'unité de l'énergie électrostatique en Kcal.mol⁻¹. Les constantes diélectriques représentent respectivement les charges des atomes du ligand et ceux de la protéine. Pour cela nous avons donné des exemples de charges d'atomes (tableau 4).

Tableau 4. Exemples de charges d'atomes. Modèle de charges prisent en compte par MolDock Score [Grid] pour différents ligands et protéines.²⁷

Atomes de ligand	Atomes de protéine	Charge
Les atomes de l'azote (N) dans -C(NH ₂) ₂	His (atomes N) ; Arg (atomes N)	0,5
Les atomes de l'azote (N) dans -N(CH ₃) ₂ et -(NH ₃)	Lys (atomes N)	1,0
Les atomes de l'oxygène (O) dans -COO ; -SO ₄ ; -PO ₂	Asp (atomes O) ; Glu (atomes O)	-0,5
Les atomes de l'oxygène (O) dans -PO ₃		-0,66
Les atomes de l'oxygène (O) dans -SO ₃		-0,33
Les atomes de l'azote (N) dans -SO ₂ NH		-0,1
Ions métalliques tel que Na ⁺		+1,0
Ions métalliques tel que Ca ²⁺ et Fe ²⁺		+2,0

Le potentiel prend en considération principalement deux paramètres :

- le premier concerne l'optimisation des interactions stériques (Van der Waals) entre les atomes.
- Le second correspond à l'optimisation des potentiels des liaisons hydrogènes.

Une liaison est considérée comme une liaison hydrogène si l'un des atomes peut donner un atome d'hydrogène et l'autre atome peut l'accepter. L'équation suivante représente l'énergie d'interaction intramoléculaire du ligand :

$$E_{intra} = \sum_{i \in \text{ligand}} \sum_{j \in \text{protein}} E_{PLP}(r_{ij}) + \sum_{\text{flexible bonds}} A [1 - \cos(m \cdot \theta - \theta_0)] + E_{clash} \quad \text{Eq.7}$$

Le premier terme de cette équation correspond à la somme d'énergies d'interaction entre toutes les paires d'atomes du ligand. Le second terme représente l'énergie de torsion.

b) PlantScore

La fonction de score, PlantScore, est définie par les termes d'énergie suivants (Eq.8):

$$E_{plantscore} = f_{PLP} + f_{clash} + f_{tors} + f_{site} - 20 \quad \text{Eq.8}$$

Le potentiel PLP est similaire à celui utilisé par MolDock Score, sauf qu'il prend en compte plusieurs types d'interactions (répulsifs, enterrés, non polaires, liaisons hydrogène et métaux). Le choc de ligand et les potentiels de torsion, f_{clash} et f_{tors} prennent en compte les conflits des ligands internes et les contributions de torsion pour les liaisons flexibles dans le ligand. Le terme f_{site} spécifie une pénalité qui est calculée dans le cas d'une conformation de ligand (pose) est située à l'extérieur du site de liaison. Le décalage énergétique de -20 est initialement requis pour l'algorithme de recherche PLANTS. Ce décalage énergétique est inclus ici afin que les scores PLANTS soient comparables à celle standard.¹⁴

Les paramètres de calculs, indépendants entre eux, sont choisis pour augmenter l'efficacité d'amarrage. Deux scores basés sur des contributions énergétiques ont été sélectionnés pour l'analyse des résultats de calculs, MolDock et Rerank, ainsi que le RMSD. Le score MolDock, exprimé en unité arbitraire, correspond à une composante énergétique calculée

(cf. Eq.5) avant et après minimisation. Le score Rerank, aussi exprimé en unité arbitraire, correspond à refaire une pondération des composantes du score MolDock. Ce score est sensé mieux reproduire les corrélations entre les affinités expérimentales et celles calculées, obtenues à partir des tests effectués pour les complexes protéine-ligand. Le filtre *in silico* est basé sur les règles suivantes:

- Les poses présentant les meilleurs scores MolDock et Rerank sont sélectionnées (valeur négative la plus faible). Cette sélection correspond à environ 20% de la totalité des poses.
- Chaque pose est analysée visuellement de façon individuelle et les poses conformes sont sélectionnées. La conformité d'une pose est définie par la similarité avec des structures cristallines des complexes au niveau de la disposition de la tête polaire du ligand et la trajectoire des chaînes aliphatiques. En effet, les chaînes aliphatiques ne se positionnent pas de façon aléatoire dans le site, mais progressent le long de trajectoires au nombre limité.

Lorsqu'une pose correspond à la fois au meilleur score MolDock et au meilleur score Rerank, elle est considérée comme pose forte. La pose forte est conforme, celle-ci retenue est aussi notée pose limite²⁸. Dans le cas contraire, si la pose forte n'est pas conforme, celle-ci retenue correspond à la pose conforme ayant les meilleurs scores MolDock et Rerank. Ce descripteur, (pose limite ou retenue) est corrélé avec de bonnes valeurs de RMSD ($< 2\text{\AA}$). Dans notre cas, les fluctuations induites par les chaînes aliphatiques des ligands ne permettent pas d'atteindre bien ce dernier objectif.

A partir de ce protocole établi pour le processus d'amarrage, nous avons réalisé notre travail concernant la recherche et l'étude des interactions Enzyme-Ligand. Le ligand choisi dans ce cas est l'Epigallocatechine gallate (EGCG) qui un polyphénol extrait du thé vert et trois lipases digestives qui constituent la cible. Les résultats et la discussion de cette étude seront développés dans le chapitre 4.

Références Bibliographiques

- 1: Dastmalchi S. Methods and algorithms for molecular docking-based drug design and discovery, 1ST edition, British library, 2016.
- 2: Ratnavali G, Kanaka Durga Devi N, Bhavya Sri K, Kalyan Raju j, B. Sirisha B, Kapitsa R. An attempt to screen top colorectal cancer drugs by using Molegro Virtual Docker. *Annals of Biological Research*, 2011; 2 (1): 114-126.
- 3: Araújo JK, Lima JA, Pinto A-D, Alencastro RB, Albuquerque MG. Docking of the alkaloid geissospermine into acetylcholinesterase: a natural scaffold targeting the treatment of Alzheimer's disease. *Journal of Molecular Modeling*, 2011; 17: 1401-1412.
- 4: <https://www.ebi.ac.uk/thornton-srv/software/LIGPLOT>.
- 5: Andrew CW, Roman AL, Janet MT. LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. *Protein Engineering, Design and Selection*, 1995; 8(2): 127–134.
- 6: Hermoso J, Pignol D, Kerfelec B, Crenon I, Chapus C, Fontecilla-Camps JC. Lipase activation by nonionic detergents: the crystal structure of the porcine lipase-colipase-tetraethylene glycol monoethyl ether complex. *Journal of Biological Chemistry*, 1996; 271: 18007-18016.
- 7: Van Tilbeurgh H, Sarda L, Verger R, Cambillau C. Structure of the pancreatic lipase procolipase complex. *Nature*, 1992; 359:159-62.
- 8: Van Tilbeurgh H, Egloff MP, Martinez C, Rugani N, Verger R, Cambillau C. Interfacial activation of the lipase-procolipase complex by mixed micelles revealed by X-Ray crystallography. *Nature*, 1993; 362 : 814-20.
- 9: Egloff MP, Marguet F, Buono G, Verger R, Cambillau C, Van Tilbeurgh H. The resolution structure of the pancreatic lipase-colipase complex inhibited by C11 alkyl phosphonate. *Biochemistry*, 1995; 34(9): 2751-2762.
- 10: Chahinian H, Sias B, Carriere F. The C-terminal domain of pancreatic lipase: functional and structural analogies with C2domains. *Current Protein & Peptide Science*, 2000; 1(1): 91-103.

- 11:** Bourbon Freie A, Ferrato F, Carriere F, Lowe ME. Val 407 and Ile 408 in the 5' Loop of Pancreatic Lipase Mediate Lipase-Colipase Interactions in the Presence of Bile Salt Micelles. *Journal of Biological Chemistry*, 2006; 281: 7793-800.
- 12:** Carrière F. Soixante ans de recherche sur la lipolyse enzymatique des corps gras à Marseille. OCL 2008; 15 (3): 196-207.
- 13:** Roussel A, Miled N, Berti-Dupuis L, Rivière M, Spinelli S, Berna P, Gruber V, Verger R, Cambillau C. Crystal structure of the open form of dog gastric lipase in complex with a phosphonate inhibitor. *Journal of Biological Chemistry*, 2002; 277(3): 2266-2274.
- 14:** MolegroVirtual Docker, User manual 2011, 273.
- 15:** Eduardo B. thèse doctorat, simulation moléculaire appliqué à l'acylation de flavonoïde catalysé par lipase, l'institut polytechnique Lorrain, 2009, 22.
- 16:** Venkatachalam CM, Jiang X, Oldfield T, Waldman M. Ligand fit: a novel method for the shape-directed rapid docking of ligands to protein active sites. *Journal of Molecular Graphics and Modelling*, 2003; 21(4): 289-307.
- 17:** Férey N, Bouyer G, Martin C, Drif A, Bourdot P, Ammi M, Nelson J, Burkhart JM, Autin L. Docking De Protéines En Réalité Virtuelle : Une Approche Hybride Et Multimodale. 2^e Soumission A Rsti 2008: 10.
- 18:** Aurélien G, Conception d'un logiciel de docking et applications dans la recherche de nouvelles molécules actives. Thèse de Doctorat, Université Joseph Fourier, Faculté de pharmacie de Grenoble 2007: 20.
- 19:** Thomsen R, Christensen M-H. MolDock: a new technique for high-accuracy molecular docking. *Journal medicinal chemistry*, 2006; 49: 3315-332.
- 20:** Michalewicz Z, Foguel David B, How to Solve It: Modern Heuristics. Springer-Verlag, Berlin 2000.
- 21:** Venkatachalam CM, Jiang X, Oldfield T, Waldman M. LigandFit: a novel method for the shape-directed rapid docking of ligands to protein active sites. *Journal of Molecular Graphics and Modelling*, 2003; 21: 289-307.

22: Storn R, Price K, Differential Evolution : A Simple And Efficient Adaptive Scheme for Global Optimization over Continuous Spaces. Tech-report, International Computer Science Institute, Berkley 1995.

23: Nelder JA, Mead R. A Simplex-Method for Function Minimization. *Computer journal*, 1965; 7: 308-313.

24: Gehlhaar DK, Verkhivker G, Rejto PA, Fogel DB, Fogel LJ, Freer ST. Docking Conformationally Flexible Small Molecules Into a Protein Binding Site Through Evolutionary Programming. Proceedings of the Fourth International Conference on Evolutionary Programming 1995: 615-627.

25: Gehlhaar DK, Bouzida D, Rejto PA. Fully Automated And Rapid Flexible Docking of Inhibitors Covalently Bound to Serine Proteases. Proceedings of the Seventh International Conference on Evolutionary Programming 1998: 449-461

26: Yang JM, Chen CC. Gemdock: A generic evolutionary method for molecular docking. *Proteins: Structure, Function and Bioinformatics*, 2004; 55: 288-304.

27: Lakhel S. Etude docking et synthèse de dérivés de xanthone : voie d'accès à de nouveaux inhibiteurs de l' α -glucosidase, thèse doctorat université Annaba, 2016, 67.

28: Safi R. Etude de la férutinine et de ses analogues: hémisynthèse et activité anticancéreuse vis-à-vis des lignées cellulaires hormono-dépendantes. Thèse doctorat Université Toulouse 3 Paul Sabatier 2015.

Chapitre IV

Résultats et discussion

Les enzymes sont des macromolécules sélectives qui se caractérisent par un énorme pouvoir catalytique, pouvant accélérer de façon spécifique les réactions chimiques de la cellule à des taux de plus de 10^{16} fois ceux des niveaux non catalysés sans être elles-mêmes modifiées par la réaction. Elles sont les protéines responsables des transformations biochimiques au sein des cellules des organismes vivants et au centre de l'organisation du métabolisme et de la régulation des processus physiologiques.¹

La spécificité des enzymes à catalyser des réactions chimiques précises est à la base de cette organisation, faisant de ces molécules biologiques des catalyseurs de choix pour différentes applications biotechnologiques modernes. Les enzymes se divisent en six classes distinctes sur la base du type de réaction catalysée selon l'« Enzyme Commission ». Les hydrolases forment la troisième classe (EC3).²

Ces enzymes catalysent l'hydrolyse ou la synthèse de liaisons C-O, C-N ou C-C dans différents types de substrats. La famille des hydrolases englobent plusieurs types d'enzymes qui sont conçues pour la catalyse de divers réactions. Parmi ces biocatalyseurs on cite les lipases. Ces dernières sont des hydrolases qui catalysent l'hydrolyse ou la synthèse de liaisons esters (C-O) dans les huiles et autres esters aliphatiques, elles sont impliquées dans le métabolisme des lipides. Chez les mammifères par exemple, les enzymes lipolytiques sont notamment impliquées dans la digestion, l'absorption et l'utilisation des lipides et des vitamines, dans la transformation métabolique de médicaments naturels et artificiels ainsi qu'au niveau de la détoxification hépatique et de l'utilisation de neurotransmetteurs.³

Chez les microorganismes, elles jouent un rôle dans l'utilisation de sources de carbone et dans la détoxification. Cependant le rôle métabolique que jouent ces enzymes chez les microorganismes reste encore obscur.⁴

Une hypothèse a été soulevée, indiquant que la lipoprotéine lipase (LPL) pourrait être à l'origine d'un signal de faim, notamment chez des rongeurs obèses. Des études ont montré que l'activité de la LPL au cours des obésités humaines augmente de façon significative.⁵

La majeure partie de la digestion des graisses s'effectue dans l'intestin grêle. Les triglycérides alimentaires ne sont pas absorbables en tant que tels. La digestion des lipides requiert l'intervention des acides biliaires et de leurs sels pour émulsionner les gouttelettes de lipides qui seront ensuite hydrolysées principalement par le suc pancréatique et à moindre mesure par les lipases présentes dans le suc gastrique.⁶

L'obésité a aujourd'hui atteint le stade de pandémie. Pour faire face à ce fléau qui porte atteinte à la santé publique, les différents acteurs du système de santé se mobilisent. Ainsi, les efforts récents de la communauté scientifique ont permis de mieux appréhender l'obésité dans son ensemble: ses origines multiples, ses mécanismes complexes, ses différentes complications et ses traitements potentiels. Parmi les récentes découvertes figure l'identification de nouvelles molécules impliquées dans cette morbidité. Ceci s'inscrit dans un cadre plus large d'études susceptibles d'ouvrir la voie à de prochaines innovations tant sur le plan préventif que curatif.

