



République Algérienne Démocratique et Populaire

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

Université Med Khider Biskra

Faculté des Sciences Exactes et des Sciences de la Nature et de la Vie

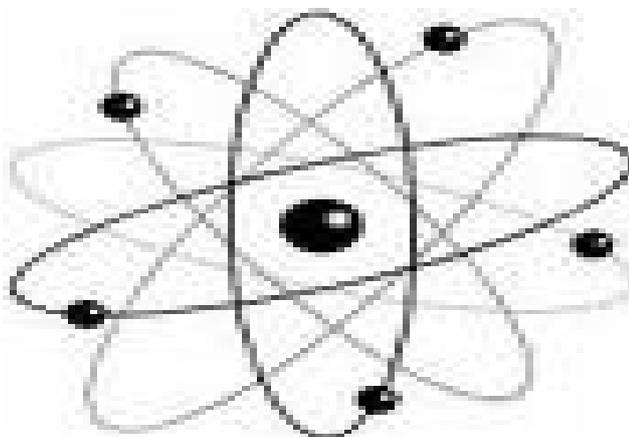


Département des Sciences de la Matière

Domaine : Sciences de la Matière

Filière : Chimie

Spécialité : Chimie Pharmaceutique



Mémoire de fin d'étude en Master

Intitulé :

**Conception par modélisation moléculaire et
Criblage virtuel d'une série bioactive Phtalazine**

Présenté par :

Abdelaziz Nor el Imane

Devant le jury :

Mme. Mellaoui Malika	(M.C.B)	Université de Biskra	Présidente
Mme. Khamouli Saida	(M.A.A)	Université de Biskra	Rapporteur
Mme. Djouama Hanane	(M.A.A)	Université de Biskra	Examinatrice

REMERCIEMENTS

C'est à Dieu le tout puissant que je dois adresser mes remerciements les plus sincères pour m'avoir donné la santé, la volonté et le courage sans lesquels ce travail n'aurait pu être réalisé. Malgré les apparences et les modestes mots qu'elle contient, cette page de ce mémoire de master est la plus difficile à rédiger. Réellement, comment exprimer en si peu de mots toute ma reconnaissance aux personnes qui m'ont encouragée. Ces quelques lignes sont une occasion pour moi afin de remercier les personnes qui de près ou de loin, m'ont soutenue par diverses manières et sans qui, ce travail n'aurait pas abouti.

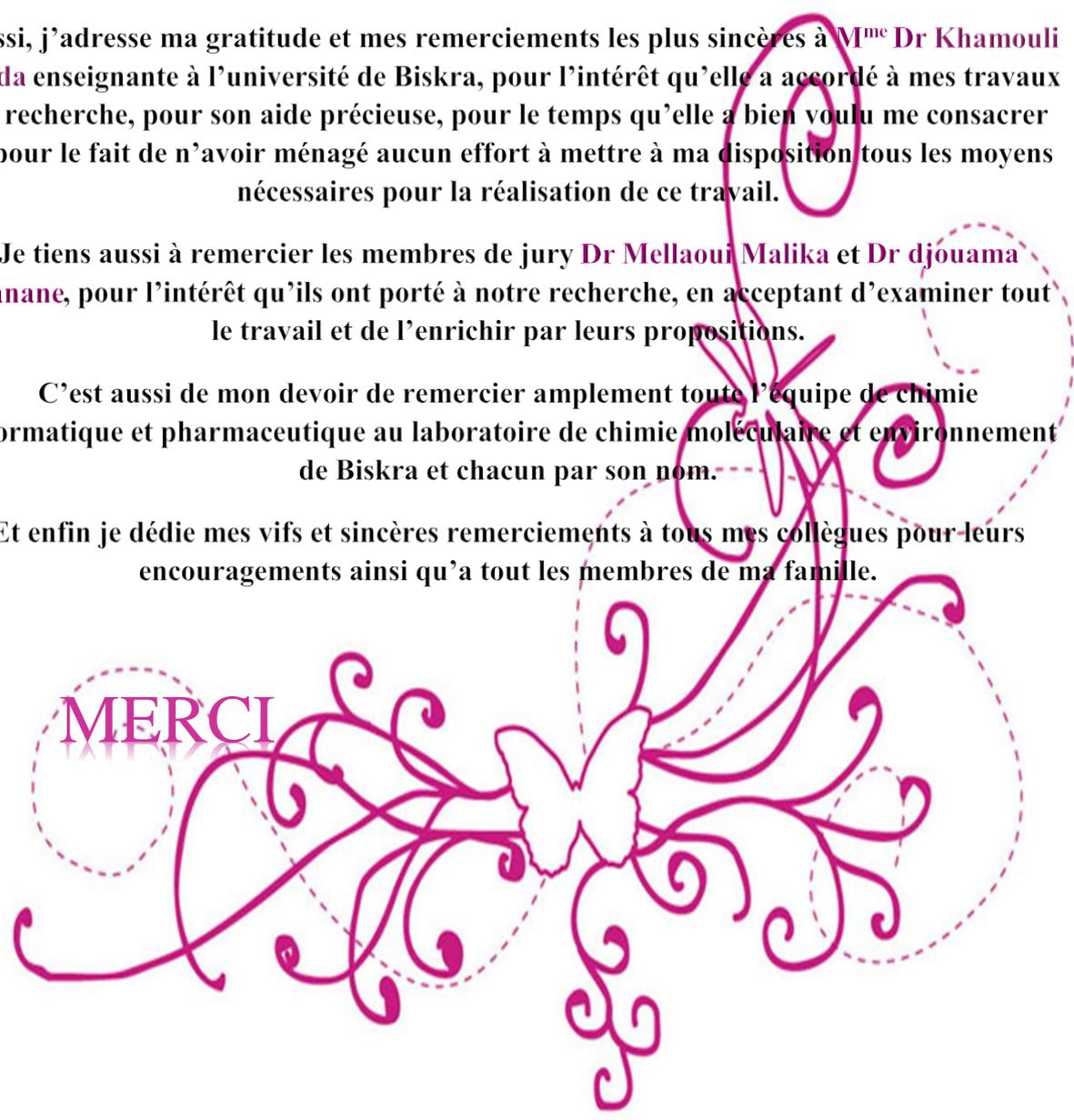
Aussi, j'adresse ma gratitude et mes remerciements les plus sincères à **M^{me} Dr Khamouli Saida** enseignante à l'université de Biskra, pour l'intérêt qu'elle a accordé à mes travaux de recherche, pour son aide précieuse, pour le temps qu'elle a bien voulu me consacrer et pour le fait de n'avoir ménagé aucun effort à mettre à ma disposition tous les moyens nécessaires pour la réalisation de ce travail.

Je tiens aussi à remercier les membres de jury **Dr Mellaoui Malika** et **Dr djouama Hanane**, pour l'intérêt qu'ils ont porté à notre recherche, en acceptant d'examiner tout le travail et de l'enrichir par leurs propositions.

C'est aussi de mon devoir de remercier amplement toute l'équipe de chimie informatique et pharmaceutique au laboratoire de chimie moléculaire et environnement de Biskra et chacun par son nom.

Et enfin je dédie mes vifs et sincères remerciements à tous mes collègues pour leurs encouragements ainsi qu'à tout les membres de ma famille.

MERCI

A decorative graphic at the bottom of the page features the word 'MERCI' in a stylized, pink, serif font. To the right of the text is a large, intricate pink butterfly with its wings spread, surrounded by swirling, vine-like patterns. The entire graphic is set against a background of faint, dashed pink circles.

Dédicace

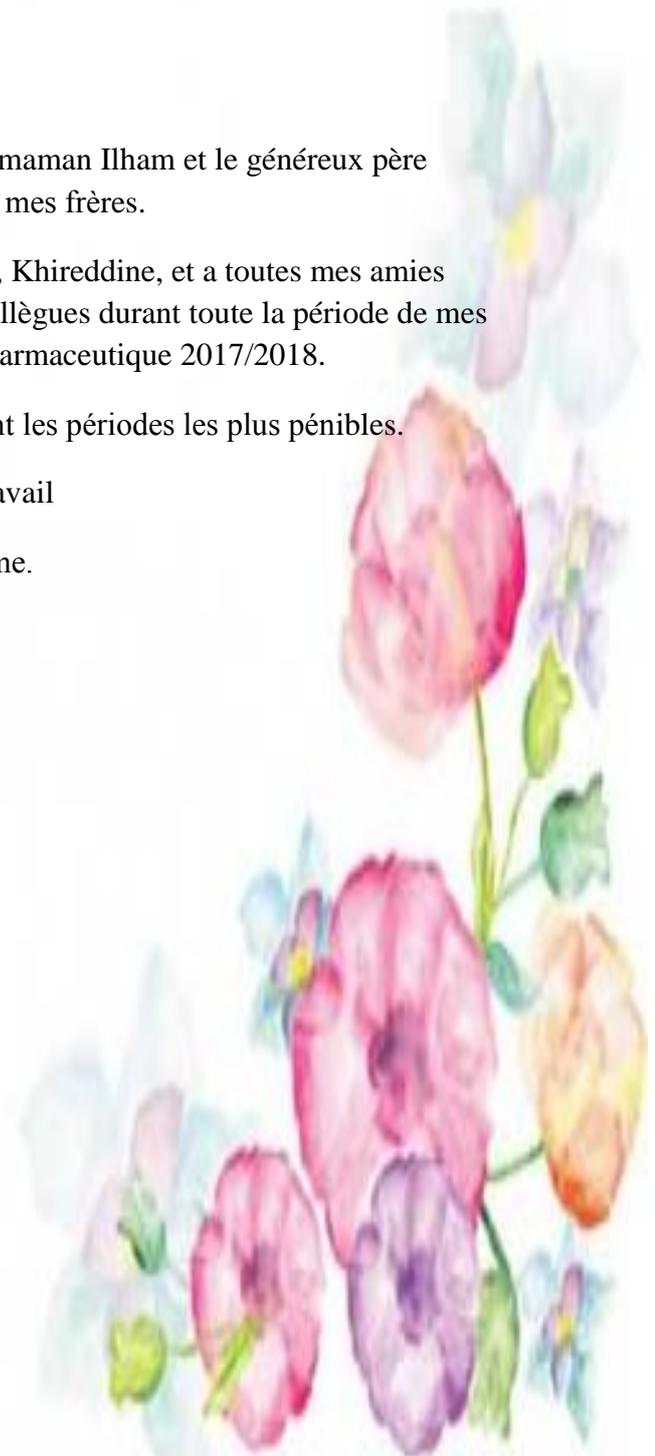
Je dédie le fruit de mon modeste travail à ma chère maman Ilham et le généreux père Abdelmadjid. A mes sœurs et mes frères.

A tous les membres de ma famille Abdelaziz, Laloui, Khireddine, et a toutes mes amies Nabila, Wafa, Hind, surtout « Les halogènes » et mes collègues durant toute la période de mes études, et a toute la promotion de chimie pharmaceutique 2017/2018.

A tous ceux qui m'ont encouragé et soutenu durant les périodes les plus pénibles.

Je dédie ce modeste travail

A tous ceux que j'aime.



Sommaire

SOMMAIRE

Remerciements	
Liste des tableaux	
Liste des figures	
Liste des abréviations	
INTRODUCTION GENERALE.....	1
Référence.....	3
Chapitre I : Généralités sur le Cancer de la Peau et les Composés Hétérocycliques (Phtalazine)	
I.1. Le Cancer.....	4
I.1.1.Définition.....	4
I.1.2. Mécanismes fondamentaux de la carcinogenèse.....	4
I.1.2.1.Tumeurs bénignes	4
I.1.2.2.Tumeur maligne.....	4
I.1.3.Les étapes de la carcinogenèse :	5
I.1.3.1. Initiation	5
I.1.3.2. Promotion	5
I.1.3.3. Progression	5
I.2.La peau humaine	6
I.2.1. Structure et fonction de la peau	6
I.2.2.Les différentes couches de la peau	7
I.2.3.Les mélanocytes : des cellules de l'épiderme	8
I.3.Cancer de la peau	9
I.3.1.Définition.....	9
I.3.2.Les différents types de cancer de la peau	9
I.3.2.1. Les carcinomes basocellulaires	9
I.3.2.2. Les carcinomes épidermoïdes.....	9
I.4. EGFR (Epidermal Growth Factor Receptor).....	10
I.4.1. Structure de l'EGFR.....	10

Sommaire

SOMMAIRE

I.4.2. L'activation de l'EGFR.....	11
I.4.3.L'EGFR et le cancer.....	12
I.4.4. Les modifications de l' EGFR impliquées dans le cancer.....	12
I.4.5. Le traitement anti-EGFR.....	13
I.4.6. Conséquences des interactions ligands-récepteurs.....	13
I.5. Les Hétérocycles	14
I.5.1. Définition des hétérocycles à six chaînons	14
I.5.2. Classification	14
I.6. phtalazine	15
I.6.1. Structure et nomenclature de phtalazine.....	15
I.6.2. Préparation de phtalazine	15
I.6.3. L'activité biologique de phtalazine	16
Référence.....	17

Chapitre II : Base Théorique des Méthodes de Calculs Utilisées

II.1.Introduction	19
II.2.Chimie quantique.....	20
II.2.1.Bases de la chimie quantique.....	20
II.2.1.1 Equation de Schrödinger.....	20
II.2.1.2.Approximation de Born-Oppenheimer.....	21
II.2.2.Théorie de la fonctionnelle de la densité (DFT).....	21
II.2.2.1. Aperçu historique.....	21
II.2.2.2. Théorème de Hohenberg et Kohn	22
II.2.2.3. Approche Kohn-Sham.....	23
II.2.2.4. Fonctionnelles d'échange-corrélation.....	24
II 2.2.4.1. Approximation de la densité locale (LDA).....	24
II.2.2.4.2. Approximation du gradient généralisé (GGA).....	25
II.2.3. Méthodes semi-empiriques	26
II.3. Etude quantitative des relations structure-activité (QSAR).....	26

Sommaire

SOMMAIRE

II.3.1. Introduction	26
II.3.2. Généralités sur la modélisation QSPR/QSAR.....	27
II.3.3.Principe	28
II.3.4. Importance de la base de données.....	29
II.3.5. Descripteurs Moléculaires	29
II.3.5.1. Descripteurs constitutionnels.....	29
II.3.5.2. Descripteurs topologiques.....	29
II.3.5.3. Descripteurs géométriques.....	30
II.3.5.4. Descripteurs électrostatiques.....	30
II.3.5.5. Descripteurs quantiques.....	30
II.3.5.6. Descripteurs thermodynamiques.....	31
II.3.6. Régression linéaire multiple (MLR).....	31
II.3.6.1.Estimation des paramètres statistiques du modèle.....	32
II.3.6.2.Tests sur le modèle linéaire.....	32
II.3.6.3.Test de la signification globale de la régression (F-Fisher).....	33
II.3.6.4.Interprétation et validation d'un modèle QSPR/QSAR.....	33
II.3.6.5.Validation interne.....	33
II.3.6.6.Validation externe.....	33
II.3.6.7.Critères de validation Externe (série de test).....	34
II.3.6.8.Critère de validation croisée : PRESS	34
II.3.6.9.Validation d'un modèle.....	35
II.4.Docking moléculaire	35
II.4.1.Introduction.....	35
II.4.2.Principe de docking.....	36
II.4.3.Les différents types de docking moléculaire.....	36
II.4.4.Interaction protéine-ligand	37
Référence.....	39
Chapitre III : Étude Qualitative et Quantitative des Propriétés QSAR d'une Série de Dérivés de Phtalazine	

Sommaire

SOMMAIRE

III.1.Introduction	43
III.2. Etude qualitative des propriétés QSAR des dérivés de phtalazine	44
III.2.1. Structure chimique des dérivés de phtalazine.....	44
III.2.2. Etudes des paramètres amphiphiliques de la série phtalazine	47
III.2.2.1. Volume et Surface Moléculaire.....	47
III.2.2.2. Polarisabilité	47
III.2.2.3. Energie d'hydratation	48
III.2.2.4. Réfractivité Molaire.....	48
III.2.2.5. Hydrophobicité	49
III.2.2.6. Lipophilicité	49
III.2.2.7. Coefficient de partage (P).....	49
III.3.Représentation des calculs « drug-like » sur la base de Lipinski	51
III.4.Les règles de Veber	53
III.5. Efficacité de ligand « LE ».....	55
III.6.Efficacité lipophile de ligand (LLE).....	57
III.7.Etude quantitative de la relation structure-activité	59
III.7.1. Descripteurs quantique/électronique	59
III.7.1.1. Energies des orbitales HOMO et LUMO	59
III.7.1.2. Moment dipolaire	60
III.7.1.3. Energie totale	61
III.7.2. Le modèle QSAR	62
III.7.2.1. Validation de modèle.....	64
III.7.2.2. Validation externe.....	65
III.8.Conclusion	68
Référence.....	70

Sommaire

SOMMAIRE

Chapitre IV : Etude par Doking Moléculaire des Interactions entre l'enzyme 1M17 et des Dérivés de Phtalazine

IV.1.Introduction	73
IV.2. Les étapes de Docking moléculaire	74
IV.2.1. Structure de la macromolécule (l'enzyme)	74
IV.2.2. Structure des ligands.....	74
IV.3. Résultats de docking moléculaire.....	76
IV.3.1. Détection des cavités	76
IV.3.2. Énergie d'interaction	77
IV.4.Conclusion	82
Référence.....	83
CONCLUSION GENERALE.....	84

LISTE DES TABLEAUX

LISTE DES TABLEAUX

Chapitre III :

Tableau III.1 : Structures chimiques des dérivés de phtalazine	44
Tableau III.2 : Paramètres physicochimiques des dérivés de phtalazine	49
Tableau III.3 : Les paramètres de Lipinski des dérivés de phtalazine.....	52
Tableau III.4 : Les règles de Veber appliquée aux phtalazine étudiés.....	54
Tableau III.5 : L'efficacité de ligand de la série phtalazine.....	56
Tableau III.6 : L'efficacité lipophile de ligand de la série phtalazine.....	58
Tableau III.7 : Les valeurs des descripteurs moléculaires utilisés dans l'analyse de régression.....	61
Tableau III.8 : Matrice de la corrélation du modèle.....	63
Tableau III.9 : Paramètres de validation croisée.....	64
Tableau III.10 : Les Valeurs observées, prédites et les résidus ordinaires de pIC ₅₀ par MLR pour l'ensemble de validation.....	65

Chapitre IV :

Tableau IV.1 : Volume et surface des cavités détectées.....	76
Tableau IV.2 : Résultats de docking moléculaire des dérivés de phtalazine avec la cavité 1 de l'enzyme 1M17.....	77
Tableau IV.3 : Meilleur complexes formés par L ₂ , L ₂₄ , L ₂₅	78
Tableau IV.4 : Distance (stérique et Hydrogènes) mesuré en Å entre les résidus du site actif et les atomes de ligand L ₂	80
Tableau IV.5 : Distance (stérique et Hydrogènes) mesuré en Å entre les résidus du site actif et les atomes de ligand L ₂₄	80
Tableau IV.6 : Distance (stérique et Hydrogènes) mesuré en Å entre les résidus du site actif et les atomes de ligand L ₂₅	81

Liste des figures

Chapitre I :

Figure I.1 : Différence entre cellule normale et cellule cancéreuse.....	6
Figure I.2 : Structure de la peau humaine : l'hypoderme, le derme et l'épiderme.....	8
Figure I.3 : Cancer de la peau (mélanome).....	10
Figure I.4 : Schématisation de la structure d'un récepteur HER.....	11
Figure I.5 : Voies d'activation de l'EGFR et niveaux d'action possibles.....	12
Figure I.6 : Exemples d'interactions ligand-récepteur à la surface de cellules.....	14
Figure I.7 : Noyau de base phtalazine.....	15
Figure I.8 : Synthèse pour la préparation de la phtalazine.....	15

Chapitre II :

Figure II.1 : Principe de la méthode QSPR.....	28
Figure II.2 : Représentation schématique du docking de deux molécules. (a)molécules de départ. (b) complexe moléculaire.....	36
Figure II.3 : Principe générale d'un programme de docking.....	37

Chapitre III :

Figure1 : Structure générale de phtalazine.....	44
FigureIII.2 : Les orbitales frontières HOMO/LUMO.....	60
Figure III.3 : Moment dipolaire entre deux atomes.....	61
Figure III.4 : La courbe des valeurs prédites en fonction des valeurs expérimentales de log (I/IC ₅₀).....	67
Figure III.5 : La courbe des valeurs résiduelles par rapport à l'observé expérimentalement.....	68

Chapitre IV :

Figure IV.1 : Structure de récepteur du facteur de croissance épidermique tyrosine kinase 1M17.....	74
--	----

Liste des figures

Figure IV.2 : Protocole de calcul.....	75
Figure IV.3 : Interface graphique de la cavité 1 indiquée par Mol Dock	77
Figure IV.4 : Les Interaction hydrogène et stérique entre les ligands et l'enzyme 1M17.....	79

Liste des abréviations :

ONU	Organisation des Nations unies.
EGF	Epidermal Growth Factor.
kDa	Kilodalton.
TGF-α	Transforming Growth Factor α .
HER	Human Epidermal Growth Factor Receptor.
MAP	Mitogen Activated Proteins.
RAF	Rapidly Accelerated Fibrosarcoma.
Pi3K	phosphatidylinositol-3-kinase.
PTEN	Phosphatase and tensin homolog.
mTOR	Mechanistic target of rapamycin.
EGFR	Epidermal Growth Factor Receptor.
UVA, UVB	UltraViolet A, UltraViolet B.
ADN	Acide désoxyribonucléique.
RTK	Receptor tyrosine kinases.
ATP	Adénosine-triphosphorique.
VEGFR-2	vascular endothelial growth factor receptor.
IC50	Concentration inhibitrice 50.
nM	Nanomètre.
hCA	Heterocyclic amine.
μM	Micromètre.
Ψ	Fonction d'onde de la molécule.
H	Hamiltonien totale.
E_{TOT}	Energie totale.
DFT	Théorie de la fonctionnelle de la densité.

Liste des abréviations :

LDA	Local Density Approximation.
GGA	Generalized Gradient Approximation.
B3LYP	Becke 3-paramètres Lee-Yang-Parr.
CNDO/2	Complete Neglect of Differential Overlap.
INDO	Intermediate Neglect of Differential Overlap.
NDDO	Neglec of Diatomic Differential Overlap.
MNDO	Modified Neglected of Differential Overlap.
AM1	Austin Model 1.
PM3	Parametric Method 3.
HOMO	Highest Occupied Molecular Orbital.
LUMO	Lowest Unoccupied Molecular Orbital.
QSPR	Quantitative Structure-Property Relationships.
QSAR	Quantitative Structure-Activity Relationships.
SAR	Relation structure-activité.
MLR	Régressions multi-linéaires.
PLS	Partiel least squares regression.
2D, 3D	Deux dimension, Trois dimension.
CSPA	Charged partial surface area.
LOO, LMO	Leave-one out, Leave-many out.
PRESS	prediction sum of squares.
RMN	Résonance magnétique nucléaire.
Å°	Ångström (unité de longueur).
UMA	unité de masse atomique.
d	Densité.
Kcal	Kilocalorie.

Liste des abréviations :

LF	Liaisons retables.
HBD	Liaisons hydrogène donneurs.
HBA	Liaisons hydrogène accepteurs.
PSA	Surface polaire atomique.
%ABS	pourcentage d'absorption.
LE	Efficacité de ligand.
LiPE	Efficacité lipophile.
NH	Nombre des atomes lourds.
OM	Les orbitals Moléculaires.
eV	Electron volt.
ADME	Absorption, Distribution, Métabolisme et Excrétion.
PDB	Protein Data Bank.
MVD	Molegro Vertual Docker.

Introduction Générale



La découverte des médicaments est un processus long et complexe [1]. Tant au niveau du temps que de l'argent investi [2]. Le domaine du « drug design » ou « conception de médicaments » peut être exploré par la modélisation moléculaire et la conception de nouvelles molécules à l'aide des outils informatiques [3]. Il se résume essentiellement à identifier de nouveaux composés (molécules naturelles ou synthétiques) qui vont idéalement évoluer en des médicaments agissant sur des cibles biologiques spécifiques responsable de dysfonctionnements. L'identification des cibles thérapeutiques est liée à la connaissance du fonctionnement moléculaire, des voies métaboliques, des systèmes biologiques en général et à la cause des maladies [4].

La modélisation offre également de nombreuses opportunités en biologie expérimentale. Ainsi, la modélisation moléculaire, permet d'obtenir par exemple la structure de l'état de transition d'une réaction chimique donnée, ce qui est difficile, voire impossible, pour la chimie expérimentale. Le nombre d'études théoriques utilisant ces techniques, pour guider ou expliciter des observations expérimentales [5].

Les relations entre les structures des molécules et leurs propriétés ou activités sont généralement établies à l'aide de méthodes de modélisation par apprentissage statistique. Les techniques usuelles reposent sur la caractérisation des molécules par un ensemble de descripteurs, nombres réels mesurés ou calculés à partir des structures moléculaires. Il est alors possible d'établir une relation entre ces descripteurs et la grandeur modélisée [6]. La prédiction de l'activité biologique est basée sur des calculs mathématiques et statistiques, c'est pour cette raison on met le choix d'une méthode statistique, appelée la régression multilinéaire (RML).

Les interactions entre molécules sont à la base de la plupart des mécanismes biologiques. Etudier les détails d'interactions, à l'échelle moléculaire, entre un ligand et la cible thérapeutique s'avère donc d'un très grand intérêt. Actuellement, les approches de criblage virtuel basé sur la structure de la cible, et plus précisément le docking moléculaire, permettent de prédire les modes d'interactions possibles et fournissent un moyen d'étudier les interactions au niveau moléculaire et sont, de ce fait, une indication de l'activité biologique de nouvelles molécules ne tenant compte que des critères structuraux.[7]

Le docking moléculaire est devenu, un outil indispensable dans le domaine de la conception de molécules médicamenteuses qui pourraient agir le plus favorablement, ce qui est considérablement plus facile à mettre en œuvre que les méthodes classiques.[8]

L'objectif de ce travail est l'application de différentes méthodes de la modélisation moléculaire pour étudier des relations structure-activité sur les dérivés de phtalazine et de développer un modèle QSAR pour ces composés par rapport à leur activité anticancéreuse, et criblage virtuel par docking moléculaire.

Le présent travail comporte quatre chapitres contenant :

Le premier chapitre est scindé en deux parties : Dans la première partie, nous allons présenter des généralités sur le cancer de la peau. Dans la deuxième partie, nous allons parler sur le noyau de base phtalazine et leur activité biologique.

Dans le deuxième chapitre, nous allons décrire le formalisme de la méthodologie choisie dans la modélisation moléculaire qui comporte les différentes méthodes de calcul utilisées et engagées dans notre travail.

Le troisième chapitre comporte une étude qualitative et quantitative des propriétés QSAR, et décrire la relation entre les propriétés physicochimiques, quantiques et l'activité biologique de cette série bioactives par l'utilisation d'un Modèle QSAR.

Dans le quatrième chapitre, Nous présenterons également une étude par docking moléculaire des différents dérivés de phtalazine.

Référence

- [1] Monge A., Thèse de doctorat, Création et utilisation de chimiothèques optimisées pour la recherche « in silico » de nouveaux composés bioactifs, Université d'Orleans , 2006.
- [2] Guillemain H., thèse de doctorat, Evaluation et application de méthodes de criblage insilico, Universite de Cnam, 2012.
- [3] Asses Y., Thèse de doctorat, Conception par modélisation et criblage in silico d'inhibiteurs du récepteur c-Met, Université Nancy, 2011.
- [4] Saadi A., A.Cowe, Modélisation moléculaire 3D et criblage virtuel, Séminaire interdisciplinaire de bioinformatique, 2007.
- [5] Archer S., Weir E., Wilkins M., Am. J. Physiol. Heart .Circ. Physiol, (2007), 294, 570
- [6] Aurélie G.S.A., Thèse de doctorat, Université de Paris 6, (2008).
- [7] Cheng A.C., Coleman R.G., Smyth K.T., Nat Biotechnol, 2007, 25, 71-75.
- [8] Yuriev E., Agostino M., Ramsland P.A., Mol J. Recognit, 2011, 24, 149-164.

Chapitre I

Généralités sur le Cancer de la Peau et les Composés Hétérocycliques (Phtalazine)



I.1.Le Cancer

I.1.1.Définition

Le cancer est une maladie caractérisée par la prolifération anarchique et incontrôlée de certaines cellules de l'organisme ayant échappé aux mécanismes normaux de contrôle de leur régulation et de leur différenciation. Ces cellules anormales dérivent toutes d'une cellule initiatrice, provenant d'un tissu normal de l'organisme, ayant subi une mutation et ayant acquis les caractéristiques nécessaires pour se multiplier indéfiniment (tumeur), envahir les tissus avoisinants et essaimer vers d'autres organes formant ainsi des métastases.

I.1.2. Mécanismes fondamentaux de la carcinogenèse

Le processus tumoral s'explique par un dérèglement du programme génétique de la cellule qui acquiert la capacité à se diviser de manière anarchique pour aboutir au phénomène d'hyperplasie. Puis, ces cellules anormales subissent une anomalie de développement conduisant au phénomène de dysplasie amenant une différenciation fonctionnelle et structurale. On parle alors de processus tumoral caractérisé par l'apparition de tumeurs :

I.1.2.1.Tumeurs bénignes

Les tumeurs bénignes ne sont pas cancéreuses, c'est à dire qu'elles n'envahissent pas les organes voisins et ne font que les repousser ; elles ont un développement généralement limité ; elles n'essaient pas leurs cellules ailleurs, ce qui signifie qu'elles ne font pas de métastases. Les tumeurs bénignes peuvent malgré tout poser des problèmes selon l'endroit où elles se situent (le tympan par exemple qu'elles peuvent détruire, ou les intestins qu'elles peuvent boucher).

I.1.2.2.Tumeur maligne

La tumeur maligne qui a une évolution rapide et se fait vers l'atteinte du reste de l'organisme. Le cancer est le nom donné à tout type de cette tumeur.

Le cancer est dû à une anomalie de fonctionnement des cellules qui entraîne un développement anarchique de celles-ci. Il se développe d'abord dans un organe et forme une ou des tumeurs, mais il arrive que les cellules cancéreuses atteignent d'autres parties du corps. Ce sont les métastases. On parle alors de cancer primitif et de cancer secondaire. [3]

Le mécanisme de la carcinogenèse se caractérise par l'implication de trois types de gènes : les oncogènes (ou proto-oncogènes), les gènes suppresseurs de tumeurs (ou anti-oncogènes) et les gènes de maintien de l'intégrité. Une centaine d'oncogènes, tels que le gène ras, et une dizaine de gènes suppresseurs de tumeurs, tels que le gène p53, sont répertoriés et ceci implique la multiplicité des types de cancer :

- **les oncogènes** : ces gènes augmentent l'activité proliférative. Ils codent pour des protéines impliquées dans le processus de prolifération cellulaire (facteurs de croissance, récepteurs membranaires,...). Lors du processus de carcinogenèse, ces oncogènes subissent des modifications génétiques (mutation, amplification ou translocation chromosomique) ce qui conduit à leur hyper activation et de ce fait à une stimulation ininterrompue de la division cellulaire.

- **les gènes suppresseurs de tumeurs** : ce sont des régulateurs négatifs de la multiplication cellulaire. Certains anti-oncogènes, comme le p53, favorise l'apoptose (la mort cellulaire programmée) par l'arrêt du cycle cellulaire. Lors du processus tumoral, ces gènes perdent leur fonction (par altération ou mutation ponctuelle chromosomique) et favorisent ainsi la transformation et la prolifération cellulaire.

- **les gènes de réparation** : qui codent pour un complexe multifonctionnel capable de réparer des anomalies de l'ADN. En cas d'échec, la cellule lésée est amenée à une mort cellulaire programmée. Lors du processus cancéreux, ces gènes sont moins exprimés et/ou inactivés ce qui conduit à une prolifération des cellules anormales.

I.1.3. Les étapes de la carcinogenèse :

Un cancer met habituellement plusieurs années à se former, chez les adultes du moins. On distingue 3 étapes :

I.1.3.1. Initiation

Les gènes d'une cellule sont endommagés ; cela se passe fréquemment. Par exemple, les substances cancérigènes présentes dans la fumée de cigarette peuvent causer de tels dommages.

On parle alors d'apoptose ou de «suicide» cellulaire. Mais lorsque la réparation ou la destruction de la cellule ne se fait pas, la cellule reste endommagée et passe à l'étape suivante.

Remarque 1

La plupart du temps, la cellule répare l'erreur automatiquement. Si l'erreur est irréparable, la cellule meurt.

I.1.3.2. Promotion

Des facteurs extérieurs vont stimuler ou non la formation d'une cellule cancéreuse. Il peut s'agir des habitudes de vie, comme le tabagisme, le manque d'activité physique, une mauvaise alimentation, etc.

I.1.3.3. Progression

Les cellules prolifèrent et la tumeur se forme. Dans certains cas, elles peuvent envahir d'autres parties du corps.

Remarque 2

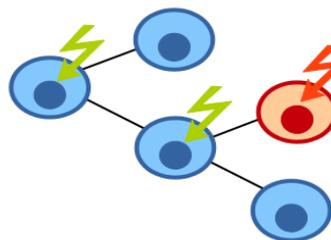
Dans sa phase de croissance, la tumeur commence à provoquer des symptômes : des saignements, de la fatigue, etc. [4]

Le potentiel de prolifération des cellules cancéreuses est illimitées ; à l'inverse des cellules saines, les cellules tumorales sont potentiellement immortelles et ont la particularité d'échapper aux défenses immunitaires. De ce fait, leur propagation dans l'organisme est très rapide. Pour permettre l'accroissement de ces cellules anormales, et de ce fait la généralisation du cancer par leur envahissement progressif des tissus adjacents puis des organes à distance, la vascularisation de la tumeur est essentielle. Ce processus invasif s'effectue par le mécanisme de l'angiogenèse assurant la survie de la tumeur par son alimentation en oxygène et en nutriments. L'intervention de facteurs dits angiogéniques contraignent alors les tissus avoisinants à former de nouveaux vaisseaux sanguins pour approvisionner la tumeur. Une fois en surnombre, ces cellules tumorales sont en mesure de quitter leur tissu d'origine et de coloniser par voie sanguine ou lymphatique d'autres tissus pour former les métastases.

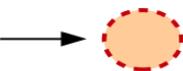
La thématique du laboratoire E.A.2692 est orientée vers l'inhibition de la croissance cellulaire Cancéreuse prostatique. Le but de notre travail consiste ainsi à développer de nouveaux agents cytostatiques permettant de limiter ce processus tumoral en visant différentes cibles pharmacologiques. [5]

Cellule normale

Sous l'action d'un **facteur de croissance**, une cellule mère se divise en deux cellules filles (*mitose*)



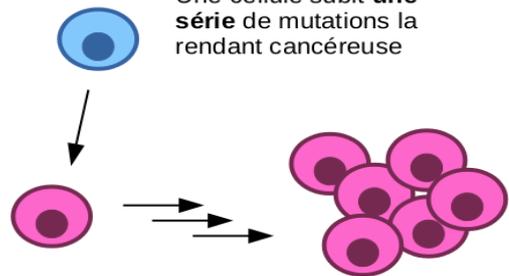
Si une division s'est mal passée, la cellule reçoit un **signal de mort** et se suicide (*apoptose*)



En l'absence de signal, la cellule ne se divise pas

Cellule cancéreuse

Une cellule subit une **série de mutations** la rendant cancéreuse



Les cellules prolifèrent sans contrôle et forment une tumeur

Figure I.1 : Différence entre cellule normale et cellule cancéreuse.

I.2. La peau humaine

I.2.1. Structure et fonction de la peau

La peau constitue l'enveloppe corporelle qui assure l'interface physique entre les milieux extérieur et intérieur. Elle joue le rôle de barrière de protection contre les agressions externes multiples et évite la déshydratation de l'organisme. Elle est le siège de nombreux processus métaboliques et assure des fonctions sensorielles, immunitaires et thermo régulatrices. Cet organe vital dont l'épaisseur varie de 1,5 à 4 mm chez l'homme est constitué de trois compartiments distincts d'origine embryologique différente : **l'épiderme** (d'origine ectodermique), **le derme** et **l'hypoderme** (d'origine mésodermique) (Figure I.2). [6]

I.2.2. Les différentes couches de la peau

La peau est constituée de trois couches : l'épiderme, le derme et l'hypoderme :

-**L'épiderme** est l'enveloppe externe de la peau. Il contient différents types de cellules, dont les mélanocytes. Il est lui-même composé de quatre étages : la couche cornée, la couche granuleuse, la couche épineuse et la couche basale. Dépourvu de vaisseaux sanguins, l'épiderme doit compter sur le derme pour se nourrir.

-**Le derme**, couche intermédiaire, se situe sous l'épiderme. Principal tissu de soutien de la peau, il est responsable de sa solidité. Il est composé principalement de fibroblastes, cellules responsables de la production de fibres de collagène et de fibres élastiques, à l'origine de la souplesse et de l'élasticité de la peau.

-Le derme contient également des vaisseaux sanguins, des vaisseaux lymphatiques et des terminaisons nerveuses. Il est également le siège des follicules pileux (qui forment la base du poil), des glandes sudoripares (qui produisent la sueur) et sébacées (qui produisent le sébum). Tous ces constituants assurent l'hydratation et la nutrition de la peau et participent à la protection du corps contre les agressions.