C'est ainsi que nous avons abordé cette problématique par une approche théorique qui consiste à étudier l'inhibition des enzymes impliquées dans la maladie de l'obésité, avec un inhibiteur naturel. Pour cela nous avons fait appel à l'Epigallocatechin-3-gallate (EGCG). Ce composé, qui est un polyphénol dérivé des flavonoïdes que l'on trouve en abondance principalement dans le thé vert. La caractérisation physico-chimique, ainsi que celle thérapeutique de cet 'inhibiteur, ont été détaillé dans le chapitre 1.

L'objectif de cette étude est de minimiser la formation des complexes et par la suite retarder leur progression. Afin de rationaliser les propriétés de cet 'inhibiteur et de déterminer le processus réactionnel impliquant ce composé, nous avons essayée d'optimiser et de modéliser les complexes formés par des méthodes théorique utilisant l'outil informatique. La technique de modélisation adoptée dans cette étude est le docking moléculaire.

IV.1. Description de l'EGCG

l'Epigallocatechin-3-gallate est une molécule hautement fonctionnalisée possédant dans sa structure des cycles aromatiques et des fonctions hydroxyle et carbonyle (fig.1). Grâce à ces éléments structuraux, cette molécule peut établir plusieurs types d'interaction avec les enzymes lipolytiques. En particulier celles entre l'EGCG et la lipase pancréatique. Pour expliquer sur un plan moléculaire, l'anti-obésité de cet inhibiteur, la compréhension de cette activité est importante. Les éléments clés de cette étude, sont l'identification de la cavité de fixation et la connaissance des interactions susceptibles de se former entre les acides aminés de ces lipases et l'inhibiteur impliqué.

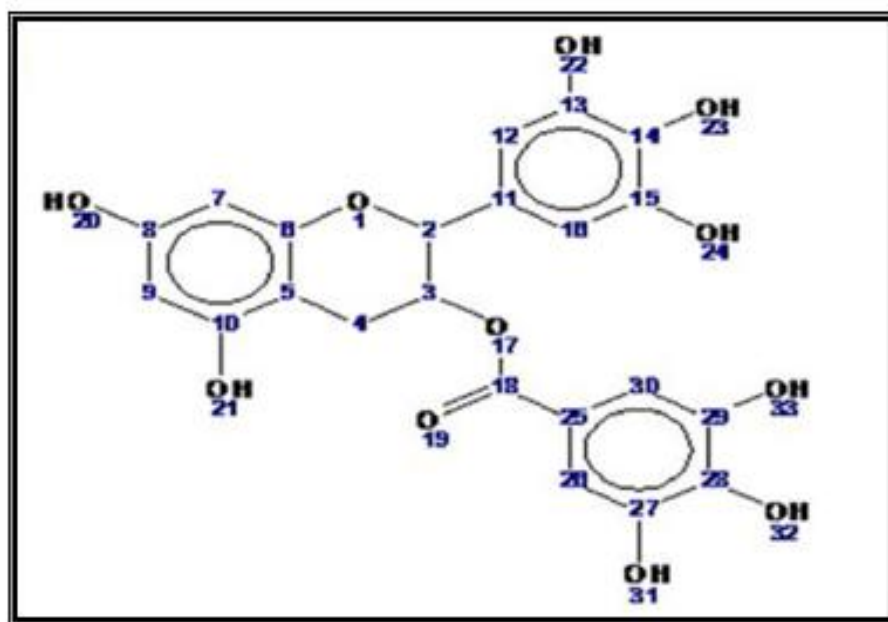


Figure 1. Structure chimique (2D) de l'Epigallocatechin-3-gallate.

Une étude cinétique a été également réalisée, suggérant l'inhibition non compétitive de ce composé. Les paramètres thermodynamique indiquent que l'association entre la lipase de pancréas et l'EGCG est gouvernée par des interactions de type spontanée, tel que la liaison hydrogène, interaction hydrophobe. Le résultat de dichroïsme circulaire suggère un changement dans la conformation protéique lors de la fixation d'EGCG.⁷

Dans ce chapitre, nous présenterons les résultats obtenus et leur discussion. Il s'agit de l'étude des interactions entre l'Epigallocatechin-3-gallate et trois lipases digestives par la méthode du docking moléculaire. La démarche effectuée dans cette étude théorique sera basée sur les énergies d'interactions, la fonction de score et les différents types d'interactions présentes entre certains acides aminés de la protéine étudiée et celle de l'inhibiteur.

IV.2. Fiabilité du programme MVD

La validation de la performance du logiciel choisi pour développer une étude théorique est une étape importante. Cette étude préliminaire est réalisée afin de vérifier les performances du logiciel choisi avec les systèmes étudiés. La méthode la plus simple pour évaluer la fiabilité du programme utilisé dans une procédure de docking est de déterminer si la pose la mieux classée à l'issue du docking est structuralement proche de la conformation du ligand natif indiquée par les données cristallographiques.

Ce test de similarité entre les deux ligands (testé et natif), montre selon la littérature, que la prédiction est dite réussie si la valeur de l'écart quadratique moyen (RMSD) de la meilleure conformation est inférieure à 2 Å par rapport aux données expérimentales. La précision de MVD a été évaluée par le docking de deux inhibiteur phosphonate : ($C_{12}H_{27}O_3P$) et ($C_{15}H_{33}O_2P$) dans le site actif de la LPH et LGC, respectivement. Les structures cristallographiques utilisées dans le cadre de cette validation est celle correspondant respectivement aux codes PDB: 1LPB et 1K8Q. Le complexe 1ETH a été exclus de ce teste à cause de la non existence d'un inhibiteur au sein du son site actif. Le substrat co-cristallisé existant se présente sous forme d'une molécule de détergent: éther de mono-octyl tétraéthylène glycol (TGME).

L'évaluation de la performance du programme Molegro virtuel Docker (MVD) s'effectue par les deux tests suivants :

- L'écart quadratique moyen ou le RMSD (root mean square deviation)
- La superposition des ligands.

IV.2.1. Ecart quadratique moyen (RMSD)

La valeur de la RMSD du modèle conçu par le logiciel vis-à-vis de la structure du cristal ne dépasse pas 2Å dans le complexe 1K8Q. Certains paramètres ne peuvent pas être pris en compte par notre algorithme, comme le nombre élevé des liaisons rotatifs (liaisons simples ou liaisons sigma pouvant tourner librement) (fig.2) et la présence d'une liaison covalente entre le ligand et le récepteur, grâce aux conditions initiales défavorables.

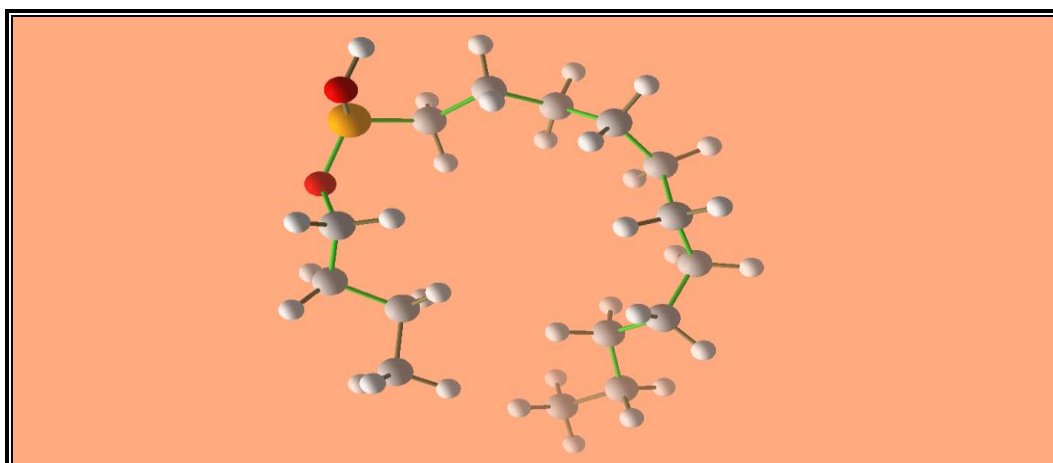


Figure 2. Structure 3D du l'inhibiteur phosphonate ($C_{15}H_{33}O_2P$). Les liaisons rotatives sont représentées en vert.

Dans le cas du complexe 1LPB (fig.3), on constate que la valeur de RMSD correspondant à la structure cristalline de 2,27 Å, est légèrement supérieure à celle du seuil (2Å). Cette différence négligeable est attribuée à la flexibilité de ligand (11 liaison rotatives), due son faible poids moléculaire, et aux échecs cités précédemment.

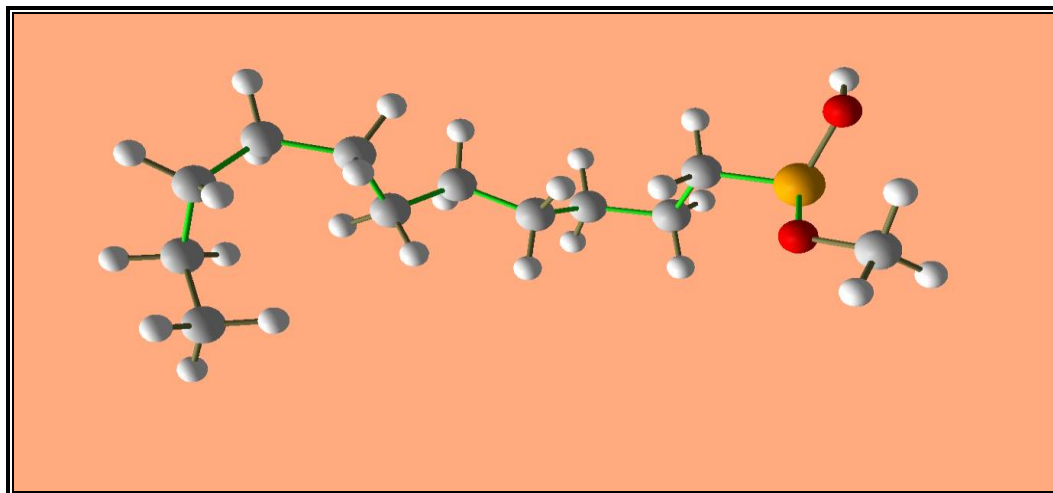


Figure 3. Structure 3D du l'inhibiteur phosphonate (C₁₂H₂₇O₃P). Les liaisons rotatives sont représentées en vert.

IV.2.2. Superposition des ligands

Les résultats obtenus (fig.4), montrent que la conformation obtenue pour l'inhibiteur phosphonate (C₁₅H₃₃O₂P) se superpose parfaitement avec celle présente dans la structure cristallographique. La valeur de RMSD des atomes lourds est de 1,43 Å, et ceci quand on prend l'inhibiteur dans la structure 1K8Q comme référence. Ces résultats mettent en évidence la bonne précision et les performances de MVD pour les simulations de docking. La superposition des géométries de l'inhibiteur phosphonate (C₁₅H₃₃O₂P), donnée par rayons X (coloré par type d'atomes) et par docking avec MVD (coloré en vert), est représentée dans la figure 4.

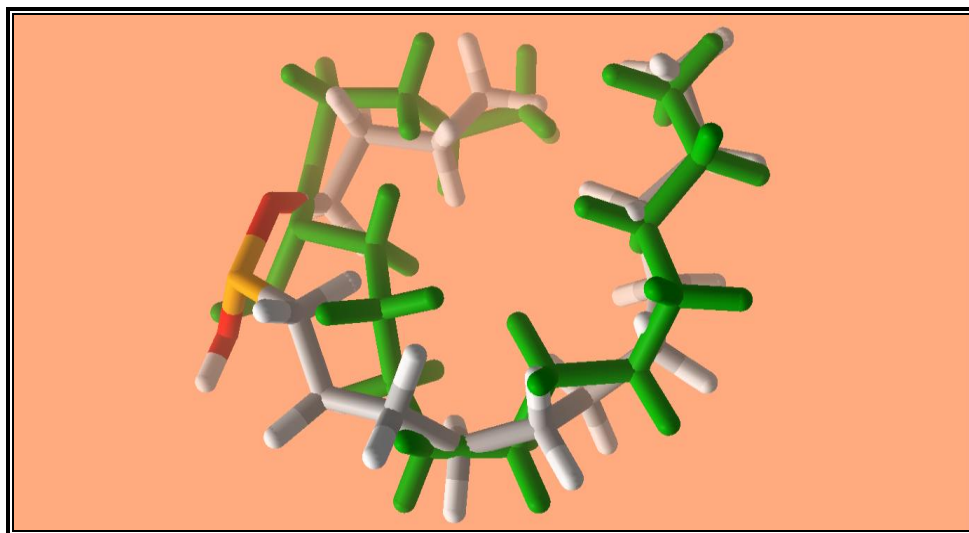


Figure 4. Superposition des géométries d'inhibiteur phosphonate ($C_{15}H_{33}O_2P$).

Le RMSD entre le mode d'interaction prédit et la structure cristalline est égale à 2,27Å, indiquant l'existence d'une petite différence conformationnelle. Les résultats obtenus (fig.5) et qui mettent en évidence la superposition des de deux conformations (réelle et prédite) montrent que la déviation des atomes lourds a commencée au niveau de groupement phosphonate. Celle-ci établit une liaison covalente avec la serine du récepteur, qui ne peut pas être prise en compte par notre algorithme. Les autres échecs peuvent s'expliquer par la présence de contacts cristallins entre le ligand et un complexe voisin de la maille cristalline.