-**L'hypoderme**, couche la plus profonde, est situé sous le derme. Il est constitué de cellules grasses, les adipocytes. Cette couche de graisse stocke l'énergie, assure l'isolation pour conserver la chaleur du corps et protège les organes recouverts. [7]

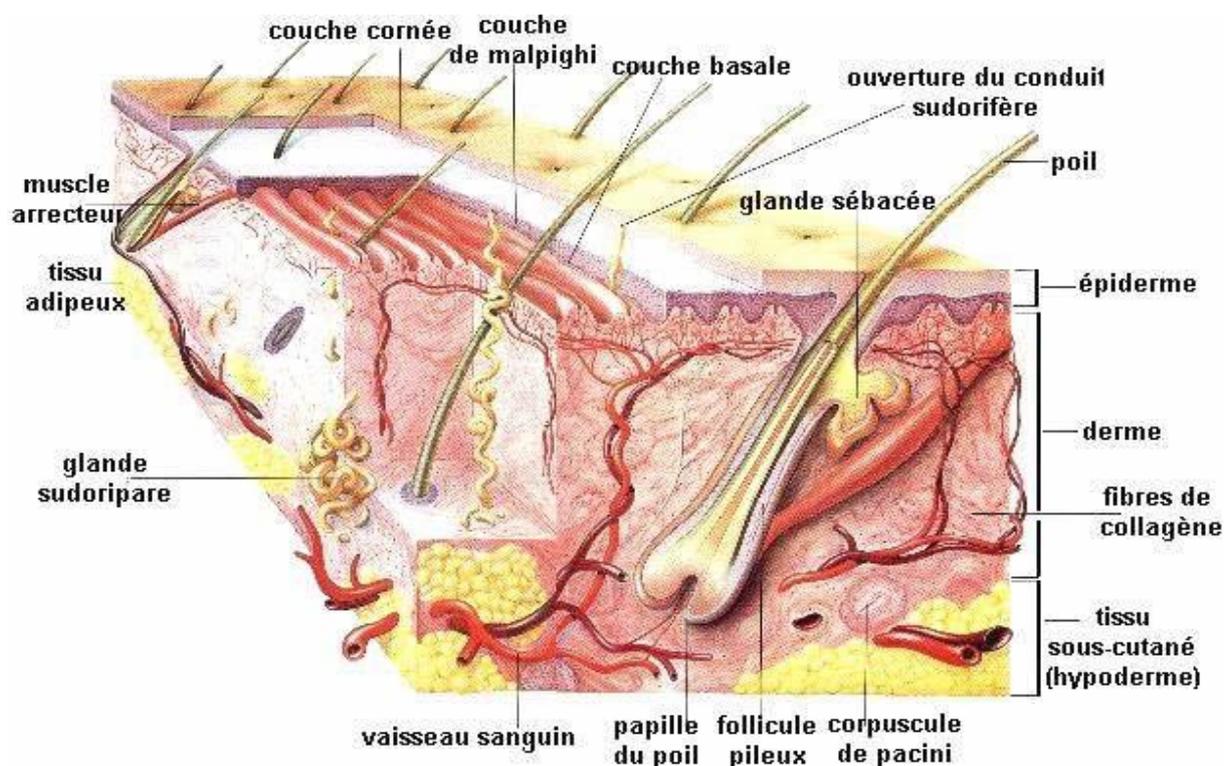


Figure I.2 : Structure de la peau humaine : l'hypoderme, le derme et l'épiderme.

I.2.3. Les mélanocytes : des cellules de l'épiderme

Les mélanocytes, cellules à l'origine du développement du mélanome de la peau, se situent dans l'épiderme. Ces cellules sont responsables de la production de mélanine, un pigment qui donne sa couleur à la peau. Elles siègent au niveau de la couche basale de l'épiderme. Lorsque la peau est exposée aux rayons ultraviolets, le nombre de mélanocytes augmente et la mélanine est plus abondante. Les mélanocytes produisent parfois une plus grande quantité de mélanine à certains endroits. C'est ce qui explique les taches de rousseur et certaines taches de vieillesse.

L'épiderme comporte aussi trois autres types de cellules :

-**les kératinocytes**, cellules les plus nombreuses. Elles contiennent une protéine, la kératine, qui aide à protéger la peau contre les agressions extérieures. La kératine, qui se dépose progressivement en superficie, est le constituant principal de la couche cornée.

-**les cellules de Langerhans**, spécialisées dans l'activation du système immunitaire. Elles font partie des cellules dendritiques qui participent à la défense de l'organisme vis-à-vis des agressions extérieures, en permettant aux lymphocytes de détecter les corps étrangers comme les bactéries et les virus.

-**les cellules de Merkel** qui jouent un rôle dans la sensibilité de la peau au toucher grâce à des terminaisons nerveuses. Elles permettent d'identifier la forme et la matière des objets au contact de notre peau. [8]

I.3. Cancer de la peau

I.3.1. Définition

Ce type de cancer apparaît au départ des mélanocytes ou cellules pigmentaires. Celles-ci sont présentes sous forme disséminée dans toute la peau. Elles peuvent également être réunies en petits amas et forment alors des taches pigmentées, comme les grains de beauté (nævus). On parle de mélanome lorsque, à un endroit de la peau, les cellules pigmentaires se transforment en cellules cancéreuses. La tumeur maligne qui en résulte va se présenter comme une tache pigmentée qui grandit, grossit, ou change de couleur...

Un mélanome se développe parfois au départ d'une tache pigmentée (nævus) mais le plus souvent au niveau d'une zone de peau "saine", sans tache préexistante. Les mélanomes peuvent se développer à n'importe quel endroit de la peau. Chez les femmes, ils apparaissent un peu plus fréquemment au niveau des jambes, chez les hommes au niveau du tronc.

Les zones recouvertes de poils, comme le cuir chevelu, peuvent également être atteintes, de même que la peau sous les ongles. Exceptionnellement, un mélanome peut apparaître sur une muqueuse (bouche, cavités nasales, gorge, vagin, anus...), voire à l'intérieur des paupières ou dans l'œil proprement dit.

I.3.2. Les différents types de cancer de la peau

Outre les mélanomes qui représentent 10 % des cancers de la peau, il en existe d'autres types. Les plus fréquents sont les carcinomes : ils représentent 90 % de l'ensemble des cancers cutanés. Ils surviennent généralement après l'âge de 50ans, sur les zones découvertes du corps. Ils sont le plus souvent dus à une exposition au soleil excessive et répétée. Contrairement au mélanome de la peau, le point de départ de ces cancers cutanés n'est pas les mélanocytes, mais d'autres cellules, les kératinocytes, également situées dans l'épiderme. On en distingue deux types :

I.3.2.1. Les carcinomes baso-cellulaires

Ce sont les plus fréquents et les moins graves. Ils évoluent lentement et n'entraînent pas de métastases à distance. Cependant, ils peuvent s'étendre localement et entraîner une destruction des tissus sous-cutanés (c'est-à-dire situés sous la peau). Il suffit, le plus souvent, de les retirer chirurgicalement pour en assurer la guérison.

I.3.2.2. Les carcinomes épidermoïdes (anciennement appelés carcinomes spinocellulaires)

Ils sont plus rares et peuvent entraîner des métastases. Ils sont toutefois facilement guérissables dans la plupart des cas, grâce à une détection précoce permettant un traitement chirurgical.

Les traitements proposés étant spécifiques à chaque type de cancer de la peau, il est important de connaître le nom précis de la maladie pour accéder à une information adaptée à votre situation. [9]

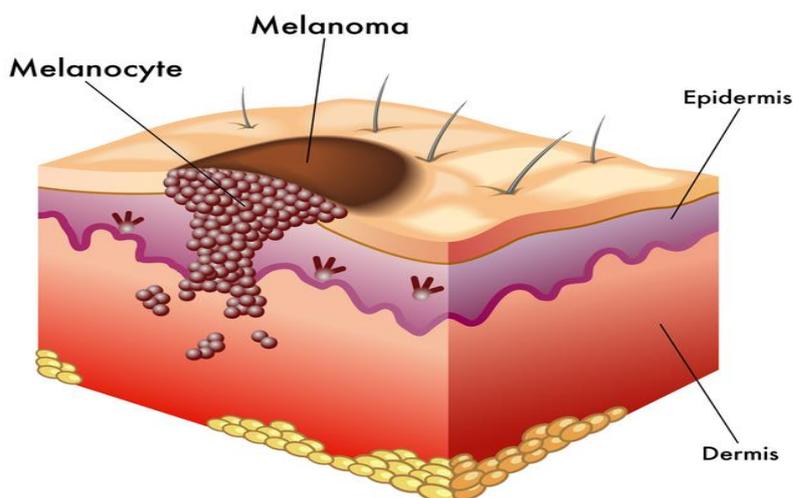


Figure I.3 : Cancer de la peau (mélanome).

I.4. EGFR (Epidermal Growth Factor Receptor)

Le premier membre de la famille des récepteurs à l'EGF (Epidermal Growth Factor) est l'EGFR, une glycoprotéine de 170 kDa qui correspond au récepteur original de l'EGF. Il est surexprimé à la surface des cellules tumorales de nombreux cancers, tels que le cancer de la prostate, du cou, du rein et du sein. [10] Dans les cellules saines, on peut compter entre 40 et 100000 récepteurs par cellule alors que dans les cellules tumorales, ce nombre peut atteindre 2 millions. Cette protéine, formée de 1186 acides aminés, possède une partie extracellulaire constituée de 4 domaines I-IV (622 résidus) permettant la fixation de nombreux ligands notamment l'EGF et le TGF- α (Transforming Growth Factor α). Ce récepteur représente le principal médiateur de l'activité tumorale en induisant la prolifération, le blocage de l'apoptose et l'effet angiogénique [11].

I.4.1. Structure de l'EGFR

L'EGFR, comme les autres récepteurs de sa famille, est composé de (figure 4) :

- d'un domaine extracellulaire, auquel peuvent se fixer différents ligands (EGF, TG Fa, amphiréguline, épiréguline.) .
- d'un domaine transmembranaire d'ancrage.

– d'un domaine intracellulaire qui possède une activité tyrosine-kinase, c'est-à-dire capable de « Phosphoryler » ses cibles sur des résidus tyrosine.[12]

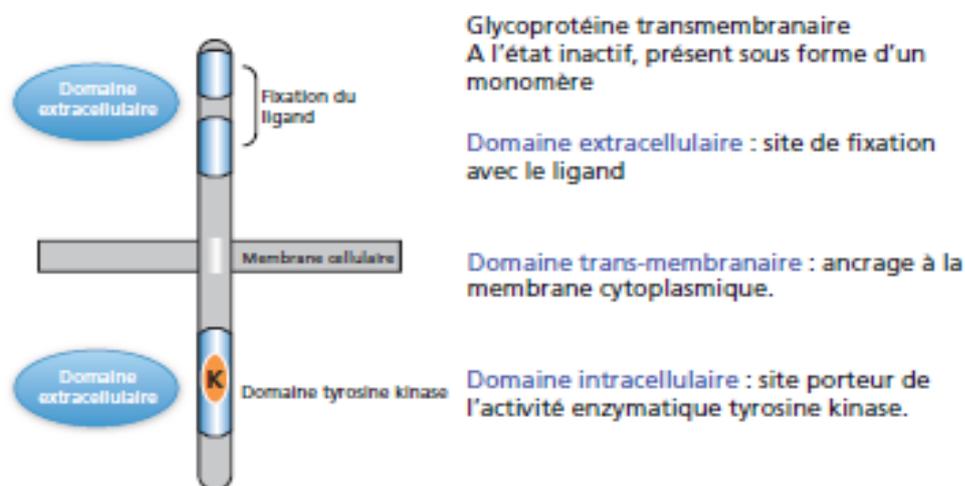


Figure I.4 : Schématisation de la structure d'un récepteur HER.

I.4.2. L'activation de l'EGFR

L'activation de l'EGFR peut se schématiser en plusieurs étapes. La fixation du ligand sur la partie extracellulaire du récepteur entraîne une dimérisation du récepteur, puis l'autophosphorylation de son domaine intracellulaire, le rendant à son tour capable de phosphoryler ses cibles moléculaires intracellulaires. L'activation du domaine intracellulaire du récepteur aboutit à la phosphorylation en cascade de médiateurs intracellulaires, qui jouent le rôle « d'interrupteurs » moléculaires. Ces voies de signalisation aboutissent, enfin, à la libération de facteurs de transcription qui régulent l'expression de nombreux gènes-cibles impliqués notamment dans la prolifération, la différenciation et la survie cellulaires.

Plusieurs voies d'aval, activées par les récepteurs de la famille HER, sont décrites (figure 5) :

- la voie des MAP-kinases (pour Mitogen Activated Proteins), qui fait intervenir successivement RAS (dont KRAS est une isoforme), RAF (dont BRAF est une isoforme), MEK, ERK. Elle aboutit à la libération de facteurs transcriptionnels tels que Jun, Fos, Myc.
- la voie Pi3K/AKT (Pi3K = phosphatidylinositol-3-kinase), impliquée dans la résistance à l'apoptose et la prolifération cellulaire. Cette voie est négativement régulée par PTEN, par conséquent considéré comme un gène suppresseur de tumeur. AKT possède plusieurs cibles d'aval, qui peuvent être activées (notamment par mTOR) ou inhibées.
- la voie protéine kinase C/Jak/STAT.[13]

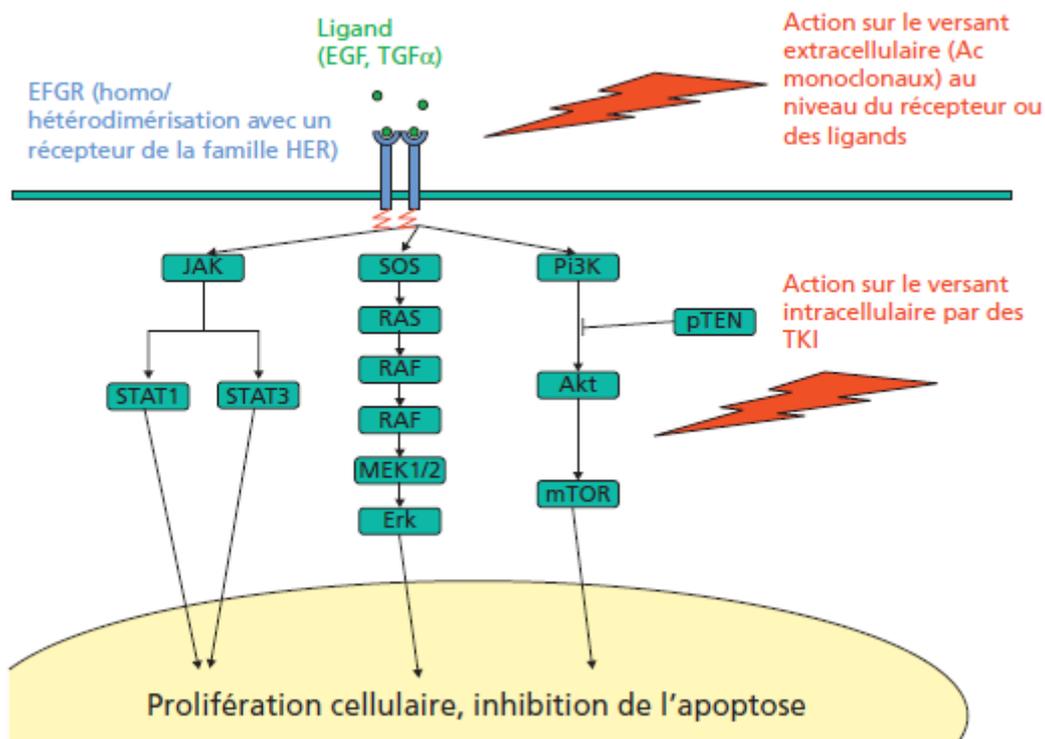


Figure I.5 : Voies d'activation de l'EGFR et niveaux d'action possibles.

I.4.3. L'EGFR et le cancer

Le cancer étant défini comme une hyper prolifération cellulaire incontrôlée et un défaut au niveau de la voie de différenciation, l'intérêt des chercheurs s'est focalisé sur le rôle de l'EGFR dans cette prolifération anarchique de cellules.

La formation des tumeurs de la peau est divisée en trois étapes : l'initiation, la promotion et la progression [14] toutes ces étapes peuvent être induites par les UVA et les UVB, causant des dommages à l'ADN liés à des stress oxydatifs ainsi qu'une augmentation de la transcription de gènes par l'activation de la voie de MAPK [15]. EGFR ne semble pas essentiel pour l'initiation de la tumeur, mais il facilite son développement. Des études supplémentaires montrent qu'EGFR est nécessaire pour maintenir la population proliférative dans le compartiment de cellules basales. La déficience en EGFR induit la migration des cellules prolifératives dans les couches supra basales et leur différenciation, associée à un arrêt du cycle cellulaire.

I.4.4. Les modifications de l'EGFR impliquées dans le cancer

EGFR a été le premier récepteur tyrosine kinase à avoir été mis en cause dans des tumeurs humaines. EGFR est impliqué dans le développement de nombreux cancers humains [16]. Il a été montré que les patients avec des altérations au niveau des récepteurs ont les maladies les plus agressives.

Dans les cellules cancéreuses, l'EGFR est constitutivement actif par plusieurs mécanismes, qui régulent normalement l'activité du récepteur (augmentation de l'expression des ligands, voie en aval de l'EGFR déficiente ou des interactions avec les Récepteurs Couplés aux Protéines G, les RTK ou des molécules d'adhésion).

L'EGFR est donc très important et est malheureusement répandu dans l'ensemble des cancers épithéliaux d'où la nécessité de comprendre comment il est modifié dans ces cancers et comment ces modifications peuvent induire le développement de cancer. Des études sur ce sujet sont en cours mais aucune ne permet de comprendre parfaitement les mécanismes mis en jeu.

I.4.5. Le traitement anti-EGFR

EGFR peut être considéré comme un facteur de survie dans les cellules tumorales, rendant cette voie de signalisation une cible intéressante pour les interventions thérapeutiques [17]. De nombreuses équipes cherchent à développer des thérapies ciblant ce récepteur afin d'éradiquer les tumeurs mais les résultats ne sont pas encore satisfaisants. Deux classes principales de thérapie anti ErbB sont utilisées : les anticorps bloquant la conformation stabilisée du domaine extracellulaire, induite par la liaison du ligand au récepteur et les petites molécules inhibitrices de tyrosine kinase, qui entrent en compétition avec l'ATP pour le domaine tyrosine kinase intracellulaire.

I.4.6. Conséquences des interactions ligands-récepteurs

L'interaction ligand-récepteur génère une réponse cellulaire, c'est-à-dire des modifications du comportement cellulaire, qui sont de trois ordres :

- a) Changements de perméabilité membranaire vis-à-vis des ions et de l'eau.
- b) Modifications des activités enzymatiques à la surface et à l'intérieur de la cellule.
- c) Modifications des activités transcriptionnelles. [18]

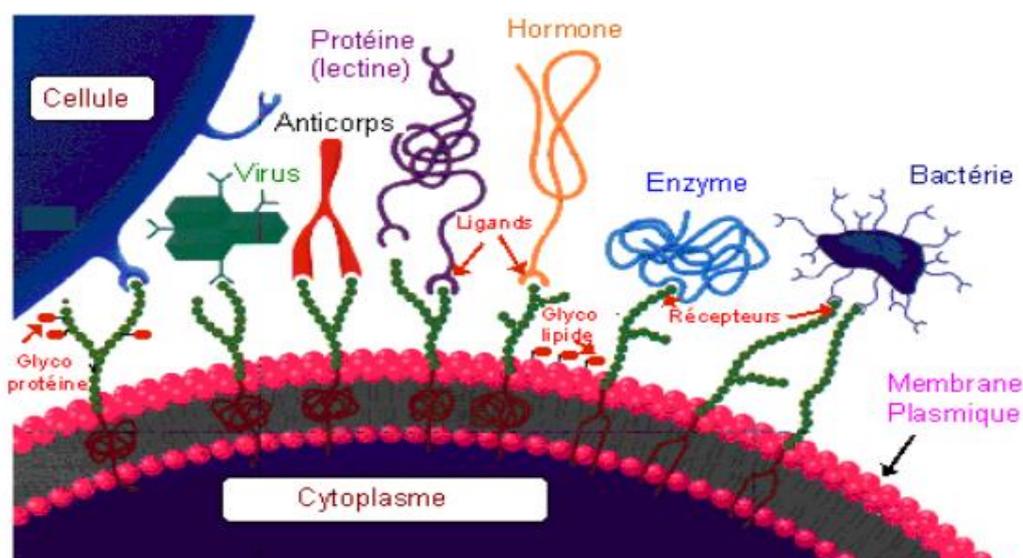


Figure I.6 : Exemples d'interactions ligand-récepteur à la surface de cellules.