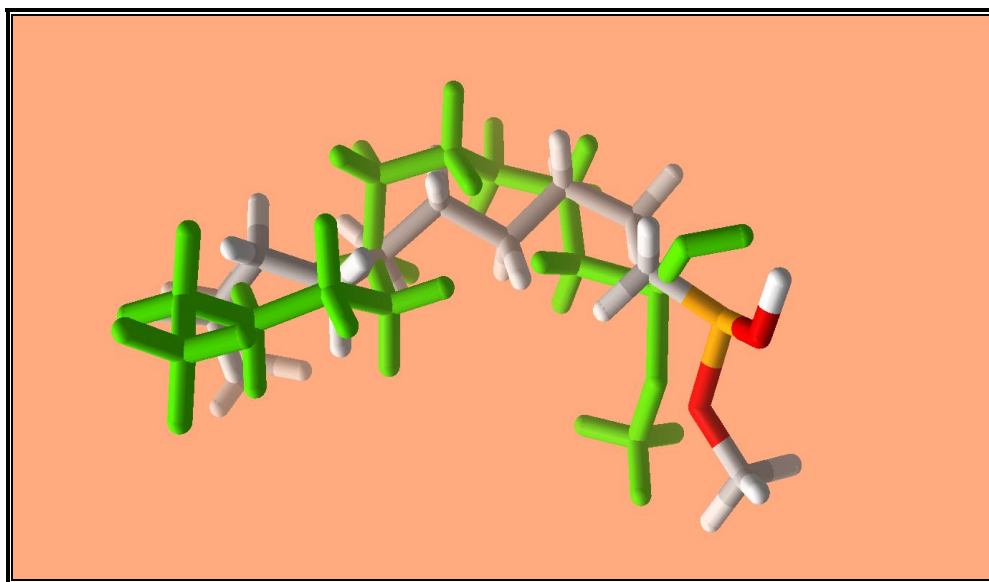


Figure 5. Comparaison de la conformation expérimentale de l'inhibiteur phosphonate ($C_{12}H_{27}O_3P$). (colorée par type d'atome) et sa conformation optimale simulée par MVD (colorée en vert).

Les poses montrent une superposition acceptable, ce qui indique que les poses calculées par MVD sont extrêmement reproductibles. Cette notion est essentielle car elle met en évidence la grande convergence de résultats obtenue avec MVD, ainsi que le calcul du RMSD, qui est le plus souvent utilisé comme indicateur pour la reproduction de poses. Cette fonction accentue les défauts du docking, sachant qu'un ligand de faible poids moléculaire, même placé de façon aléatoire, peut présenter un RMSD plus faible en comparaison avec une molécule possédant un poids moléculaire important.

Une valeur de RMSD, n'est qu'un élément de conformité lorsque la molécule calculée est la même que la molécule de référence. Dans le cas contraire, un autre test de fiabilité mis en jeu. Ce sont les indicateurs basés sur des critères bio-structuraux qui seront de plus en plus utilisés dans la comparaison des structures moléculaires qui divergent. Dans ce dernier cas, la conformité sera plus significative lorsque nous aurons des données structurales cohérentes.

En conclusion, l'étude préliminaire nous a permis de valider la performance du logiciel de docking MVD, montrant ainsi la compétence de ce programme vis-à-vis de l'échantillonnage de l'espace de recherche. A partir de ce screening élémentaire sur la fiabilité du logiciel MVD, adopté pour la modélisation et le docking choisis dans cette étude, nous avons utilisé ce programme pour générer des modes d'interactions fiables et ceci en travaillons avec des ligands moins flexibles.

Dans notre étude le ligand (EGCG), choisi pour le docking, est caractérisé par un poids moléculaire moyen (458,32 g/mol) et un nombre restreint des liaisons flexibles ($n_{\text{RotB}}=3$) (fig.6).

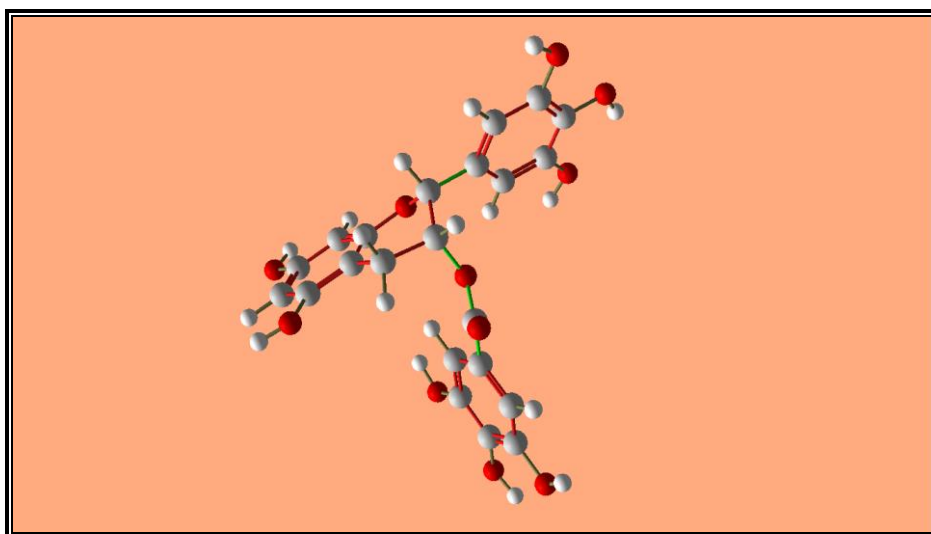


Figure 6. Structure 3D du l'égigallocatechin-3-gallate (EGCG). Les liaisons rotatives sont représentées en vert.

L'étape de screening réalisée précédemment, a pu évaluer les conditions qui nous ont permis de mettre au point notre protocole de docking moléculaire adopté dans ce travail. A partir de ces résultats nous exposerons dans ce qui suit les étapes de notre étude concernant la modélisation de l'effet inhibiteur de l'EGCG sur les lipases digestives.

IV.3. Docking de l'épigallocatechin-3-gallate (EGCG)

IV.3.1. Cas de la lipase pancréatique de porc (LPP)

La fixation de l'EGCG sur un site autre que le site actif de l'enzyme, fait de ce composé un inhibiteur non compétitif très puissant de la lipase pancréatique⁸. Dans ce cas l'inhibiteur se lie à l'enzyme à un site autre que le site actif, causant ainsi un changement réversible dans la structure tertiaire de l'enzyme qui interfère la vitesse de conversion du substrat en produit.

L'inhibiteur peut interagir avec l'enzyme libre ou avec le complexe [ES]. Dans ce type d'inhibition, la quantité d'enzyme active semble diminuer lorsque [I] augmente; V_{max} de la réaction diminue (schéma 1).

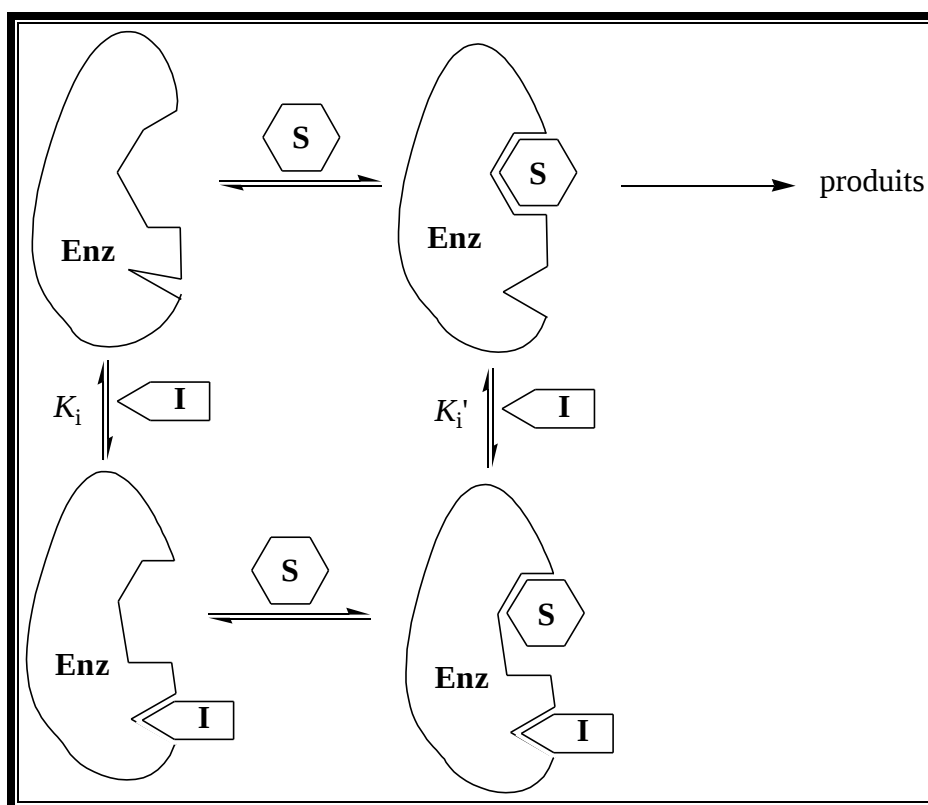


Schéma 1. Mode d'action d'un inhibiteur non compétitif.⁹

A partir de cette hypothèse, nous avons entamé l'étude du criblage virtuel dans toutes les cavités probables afin de localiser le site de fixation de l'inhibiteur, en employant la méthode de docking moléculaire. Plusieurs cavités sont détectées à cause de la topologie de surface des enzymes qui est assez irrégulière. Parmi les cavités détectées, nous avons identifié les sites d'interaction actifs. Pour cela nous nous sommes basé sur notre connaissance de la nature de l'enzyme étudiée. L'identification de ces sites va nous permettre par la suite de faire le docking du ligand dans les étapes suivantes.

Compte-tenu de la taille de l'inhibiteur, les cavités au volume petit sont exclues d'amarrage. Les cavités qui se trouvent proche l'une de l'autre, ont été fusionnées, et ceci en se basant sur les données cristallographiques.

Pour effectuer le docking de l'EGCG, nous avons utilisé le complexe 1ETH qui est déjà préparé pour l'amarrage. La figure 7 montre un zoom sur le site actif de la lipase pancréatique du porc (LPP) où sont représentés les résidus principales ; Ser152, His264, et Asp177.

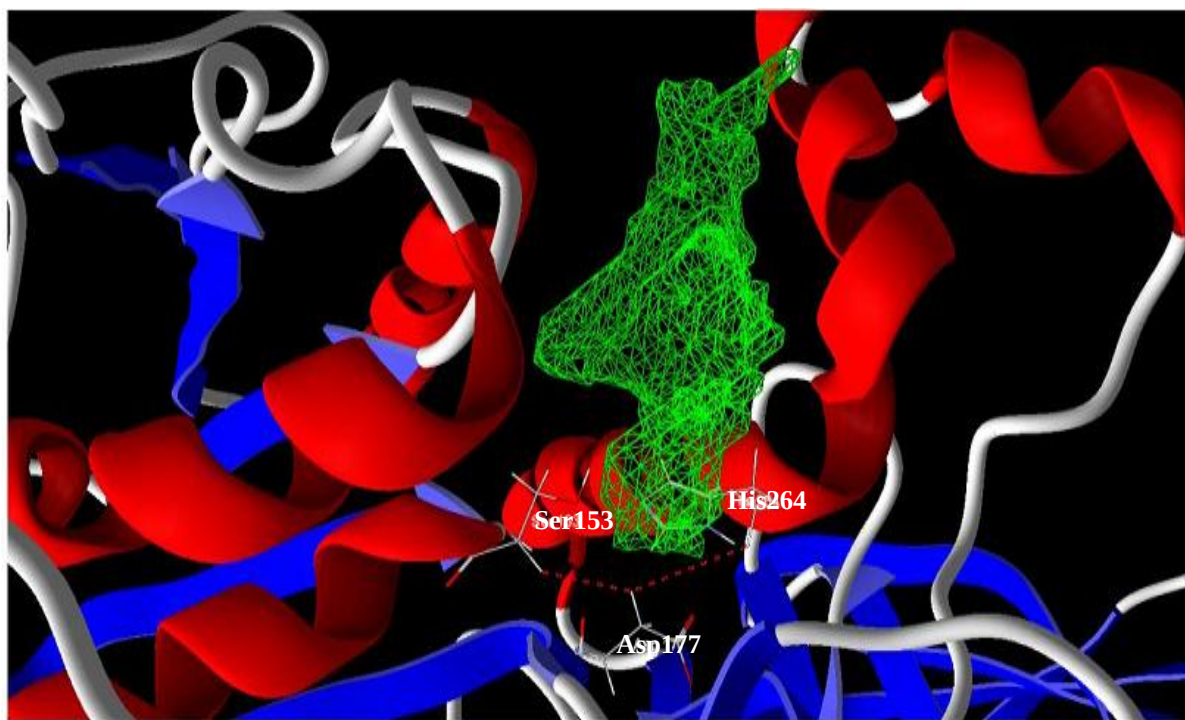


Figure 7. Illustration du site catalytique de la LPP avec les résidus principaux (Ser153, His264, et Asp177).

L'amarrage débute par l'importation de la cible enzymatique ainsi que le ligand, dans l'espace de recherche du MVD. Cette procédure vient après une étape de préparation préliminaire, qui permet de vérifier et de s'assurer que les propriétés structurales de chaque espèce sont prises en compte par le logiciel.

Au cours de l'amarrage moléculaire l'inhibiteur sera orienté vers la cavité qui est déjà localisée, à partir du protocole MDV décrit ci-dessus. Le «scoring» des poses est obtenue après 30 tours de recherche. Ce résultat permet de classer les meilleures poses en se basant sur leurs énergies d'interactions obtenues par le «Moldock score». Après que le logiciel prédit les poses en utilisant la fonction Moldock score plusieurs termes d'énergie supplémentaires peuvent être calculés. Tous ces termes sont par la suite stockés pour une étude ultérieure.

Chacun des complexes optimisés a été soumis à une évaluation en utilisant la fonction de score ReRankScore présente dans le logiciel MVD. Cette fonction de re-évaluation, est une combinaison linéaire des termes énergétiques stockés avec d'autres représentants l'interaction de Lennard-Jones et l'énergie de l'approche stérique. La plupart des poses sélectionnées correspondent à des poses fortes. On note que les scores MolDock et Rerank sont les meilleurs scores pour la même pose pour lesquels les calculs semblent stables (tableau 5).

Tableau 5. Energies d'interaction (kcal mol^{-1}) des meilleures poses générées par MolDock Score, Rerank Score et H-bond Score, obtenues pour chaque cavité.

Cavité	Moldock score	Rerank Score	Hbond Score
2	-128,64	-91,28	-17,27
3	-139,69	-115,27	-13,43
4	-112,04	-87,66	-24,11
7	-128,69	-98,47	-12,24

La plus basse énergie d'interaction entre l'EGCG et la LPP qui s'évalue à $-112,04 \text{ kcal mol}^{-1}$ a été observée au sein de la quatrième cavité. Cette cavité est éloignée de la cavité principale d'environ $23,14 \text{ \AA}$, donnant ainsi beaucoup d'espace. Par conséquent l'influence de l'interaction spontanée avec sa faible énergie n'atteint pas la cavité principale jusqu'elle se déforme. Bien que les cavités 2 et 7 soit plus proche de la poche principale à une distance de $8,47 \text{ \AA}$, mais le résultat de l'amarrage obtenu pour ce cas n'a pas donné de valeurs d'énergies significatives. Ce résultat négatif est dû probablement à l'inadéquation de la forme et du volume de cette poche avec la taille de l'inhibiteur (EGCG).

Les orientations des meilleures poses de docking dans les cavités : 2, 3, 4, 7 sont illustrés par la figure 8. Les grilles vertes représentent les cavités détectées.

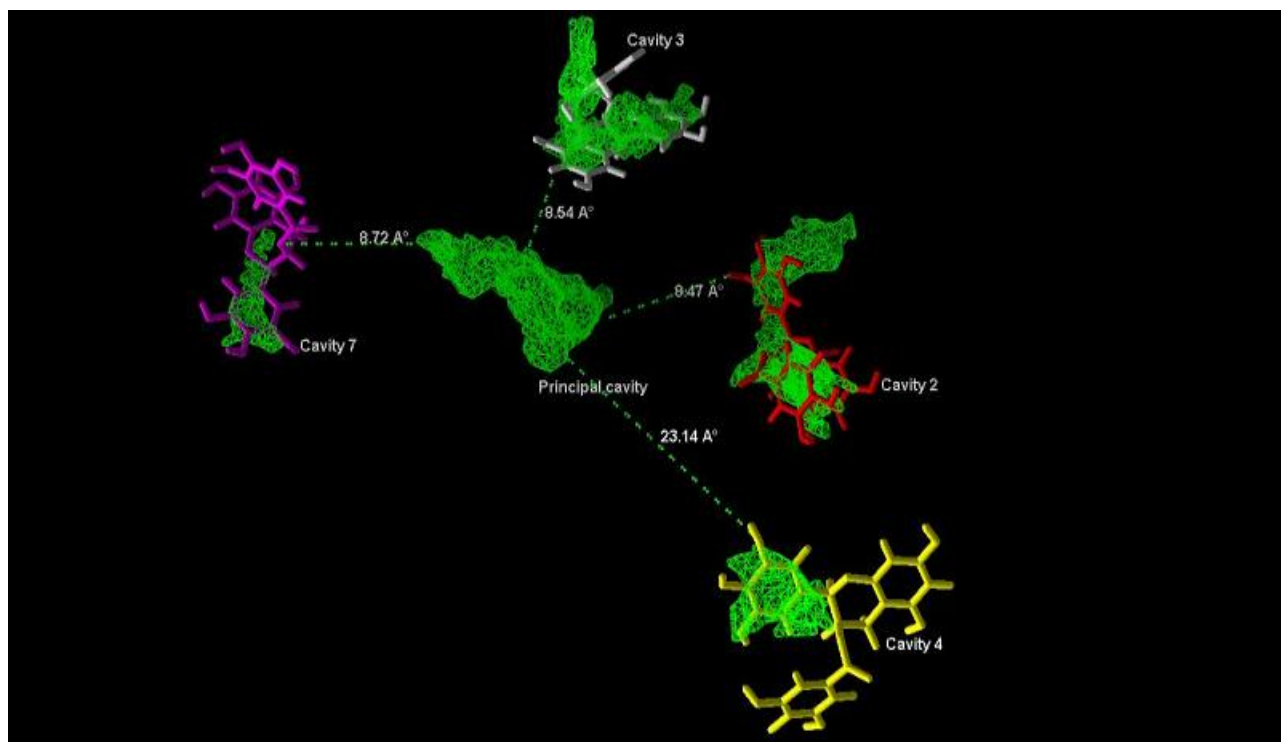


Figure 8. Orientation des meilleures poses de docking dans les cavités : 2, 3, 4, 7.

L'analyse visuelle de la meilleure pose montre que la meilleure énergie est obtenue quand l'EGCG est docké dans la cavité 3 de la LPP, avec une affinité de $-139,69 \text{ kcal mol}^{-1}$ et $-115,21 \text{ kcal mol}^{-1}$ basé sur les deux fonctions de score : MolDock Score et Rerank, respectivement. Il convient de noter que cette cavité est située au voisinage du site actif à une distance de $8,54 \text{ \AA}$. Celle-ci entourée de plusieurs Helices α , de sorte que la fixation de l'inhibiteur dans cette poche diminue significativement sa valeur énergétique.

L'analyse des résultats des simulations de docking, montrent que de point de vue énergétique le gallate épigallocatechine est plus stable quand il est au sein de la cavité 3. Pour vérifier cette hypothèse, nous avons complété cette étude par des simulations de docking en utilisant d'autre fonction de score ainsi que d'autre algorithme. Ce type d'approche croisée est nécessaire car il s'agit d'un critère qui est lié à l'efficacité. Pour estimer l'affinité et l'orientation du ligand dans les cavités sélectionnées, nous avons effectué un calcul MVD en utilisant les fonctions de score : MolDock score, Plant Score et les algorithmes suivants : SE, Simplex, Optimiseur. Les résultats de l'ancrage sont résumés dans le tableau 6.

Tableau 6. Les énergies d'interaction (kcal.mol⁻¹) des meilleures poses obtenues pour chaque fonction du score et algorithme.

Fonction de score	Algorithme	Affinité énergétique			
		Cavité	Moldock-Score	Reranck-Score	HBond-Score
Moldock score Grid	SE	2	-128,75	-90,90	-17,60
		3	-139,15	-115,01	-17,28
		4	-104,38	-85,61	-17,45
		7	-128,35	-94,21	-10,72
Moldock Score Grid	Simplex	Cavité	Moldock Score	Reranck Score	H-Bond Score
		2	-128,31	-91,29	-17,69
		3	-139,55	-114,92	-13,40
		4	-111,86	-87,49	-24,23
		7	-130,27	-98,85	-12,34
PlantScore Grid	Optimizeur	Cavité	Plant-Score		
		1	-65,79		
		3	-72,23		
		4	-67,45		
		7	-65,39		
		Cavité	Plant-Score		
		2	-55,81		
		3	-74,82		
		4	-65,68		
		7	-67,14		

A partir des résultats énergétiques obtenus (tableau 6), en utilisant d'autres fonctions de score et d'autres algorithmes, nous avons pu identifier aussi la cavité 3 comme site préférentiel de l'EGCG (fig.9). Ce résultat est également corrélé avec celui de dichroïsme circulaire (CD) observés par Wang *et al*⁷. Ces derniers ont constaté au cours de leur étude, une diminution très forte de l'hélice α trouvée autour du site catalytique lors de la formation du complexe EGCG-LPP. Il semble que la fixation d'épigalocatéchine gallate dans la cavité 3 engendre des interactions physiques responsables d'une déformation conformationnelle au niveau de la cavité principale de la LPP.

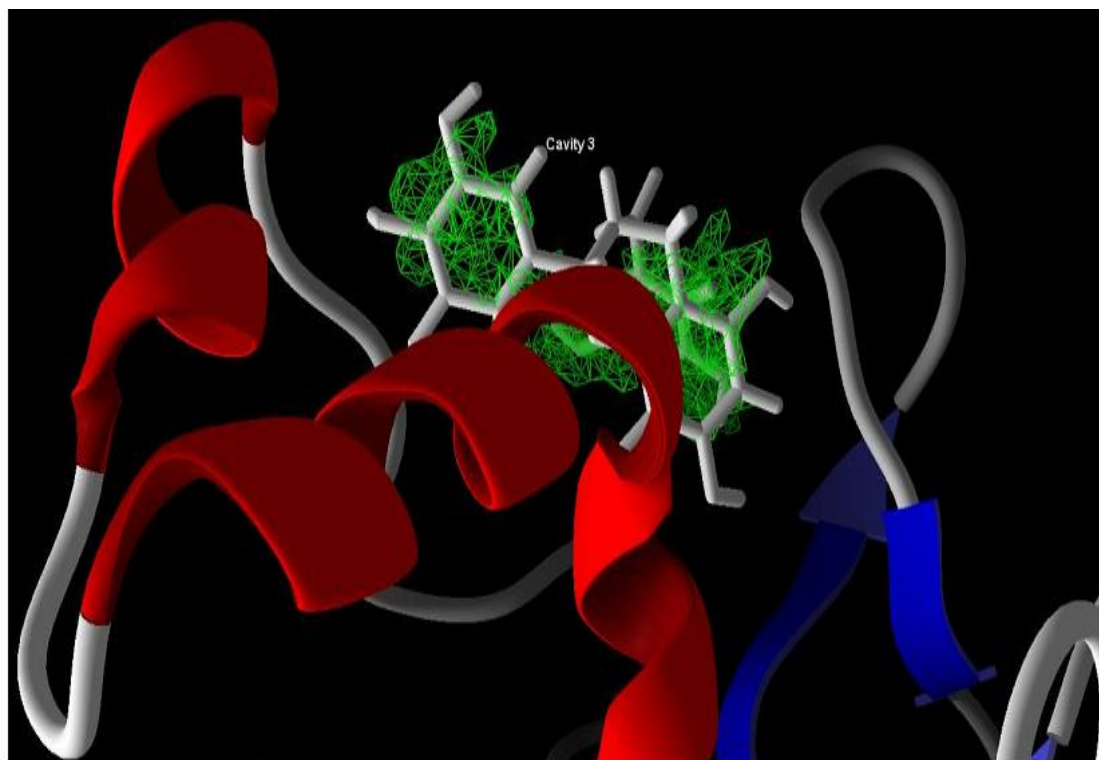


Figure 9. Positionnement du ligand EGCG (coloré en blanc) au sein de la cavité N°3 (vert).

IV.3.1.1. Etude des interactions entre EGCG et LPP

L'Epigallocatechine se caractérise par une forte activité inhibitrice via la lipase pancréatique de porc ($IC_{50} = 0,8 \mu\text{mol}$)¹⁰. Le programme Molegro virtuel Docker confirme parfaitement ce résultat en fournissant une énergie d'interaction égale à -139,69 Kcal/Mole. Cette énergie résulte de la mise en place de cinq liaisons hydrogène (fig.10) et plusieurs interaction hydrophobique. Les longueurs et les orientations spatiales des différentes liaisons hydrogène obtenues sont résumées dans le tableau 7:

Tableau 7. Longueurs et orientations spatiales des différentes liaisons hydrogène obtenues.

N° de liaisons H	Résidus impliqué	Groupe du ligand	Distance Å°
1	His264	OH(20)	2,59
2	Leu214	OH(21)	2,56
3	Gln 245	OH(22)	2,78
4	Asp 206	OH(33)	2,58
5	Asp 206	OH(20)	3,09

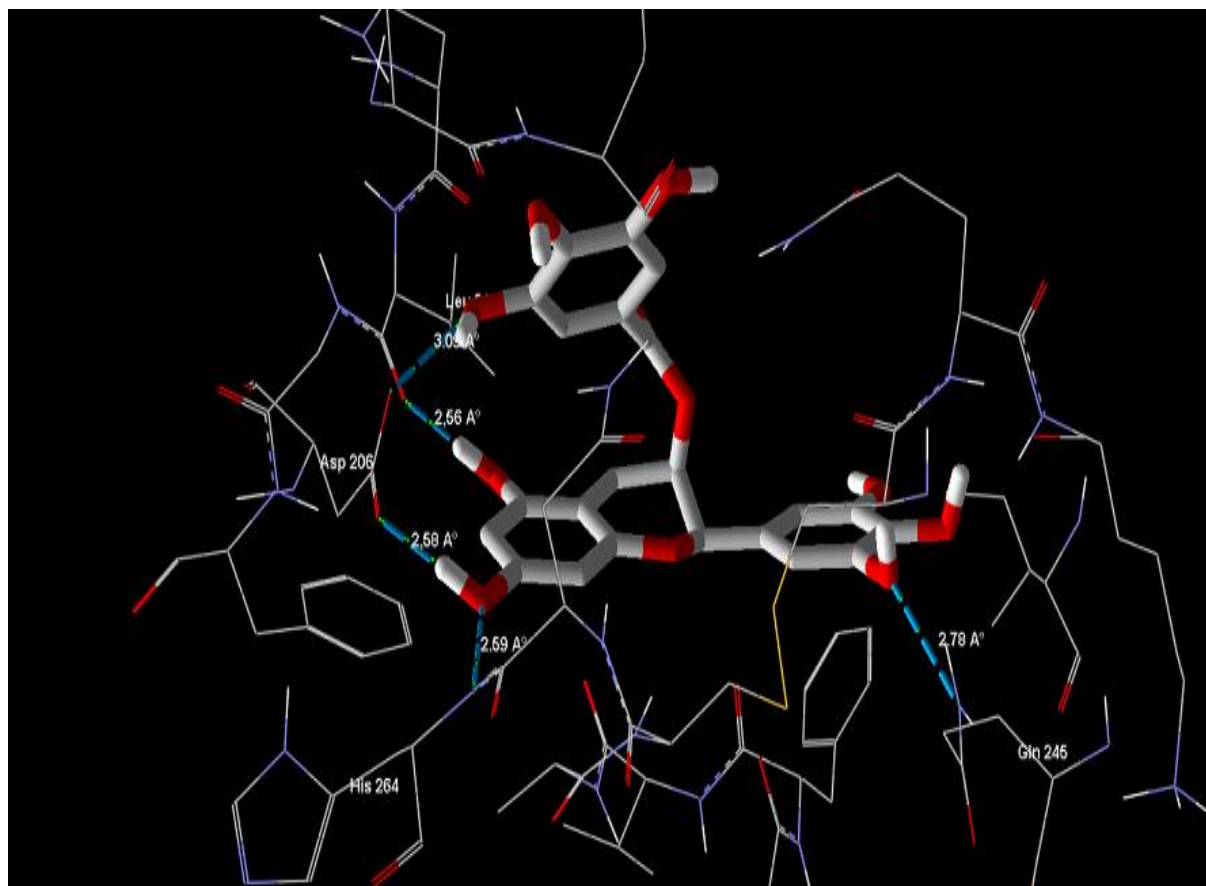


Figure 10. Interactions de type liaison hydrogène entre EGCG et les résidus du site d'interaction. Les liaisons hydrogène présentées par un trait bleu discontinue.