I.5. Les Hétérocycles

En chimie organique, la classification des molécules est basée sur le nombre et la diversité des atomes qui les composent, mais aussi sur les différents types de liaisons qui constituent leur structure.

Dans le cas où les atomes forment une chaîne, les composés correspondants sont dits **acycliques**. Au contraire, si l'enchaînement des atomes forme un cycle, on dira qu'il s'agit de composés **cycliques**. Si le cycle est entièrement composé d'atomes de carbone, il s'agit de **carbocycle**.

D'une manière générale, si le cycle est entièrement constitué d'un seul type d'atome, carbone ou autre, on ajoute le préfixe **iso**, **composés iso cycliques**. Un cycle qui est constitué d'au moins deux types d'atomes est un **hétérocycle** [19].

I.5.1. Définition des hétérocycles à six chaînons

La déficience électronique des hétérocycles aromatiques à six membres s'explique par la présence d'un hétéroatome plus électronégatif que le carbone. Le cycle peut donc subir une attaque d'un nucléophile et la charge négative pourra être mieux accommodée par l'hétéroatome que par un carbone, d'où la réactivité de ces cycles comme électrophiles. Comme la paire d'électron sur l'azote dans les aromatiques pauvres en électrons n'est pas impliquée dans l'aromaticité, il est possible d'activer davantage le cycle par N-alkylation [20].

I.5.2. Classification

- par la nature de l'hétéroatome : les hétérocycles oxygénés, soufrés et azotés sont les plus importants.

- par le nombre de chaînons : les hétérocycles à 5 et 6 chaînons sont les plus stables et les plus fréquents.

- par le degré d'insaturation (analogie avec les composés carbocycliques): hétérocycles saturés, hétérocycles éthyléniques (partiellement saturés), hétérocycles aromatiques (insaturés) [21].

I.6. phtalazine

I.6.1. Structure et nomenclature de phtalazine

La phtalazine connue comme benzo-orthodiazine ou même benzopyridazine, est un composé organique hétérocyclique de formule chimique brute $C_8H_6N_2$. Il est isomérique avec la quinoxaline, la cinnoline et la quinazoline.

La phtalazine possède des propriétés de base et des produits d'addition avec les formes les iodures d'alkyle.

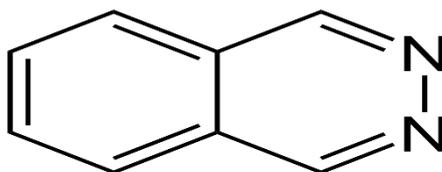


Figure I.7 : Noyau de base phtalazine.

I.6.2. Préparation de phtalazine

Etant donné que la phtalazine est raisonnablement peu coûteuse à l'achat, il y a eu peu d'incitation à développer des procédures de synthèse nouvelles ou améliorées. Cependant, la pratique d'une synthèse primaire a été rapportée (voir figure 6) ; 1,4-dihydrazinophtalazine(1) a été oxydé en phtalazine (2) ($NaOH$, $EtOH$, H_2O , O_2 , 20 C° , 3-4 h $\approx 70\%$), et la phtalazine peuvent être récupérés à partir de son 2-méthiodure par traitement avec du chlorhydrate de pyridine pur à reflux. [22]

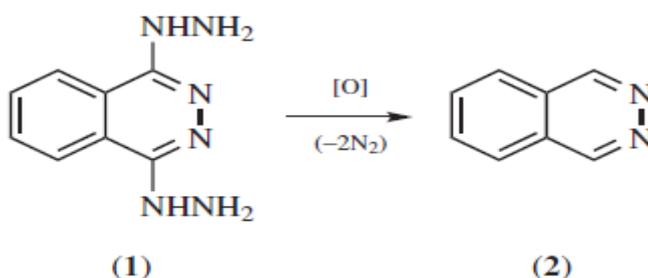


Figure I.8 : Synthèse pour la préparation de la phtalazine.

I.6.3. L'activité biologique de phtalazine

La phtalazine est un noyau bioactif intéressant, de nombreux composés basés sur le noyau phtalazine ont été signalés possèdent une activité anticancéreuse significative, parmi lesquels PTK787 et AAC789 qui sont des inhibiteurs prometteurs de VEGFR-2. Pour poursuivre notre étude précédente, de nouveaux composés ont été conçus et synthétisés en utilisant le noyau phtalazine comme noyau en incorporant différentes fractions 1, 3 diphénylallylidène liées à la position 1 du cycle phtalazine par l'intermédiaire d'un pont hydrazine. Visant à développer de nouveaux agents cytotoxiques. Lorsque la 1-hydrazinyphtalazine 4 a été mise en réaction avec plusieurs chalcones 9a-f préalablement préparées dans de l'éthanol absolu et une quantité catalytique d'acide acétique glacial, cette nouvelle série de composés 3a-f a été testée pour son activité anticancéreuse contre 14 lignées cellulaires cancéreuses, faisant apparaître une activité cytotoxique avec IC₅₀ dans la gamme nanomolaire. D'autres recherches ont été menées en étudiant leur activité inhibitrice enzymatique contre les kinases VEGFR-2 et EGFR, les résultats ont révélé une inhibition plus puissante de VEGFR-2 comparable à l'EGFR suggérant que cette voie soit le mécanisme principal de l'activité anticancéreuse de ces composés. Les dérivés les plus actifs étaient 10a, 10d et 10f, avec IC₅₀ = 0,42, 0,55 et 0,41 nM, respectivement contre VEGFR-2. [23]

Les dérivés de phtalazine sont des composés à potentiel biologique ayant diverses activités biologiques. Dans ce revoir, discuter brièvement de l'effet des dérivés de phtalazine sur les fonctions cardiaques et rénales. Une séries d'analogues de 4- (4-bromophényl) phtalazine et de phtalazinone reliés par un espaceur 2-propanoldes résidus de pipérazine N-substitués ont été testés pour leur effet sur l'activité de blocage β-adrénergique. Plus les composés présentaient une activité β-adrénolytique appréciable par rapport au propranolol. Composés 1a, 1d, 1e et 2c ont montré une inhibition appréciable de la contraction de l'anneau aortique induite par la norépinéphrine. Une autre série de l'urée substituée par la phtalazine et les dérivés de thiourée (5a p) ont été testés pour leurs actions inhibitrices activité des enzymes humaines de l'anhydrase carbonique (hCAs I et II). Tous ces composés ont inhibé l'activité des isoenzymes. Le composé 5a (CI₅₀ = 6,40 μM pour hCA I et 6,13 μM pour hCA II) a le plus activité inhibitrice. Les dérivés de la phtalazine ont montré à la fois le blocage β-adrénergique et carbonique activités d'anhydrase et utiles pour les fonctions cardiaques et rénales. [24]

Référence

- [1] Guérin S., & al., La fréquence des cancers en 2006 et les évolutions de la mortalité depuis 1950 et de l'incidence depuis 1980. *Bull. Cancer*, 2009, 96, 51-57.
- [3] site internet : santé-medecine.journaldesfemmes.
- [4] site internet : tebyan.net/index.aspx?pid=138818
- [5] Brureau, L., Multigner, L., Wallois, A., Verhoest, G., Ndong, J. R., Fofana, M., & Blanchet, P. Cancer de la prostate en Guadeloupe: incidence, mortalité, caractéristiques cliniques et anatomopathologiques. *Bulletin du Cancer*, 2009, 96, 165-170.
- [6] Kanitakis J., Structure histologique de la peau humaine. Dans : *Biologie de la peau humaine*: Schmidt D. Les éditions INSERM, 1995, 1-20.
- [7] Blanpain C., Fuchs E., *Annu.Rev., CellDev.Biol.*, Blanpain C et Fuchs E, *Nat.Rev.Mol.CellBiol.*, 2009.
- [8] Guillot B., Charles, J., Jeudy, G., Cupissol, D., Dupuy, A., Dutriaux, C., ... & Pracht, M. Traitements des patients atteints de mélanome de stade III inopérable ou de stade IV. *Société française de dermatologie*. In *Annales de Dermatologie et de Vénérologie*. Elsevier Masson, (2018, April).
- [9] site internet : fondation contre le cancer.
- [10] Leccia., M. T., Planchamp., F., & Mazeau-Woynar, V. Prise en charge thérapeutique des patients atteints d'un mélanome cutané métastatique: comment ont été élaborées les recommandations nationales françaises. *Dermato Mag*, (2014), 2(3), 172-173.
- [11] Yarden Y., & al., Untangling the ErbB signalling network. *Nat. Rev. Mol. Cell. Biol.*, 2002, 2, 127-137.
- [12] Herbst R.S., Review of Epidermal Growth Factor Receptor biology. *Int. J. Radiat. Oncol. Biol. Phys.*, 2004, 59, 21-26.
- [13] Arnold D., Seufferlein T. Targeted treatments in colorectal cancer: state of the art and future perspectives. *Gut* 2010, 59 , 838-58.
- [14] Bachelor MA et Bowden GT, *Semin.Cancer Biol.*, 2004.
- [15] Bode AM et Dong Z, *Sci.STKE*, 2003
- [16] Normanno N et al., *Gene*, 2006
- [17] Schneider MA et al., *Am.J.Pathol.*, 2008
- [18] Winder T., Lenz HJ. Vascular endothelial growth factor and epidermal growth factor signaling pathways as therapeutic targets for colorectal cancer. *Gastroenterology*

[19] R. MILCENT. Avec la contribution de François CHAU. « CHIMIE ORGANIQUE HÉTÉROCYCLIQUE ». © EDP Sciences 2003. 2-4, 10.

[20] BÉLANGER G. «CHIMIE ORGANIQUE HÉTÉROCYCLIQUE». COR 706.Département de chimie Faculté des sciences UNIVERSITÉ DE SHERBROOKE. 2013.20.

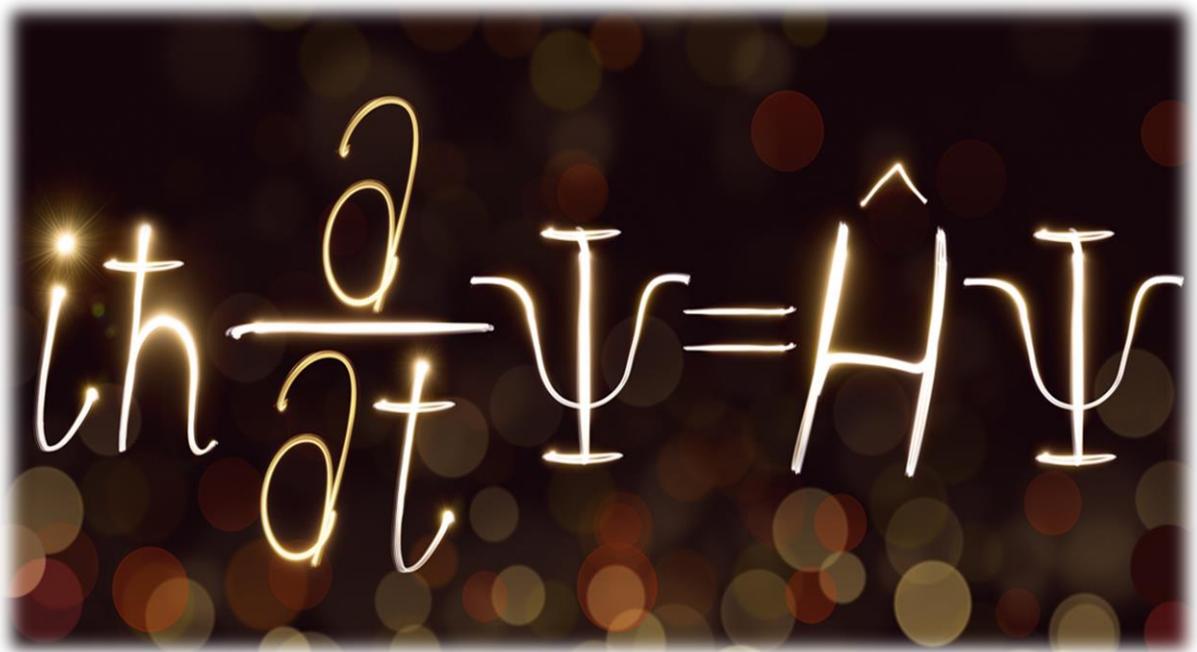
[21] site internet : <http://pharmtox.free.fr/pharmacie/chimie/poly/Heterocycles.htm>

[22] Berekoff, B. Modern Therapeutics 2012: Advances in Physiology, Pharmacology, and Pharmaceutical Sciences. An international symposium held jointly by CSPS, CSPT, and CPS, Toronto, ON, Canada, June 2012. Journal of Pharmacy & Pharmaceutical Sciences, 2012, 15(4), 1-190s.

[23] Brown., D. J. Primary Syntheses of Phthalazines. Cinnolines and Phthalazines: Supplement II, Volume 64, 109-171.

[24] Sajeshkumar., N. K., Vazhacharickal, P. J., Mathew, J. J., &Sebastin, A Synthesis of silver nanoparticles from curry leaf (*Murrayakoenigii*) extract and its antibacterial activity. CIB Tech J Pharm Sci, 2015, 4, 15-25.

Chapitre II
Base Théorique des Méthodes de Calculs
Utilisées


$$i\hbar \frac{\partial}{\partial t} \Psi = \hat{H} \Psi$$

II.1.Introduction

La recherche et la synthèse de nouveaux composés chimiques sont aujourd'hui souvent associées à une étude par modélisation moléculaire. La modélisation moléculaire est une technique permettant, non seulement de représenter les propriétés et les réactions chimiques mais aussi de manipuler les modèles des structures en deux ou trois dimensions.

Les premiers essais de modélisation d'activités de molécules datent de la fin du **19^{ème}** siècle, lorsque Crum-Brown et Frazer [1] postulèrent que l'activité biologique d'une molécule est une fonction de sa constitution chimique. Mais ce n'est qu'en **1964** que furent développés les modèles de "contribution de groupes", qui constituent le début réel de la modélisation QSPR avec les travaux de Hansch [2]. Depuis, l'essor de nouvelles techniques de modélisation par apprentissage, linéaires d'abord, puis non linéaires, ont permis la mise en place de nombreuses méthodes ; elles reposent pour la plupart sur la recherche d'une relation entre un ensemble de nombres réels, descripteurs de la molécule, et la propriété ou l'activité que l'on souhaite prédire. Dans les dernières années, l'utilisation des méthodes QSAR/QSPR n'a cessé de progresser. Elle est même devenue indispensable en chimie pharmaceutique et pour la conception de médicaments [3-4]. Leur développement dans une gamme plus large d'applications leur ouvre d'ailleurs de grandes perspectives [5]. Les informations extraites à partir des résultats d'études QSAR/QSPR peuvent être utilisées pour obtenir une meilleure connaissance des structures moléculaires et probablement le mode d'action au niveau moléculaire. Ces informations peuvent alors être utilisées pour prédire les propriétés physicochimiques et les activités biologiques de nouveaux composés ainsi que pour concevoir de nouvelles structures.

Le docking moléculaire est très utile en biologie, pharmacie et médecine, car la plupart des principes actifs sont de petites molécules (ligand) qui interagissent avec une cible biologique d'intérêt thérapeutique.

Le récepteur macromoléculaire étant le plus souvent une protéine (cible), le terme Docking seul est couramment employé pour désigner un « docking protéine-ligand ».

II.2.Chimie quantique

Différentes approches sont envisageables dans le cadre des outils de modélisation moléculaire. Si ceux de mécanique classique, économiques en termes de temps de calcul, permettent de traiter des systèmes moléculaires de grande taille, les méthodes quantiques (ab initio, semi-empiriques ou théorie de la fonctionnelle de la densité) sont, quant à elles, capables de calculer les propriétés électroniques des systèmes. C'est pour cette raison que ces approches ont été employées dans le cadre de cette étude. Des compléments sur les méthodes décrites ici peuvent être trouvés dans différents ouvrages de références [6-7].

II.2.1.Bases de la chimie quantique

II.2.1.1 Equation de Schrödinger

La chimie quantique définit la structure moléculaire comme un noyau autour duquel gravitent des électrons, qui sont décrit par leur probabilité de présence en un point et représentés par des orbitales [1]. Les équations de la chimie quantique sont basées sur la résolution de l'équation de SCHRÖDINGER qui s'écrit pour les états stationnaires [2] :

$$E \Psi = H \Psi$$

Où :

Ψ : Fonction d'onde de la molécule.

H : Hamiltonien totale d'une molécule comportant N noyaux et n électrons. Il est défini par la somme de cinq termes (cinétique des électrons, cinétique des noyaux, répulsion entre électrons, répulsion entre noyaux et le terme d'attraction électrons-noyaux).

E : Energie totale d'une molécule.

La résolution exacte de cette équation n'est possible que pour l'atome d'hydrogène et les systèmes mono-électroniques. Pour les systèmes poly électroniques, on fait appel aux méthodes d'approximation, pour la résolution approchée de l'équation de SCHRÖDINGER.

L'Hamiltonien rend compte des différentes contributions à l'énergie totale du système à partir d'opérateurs pour les énergies cinétiques des électrons et des noyaux ainsi que les interactions noyau-électron, électron-électron et noyau-noyau.

$$\hat{H} = -\frac{1}{2} \sum_i \nabla_i^2 - \sum_k \frac{1}{2} \nabla_k^2 - \sum_i \sum_k \frac{Z_k}{r_{ik}} + \sum_i \sum_{i < j} \frac{1}{r_{ij}} + \sum_k \sum_{k < l} \frac{Z_k Z_l}{r_{kl}}$$

Où : ∇^2 est l'opérateur d'énergie cinétique, Z_k est le numéro atomique de l'atome k , r_{ik} est la distance entre un électron i et un noyau k , r_{ij} est la distance entre deux électrons i et j et r_{kl} est la distance entre deux noyaux k et l .

III.2.1.2. Approximation de Born-Oppenheimer

Une première simplification du problème est l'approximation de Born-Oppenheimer. Cette approximation consiste à considérer que les deux types de particules, électrons et noyaux, ont des vitesses d'évolution caractéristiques très différentes ; pour être précis, les électrons se déplacent beaucoup plus rapidement que les noyaux. [8] On peut alors considérer le comportement des électrons dans une molécule en supposant que les noyaux occupent les positions fixes dans l'espace. On aboutit alors à l'équation de Schrödinger électronique : [9]

$$E_{el} \Psi_{el} = \hat{H}_{el} \Psi_{el}$$

Où \hat{H}_{el} est l'hamiltonien électronique : $\hat{H}_{el} = T_e + V_{e-e} + V_{ext}$. La variable scalaire E_{el} est l'énergie propre, associée à l'hamiltonien \hat{H}_{el} . [10]

II.2.2. Théorie de la fonctionnelle de la densité (DFT)

Dans la théorie de la fonctionnelle de la densité, les propriétés de l'état fondamental d'un système de particules interagissant entre-elles, sont exprimées en fonction de la densité électronique.

II.2.2.1. Aperçu historique

La théorie fonctionnelle de la densité a pour objet de décrire un système en considérant la densité $\rho(r)$ comme variable de base. Ainsi le problème à n électrons est étudié dans l'espace de $\rho(r)$ qui est de dimension 3 au lieu de l'espace de dimension $3n$ de la fonction d'onde ψ .

Les premiers à avoir exprimé l'énergie en fonction de la densité furent L. H. Thomas et E. Fermi en 1927 [11]. Dans leur modèle, les interactions électroniques sont traitées classiquement et l'énergie cinétique est calculée en supposant la densité électronique homogène. Ce modèle a été amélioré par P. A. Dirac en 1930 [11] avec un terme d'échange.

Un peu plus tard, en 1951 J. C. Slater [11] proposa un modèle basé sur l'étude d'un gaz uniforme améliorée avec un potentiel local. Cette méthode, appelée Hartree-Fock- Slater ou $X\alpha$, fut essentiellement utilisée en physique du solide dans les années 70.

Mais la DFT a véritablement commencé avec les théorèmes fondamentaux de Hohenberg et Kohn en 1964 [12] qui établissent une relation fonctionnelle entre l'énergie de l'état fondamental et sa densité.

Alors que le premier succès applications de DFT pour la recherche sur la structure électronique moléculaire a commencé à apparaître dans les années 90 avec le développement

des fonctionnels d' échange et de corrélation [11-13-14], les plus précises et les plus rapides pour le calcul des propriétés électroniques de grands systèmes moléculaires ou a été introduite dans le code Gaussienne. Enfin nous signalons par le prix Nobel qui a été attribué à Kohn et à Pople [15-16] en 1998 dans le cadre de développement de cette méthode.

II.2.2.2. Théorème de Hohenberg et Kohn

Si l'utilisation de la densité électronique comme variable de base a été envisagée par Thomas et Fermi dès 1927 puis Slater dans les années 1950 [17], le véritable formalisme de la DFT naît des théorèmes de Hohenberg et Kohn dans les années 1960 [18].

Premier théorème : théorème d'existence

Ce théorème de Hohenberg et Kohn montre que la densité électronique est la seule fonction nécessaire pour obtenir toutes les propriétés électroniques d'un système. Si on considère un gaz d'électrons, le potentiel externe agissant sur ces particules détermine l'état fondamental de ce système et la densité de charge correspondante [19]. Ainsi, toutes les propriétés concernant cet état sont des fonctionnelles du potentiel externe. Comme cela a été démontré initialement par Hohenberg et Kohn, en raison de la correspondance biunivoque existant entre le potentiel externe v_{ext} et la densité électronique $\rho(r)$ L'expression de l'Hamiltonien électronique d'un système poly-électronique est :

$$\hat{H} = -\frac{1}{2} \sum_i^n \Delta_i + \sum_{i>j}^n \frac{1}{r_{ij}} + v_{ext}(r_i)$$

$$v_{ext}(r_i) = -\sum_{K=1}^N \frac{Z_K}{R_{Ki}}$$

Avec :

$V_{ext}(r_i)$: Potentiel externe de l'électron i .

$\rho(r)$: La densité électronique. Elle est égale à n , le nombre totale des électrons, lorsqu'elle est intégrée sur tout l'espace.

$$\int_0^\infty \rho(r) dr = n$$

L'énergie totale du système peut s'écrire comme la somme de trois fonctionnelles : celle de l'énergie potentielle noyau-électron, de l'énergie cinétique et de l'énergie potentielle électron-électron.

En posant : $E[\rho] = V_{ne}[\rho] + T[\rho] + V_{ee}[\rho]$

$$V_{ne}[\rho] = \int(\mathbf{r}) v_{ext}(\mathbf{r}) d\mathbf{r}$$

$$F[\rho] = T[\rho] + V_{ee}[\rho]$$

La fonctionnelle d'énergie s'écrit alors :

$$E[\rho] = \int(\mathbf{r}) v_{ext}(\mathbf{r}) d\mathbf{r} + F[\rho]$$

2^{ème} Théorème de Hohenberg : principe variationnel

Le second théorème applique le principe variationnel à la DFT. Ainsi, l'énergie fonctionnelle d'une densité électronique approchée est supérieure ou égale à l'énergie exacte du système dans son état fondamental [20].

$$E[\rho] \geq E[\rho_0] = E_0$$

II.2.2.3. Approche Kohn-Sham

W.Kohn et L.J.Sham (KS) [21] ont apporté des applications pratiques de la théorie de la densité. En 1965, ils ont proposé un ensemble d'équations mono-électroniques à partir desquelles, il est possible d'obtenir la densité électronique et ensuite l'énergie totale.

Les orbitales de Kohn-Sham diffèrent des orbitales Hartree-fock :

Ce sont celles d'un système fictif d'électrons sans interaction, mais elles possèdent la même densité que le système réel. L'énergie totale d'un système à n électrons peut être écrite sans approximation [19-22-23-24], en exprimant les densités en fonction des orbitales moléculaires de Kohn-Sham [20].

Le système de référence est choisi de telle façon à avoir [25] :

$$\rho_s(\mathbf{r}) = \rho_0$$

Etant donné que les électrons n'interagissent pas entre eux dans le système de référence, l'hamiltonien de système de référence s'écrit :

$$\hat{H}_s = \sum_{i=1}^n [-1/2 \nabla_i^2 + v_s(r_i)] = \sum_{i=1}^n h_i^{KS}$$

$$h_i^{KS} = -1/2 \nabla_i^2 + V_s(\mathbf{r}_i)$$

Avec :

Par conséquent, les équations de Kohn et Sham, pour l'électron i , peuvent s'écrire comme suit:

$$h_i^{KS} \theta_i^{KS} = \epsilon_i^{KS} \theta_i^{KS}$$

θ_i^{KS} : Orbitale de kohn et Sham de l'électron i .

II.2.2.4. Fonctionnelles d'échange-corrélation

Finalement, toute la difficulté de la DFT relève du terme d'échange-corrélation dont l'expression exacte reste inconnue. Aussi est-il calculé à partir de fonctionnelles d'échange-corrélation approchées qui se distinguent par les approximations qui les définissent.

II 2.2.4.1. Approximation de la densité locale (LDA)

L'approximation de la densité locale LDA (de l'anglais Local Density Approximation) proposée par Kohn et Sham, en 1965 [26] est basée sur le modèle du gaz uniforme d'électrons. L'énergie d'échange-corrélation est définie de la manière suivante :

$$E_{XC}^{LDA}[\rho] = \int_{xc} \epsilon(\rho) dr$$

Où $\epsilon_{xc}(\rho)$ est la densité d'énergie d'échange et de corrélation. Par ailleurs, si l'on partageons l'énergie d'échange et de corrélation en deux (énergie d'échange ϵ_x et énergie de corrélation ϵ_c) telle que : $\epsilon_{xc} = \epsilon_x + \epsilon_c$

On peut utiliser l'énergie d'échange proposée par Dirac [27] comme approximation de ϵ_x : $KD [\rho(r)] S_x = \int \rho(r)^{3/4}$

La fonctionnelle de corrélation la plus utilisée a été développée par Vosko, Wilk et usair en 1980 [28]. Ces auteurs ont utilisé les résultats de calculs Monte Carlo effectués par Ceperley et Alder, pour ajuster une expression analytique de l'énergie de corrélation. Cette fonctionnelle est connue sous l'abréviation « VWN ».

Depuis 1985 d'énormes efforts ont contribué à l'amélioration des fonctionnelles d'échange-corrélation. Ces travaux ont débouché sur une deuxième génération de fonctionnelles incluant l'inhomogénéité de la densité électronique : ces fonctionnelles prennent donc en compte la densité électronique ainsi que son gradient. Alors que Le X α une

approximation, déjà citée, entre dans cette catégorie, depuis la corrélation est rapproché par un pourcentage donné de l'énergie d'échange [27].

II 2.2.4.2 Approximation du gradient généralisé (GGA)

L'approche LDA se fondait sur le modèle du gaz d'électrons et supposait donc une densité électronique uniforme. Cependant les systèmes atomiques ou moléculaires sont le plus souvent très différents d'un gaz d'électrons homogène et, de manière plus générale, on peut considérer que tous les systèmes réels sont inhomogènes c'est-à-dire que la densité électronique possède une variation spatiale. On peut généraliser l'approximation LDA en faisant dépendre la fonction f , non pas uniquement de la valeur de la densité $\rho(r)$ au point r , mais aussi au voisinage de r , cette variation de la densité exprime les énergies d'échanges et de corrélation en fonction de la densité mais également son gradient. En pratique, on utilise pour cela la dérivée de la densité par rapport à la position, au point r . On parle alors d'approximations semi-locales (parfois aussi appelées méthodes non locales), ou GGA (Generalized Gradient Approximation).

En général, l'énergie d'échange-corrélation est définie dans l'approximation GGA comme :

$$E_{xc}(\rho) = \int \rho(r) \varepsilon_{xc}(\rho, \nabla\rho) dr$$

L'introduction explicite de termes gradient de la densité $\nabla\rho$ améliore très sensiblement les performances de la méthode. À titre d'exemple, la GGA estime correctement les énergies de liaison dans les molécules alors que la LDA les surestime.

L'énergie d'échange associée à une fonctionnelle GGA s'exprime sous la forme :

$$E_x(\rho) = E_x^{LDA} - \int F(S) \rho^{4/3}(r) dr;$$
$$\text{Avec : } S(r) = \frac{|\nabla\rho(r)|}{\rho^{4/3}(r)}$$

$S(r)$ est une quantité sans dimension appelée gradient de densité réduit. Deux classes de fonctions $F(S)$ sont couramment utilisées pour l'échange ajustées sur des résultats expérimentaux. [29]

Fonctionnelle hybride B3LYP

La fonctionnelle hybride B3LYP (Becke 3-paramètres Lee-Yang-Parr) est une fonctionnelle à trois paramètres combinant les fonctionnelles d'échange local, d'échange de Becke et d'échange HF, avec les fonctionnelles de corrélation locale (VWN) et corrigée du gradient de Lee, Yang et Parr :

$$E_{XC}^{B3LYP} = (1 - a_0 - a_x) E_X^{LDA} + a_0 E_X^{HF} + a_x E_X^{B88} + (1 - a_c) E_C^{VWN} + a_c E_C^{LYP}$$

Où les paramètres a_0 , a_x et a_c ont été ajustés respectivement à 0,20, 0,72 et 0,81 [30].

II.2. 3. Méthodes semi-empiriques

Les calculs semi-empiriques sont eux développés sur la même structure générale que les calculs HF, mais certaines parties de l'information sont sujettes à approximation ou même complètement omises, afin de les rendre moins exigeants en termes de temps de calcul. [31]

En particulier, dans le cadre de ces approches, trois types de simplifications sont principalement réalisés :

- non-considération des électrons de cœur puisqu'ils ne contribuent pas à la réactivité chimique du système. Ils seront alors considérés avec le noyau au sein d'une fonction paramétrée (ex : méthode de Hückel étendu).
- utilisation d'un jeu réduit de fonctions de base (typiquement STO-3G).
- réduction du nombre d'intégrales bi-électroniques grâce à l'introduction de paramètres empiriques. [31]

Dans les différentes méthodes semi-empiriques du type CNDO/2 (Complete Neglect of Differential Overlap), INDO (Intermediate Neglect of Differential Overlap), NDDO (Neglect of Diatomic Differential Overlap), MNDO (Modified Neglected of Differential Overlap), AM1 (Austin Model 1), PM3 (Parametric Method 3), seuls les électrons de valence sont pris en considération. Ces méthodes rendent compte des phénomènes de répulsion électron-électron, mais négligent complètement les orbitales de recouvrement orbitalaire dans la procédure SCF. [32] Les termes énergétiques les plus difficiles à calculer sont estimés à partir des données Expérimentales. [33]

Par ces méthodes, nous avons accès à différentes propriétés moléculaires comme les charges atomiques, les énergies des orbitales (HOMO, LUMO) entre autres, le potentiel d'ionisation, l'enthalpie de formation, la répartition électronique, etc....

II.3. Etude des relations structure-activité (QSAR)

II.3.1. Introduction

La connaissance des propriétés et des activités est d'une importance capitale pour pouvoir classer et utiliser les composés chimiques. La caractérisation expérimentale complète est difficile, voire impossible, pour des raisons de temps, de coût, de dangerosité de certains essais ou d'éthique (limitations des essais sur les animaux). L'utilisation des méthodes alternatives à l'expérience est devenue plus qu'indispensable. Parmi ces méthodes, on trouve

les méthodes de modélisation moléculaire qui permettent de justifier les données expérimentales disponibles et prédire les propriétés/activités pour des composés nouveaux ou des composés pour lesquels les données expérimentales ne sont pas disponibles. Parmi ces méthodes de modélisation les plus utilisées, on peut citer les méthodes QSPR (Quantitative Structure-Property Relationships) et QSAR (Quantitative Structure-Activity Relationships).

Ces méthodes s'appuient sur le principe que les propriétés physico-chimiques et les activités biologiques des molécules dépendent fortement de leurs structures chimiques.

Il s'agit de présenter ici le principe des modèles QSAR ainsi que ceux des différents outils employés pour leur mise en place et leur évaluation : bases de données expérimentales, descripteurs, outils d'analyse de données.

II.3.2. Généralités sur la modélisation QSPR/QSAR

Une relation QSPR/QSAR est un modèle ou une formule mathématique qui permet de relier, d'une manière quantitative, la structure d'une molécule à une propriété ou à une activité donnée. Les méthodes QSPR/QSAR sont de plus en plus utilisées, du fait de la croissance des moyens de calculs (machines, logiciels,...). Récemment, on assiste à la mise en place d'un nouveau règlement REACH [34] (Registration, Evaluation, Authorisation and Restriction of Chemicals) qui recommande l'utilisation des méthodes alternatives pour limiter le recours à l'expérimentation. En fait, les premiers travaux QSPR/QSAR remontent au 19ème siècle. En effet, dès 1868, Crum-Brown et Fraser [35] ont postulé l'existence de relations entre les activités physiologiques et les structures chimiques en reliant les variations de l'activité biologique à des modifications structurales. Cependant, à cette époque, les structures moléculaires n'étaient pas encore connues.

Une avancée importante vers les modèles QSPR/QSAR proprement dits a été réalisée grâce au développement des équations de Hammett [36] dans lesquelles les constantes σ caractérisent de manière quantitative les vitesses de réactions pour les composés organiques.

$$\log \frac{K}{K_0} = \sigma \rho$$

Où K et K_0 sont les constantes respectives de la réaction étudiée et de celle d'une référence et ρ une constante de réaction dépendant du type de réaction [35 -37].

Les premiers travaux utilisant la méthodologie QSPR/QSAR telle qu'employée actuellement sont dus à Hansch [38] et Free et Wilson [39]. D'un côté, Hansch a proposé des modèles reliant directement l'activité biologique des composés avec les propriétés hydrophobes, électroniques et stériques à l'échelle moléculaire. D'un autre côté, Free et

Wilson ont développé des modèles empiriques, dits de contributions de groupes, pour l'étude de l'activité biologique prédictif des modèles proposés.

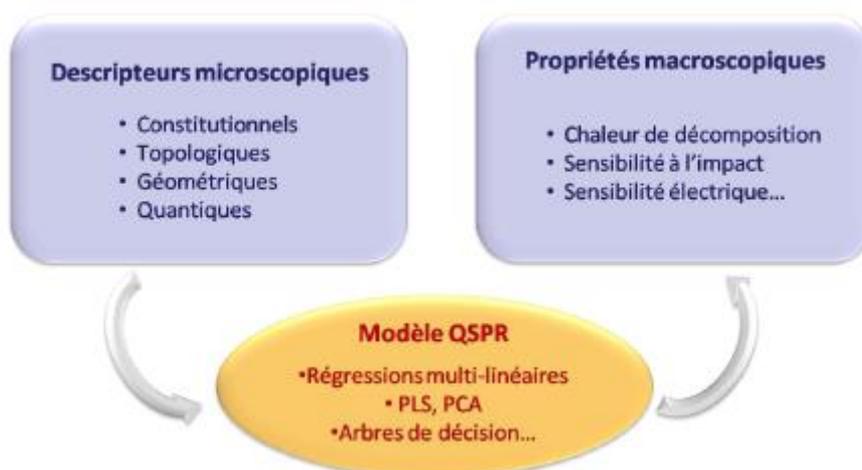
II.3.3.Principe

Le principe des méthodes QSPR est, comme leur nom l'indique, de mettre en place une relation mathématique reliant de manière quantitative des propriétés moléculaires aussi bien électroniques que géométriques, appelées descripteurs, avec une observable macroscopique (activité biologique, toxicité, propriété physico-chimique, etc.), pour une série de composés chimiques similaires à l'aide de méthodes d'analyses de données. Aussi la forme générale d'un tel modèle est la suivante :

$$\text{Propriété} = f(\text{Descripteurs})$$

L'objectif d'une telle méthode est donc d'analyser les données structurales afin de détecter les facteurs déterminants pour la propriété mesurée. Pour ce faire, différents types d'outils peuvent être employés : régressions multilinéaires (MLR) [40], régressions aux moindres carrés partiels (PLS) [41], arbres de décision [42], réseaux de neurones [43-44], algorithmes génétiques [45].

Une fois cette relation mise en place et validée sur un jeu validation, elle peut alors être employée pour la prédiction de la propriété de nouvelles molécules, pour lesquelles la valeur expérimentale n'est pas disponible, voire pour des molécules encore non synthétisées. De tels modèles peuvent également, dans certains cas, être utilisés pour mieux appréhender les phénomènes moléculaires mis en jeu dans la propriété d'intérêt.



FigureII.1 : Principe de la Méthode QSPR

II.3.4. Importance de la base de données

Un modèle QSPR est dépendant des données expérimentales de référence, le choix de la base de données est un point critique dans le développement de ces modèles. Dans la plupart des cas, les données expérimentales sont issues de la littérature, et pour être de qualité, une base de données doit être composée de données expérimentales aussi fiables que possible, puisque les barres d'erreurs sur celles-ci se propageront dans le modèle final, étant donné que les paramètres de ce dernier sont ajustés par rapport à ces données. Il est donc important de choisir des données présentant des incertitudes faibles, afin de limiter les barres d'erreur expérimentales. De plus, les données doivent être obtenues suivant un protocole expérimental unique. En effet, les conditions expérimentales ont une forte influence sur les valeurs obtenues. La définition de la propriété en termes de conditions expérimentales est d'ailleurs un point important de la démarche.

II.3.5. Descripteurs Moléculaires

II.3.5.1. Descripteurs constitutionnels

Les descripteurs constitutionnels sont directement liés à la formule brute de la molécule, à l'aide de la composition moléculaire, c'est-à-dire les atomes qui le constituent, Il s'agit de :

- La masse molaire.
- Les nombres absolus et relatifs d'atomes (C, H, O, S, N, F, Cl, Br, I, P. . .).
- Les nombres absolus et relatifs de groupes fonctionnels (NH₂, COOH, OH. . .).
- Les nombres absolus et relatifs de liaisons (simples, doubles, aromatiques. . .).
- Les nombres absolus et relatifs de cycles (aromatiques ou non).