La plus importante liaison hydrogène est observée entre la fonction hydroxyle de EGCG en position 21 et la fonction amine du résidu His264. Cette liaison est présentée par l'un des acides aminés de la triade catalytique. Cette liaison induit des changements conformationnels au niveau spatial des résidus catalytiques, notamment His264 dont l'orientation change. Ce résultat conforte l'idée que la fixation de l'inhibiteur sur ce résidu est bien responsable de l'inactivation de l'enzyme. Le complexe formé est stabilisé au sein de la cavité 3 de la LPP par des interactions hydrophobiques par les résidus suivants : Val 260, Leu 214 et Ile 242.

Le composé EGCG établit également des interactions hydrophobiques de type π - π entre le ligand et les cycles benzéniques des acides aminés Phe 216, Phe 259 (fig.11).

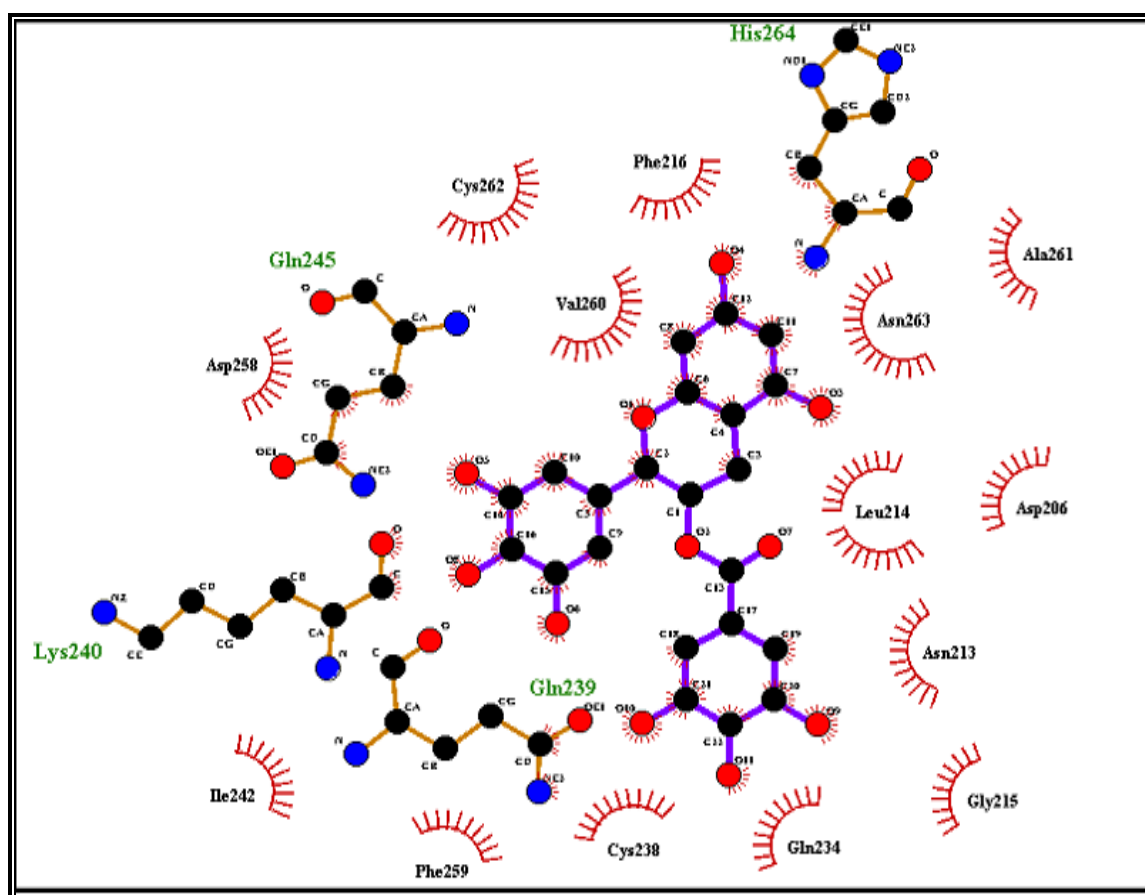


Figure 11. Interactions hydrophobes entre EGCG et PPL, produites par le programme LIGPLOT +. Les pointes rouges sur les arcs représentent les interactions hydrophobes avec les résidus (en noir).

En conclusion, l'étude structurale a permis de définir le groupement galloyle présent dans l'EGCG comme étant important. Ce groupement se trouve dans une orientation spécifique qui favorise une certaine hydrophobicité.

IV.3.2. Cas de la lipase pancréatique humaine (LPH)

Un docking moléculaire de l'Epigallocatechine gallate (EGCG) dans La lipase pancréatique humaine (LPH), présente à son tour une similarité de séquences avec la LPP. Le complexe utilisé pour l'amarrage est 1LPB. Dans ce cas aussi on adopte la même procédure préliminaire de préparation de la cible comme il a été défini dans le cas de LPP. Après une préparation automatique de la cible enzymatique par le logiciel MVD, Trois cavité importante ont été détecté pour la LPH. La cavité n°1 présente la poche principale par rapport aux deux autres cavités (2,3). Pour cela un processus de docking a été lancé pour l'inhibiteur (EGCG) en utilisant l'algorithme de recherche Moldock Optimizer, et la fonction de score Moldock Score [Grid].

Notre choix de cet algorithme est justifié par sa meilleure performance énergétique, ainsi que sa rapidité pour étudier et prédire le mode d'interaction ligand-récepteur en comparaison avec d'autres algorithmes. La meilleure énergie d'interaction observée avec la cible enzymatique est obtenue quand l'EGCG est orienté vers la troisième poche (tableau 8).

Tableau 8. Les énergies d'interaction Moldock Score, Rerank Score, et H-Bond Score des meilleures poses obtenues pour chaque cavité.

Cavité	Moldock score (kcal/mol)	Rerank Score (kcal/mol)	H-bond Score (kcal/mol)
2	-122,20	-94,17	-13,88
3	-157,25	-128,24	-21,56

IV.3.2.1. Etude des interactions EGCG-LPH

En complément de l'examen des résultats obtenus ; une analyse visuelle des interactions de EGCG avec la LPH a été réalisée. Les résultats obtenus mettent en évidence neuf liaisons hydrogène entre l'Epigallocatechine et les résidus du site de fixation (fig.12), ce qui explique la valeur énergétique significative des liaisons H, obtenue pour ce composé et qui égale à - 21,56 kcal mol⁻¹. Ces liaisons sont caractérisées ci-dessous :

- le groupe hydroxyle en position 21 établit deux liaisons hydrogènes. La première avec le résidu Asp205 (O32—N: d=2,61 Å), et l'autre avec le phe215 (O32—O: d=3,29 Å).
- une autre liaison hydrogène est formée entre l'oxygène de cycle oxane du ligand et le résidu Lys 238 (O1—N : d=3,20).
- Le groupe carbonyle de résidu Lys239 forme une liaison hydrogène de longueur 3,20Å avec l'oxygène en position 23 du ligand (O23—O).
- Une autre liaison hydrogène est formée entre le Asn212 et le ligand (O20—O: d=3,03 Å) Deux liaisons hydrogènes sont observées entre le carbonyle du groupe galloyl du ligand et les résidus Cys261 et Phe285 avec une longueur 2,84 Å et 3,01 Å, respectivement.

- le groupe hydroxyle en position (31) du ligand forme une bonde hydrogène avec la Leu 213 (O31—O: 2,71 Å).
- Finalement, une importante liaison hydrogène détecté entre l'épigalocatécine-3-gallate et le résidu His 263 (O33—N: d=2.59 Å), qui présente une partie de triade catalytique.

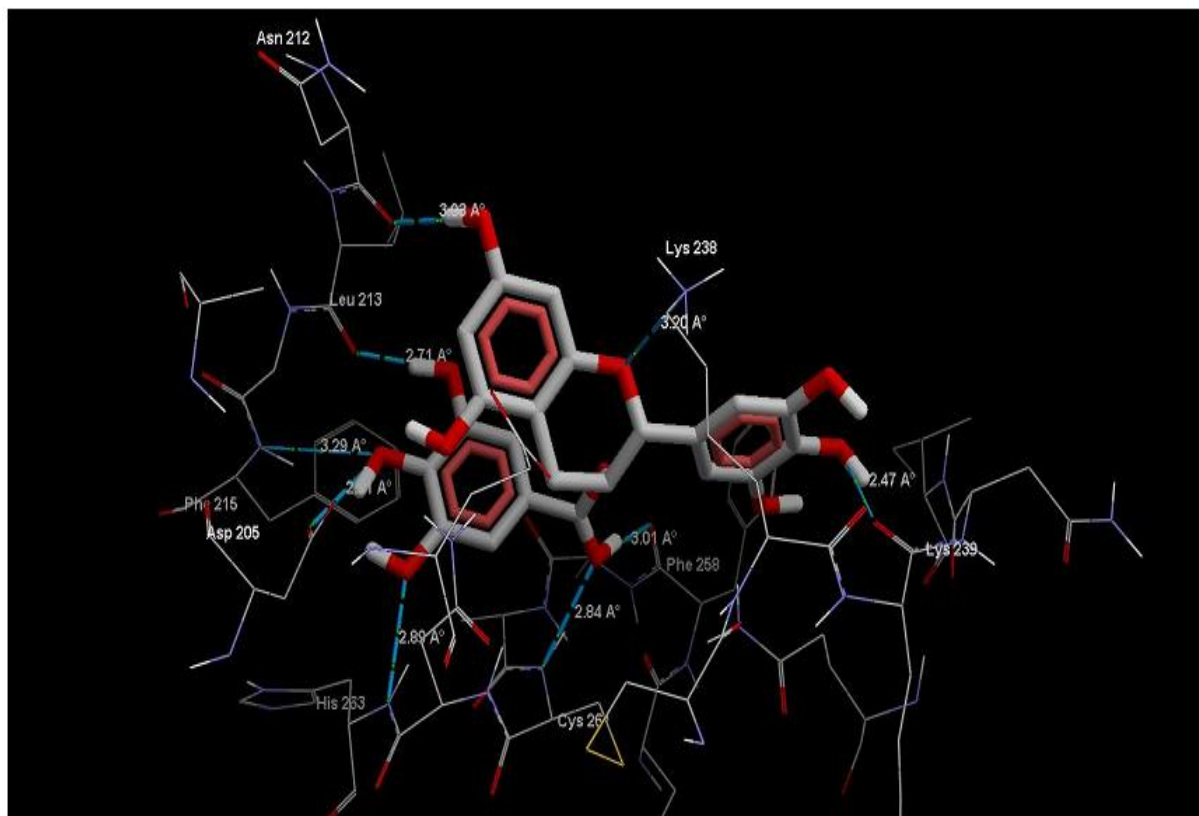


Figure 12. Représentation des interactions EGCG-LPH. Les liaisons hydrogène sont présentées par un trait bleu discontinue.

La visualisation des résultats du docking (fig.13) montre que le complexe formé EGCG-LPH est stabilisé au sein de la cavité 3 par des interactions hydrophobiques avec les résidus: Ile 241, Leu 213, Ala 259, Gly 214 et Lys 132,

Le composé EGCG établit également des interactions hydrophobiques de type π - π avec les cycles benzéniques des acides aminés: Phe 215 et Phe 258.

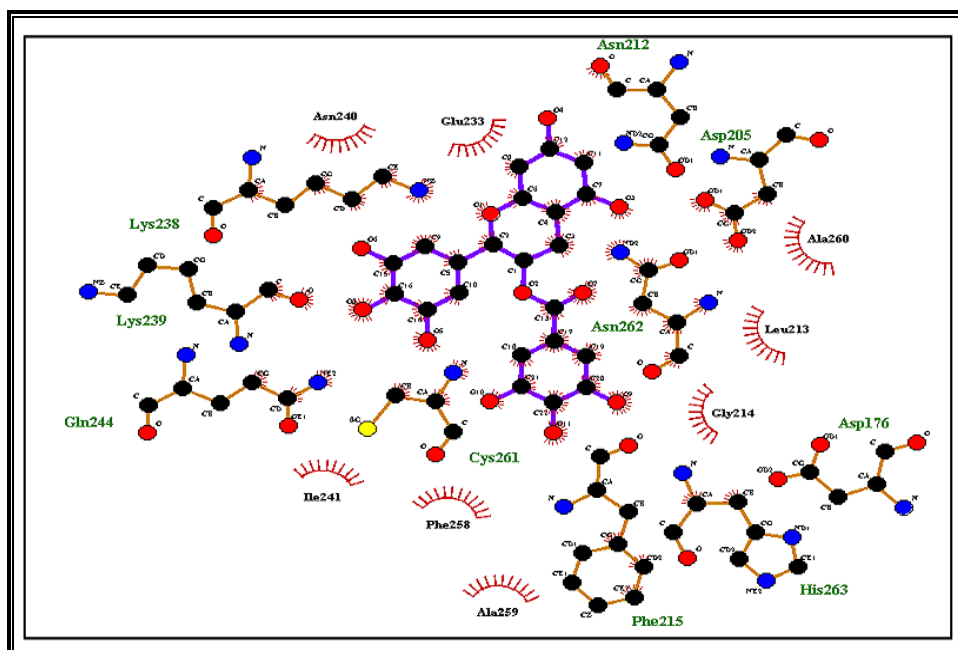


Figure 13. Représentation 2D des interactions hydrophobes.

IV.3.3. Comparaison des interactions entre LPP-EGCG et LPH-EGCG

La superposition de La lipase pancréas du porc (PDB ID: 1ETH) et la lipase pancréatique humaine (PDB ID: 1LPB) révèle une meilleure identité avec un pourcentage de 85% et une similarité de 90%. La comparaison visuelle des structures 3D (fig.14) montre une grande similarité, ce qui explique pourquoi l'EGCG a occupé la même poche pour les deux lipases. Cela est également confirmé par l'alignement représenté en format texte (fig.15).

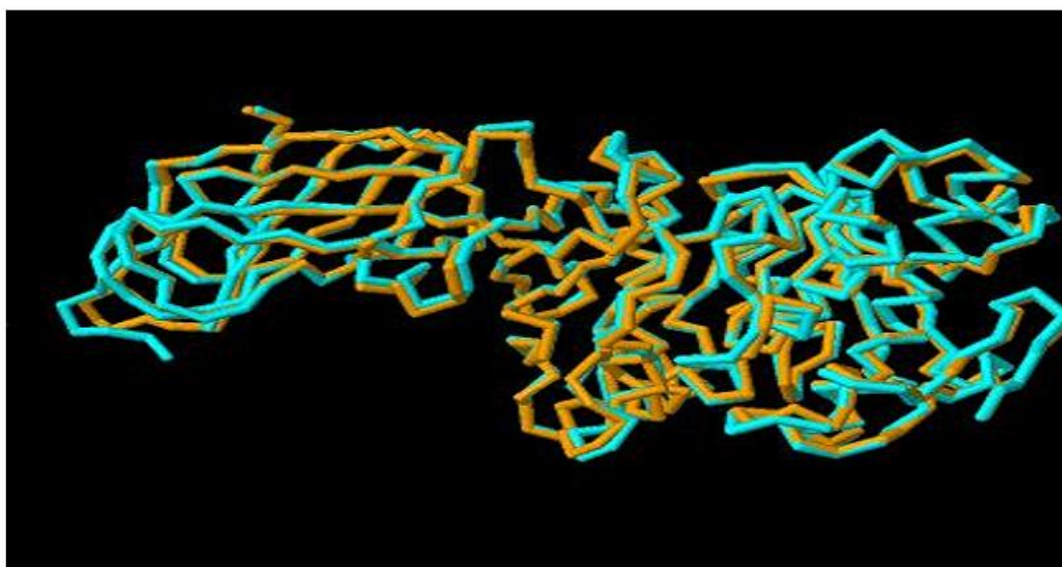


Figure 14. La superposition de la lipase pancréatique du porc (en orange) et la lipase pancréatique humaine (en bleu).

L'alignement séquentiel effectué été réalisé par l'Algorithme JCE¹¹. Il représente deux séquences : 1ETH.A.PDB et 1LPB.B.PDB. Ces deux séquences sont positionnées l'un sous l'autre, de manière à en faire ressortir les régions homologues ou similaires. L'objectif de cet alignement est de disposer les composants pour identifier les zones de concordance.

Les colonnes alignées contiennent des caractères identiques ou similaires, indiqués par un système cohérent de symboles (fig.15). Cet alignement met en évidence les résidus structurellement équivalents et identiques, ceux structurellement équivalents et similaires et les résidus structurellement équivalents. Les résidus similaires ne sont pas indiqués ici.

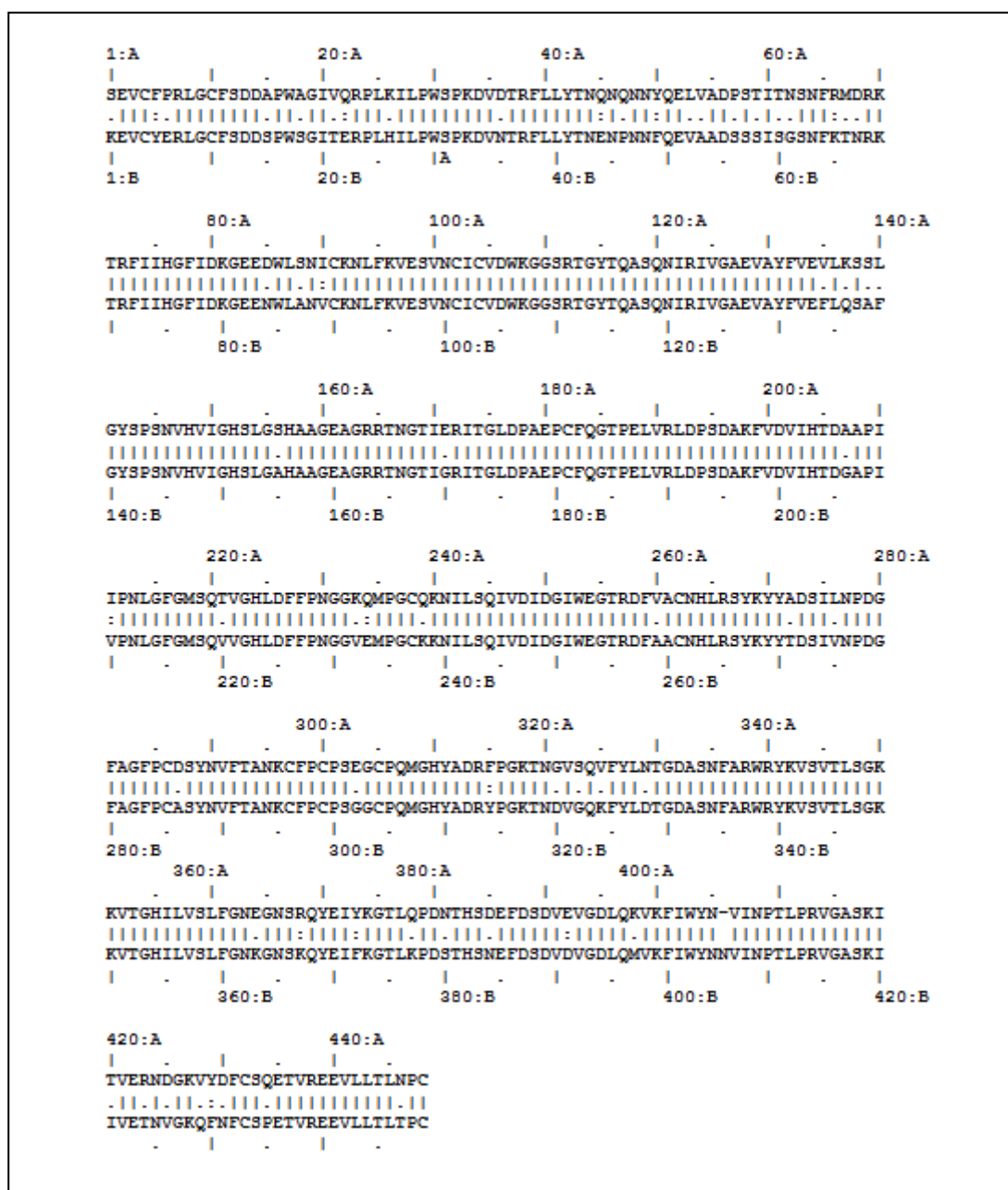


Figure 15. Alignement des séquences 1ETH.A.PDB et 1LPB.B.PDB.¹¹

La poche 3 située dans le domaine N-terminal est du type α / β . Elle comprend 200-270 résidus. Dans ce domaine, l'alignement des séquences de 1ETH.A.PDB et 1LPB.B.PDB (fig.15) montre une haute homogénéité. Cependant, l'existence de quelques résidus non similaire peut influencer les résultats de docking (tableau 9).

Ce réarrangement peut probablement altérer l'orientation d'inhibiteur (EGCG). Ce qui augmente le nombre d'interaction avec LPH par rapport à la lipase pancréatique porcine.

Tableau 9. Les résidus non similaires obtenus lors de l'alignement dans l'intervalle [200,270].

Séquences de 1ETH.A.PDB	Séquences de 1LPB.B.PDB
Ala 207	Gly206
ILE 211	Val 210
THR 221	Val 220
Lys233	Val 233
Gln 234	GLU 233
Gln 239	Lys 238
Val 260	ALA 259

IV.3.4. Cas de la lipase gastrique (LG)

La lipase gastrique humaine est synthétisée par les cellules principales de l'estomac, et sa sécrétion est stimulée par la gastrine et son équivalent synthétique la pentagastrine, sans modification de sa concentration dans le suc gastrique.^{12, 13} Cette lipase est sécrétée sous forme directement active, à la concentration moyenne d'environ 80 à 100 μg par millilitre de suc gastrique dans les conditions physiologiques normales basales.¹⁴

La lipase gastrique humaine (LGH) est une enzyme lipolytique sécrétée par les cellules principales situées dans la partie fundique de l'estomac. LGH joue un rôle important dans la digestion des lipides, car il favorise l'action hydrolytique ultérieure de la lipase pancréatique dans la lumière duodénale.¹⁵

De nombreuses études épidémiologiques ont démontré que l'extrait de thé vert conduisait à une inhibition des lipases digestives. Une étude in vitro, a en effet permis de démontrer que

l'extrait de thé vert à la dose de 6 mg d'extrait pour 100 mg de lipides, permettait de supprimer partiellement l'émulsification des lipides, tant en milieu gastrique qu'intestinal¹⁶. Lorsque l'on sait que l'émulsification des lipides est l'étape indispensable à l'action des lipases sur les lipides alimentaires, ces résultats peuvent expliquer la capacité d'inhibition des lipases digestives.

Une autre étude *in vitro*, réalisée et qui reproduise les conditions physiologiques sous forme d'action successive sur la trioléine de la lipase gastrique puis de la lipase pancréatique, a démontré que l'extrait de thé vert, à la dose de 6mg/100mg de lipides, permettait une inhibition presque totale de la lipase gastrique (89% d'inhibition) et partielle de la lipase pancréatique (32% d'inhibition) soit une inhibition de la lipolyse totale de près de 40%¹⁷. Ces extrait de thé vert riche en catéchols, en particulier contenant de 20 à 50 % en masse de catéchols sont constitués de l'épigallocatechine gallate (EGCG).

Cette partie de travail concerne l'étude du mécanisme d'action de l'épigallocatechine gallate en tant qu'inhibiteur de la lipase gastrique, qui à notre connaissance n'a jamais été décrite, en utilisant le docking moléculaire.

Dans cette étude nous avons choisis comme récepteur la lipase gastrique du chien (LGC) qui présente un bon modèle de la lipase gastrique humaine (LGH). La LGC et la LGH, ont été exprimées sous formes recombinantes¹⁸. Ce qui a permis l'obtention de leur structure cristallographique¹⁹. Elles sont composées d'une structure globulaire dont le repliement α/β , au centre, et entourée par six hélices α . Le site actif, composé d'une triade catalytique Ser-Asp-His, est recouvert d'un volet bloquant son accès. La LGH a été résolue sous sa forme fermée à une résolution de 3 Å, tandis que la LGC a été obtenue sous sa forme ouverte, à une résolution de 2,7 Å, avec un inhibiteur alkylphosphonate dans son site actif²⁰. Dans le cadre de ce travail, l'étude de l'ancrage du ligand dans le site actif de l'enzyme oriente le choix vers une forme ouverte de la lipase. Pour cette raison nous avons choisis comme récepteur LGC qui présente un bon modèle de la lipase gastrique humaine.

Pour cette étude nous avons suivi le même protocole de docking adopté précédemment. Afin de déterminer le site d'interaction nous avons effectué l'ancrage de l'épigallocatechine dans toutes les cavités existantes du récepteur. Une fois cette étape de fixation a été réalisée nous avons validé l'identification du site actif à partir de la meilleure valeur d'énergie d'interaction avec la cible. Etant donné qu'aucun travail expérimental n'a décrit le mode

d'inhibition de la lipase gastrique par l'EGCG, la cavité principale sera engagée dans ce cas pour la simulation (fig.16).

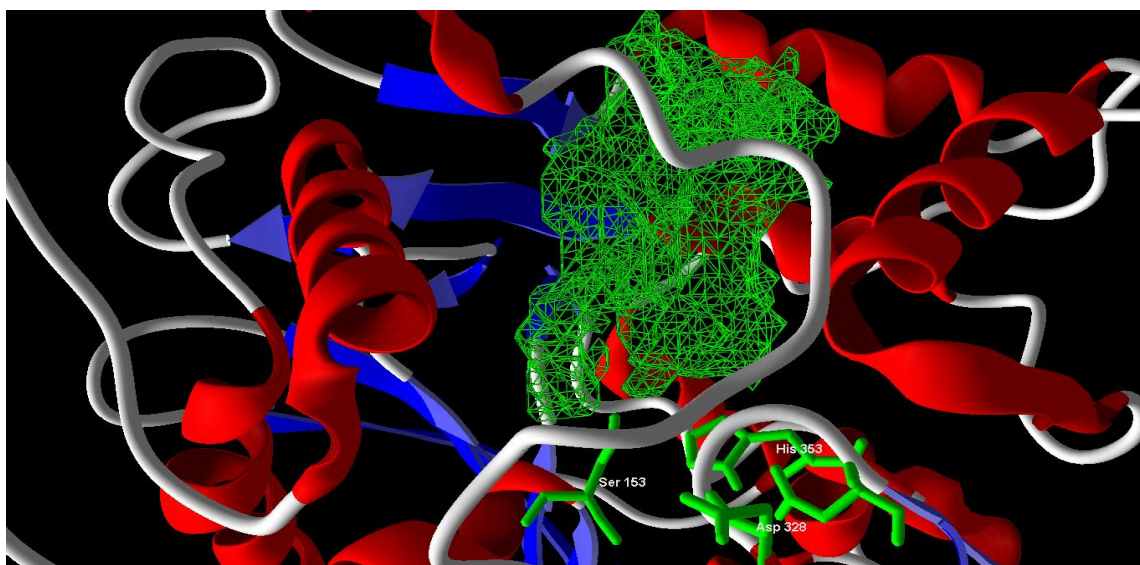


Figure 16. Zoom sur le site catalytique de la LGC avec les résidus principaux (Ser153, His353, et Asp328).