Ces descripteurs sont très utilisés du fait de leur extrême simplicité non seulement d'un point de vue conceptuel mais surtout calculatoire. On peut remarquer que ces descripteurs ne permettent pas de distinguer les isomères de constitution. C.-à-d., si on développe des modèles avec ce type de descripteurs seulement, alors que ces derniers peuvent poser problème pour l'interprétation des mécanismes d'interaction mis en jeu pour la propriété étudiée.

II.3.5.2. Descripteurs topologiques

Les descripteurs topologiques "ou indices topologiques", décrivent la connectivité atomique dans la molécule, ils sont obtenus à partir de la structure 2D de la molécule, et donnent des informations sur sa taille, sa forme globale et ses ramifications. Ces descripteurs s'inspirent de la théorie des graphes appliquée à la table de connectivité qui n'est autre qu'une

représentation compacte de la connectivité interatomique au sein de la molécule. Les indices topologiques les plus fréquemment utilisés sont l'indice de Wiener [45]. L'indice de Randić [46]. L'indice de connectivité de valence de Kier-Hall [47] et l'indice de Balaban [48].

II.3.5.3. Descripteurs géométriques

Les descripteurs géométriques d'une molécule sont évalués à partir des positions relatives de ses atomes dans l'espace, et décrivent des caractéristiques plus complexes ; leurs calculs nécessitent de connaître, la géométrie 3D de la molécule, par modélisation moléculaire empirique ou ab initio, Ces descripteurs s'avèrent donc relativement coûteux en temps de calcul, mais apportent davantage d'informations, et sont nécessaires à la modélisation de propriétés qui dépendent de la structure 3D. On distingue plusieurs descripteurs importants, le volume moléculaire, la surface accessible au solvant, le moment d'inertie. Les distances, les angles de liaisons ou angles dièdres dans la molécule. [49-50].

II.3.5.4. Descripteurs électrostatiques

Ces descripteurs reflètent les caractéristiques de la distribution de charge de la molécule. Les charges partielles empiriques dans la molécule sont calculées en utilisant l'approche proposée par Zefirov [51-52]. Cette méthode est basée sur l'échelle d'électronégativité, Sur la base de ces charges partielles les descripteurs électrostatiques suivantes sont calculés :

- Les charges partielles minimales et maximales dans la molécule (q_{\min} , q_{\max})
- Les charges partielles minimales et maximales pour les atome (C, N, O, ...)
- Les indices électroniques topologiques.
- Les charges partielles de la zone de la surface (Charged partial surface area (CSPA))

II.3.5.5. Descripteurs quantiques

Les descripteurs de chimie quantique donnent des informations importantes pour la molécule. Ces descripteurs permettent de quantifier différents types d'interactions inter- et intramoléculaires, de grande influence sur des propriétés physicochimiques de molécules. Le calcul de la plupart de ces descripteurs nécessite la recherche de la géométrie pour laquelle l'énergie est minimale, et fait appel aux approches de chimie quantique. On cite toutes les données énergétiques, vibrationnelles et orbitales du système par exemple, les énergies de la plus haute orbitale moléculaire occupée HOMO et de la plus basse vacante LUMO, le moment dipolaire, la polarisabilité, le potentiel d'ionisation. . . .etc.

Les descripteurs quantiques sont classés comme suit [53] :

- Descripteurs lie à la distribution de charge

- Descripteurs liés à l'énergie.

II.3.5.6. Descripteurs thermodynamiques

Les descripteurs thermodynamiques sont calculés sur la base de la fonction de partition totale Q de la molécule [54–55], La fonction de partition commode la façon avec laquelle l'énergie d'un système de molécules est répartie parmi les individus moléculaires. Sa valeur dépend du poids moléculaire, de la température, du volume moléculaire, des distances inter nucléaires, des mouvements moléculaires et des forces intermoléculaires. La fonction de partition est le point le plus commode entre les propriétés microscopiques des molécules individuelles (niveaux d'énergie, moments d'inertie) avec les propriétés macroscopiques (chaleur spécifique, entropie). La molécule peut accroître son énergie de translation, de vibration, de rotation de façon pratiquement indépendante.

II.3.6. Régression linéaire multiple (MLR)

Tout comme en régression linéaire simple, la régression multiple cherche approximer une relation trop complexe en général, par une fonction mathématique simple. Elle repose sur l'hypothèse qu'il existe une relation linéaire entre une variable dépendante (à expliquer) Y (ici, la toxicité) et une série de p variables indépendantes (explicatives) X_i (ici, les descripteurs). L'objectif est d'obtenir une équation de la forme suivante :

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p + \varepsilon$$

Cette équation est linéaire par rapport aux paramètres (coefficients de régression) $\beta_0, \beta_1, \dots, \beta_p$. La détermination de l'équation [56] se fait alors à partir d'une base de données de n échantillons pour laquelle à la fois les variables indépendantes et la variable dépendante sont connues. Il s'agit donc de considérer un système d'équations [57] :

$$Y_1 = \beta_0 + \beta_1 X_{1,1} + \beta_2 X_{2,1} + \dots + \beta_p X_{n,1} + \varepsilon_1$$

$$Y_2 = \beta_0 + \beta_1 X_{1,2} + \beta_2 X_{2,2} + \dots + \beta_p X_{n,2} + \varepsilon_2$$

$$Y_n = \beta_0 + \beta_1 X_{1,p} + \beta_2 X_{2,p} + \dots + \beta_p X_{n,p} + \varepsilon_n$$

Où les résidus ε_i représentent l'erreur du modèle, constituée par l'incertitude sur la variable dépendante Y_i d'une part, sur les variables indépendantes X_i d'autre part, mais aussi par les informations contenues dans les variables indépendantes mais non exprimées via les variables dépendantes.

Ce système d'équation peut être donné sous la forme matricielle suivante :

$$\begin{matrix} Y = \begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \\ \dots \\ Y_n \end{pmatrix} & X = \begin{pmatrix} 1 & X_{11} & X_{12} & \dots & X_{1p} \\ 1 & & & & \\ 1 & & & & \\ \dots & & & & \\ 1 & X_{n1} & X_{n2} & \dots & X_{np} \end{pmatrix} & \beta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \dots \\ \beta_p \end{pmatrix} & \varepsilon = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \dots \\ \varepsilon_n \end{pmatrix}
 \end{matrix}$$

II.3.6.1. Estimation des paramètres statistiques du modèle

Dans le cas d'un modèle à p variables régresseurs, le critère des moindres carrés s'écrit :

$$S(\beta_0, \dots, \beta_p) = \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 = \sum_{i=1}^n \varepsilon_i^2 = \sum_{i=1}^n (Y_i - \beta_0 - \beta_1 X_{1i} - \dots - \beta_p X_{pi})^2$$

Les valeurs des β qui minimisent ce critère seront les solutions $\beta_0, \beta_1, \dots, \beta_p$ du système linéaire de (p+1) équations à (p+1) inconnues.

La méthode consiste alors à choisir les coefficients du vecteur b en faisant en sorte de minimiser la somme des carrées des écarts entre les valeurs prédites et les valeurs réelles sur l'intégralité de la base de données et ceci sous couvert de certaines hypothèses de départ.

En premier lieu, les variables indépendantes X_i , comme leur nom l'indique, sont supposées indépendantes entre elles et leur incertitude est négligeable. Ensuite, les différents échantillons Y_i sont supposés indépendants entre eux. Enfin, par nature, la dépendance de Y vis-à-vis des X_i est supposée linéaire. La valeur prédite de la variable dépendante Y (estimé par le modèle de régression) s'écrit :

$$\hat{Y} = X \hat{\beta} = XB$$

Les résidus peuvent donc être définis comme la différence entre les valeurs prédites et observées de Y.

$$Y_i - \hat{Y}_i = \varepsilon_i$$

II.3.6.2. Tests sur le modèle linéaire

Comme pour le modèle linéaire simple, les hypothèses de régression linéaire doivent être vérifiées pour un modèle de régression multiple.

II.3.6.3. Test de la signification globale de la régression (F-Fisher)

Ce test permet de connaître l'apport global de l'ensemble des variables X_1, \dots, X_p à la détermination de Y .

On veut tester l'hypothèse nulle :

$H_0 : \beta_1 = \dots = \beta_p = 0$ contre H_a : il existe au moins un β_j parmi β_1, \dots, β_p non égal à 0. On calcule la statistique de test F :

$$F = \frac{MS \text{ model}}{MS_{\text{error}}}$$

II.3.6.4. Interprétation et validation d'un modèle QSPR/QSAR

Une fois développé, le modèle doit être interprété en analysant tous les paramètres statistiques de ce modèle, sa qualité doit être aussi étudiée, cette qualité est vérifiée par ce que l'on appelle validation. Sa robustesse, c'est-à-dire l'influence des composés de la série d'apprentissage sur le modèle, est estimée par des méthodes de validation interne. Afin d'estimer son pouvoir prédictif, des données expérimentales supplémentaires sont nécessaires afin de déterminer la capacité du modèle à prédire ces valeurs c'est ce que l'on appelle validation externe. Enfin, il est important de savoir quel type de molécules utilisées avec quel modèle. On parle alors de domaine d'applicabilité.

II.3.6.5. Validation interne

Dans le passé, la validation interne d'un modèle QSPR/QSAR a été réalisée en utilisant la validation croisée LOO (leave-one out) ou LMO (leave-many out) qui est quantifiée par le coefficient R^2_{cv} . Ce processus consiste à extraire un certain nombre k des molécules du jeu initial à N molécules et à construire un nouveau modèle avec les $(N-k)$ molécules restantes à l'aide des descripteurs choisis (seules les constantes de la régression changent). Ce processus est ensuite réitéré pour retirer et prédire les valeurs de toutes les molécules de la série d'apprentissage. En fonction du nombre de molécules retirées à chaque itération, on parlera de Leave-One-Out (LOO) ou de Leave-Many-Out (LMO) selon qu'une ou plusieurs molécules est (sont) retirée(s) [51].

II.3.6.6. Validation externe

Cette méthode consiste à prédire la propriété/activité d'une série de molécules appelée généralement série de test qui ne sont pas dans la série de développement du modèle, cette validation est caractérisée par les paramètres R^2 (test) R^2_{cv} (test). Récemment plusieurs études [52] ont montré l'insuffisance des paramètres R^2 , R^2_{cv} pour vérifier le pouvoir prédictif des modèles QSAR. Par conséquent, d'autres paramètres doivent être vérifiés pour cet objectif.

Ces paramètres sont connus sous le nom « critères de validation externe » ou souvent appelés « critères de Trophsa » (Trophsacriteria) [52].

II.3.6.7. Critères de validation Externe (série de test)

- $R^2 > 0.7$ (critère 1)
- $R^2_{cv} > 0.6$ (critère 2)

$$\frac{R^2 - R_0^2}{R^2} < 0.1 \text{ et } 0.85 \leq k \leq 1.15 \text{ (critère 3)}$$

- $\frac{R^2 - R_0'^2}{R^2} < 0.1 \text{ et } 0.85 \leq k' \leq 1.15 \text{ (critère 4)}$

- $|R^2 - R_0^2| \leq 0.3 \text{ (critère 5)}$

Avec :

R^2 Coefficient de corrélation pour les molécules de la série de test.

R_0^2 coefficient de corrélation entre les valeurs prédites et expérimentales pour la série de test.

$R_0'^2$ coefficient de corrélation entre les valeurs expérimentales et prédites pour la série de test.

k : est la constante de la droite (à l'origine) de corrélation (valeurs prédites en fonction des valeurs expérimentales)

k' : est la constante de la droite (à l'origine) de corrélation (valeurs expérimentales en fonction des valeurs prédites)

II.3.6.8. Critère de validation croisée : PRESS

La somme des erreurs quadratiques de prédiction « predictions of squares » (PRESS) est définie par :

$$\text{PRESS} = \sum_{i=1}^n \varepsilon(i)^2$$

Ce critère permet de sélectionner les modèles ayant un bon pouvoir prédictif. (On cherche toujours le PRESS le plus petit).

Ces informations permettent d'interpréter les tables d'analyse de variance complètes données par tout logiciel mettant en œuvre la régression linéaire.

II.3.6.9. Validation d'un modèle

Cross validated R²

La procédure statistique cross-validation peut être utilisée pour évaluer le pouvoir prédictif des modèles QSAR. Par exemple la procédure « Leave-One-Out » retire successivement une molécule de la série d'apprentissage contenant n molécules. Un modèle QSAR/QSPR est construit sur un ensemble n-1 de composés et la molécule retirée est prédite par le modèle. Cette procédure est répétée « n » fois afin de prédire les activités/propriétés de toutes les molécules.

Le coefficient qui décrit la validation est donné par l'équation ci-dessous

$$R_{CV}^2 = 1 - \frac{\sum_i (Y_i^{pred} - Y_i^{obs})^2}{\sum_i (Y_i^{obs} - Y^{mean})^2}$$

Ce coefficient peut être calculé à partir de PRESS comme suit :

$$R_{CV}^2 = 1 - \frac{PRESS}{SS_{total}}$$

II.4. Doking moléculaire

II.4.1. Introduction

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

Il y a plus de 30 ans, l'utilisation des méthodes de docking moléculaire in silico dans le processus de conception de nouvelles molécules actives a débuté, afin d'étudier les modes d'interactions des ligands avec la cible biologique pour l'inhiber ou l'activer. Ce qui est considérablement plus facile à mettre en œuvre, moins cher et plus rapide que l'utilisation d'une méthode expérimentale [53]. Ces méthodes de docking visent à prévoir la capacité ou non d'un ligand à se lier au site actif du récepteur et comprendre comment il interagit avec, en se basant pour cela sur la prédiction de la conformation et de l'orientation de la molécule lors de sa liaison au récepteur [54]. Les méthodes de docking s'attachent donc, d'une part, à identifier les molécules qui sont des ligands véritables du récepteur parmi toutes celles étudiées, mais aussi d'autre part, à déterminer les poses correctes soit les conformations

adoptées par les ligands lors de la liaison au récepteur [55]. A cet effet, les méthodes de docking combinent l'utilisation d'un algorithme de recherche, permettant de générer des modes d'interactions putatifs du ligand dans le récepteur, ou « poses », et d'une fonction de score, employée pour classer les différentes poses selon un score prédit d'affinité.

II.4.2.Principe de docking

Le docking est utilisé pour prédire la structure du complexe intermoléculaire résultant de l'association entre au moins deux molécules. Quand il s'agit de deux protéines, on parle de docking protéine-protéine [56], par opposition au docking protéine-ligand que nous avons utilisé pour ce travail.

Le processus du docking est itératif et chaque passe de calcul s'articule en deux étapes. Tout d'abord, une portion limitée de l'espace des conformations du complexe protéine-ligand est explorée afin d'améliorer la *pose* (orientation et/ou conformation) du ligand dans le site actif de la protéine. L'étape suivante fait intervenir une fonction de score qui évalue la qualité de la pose générée [57]. Cette estimation *in silico* de l'affinité du ligand pour la cible est basée sur un examen simplifié des interactions entre les deux partenaires. La répétition du cycle de recherche associé à l'estimation du score guidant l'exploration de l'espace doit assurer la convergence de l'algorithme vers un état représentatif du minimum global de l'énergie libre d'association.

II.4.3.Les différents types de docking moléculaire

Le docking moléculaire est l'étude *in silico* des différents mécanismes et interactions qui interviennent lorsque plusieurs structures moléculaires s'assemblent, l'objectif étant de parvenir à déterminer comment ces molécules vont s'agencer les unes par rapport aux autres (figure II.4) [58].

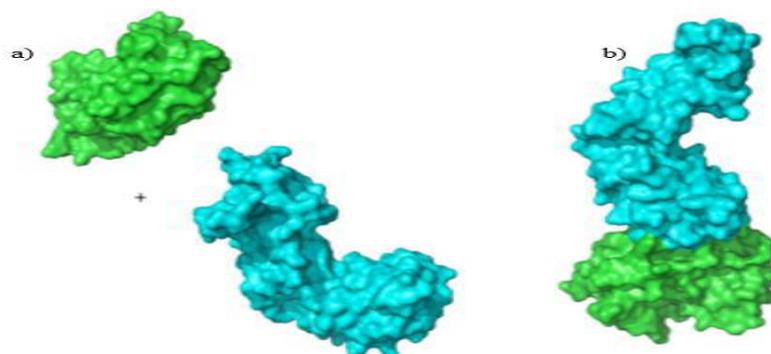


Figure II.2 : Représentation Schématique du Docking de deux Molécules. (a) Molécules de Départ. (b) Complexe Moléculaire.

Il existe deux grands types de docking moléculaire. Le docking rigide consiste à obtenir la conformation préférentielle d'un système protéine-ligand en considérant que chacune des deux molécules conserve une géométrie interne fixe, c'est à dire que le ligand peut subir des translations et des rotations dans chacune des trois dimensions. Dans ce cas, la relaxation de la géométrie interne de chaque entité, en interaction dans le complexe, n'est pas prise en compte.

Cependant, il est tout à fait concevable que les structures de la protéine et du ligand soient modifiées durant le processus de docking moléculaire afin d'optimiser au mieux l'interaction entre les deux entités. On parle, dans ce cas, de docking flexible [59].

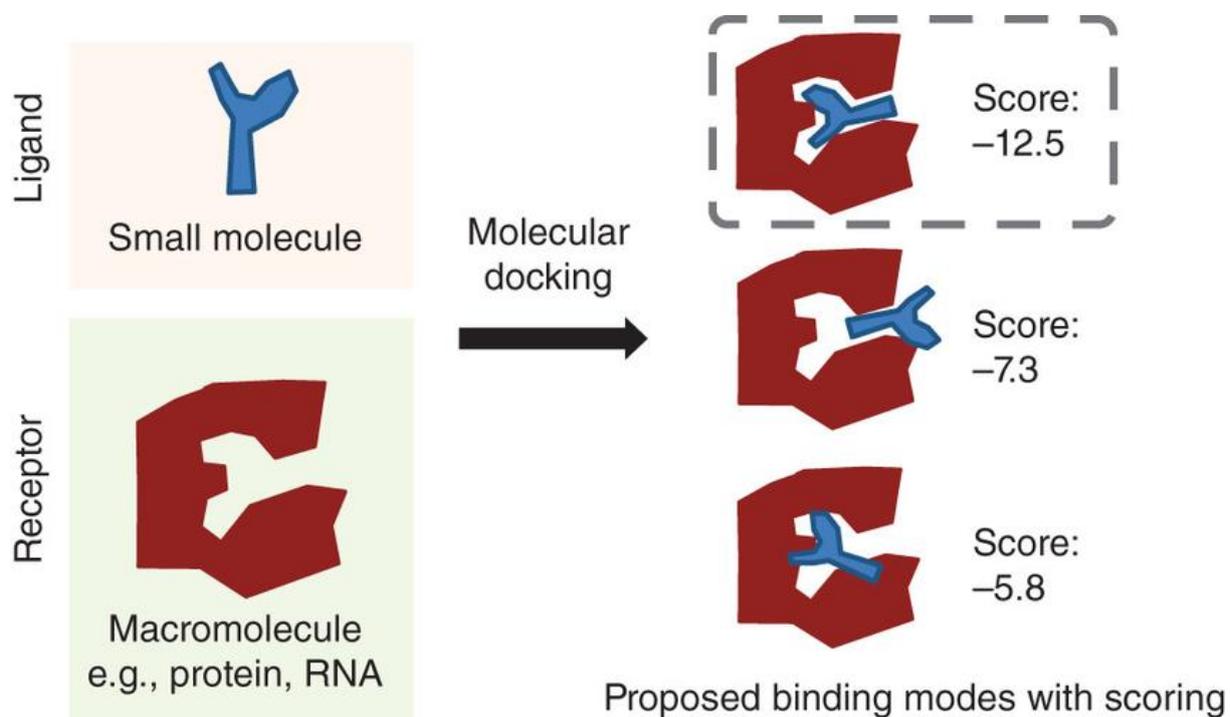


Figure II.3 : Principe générale d'un programme de docking.