La sphère définissant la RI s'est limitée à un rayon de 10 Å afin d'éviter l'échantillonnage de l'extérieur du site de fixation. La modélisation au sein de la cavité 3 et 5 est exclue puisque elles situent très loin de site actif. Les simulations de docking ont été réalisées au niveau des cavités avec le programme Molegro Virtual Docker (MVD). La structure de la lipase gastrique est importée dans son espace de recherche et les cavités sont détectées par l'algorithme de détection des cavités du MVD, après une préparation de la protéine.

Dans l'espace de recherche défini, un processus de docking est lancé pour l'épigallocatechine-3-gallate en utilisant l'algorithme de recherche Moldock Optimizer, et la fonction de score Moldock Score [Grid]. Les résultats du docking sont représentés dans le tableau 10.

Tableau 10. Valeurs des énergies obtenues pour chaque cavité.

Cavité	Moldock Score (kcal mol ⁻¹)	Rerank Score (kcal mol ⁻¹)
1	-150,27	-127,07
2	-123,92	-89,38
4	-128,65	-105,33

Les valeurs d'énergies d'interaction (tableau 10) générées par la fonction de score «Moldock Score», lors de l'amarrage de l'épigallocatechine-3-gallate dans les cavités 2 et 4, sont égales à -123,92 et 128,65 kcal mol⁻¹, respectivement.

Tandis que celle obtenue avec la cavité 1 est égale à -150,27 kcal mol⁻¹, ce qui indique que cet inhibiteur pourrait être d'une grande affinité pour les résidus de la cavité principale. Cette étude théorique préliminaire sur l'épigallocatechine-3-gallate, montre que ce dernier agit comme inhibiteur compétitif pour la lipase gastrique. Afin de valider cette hypothèse, nous avons décidé de tester un autre protocole. Ce dernier sera détaillé ci-dessous.

Pour ce faire, nous avons commencé par définir un espace de recherche. Ce dernier représenté par une sphère de rayon de 25Å (fig.17) centrée sur le centre de la cavité 1, couvre tout l'espace du récepteur, y compris le site actif et les sites allostériques. Dans cet espace de recherche, un processus de docking est lancé pour l'EGCG dans les mêmes conditions du protocole cité précédemment.

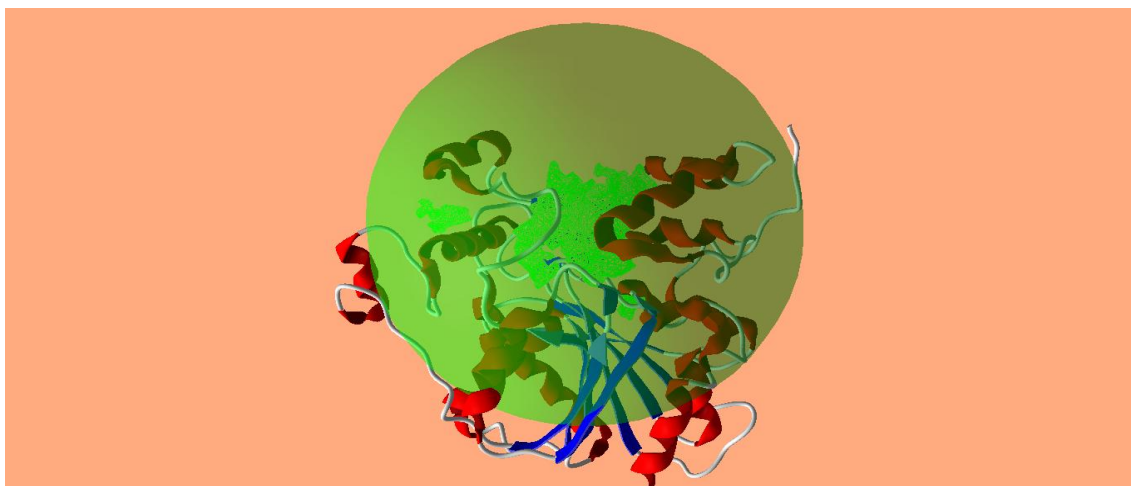


Figure 17. Illustration de l'espace de recherche qui couvre la structure 3D de la cible centré sur la cavité 1.

Les résultats obtenus (fig.17) montrent que l'orientation de la pose obtenue vers la cavité principale est similaire à celle déterminée dans la première méthode. Cette similarité est exprimée par la même valeur d'énergie obtenue dans les deux cas et qui est égale à -150,26 kcal mol⁻¹.

Dans une dernière tentative nous avons fait un deuxième balayage de l'espace couvrant la structure 3D du récepteur. La RI a été définie par une sphère de 25 Å de rayon centrée sur le

ligand de référence. L'illustration de ce deuxième balayage est représentée par la figure 18.

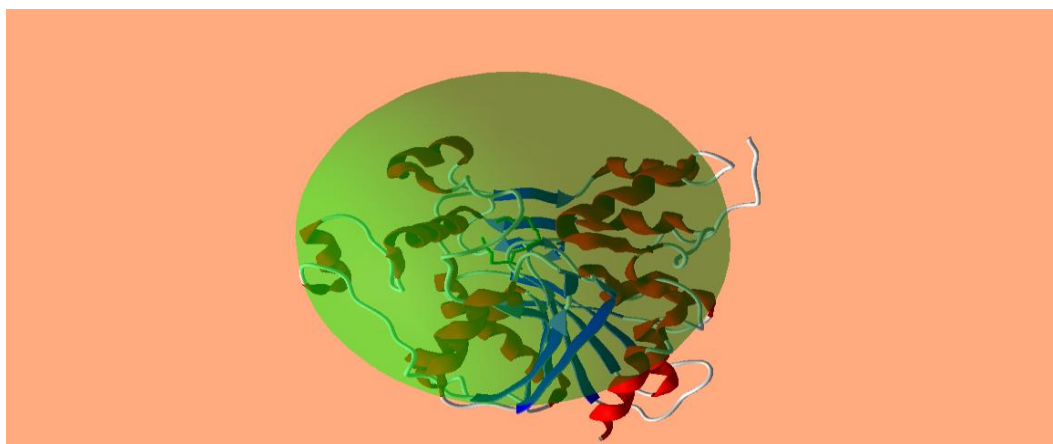


Figure 18. Illustration de l'espace de recherche qui couvre la structure 3D de la cible centrée sur le ligand de référence.

La figure 18 montre que l'épigallocatechine-3-gallate est toujours orienté vers la cavité principale avec la même valeur d'énergie. Ces résultats avec les trois protocoles de docking sont résumés dans le tableau 11.

Tableau 11. Protocoles de docking moléculaire.

Espace de recherche	Moldock Score
Un espace de recherche limitée à un rayon de 10Å couvrant la cavité 1	-150,27
Une sphère de 25 Å de rayon centrée sur la cavité 1	-150,26
Une sphère de 25 Å de rayon centrée sur le ligand de référence	-150,24

L'analyse visuelle de la meilleure pose de l'EGCG, montre que ce ligand est bien ancré au fond du site actif de la lipase gastrique (fig.19). Cette stabilisation résulte de la mise en place des cinq liaisons hydrogènes ainsi que plusieurs interactions hydrophobes. L'inhibiteur EGCG forme cinq liaisons hydrogène, dont quatre sont établies par les groupements hydroxyles et les résidus du site d'interaction (fig.20). Ces liaisons se présentent comme suit :

- 1) EGCG avec Ser153.
- 2) EGCG avec Leu 68.
- 3) EGCG avec Gly66.

4) EGCG avec Asn 263.

5) EGCG avec Asn 73.

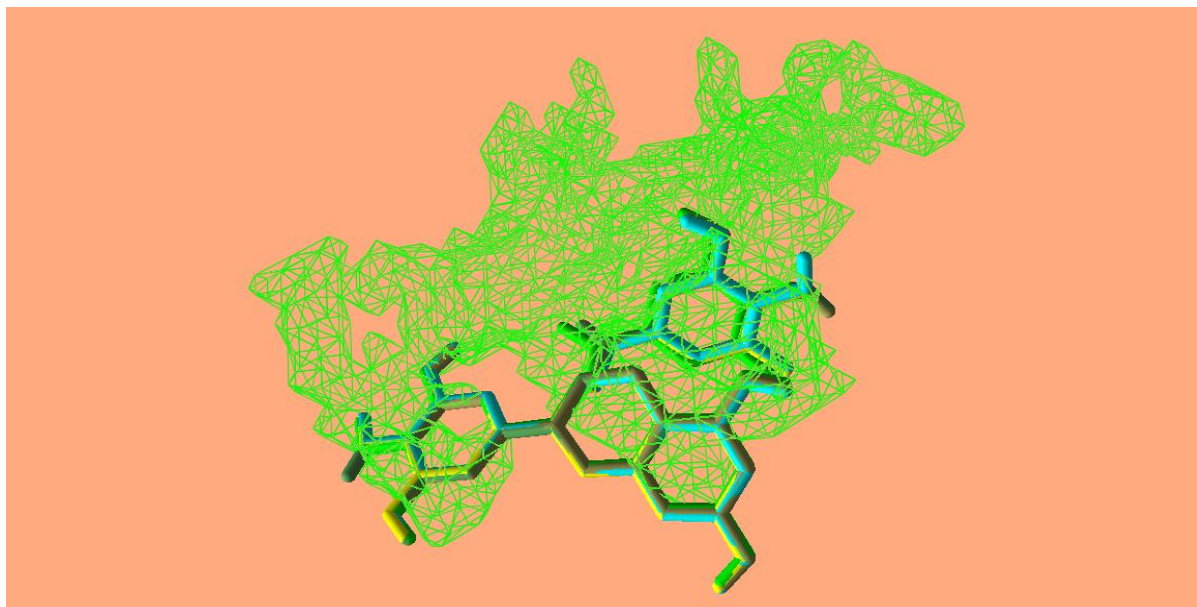


Figure 19. Orientation des trois poses de EGCG obtenues, vers la cavité principale de lipase gastrique.

La distance et l'énergie de chaque liaison hydrogène sont présentées dans le tableau 12:

Tableau 12. Valeur d'énergie et la distance des liaisons hydrogène formées.

Liaison hydrogène d'EGCG	La distance (Å°)	Energie (kcal)
Oxygène de cycle pyrane avec Ser153 (O-O)	2,81	2,5
L'hydroxyle en position (20) avec Leu 68 (O-O)	2,67	-2,5
L'hydroxyle en position (20) avec Asn 263 (O-N)	3,22	-1,89

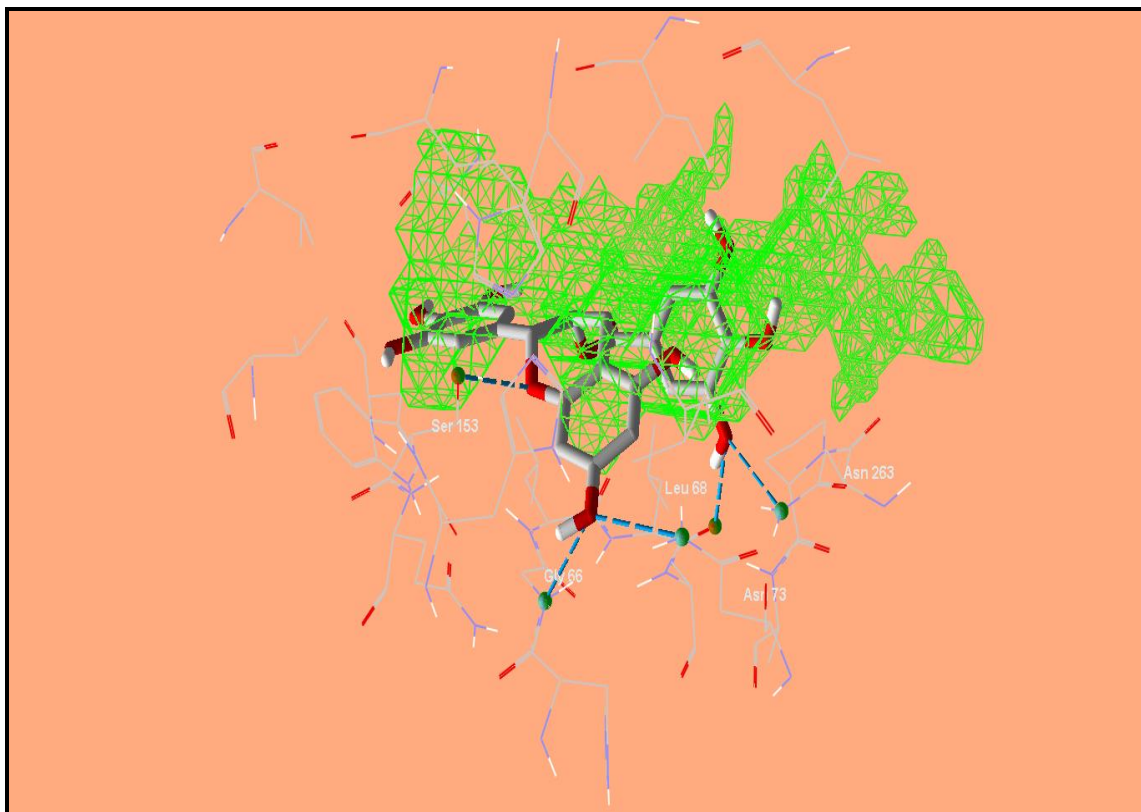


Figure 20. Représentation des interactions EGCG –LGC avec les liaisons hydrogène présentées par un trait bleu discontinue.

Le programme MVD (fig.21) nous a permis de prédire l'énergie d'interaction entre EGCG et la LGC qui s'évalue à $-150,27 \text{ kcal mol}^{-1}$. Le complexe formé est stabilisé au sein de la cavité N°1 qui représente la poche principale de la LGC. Ces interactions type hydrophobe sont définies avec les résidus suivants : Val182, Leu326, Ile243, Val185, Leu 240, Leu67, Leu68, Ala265, Phe75. Dans ce cas le ligand établit également des interactions hydrophobiques de type π - π entre le ligand et le cycle benzénique du résidu Trp275.