II.4.4. Interaction protéine-ligand

Pour attacher un ligand à une protéine (docking) le problème est de prédire la conformation et l'orientation du ligand relative au site actif de la protéine cible. Le docking est la base de la reconnaissance moléculaire et du type de l'interaction. À chaque protéine cible de structure connue le docking se révèle être la clé dans le design de nouveaux médicaments [60-61-62]. Le ligand est considéré comme une petite molécule organique. Les interactions qui peuvent exister entre cette petite molécule (le ligand) et la protéine sont : des interactions polaires (liaisons hydrogène, liaisons ioniques) et des interactions hydrophobiques (contact entre des groupes hydrophobes).

Les interactions entre la protéine et le ligand suivent le concept de la complémentarité stérique. Récemment d'autres algorithmes, se basant sur le « matching of interacting group », ont été développés [63]. Ces algorithmes sont moins performants car ils modélisent les différentes interactions de la même manière. Pour contourner cette carence la technique utilisée considère trois niveaux de types d'interaction :

- les interactions du 3^{ème} niveau concernent les liaisons hydrogène et les liaisons ioniques, elles sont les plus restrictives géométriquement à cause de leur courte distance H-accepteur ---
----- H-donneur = 1.9Å .

- les interactions du 2^{ème} niveau sont les interactions hydrophobiques 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, on peut alors le séparer des interactions purement hydrophobiques. Des études sur l'interaction protéine-ligand ont montré que ce type d'interaction est le plus fréquent.

- les interactions du 1^{er} niveau sont les 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 [63-64].

Pour le placement du ligand dans le site actif de la protéine, les interactions du 3^{ème} niveau sont utilisées en premier, ensuite celles du 2^{ème} niveau, et enfin celles du 1^{er} niveau.

Références

- [1] Donald B. Boyd. K. B. Lipkowitz, Molecular Mechanics. The Method and Its Underlying Philosophy. J. Chem. Educ, 59, pp.269, 1982.
- [2] Schrödinger E., Ann.Phys. Leipzig, 79, pp.361, 489, 734, 1926.
- [3] Dugas H., Principes de base en modélisation moléculaire, Aspects théorique et pratiques, Chapitre3 introduction aux méthodes de minimisation d'énergie, quatrième édition, Librairie de L'Université de Montréal, 1996.
- [4] Dugas H., Principes de base en modélisation moléculaire, Aspects théoriques et pratiques, Hermann Dugas, quatrième édition, librairie de l'université de Montréal, 1996.
- [5] Allinger N., Q. C. P. E., 3, 32 (1983).
- [6] Pauling L., E.B.J. Wilson, Introduction to Quantum Mechanics with Applications to Chemistry, McGraw-Hill Book Company Inc., New York, 1935.
- [7] Cramer C.J., Essentials of Computational Chemistry - Theories and Models, Wiley, Chichester, U.K., 2004.
- [8] Raphaël Roux., « Etude probabiliste de systèmes de particules en interaction. Application à la simulation moléculaire » ; thèse de Doctorat de l'Université Paris-Est, 2010, pp.15.
- [9] Latifa CHEBIL., « Acylation des flavonoïdes par les lipases de Candida antarctica et de Pseudomonas cepacia : études cinétique, structurale et conformationnelle » ; thèse de Doctorat de l'INSTITUT NATIONAL POLYTECHNIQUE DE LORRAINE ,2006, pp.88.
- [10] NGUYEN Ngoc Bich., « Modélisation de structures atomiques et électroniques. Matériaux fonctionnels nano structurés pour la micro et l'optoélectronique. » ; thèse de Doctorat de l'UNIVERSITÉ DE CAEN Basse-Normandie, 2010, pp.41.
- [11] Corminboeuf C., Tran F., Weber J., Journal of Molecular Structure: Theochem, 2006, 762.
- [12] Hohenberg P., Kohn W., Phys. Rev, 1964, 136.
- [13] Chermette H., Coordination Chemistry Reviews, 1998, 699 :178-180 ,
- [14] Roger M., L. Belkhiri, P. Thuéry, Thérèse Arliguie, M. Fourmigué, A. Boucekkine, M. Ephritikhine Organometallics, 2005, 24 : 4941.

- [15] Kohn W., Rev. Mod. Phys , 1999, 71 :1253.
- [16] Pople J.A., Rev. Mod. Phys, 1999, 71 : 1267
- [17] Allinger N., Yuh Y., Lii J., Am J .Chem. Soc., 1989, 111, 8522.
- [18] Slater J.C., Phys. Rev., 1951, 81, 385-390.
- [19] Hohenberg P., Kohn W., Phys. Rev., 1964, 136, B864-B871.
- [20] Hohenberg. P. and Kohn. W., Phys. Rev., 1964, 136, B846.
- [21] Chaquin P., Manuel de Chimie théorique ; application à la structure et à la réactivité en chimie moléculaire, Ed, Ellipses .cours en ligne « Pratique de la chimie théorique », 2000.
- [22] Kohn W., Sham L., phys. Rev. A., 1965, 140, 1133.
- [23] Thomas L.H., Proc. Cambridge Phil. Soc., 1927, 23, 542.
- [24] Fermi E., Rend. Accad.,1927, 6, 602.
- [25] Becke A., Chem J., Phys.,1993, 98, 5648.
- [26] Kohn W., Sham L.J., Phys. Rev, 1965, 140 : 1133.
- [27] Corminboeuf C., Tran F., Weber J., Journal of Molecular Structure: Theochem, 2006, 762.
- [28] Vosko S.H., Wilk L., Nusair M., Can. J. Phys, 1980, 58 : 1200.
- [29] Becke A.D., Density functional exchange energy approximation with correct asymptotic behavior, Physical Review, 1988, 38, 3098–3100.
- [30] Becke A.D., Chem J. Phys., 1993, 98, 5648.
- [31].Meyer A.Y., Forrest R.F., Am J.Chem.Soc.1981, 103, 4664.
- [32] Allinger N.L., Adv. Phys. Org. Chem, , 1976, 13, pp.1.
- [34] Pèpe G., Guiliani G., Loustalet S., Halfon P., European journal of medicinal chemistry. temps important de vérifier si le champ choisi est adapte au système à étudier, 2002, 37, 865.
- [35] Margossian N., Le règlement REACH - La règlementation européenne sur les produits chimiques, Dunod / L'usine Nouvelle, Paris, 2007.

- [36] Crum Brown A., T.R. Fraser, On the connection between chemical constitution and physiological action. Part 1. On the physiological action of salts of the ammonium bases, derived from strychnia, brucia, thebia, codeia, morphia and nicotia. *Trans. Roy.Soc. Edinburgh*, , 1868, 25 : 151-203.
- [37] Hammett L.P., The effect of structure upon the reactions of organic compounds. Benzene derivatives. *Am J. Chem. Soc.*, (1937), 59 :96-103.
- [38] Brown A.C., Fraser T.R., On the connection between chemical constitution and Physiological Action; with special reference to the physiological action of the salts of the ammonium bases derived from Strychnia, Brucia, Thebaia, Codeia, Morphia, and Nicotia, *J. Anat. Physiol* .(1868), 2 : 224-242 .
- [39] Hansch C., Leo A., Hoekmann D., Exploring QSAR: hydrophobic, electronic and steric constants. *Washington, DC: Am. Chem. Soc.*, (1995).
- [40] Free S.M., Wilson J.W., A mathematical contribution to structure–activity studies. *Med J. Chem.* (1964), 7 :395-399.
- [41] Ghasemi J, Saaidpour S., Brown S.D., *Mol J. Struct. (Theochem)*, 2007, 805, 27-32.
- [42] Geladi P., Kowalski B.R., *Anal. Chim. Acta*, 1986, 185, 1-17.
- [43] Myles A.J., Feudale R.N., Liu Y., Woody N.A., Brown S.D., *Chemom J.*, 2004, 18, 275-285.
- [44] Duprat A.F., Huynh T., Dreyfus G., *Chem J. Inf. Comput. Sci.*, 1998, 38, 586-594.
- [45] Tetko I.V., Villa A.E.P., Livingstone D.J., *Chem J. Inf. Comput. Sci.*, 1996, 36, 794-803.
- [46] Gasteiger J., Zupan J., *Angew. Chem. Int. Ed. Engl.*, 1993, 32, 503-527.
- [47] Wiener H., Structural determination of paraffin boiling points. *Journal of Chemical Information and Computer Sciences*, 1947, 69, 17.
- [48] Randic M., On characterization of molecular branching. *Journal of the American Chemical Society*, 1975, 97, 6609.
- [49] Kier L.B., L.H. Hall, Derivation and significance of valence molecular connectivity, *J. Pharm., Sci.*, 1981, 70, 583.

- [50] Balaban A.T., Highly discriminating distance-based topological index. *Chemical Physics Letters.*, 1982, 89, 399.
- [51] Momany F.A., Ronte R., *Comp J. Chem.* 1992, 13, 888.
- [52] Alder B.J., Wainwright T.E., *Chem J. Phys*, 1957, 27, 1208.
- [53] Rahman A., *Phys. Rev*, 1964, 136, A 405.
- [54] Rahman A., Stillinger F.H., *J. Chem J. Phys.* 1971, 5, 3336.
- [55] Pople J. A., Beveridge, D.L. *Approximate Molecular Orbital Theory*, McGraw Hill, New York, (1970).
- [56] McQuarrie D. A., *Statistical Thermodynamics*, Harper Row Publ., New York, (1973).
- [57] Atkins P. W. *Physical Chemistry*, Ch. 20, 2nd edition, Freeman W.H., Co, San Francisco, (1982).
- [58] Bodor N., Buchwald P. Molecular size based approach to estimate partition properties for organic solutes. *J. Phys. Chem B.* 1997, 101, 3404.
- [59] Zhang L., Zhu H., Oprea T.I., Golbraikh A., Tropsha A., *QSAR Modeling of the Blood-Brain Barrier Permeability for Diverse Organic Compounds*, *Pharm. Res.* 2008, 25, 1902–1914.
- [60] Golbraikh A., Tropsha A., Beware of q^2 ! *J. Mol. Graph. Model.* 2002, 20, 269–276.
- [61] Martin T.M., Harten P., Young D.M., Muratov E.N., Golbraikh A., Zhu H., Tropsha A., Does Rational Selection of Training and Test Sets Improve the Outcome of QSAR Modeling *J. Chem. Inf. Model.* 2012, 52, 2570–2578.
- [62] Kuntz I.D., Blaney J.M., Oatley S.J., *Mol Biol J.*, 1982, 161, 269-288.
- [63] Kitchen D.B., H. Decornez H., Furr J.R., *Nat Rev Drug Discov*, 2004, 3, 935-949.
- [64] Yuriev E., Agostino M., Ramsland P.A., *Mol J. Recognit*, 2011, 24, 149-164.

Chapitre III
Étude Qualitative et Quantitative des
Propriétés QSAR d'une Série de Dérivés de
Phtalazine



III.1.Introduction

Bien que l'étude des relations structure-activité ait commencé à la fin du 19eme siècle, ce n'est qu'au début des années 60 que les travaux de Corwin Hansch ont proposé un modèle mathématique pour corrélérer l'activité biologique et la structure chimique.

Pendant ces quarante dernières années, ce domaine a largement été étudié et les données bibliographiques disponibles sur cette approche sont maintenant importantes [1].

La méthode QSAR inclut toutes les méthodes statistiques par lesquelles des activités biologiques (le plus souvent exprimées par des logarithmes des activités molaires équipotentielles) sont reliées avec les éléments structuraux (analyse de Free Wilson), les propriétés physico-chimiques (analyse de Hansch) ou différents paramètres liés à la notion de champ aidant à la description de la structure (3D QSAR).

Les informations extraites à partir des résultats d'étude de QSAR peuvent être utilisées pour obtenir une meilleure connaissance des structures moléculaires et probablement le mode d'action au niveau moléculaire. Ces informations peuvent alors être utilisées pour prévoir les propriétés physicochimiques et les activités biologiques de nouveaux composés ainsi que pour concevoir de nouvelles structures [2].

III.2. Etude qualitative des propriétés QSAR des dérivés de phtalazine

III.2.1. Structure chimique des dérivés de phtalazine

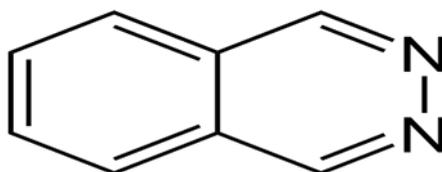
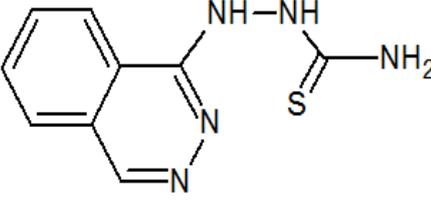
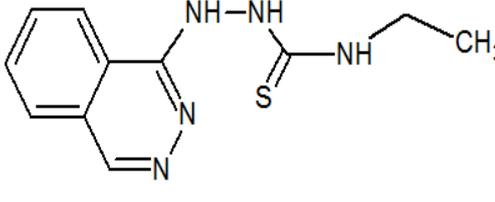
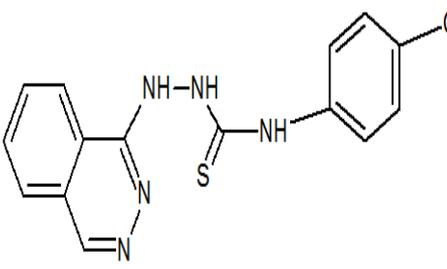
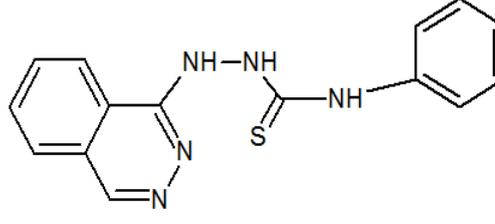
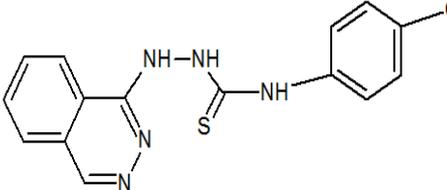
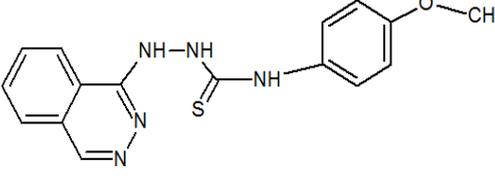
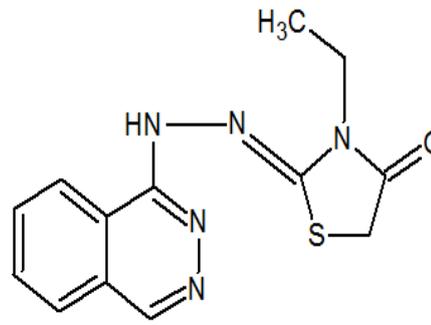
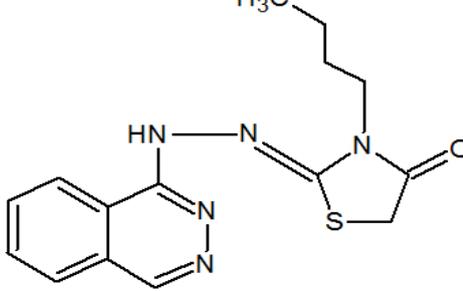
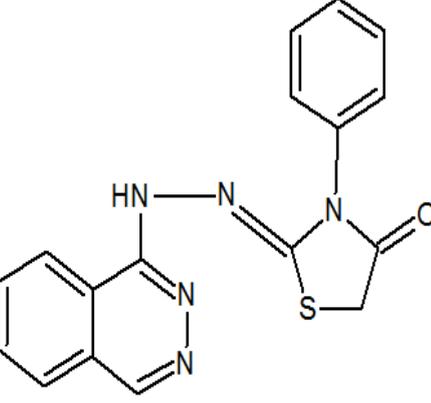
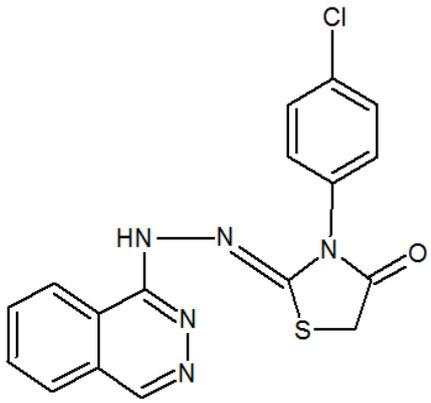


Figure III.1 : Structure générale de phtalazine

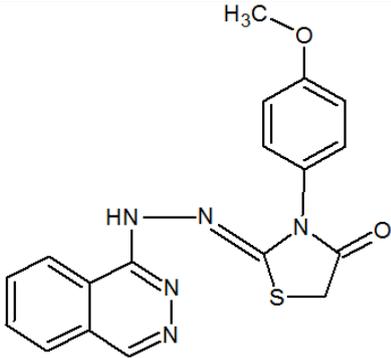
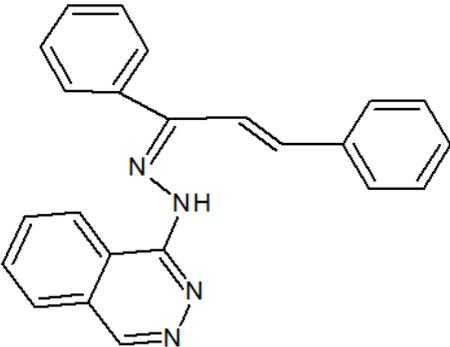
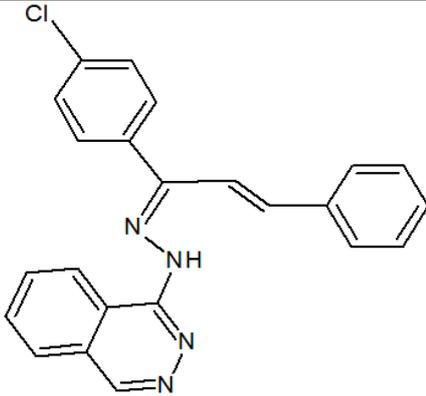
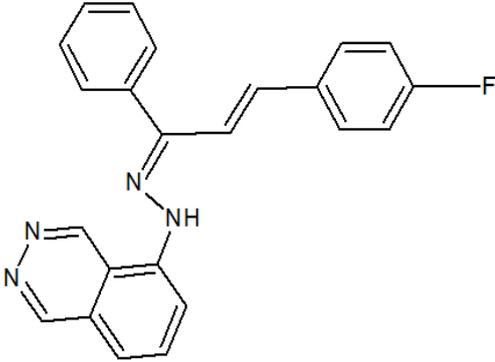
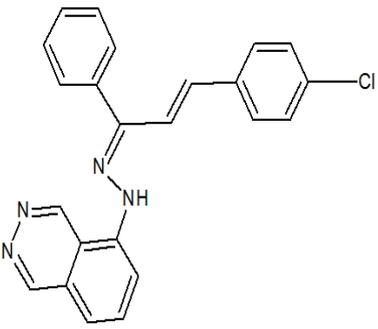
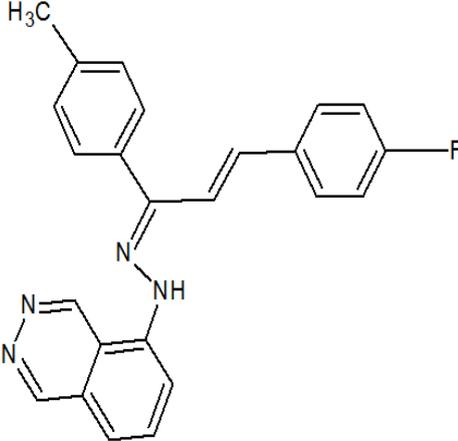
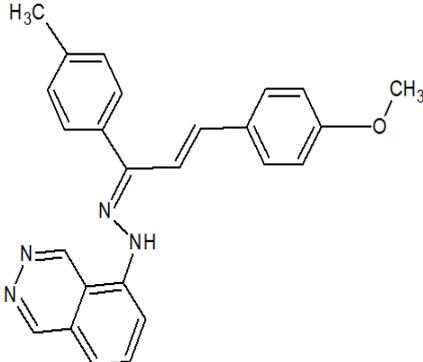
Tableau III.1 : Structures chimiques des dérivés de phtalazine désigné par molinspiration.