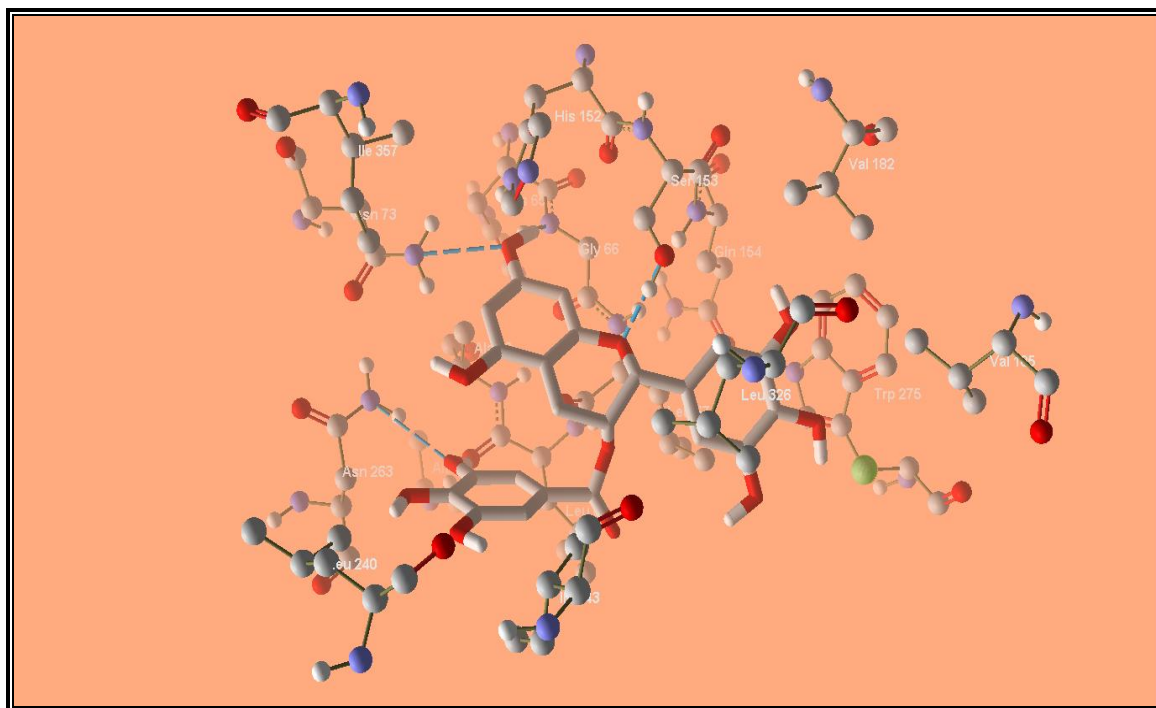


Figure 21. Interaction hydrophobe entre la lipase gastrique et l'EGCG.

On remarque ici que l'épigallocatechine se caractérise par une forte activité inhibitrice. Le programme MDV valide de façon exhaustive ce résultat en fournissant une énergie d'interaction égale à $-150,27 \text{ kcal mol}^{-1}$. La validation de ce caractère inhibiteur important est aussi constatée par la bonne pénétration de ligand au niveau du site actif. L'orientation d'épigallocatechine-3-gallate vers le site actif de la lipase gastrique nous a permis de proposer un mode d'inhibition compétitif, en attendant les résultats de la cinétique enzymatique. L'analyse des interactions mises en jeu, a montré que cette affinité est principalement due aux liaisons hydrogène favorisées par la présence de plusieurs groupements hydroxyles.

En conclusion, Les résultats obtenus ci dessus, ont illustré les différentes orientations des poses générées de l'EGCG vers le site actif de lipase gastrique (LG). Cette grande affinité avec les résidus de la cavité principale prouvée par la valeur d'énergie basse et qui est égale à $-150,27 \text{ kcal mol}^{-1}$. Cette stabilité résulte de la mise en place des cinq liaisons hydrogène et des plusieurs interactions hydrophobe vis-à-vis des résidus de la cavité principale de la protéine. Une de ces liaisons hydrogène estimée la plus importante a été formée avec le résidu Ser 153. Ce dernier faisant partie de la triade catalytique de cette protéine. Cette affinité théorique vers les résidus catalytique prédit un caractère compétitif important d'inhibition pour la lipase gastrique. Ces résultats obtenus pourront aider au développement d'un outil thérapeutique efficace pour lutter contre la croissance de la maladie d'obésité

Références Bibliographiques

- 1: Charbonneau D. Eléments structuraux essentiels à la thermostabilité de nouvelles enzymes lipolytiques. Thèse de doctorat en biophysique et biologie cellulaires. Université du Québec à Trois-Rivières 2014.
- 2: Arpigny JL, Jaeger KE. Bacterial lipolytic enzymes: Classification and properties. *Biochemical Journal*, 1999; 343: 177-183.
- 3: Hui DY, Howles PN. Carboxyl ester lipase: Structure-function relationship and physiological role in lipoprotein metabolism and atherosclerosis. *Journal of Lipid Research*, 2002, 43(12): 2017-2030.
- 4: Liu P, Wang YF, Ewis HE, Abdelal A, Lu CD, Harrison RW, Weber IT. Covalent reaction intermediate revealed in crystal structure of the *Geobacillus stearothermophilus* carboxylesterase. *Journal of Molecular Biology*, 2004; 34(230): 551-561.
- 5: Kern PA, Ong JM, Saffari B, Carty J. The effects of weight loss on the activity and expression of adipose-tissue lipoprotein lipase in very obese humans. *The New England Journal of Medicine*, 1990; 322: 1053-1059.
- 6: Tcheriatchoukine D. L'obésité: découvertes récentes relatives aux mécanismes moléculaires à l'origine de nouvelles stratégies thérapeutiques. Thèse de doctorat faculté de médecine Université Henri Poincaré Nancy 2010.
- 7: Wang S, Sun Z, Dong S, Liu Y, Liu Y. Interactions between (–)-Epigallocatechin Gallate Analogs and Pancreatic Lipase. *Plos One* 2014; 9 (11): e111143.
- 8: John S, Walther D, Phillip I, Harold B, Jeffrey R, James M, Jeffrey W. Mechanism of Action Assays for Enzymes, Eli Lilly & Company and the National Center for Advancing Translational Sciences, 2004.
- 9: Keillor JW. Inhibition des réactions enzymatiques. BCM 2504. Enzymologie 2004.
- 10: Sergent T, Vanderstraeten J, Winand J, Beguin P, Schneider YJ. Phenolic Compounds and Plant Extracts as Potential Natural Anti-Obesity Substances. *Food Chemistry*, 2012;135:68-73.

- 11:** Prlić A, Bliven S, Rose PW, Bluhm WF, Bizon CS, Godzik A, Bourne PE. Pre-calculated protein structure alignments at the RCSB PDB website. *Bioinformatics*, 2010; 26(23): 2983–2985.
- 12:** Gargouri Y, Moreau H, Verger R. Gastric lipases: biochemical and physiological studies. *Biochimica and Biophysica Acta*, 1989; 1006: 255-271.
- 13:** Hamosh M. Lingual and gastric lipases: their role in fat digestion. *CRC Press*. Boca Raton. FL1990.
- 14:** Carrière F, Barrowman JA, Verger R, Laugier R. Secretion and contribution to lipolysis of gastric and pancreatic lipases during a test meal in humans. *Gastroenterology*, 1993; 105: 876-888.
- 15:** Miled N, Canaan S, Dupuis L, Roussel A, Rivière M, Carrière F, de Caro A, Cambillau C, Verger R. Digestive lipases: from three-dimensional structure to physiology. *Biochimie*, 2000; 82(11): 973-986.
- 16:** Juhe G, Armand M, Pafumi Y, Rosier C, Vandermader J, Lairon D, Green tea extracts (AR25t) inhibits lipolysis of triglycerides in gastric and duodenal medium in vitro. *Journal of Nutritional Biochemistry*, 2000; 11: 45-51.
- 17:** Max Rombi. Composition for treating obesity and aesthetic treatment method. Patent No: WO2000041708 A1 2000.
- 18:** Max Rombi. Green tea extracts containing catechols and caffeine. Patent No; EP 1098657 B1 2002.
- 19:** Canaan S, Dupuis L, Rivière M, Faessel K, Romette JL, Verger R, Wicker-Planquart C. Purification and interfacial behavior of recombinant human gastric lipase produced from insect cells in a bioreactor. *Protein Expression and Purification*, 1998; 14(1): 23-30.
- 20:** Roussel A, Canaan S, Egloff MP, Rivière M, Dupuis L, Verger R, Cambillau C. Crystal structure of human gastric lipase and model of lysosomal acid lipase, two lipolytic enzymes of medical interest. *Journal of Biological Chemistry*, 1999; 274(24): 16995-17002.
- 21:** Roussel A, Miled N, Berti Dupuis L, Rivière M, Spinelli S, Berna P, Gruber V, Verger R, Cambillau C. Crystal structure of the open form of dog gastric lipase in complex with a phosphonate inhibitor. *Journal of Biological Chemistry*, 2002; 277(3): 2266-2274.

Conclusion générale

Conclusion et perspectives

Les approches théoriques permettant la prédiction du mode d'interaction d'un ligand avec son récepteur sont complémentaires des expériences et permettent parfois leur interprétation. Elles nous renseignent de façon significative sur les interactions au niveau moléculaire et sont, de ce fait, le socle à partir duquel une conception ou une optimisation rationnelle de molécules actives, prenant en compte des critères structuraux, peut être envisagée. La synergie entre les approches théoriques et expérimentales est cruciale pour une évolution optimale des connaissances en biologie, pharmacie et médecine. Un tel environnement de recherche est très certainement l'une des clés vers un chemin de plus en plus court pour l'identification d'un besoin thérapeutique.

La présente recherche qui rentre dans le cadre de l'axe de recherche « interactions protéines-ligands » développé au sein de notre laboratoire. Notre recherche consiste à étudier l'inhibition des enzymes impliquées dans la maladie de l'obésité, avec un inhibiteur naturel. Pour cela nous avons fait appel à l'outil informatique pour expliquer cette inhibition. La méthode employée dans cette étude est le criblage virtuel (Docking moléculaire).

Dans un premier temps, une étude bibliographique a été effectuée, rappelant des généralités sur les ligands, leurs structures, leurs caractéristiques structurales, leurs propriétés dans le domaine médicale et leur mécanisme d'action. Dans cette partie nous avons aussi présenté des généralités sur les lipases étudiés, leurs caractéristiques structurales et leur implication, particulièrement dans l'obésité. Dans cette synthèse bibliographique nous avons détaillé la technique de docking moléculaire adopté dans cette étude.

L'outil informatique utilisé pour le docking est Molegro Virtual Docker (MVD). Ce programme a été développé pour aider à la mise au point de molécules à activité thérapeutique. Toutes énergies obtenues, ainsi que les liaisons hydrogènes formées lors de l'amarrage (docking) dans le complexe protéine-ligand sont employées pour analyser les interactions entre les nouvelles structures du ligand et le site actif de la protéine cible. Les informations recueillies, après le processus de criblage virtuel, permettent de déterminer lesquelles sont les plus prometteuses.

La deuxième partie de ce travail consiste à décrire les méthodes expérimentales utilisées ainsi qu'exposé les résultats et la discussion de cette étude. Dans la première étape de cette partie, nous avons testé la fiabilité du programme Molegro Virtuel Docker utilisé dans ce

Conclusion générale

travail. Pour cela nous avons fait appel au test de RMSD. À ce titre, nous avons utilisé les ligands de deux complexes portant les codes PDB suivants: 1LPB, 1K8Q.

Les résultats obtenus dans cette étape préliminaire, ont pu élucider la performance du programme Molegro Virtuel Docker (MVD). La fiabilité du logiciel MVD adopté dans notre étude de docking, a été évaluée à partir des valeurs de la RMSD obtenus. Cette fiabilité est également visualisée par la superposition de la conformation optimale du ligand calculée par MVD et la géométrie du même ligand pris initialement de la PDB. Nous avons constaté à travers cette dernière analyse que le programme est généralement plus efficace quand le ligand est moins flexible.

La deuxième étape de cette partie de travail nous a permis d'expliquer les mécanismes d'interaction entre l'épigallocatechine-3-gallate présenté comme un inhibiteur non compétitive expérimental, vis-à-vis de deux lipases pancréatiques testées (LPP et LPH). Ces mécanismes ont été élucidés par la détermination du site de fixation et la visualisation des différents types d'interactions mis en jeu. L'analyse des résultats montrent l'existence des liaisons hydrogène importantes. Ces liaisons ont été formées avec les résidus His 263 pour la LPP et His 264 pour la LPH. On note que ces deux résidus font partie de la triade catalytique de ces enzymes. Ces interactions formées peuvent modifier la géométrie de site active de la protéine. Pour cela une étude comparative *in silico* a été également réalisée entre l'interaction des complexes EGCG-LPP et EGCG-LPH. Les résultats obtenus lors de cette étude comparative montrent que l'EGCG a une meilleure activité inhibitrice vis-à-vis de la LPH comparé avec la LPP.

Enfin dans cette partie de travail, nous avons effectuée une modélisation moléculaire, afin d'expliquer le mode d'inhibition de l'épigallocatechine-3-gallate vis-à-vis à la lipase gastrique. Cette approche théorique qui à notre connaissance n'a jamais été réalisé pour cet inhibiteur. Les résultats de docking obtenus montrent une stabilisation du ligand au niveau de la cavité principale. Cette affinité théorique vers les résidus catalytique prédit donc un caractère compétitif d'inhibition.

En conclusion, les résultats obtenus peuvent contribuer au développement d'un outil thérapeutique efficace pour lutter contre la croissance de la maladie d'obésité. A partir de ces résultats préliminaires, nous envisagerons une étude approfondie en faisant appel à d'autres programmes d'étude de Docking et de simulation *in silico*. Nous envisagerons aussi d'étudier d'autre inhibiteur naturel pour ces lipases qui possède d'autres propriétés thérapeutiques pour d'autres maladies.