N°	Composé	N°	Composé
01		02	
03		04	
05		06	
07		08	

Étude qualitative et quantitative des propriétés QSAR d'une série de dérivés de phtalazine

09		10	
11		12	
13		14	
15		16	
17		18	

Étude qualitative et quantitative des propriétés QSAR d'une série de dérivés de phtalazine

19		20	
21		22	
23		24	
25			

III.2.2. Etudes des paramètres amphiphiliques de la série phtalazine

L'étude des propriétés physico-chimiques des dérivés de phtalazine (tableau III-2) a été réalisée par la méthode de QSAR dans ce travail, en utilisant le logiciel Hyperchem (8.0.8).

Dans ce travail, nous avons étudié sept propriétés physico-chimiques de vingt-cinq des dérivés de phtalazine (tableau. III.2) par rapport à leur activité biologique. Les propriétés concernées sont : le volume molaire (V), l'énergie d'hydratation (HE), coefficient de partage octanol / eau (log P), surface moléculaire (SAG), la réfractivité molaire (MR), polarisabilité (Pol) et le poids moléculaire (MW).

III.2.2.1. Volume et Surface Moléculaire

Le volume moléculaire et la surface moléculaire sont défini par détermination du volume (ou la surface) occupé par l'enveloppe de Van der Waal de la molécule considérée.

Le volume est défini par la relation :

$$V = \frac{MW}{d}$$

Où :

MW : est la masse moléculaire.

d : est la densité. [3]

III.2.2.2. Polarisation

On appelle polarisation d'une molécule, la facilité avec laquelle son nuage électronique se déforme par un champ électrique externe. La polarisation d'une molécule peut améliorer sa solubilité aqueuse. Cette caractéristique joue un rôle très important dans la modélisation de nombreuses propriétés moléculaires [4].

La polarisation moléculaire est définie par la relation suivante :

$$P(e) = \epsilon_0 \alpha E$$

Où :

P(e) : Coefficient de polarisation.

ϵ_0 : Constante diélectrique.

α : Moment dipolaire électrique induit.

III.2.2.3. Energie d'hydratation

L'hydratation est un des facteurs clé qui déterminent la conformation, la stabilité et les interactions des protéines et donc leurs fonctions biologiques. Par exemple, dans les liaisons qui impliquent des sites d'interactions principalement hydrophobes, l'hydratation est la contribution énergétique principale.

Il a été montré que l'hydratation représente jusqu'à 50% de l'affinité et de la spécificité dans les mécanismes de reconnaissance moléculaire. L'association entre une protéine et un ligand (ou principe actif) est généralement accompagnée par un départ de molécules d'eau. Celui-ci fournit une contribution à la fois entropique et enthalpique à l'énergie qui provient de la différence d'activité de l'eau entre la solution et la sphère d'hydratation de la protéine et du ligand. Par conséquent un changement de l'activité de l'eau peut modifier les propriétés de la liaison. Ces mécanismes ont un impact important dans la conception de principes actifs qui doivent posséder une affinité importante et spécifique pour une cible thérapeutique donnée [5].

III.2.2.4. Réfractivité Molaire

Un autre moyen d'évaluer le facteur stérique est fourni par un paramètre qu'on appelle la réfractivité molaire (RM).il s'agit cette fois une mesure du volume qui est occupé par un atome ou par un groupe d'atomes.la RM s'obtient grâce à l'équation suivante :

$$MR = \frac{(n^2 - 1)}{(n^2 + 2)} \times \frac{MM}{d}$$

OU :

n : est l'indice de réfraction

MM : masse molaire moléculaire

Le terme **MM /d** définit un volume.

Tandis que **(n²-1) / (n²+2)** représente un facteur de correction qui prend en considération la facilité avec laquelle le substituant peut être polarisé. Ceci est particulièrement important si le substituant possède des électrons π ou des paires électroniques libres [6].

Étude qualitative et quantitative des propriétés QSAR d'une série de dérivés de phtalazine

III.2.2.5. Hydrophobicité

Le caractère hydrophobe d'une drogue est crucial en ce qui concerne la facilité avec laquelle elle traverse les membranes cellulaires et peut également être un facteur important lors de ses interactions avec le récepteur, il est donc important de pouvoir le quantifier.

III.2.2.6. lipophilicité

La lipophilie est l'affinité d'une substance pour les solvants apolaires comme les lipides (corps gras). Une substance lipophile (ou hydrophobe) est soluble dans un corps gras. Par contre, les substances lipophiles sont insolubles dans les solvants polaires comme l'eau. On dit qu'elles sont hydrophobes.

III.2.2.7. Coefficient de partage (P)

Il serait avantageux de n'être pas obligé de synthétiser chaque composé pour mesurer son coefficient de partage.

L'algorithme du coefficient rapport :

P = concentration de la drogue dans l'octanol / concentration de la drogue dans l'eau.

Permet d'estimer la biodisponibilité d'une molécule

- $0 < \text{Log } P < 3$: Activité biologique optimale (perméabilité, solubilité).
- $\text{Log } P < 0$: Composés trop hydrophiles (mauvaise perméabilité de bicouche lipidique).
- $\text{Log } P > 3$: Composés trop lipophiles (mauvaise solubilité aqueuse).

Dans cette partie du travail, nous nous intéressons à l'étude de la lipophilicité de phtalazine biologiquement actives à une interface hydrophobe/hydrophile.

L'optimisation des composés de phtalazine a été effectuée à l'aide du logiciel hyperchem 8.0.8 par les méthodes de la mécanique moléculaire en utilisant le champ de force MM+ et semi empirique PM3.

Tableau III.2 : Paramètres physicochimiques des dérivés de phtalazine.

Composés	Surface moléculaire (\AA^2)	Volume moléculaire (\AA^3)	Energie d'hydratation (Kcal/mol)	Log p	Réfractivité (\AA^3)	Polarisabilité (\AA^3)	Masse moléculaire (UMA)
01	487.96	805.84	-18.18	-0.56	89.40	28.70	300.33
02	537.16	927.66	-21.90	-0.12	110.27	36.64	381.41
03	549.24	941.34	-18.03	0.11	113.45	37.65	377.42
04	587.02	995.61	-19.21	-0.07	110.68	36.94	378.41
05	582.01	1012.88	-16.59	0.92	120.94	40.61	406.46

Étude qualitative et quantitative des propriétés QSAR d'une série de dérivés de phtalazine

06	425.60	684.49	-10.39	0.69	76.27	26.12	239.25
07	452.18	732.95	-14.83	1.02	81.12	28.06	246.27
08	474.80	806.22	-17.25	1.53	92.72	32.99	295.34
09	398.03	627.98	-21.92	0.92	66.02	24.33	219.26
10	460.87	738.66	-17.14	1.67	75.54	28.00	247.32
11	503.51	827.44	-15.73	2.53	84.67	31.67	275.37
12	351.07	569.83	-14.16	1.96	95.68	33.99	295.36
13	353.43	583.29	-12.92	1.60	94.10	34.31	316.79
14	387.85	638.98	-17.29	0.97	102.05	36.46	325.39
15	487.84	797.89	-11.08	1.47	82.01	30.32	287.34
16	430.99	693.22	-15.61	0.89	72.36	26.65	259.29
17	557.30	924.84	-12.38	1.62	101.85	36.31	335.38
18	535.70	924.04	-13.79	1.40	106.56	38.24	369.83
19	516.38	896.39	-15.29	0.63	108.22	38.78	365.41
20	497.88	862.92	-12.25	4.33	121.88	42.12	350.42
21	524.13	909.61	-12.11	4.11	126.60	44.04	384.87
22	498.57	863.89	-12.84	3.73	122.01	42.03	368.41
23	502.20	874.76	-12.23	4.11	126.60	44.04	384.87
24	526.40	916.12	-10.79	3.88	126.29	43.86	382.44
25	527.28	918.49	-12.46	3.49	132.54	46.42	394.48

Grande valeur : **verte** ; Petite valeur : **rouge**

Interprétation des résultats

D'après les résultats du calcul, on remarque que les valeurs du volume sont proportionnelles aux valeurs des surfaces. Le terme MM/ρ définit un volume, [7] suivant ce terme là on voit que le volume moléculaire est proportionnel aux valeurs de la masse molaire.

La polarisabilité et la réfractivité molaire augmentent relativement avec la taille et le poids moléculaire des dérivés de Phtalazine étudiés (Tableau III.2). Ce résultat est en accord avec la formule de Lorentz-Lorenz qui donne une relation entre polarisabilité, la réfractivité molaire et le volume [8].

Cette relation montre que la polarisabilité est proportionnelle avec le volume et la masse moléculaire. Aussi que la réfractivité molaire et la polarisabilité augmentent avec le volume et la masse moléculaire. Par exemple, le composé 25 porte des substituant volumineux a des valeurs importantes de la polarisabilité (46.42 \AA^3), de la réfractivité molaire (132.54 \AA^3) et de la masse moléculaire (394.48 UMA).

Le composé 9 est le petit dérivé dans la série étudiée, il possède des valeurs faibles de la polarisabilité (24.33 \AA^3), de la réfractivité molaire (66.02 \AA^3) et de la masse moléculaire (219.26 UMA).

L'énergie d'hydratation en valeur absolue, la plus importante valeur (21.90 kcal/mole) a été observée pour le composé 9. La plus faible valeur d'énergie d'hydratation (10.39

kcal/mol) a été observée pour le composé 6. En effet, dans les milieux biologiques, les molécules polaires sont entourées par les molécules d'eau où les liaisons hydrogènes peuvent être établies entre les molécules d'eau et ces molécules. Les sites donneurs de protons interagissent avec l'atome d'oxygène de l'eau et les sites accepteurs de protons interagissent avec l'atome d'hydrogène. Ces molécules hydratées sont partiellement déshydratées avant et lors de leurs interactions. Ces interactions de faible énergie et généralement réversibles, que l'on observe généralement avec les différents récepteurs biologiques [9].

La présence des groupes hydrophobes dans le radical de phtalazine provoque la diminution de l'énergie d'hydratation ; ainsi que, la présence des groupes hydrophiles augmentent l'énergie d'hydratation comme dans le composé 9 qui possède trois sites donneurs des protons (2 NH, 1NH₂) et un site accepteur de proton (S) (Tableau III-2).

La lipophile est un facteur important dans les processus de solubilité, d'absorption, de distribution, de métabolisme et d'excrétion, ainsi que, l'activité pharmacologique. Hansch et Leo ont démontrés que les molécules très lipophiles se réparties et conservées à l'intérieur des couches lipidiques des membranes cellulaires [10].

Pour une bonne biodisponibilité orale, le log P doit être supérieur à zéro et inférieur à 3. Les composés 20, 21, 22, 23, 24 et 25 possèdent des valeurs élevées de log P (4.33, 4.11, 3.73, 4.11, 3.88, 3.49) respectivement, Ces composés possèdent une faible solubilité aqueuse.

Les composés 1,2 et 4 présentent des faibles coefficients de partition (-0.59, -0.12, -0.07) respectivement. Ces composés donnent une bonne solubilité, mais une faible absorption et pénétration, en raison de leur mauvaise perméabilité par la diffusion passive. Ces composés sont une tendance élevée d'éliminer par les reins [9].

III.3.Représentation des calculs « drug-like » sur la base de Lipinski

Ainsi, Lipinski et al. Définirent dès 1997 des règles simples permettant d'identifier rapidement et à grande échelle des molécules à caractère « drug-like », plus susceptibles de présenter les caractéristiques de biodisponibilité nécessaires au développement d'un candidat médicament. [11] Ces règles, communément appelées « règles de Lipinski » ou « la règle de 5 », comportent cinq critères physico-chimiques qui décrivent la molécule :

- Une masse moléculaire (MW) maximum de 500g/mol
- Un maximum de 5 liaisons H-donneurs (DH) (OH ou NH)
- Un maximum de 10 liaisons H-accepteurs (AH) (O ou N)

Étude qualitative et quantitative des propriétés QSAR d'une série de dérivés de phtalazine

- nombre de liaisons retables (LF) ≤ 15
- Un coefficient de partition octanol / eau "logP" ou une lipophile ≤ 5 .

Particulièrement, la mesure du logP caractérise la polarité du composé (estimée par le coefficient de partition octanol/eau), permettant ainsi d'estimer la distribution du composé dans l'organisme. Les molécules hydrophobes (hautes valeurs du logP) sont principalement distribuées dans les régions hydrophobes, comme la bicouche lipidique des cellules. Inversement, les molécules hydrophiles sont retrouvées principalement dans des régions aqueuses, comme le sérum sanguin. La 5e règle de Lipinski stipule que les adjuvants et assimilés font exception aux quatre autres règles [11].

Le tableau (III.3) les paramètres de Lipinski des dérivés de phtalazine.

Composés	MW (UMA)	Log p	HBD	HBA	N° violation
01	300.33	-0.56	3	6	0
02	381.41	-0.12	2	8	0
03	377.42	0.11	2	7	0
04	378.41	-0.07	2	8	0
05	406.46	0.92	2	8	0
06	239.25	0.69	1	3	0
07	246.27	1.02	1	4	0
08	295.34	1.53	1	5	0
09	219.26	0.92	4	5	0
10	247.32	1.67	3	5	0
11	275.37	2.53	3	5	0
12	295.36	1.96	3	5	0
13	316.79	1.60	3	5	0
14	325.39	0.97	3	6	0
15	287.34	1.47	1	6	0
16	259.29	0.89	1	6	0
17	335.38	1.62	1	6	0
18	369.83	1.40	1	6	0
19	365.41	0.63	1	7	0
20	350.42	4.33	1	4	0
21	384.87	4.11	1	4	0

Étude qualitative et quantitative des propriétés QSAR d'une série de dérivés de phtalazine

22	368.41	3.73	1	4	0
23	384.87	4.11	1	4	0
24	382.44	3.88	1	4	0
25	394.48	3.49	1	5	0

HBD : liaisons hydrogène donneurs, **HBA** : liaisons hydrogène accepteurs.

Interprétation des résultats

La règle de Lipinski est la plus utilisée pour caractériser les composés « drug-like ». Cette règle a pour but d'identifier les composés posant des problèmes d'absorption et de perméabilité. Les composés ne validant pas au moins deux de leurs critères ont de très fortes chances d'avoir des problèmes d'absorption ou de perméabilité.

Tous les composés de la série ont des masses molaires inférieures à 500 Da et tous les composés sont vraisemblablement solubles et traversent facilement les membranes cellulaires et ils conformes aux règles 2 et 3 pour tous les composés. Ils sont sans doute peu polaires et facilement absorbés ; et ils répondent à la règle 4 sont solubles en solution aqueuse et donc capables d'atteindre la surface des membranes et ont une tolérance gastrique. Les médicaments les plus efficaces sont donc d'habitude le fruit d'un compromis : ils ne sont ni trop lipophiles, ni trop hydrophiles. [12]

D'après, ces critères là on peut dire que les composés de la série sont acceptables d'être administrés par voie orale. [12]

III.4. Les règles de Veber

Par la suite, d'autres critères ont été mis en place pour compléter et ajuster les règles de Lipinski dans la sélection de composés « drug-like ». Ainsi, Veber et al. choisissent d'utiliser les deux critères suivants pour une biodisponibilité orale idéale [13] :

- Les liaisons rotatives : $NRB < 10$.
- La surface polaire : $PSA < 140 \text{ \AA}^2$.

Les règles Veber suggèrent que la flexibilité moléculaire et la surface polaire (PSA) sont des déterminants importants de la biodisponibilité orale. [14]

Il est bien connu qu'une biodisponibilité orale élevée est un facteur important pour le développement de molécules bioactives en tant qu'agents thérapeutiques. La réduction de la

Étude qualitative et quantitative des propriétés QSAR d'une série de dérivés de phtalazine

flexibilité moléculaire (mesurée par le nombre de liaisons rotatives) et la faible surface polaire sont des prédicateurs importants de la bonne biodisponibilité orale. [15]

Le nombre de liaisons rotatives a été défini comme une simple liaison, non pas dans un cycle, liée à un atome non terminal (c'est-à-dire non hydrogène). Sont exclues du compte les liaisons amide C-N en raison de leur barrière à l'énergie de rotation élevée. [14]

D'autre part, la surface polaire qui est formée par des atomes polaire d'une molécule. C'est un descripteur qui montre une bonne corrélation avec le transport moléculaire passif à travers les membranes, et permet donc d'estimer les propriétés de transport des médicaments.

Egalement, le PSA a été utilisé pour calculer le pourcentage d'absorption (%ABS) selon l'équation :

$$\%ABS = 109 \pm 0,345x \text{ PSA. [16]}$$

Le tableau (III.4) suivant présente les résultats de la règle de Veber appliquée aux phtalazine étudiés.

Composés	NRB	PSA (A ^{o2})	Veber score	%ABS
1	3	97.98	0	75.197
2	5	110.01	0	71.047
3	5	96.87	0	75.580
4	5	109.76	0	71.132
5	5	109.76	0	71.132
6	2	37.81	0	95.960
7	2	61.60	0	87.750
8	5	64.12	0	86.880
9	3	75.86	0	82.830
10	5	61.87	0	87.654
11	7	61.87	0	87.654
12	5	61.87	0	87.654
13	5	61.87	0	87.654
14	6	71.10	0	84.470
15	3	70.48	0	84.684
16	5	70.48	0	84.684
17	3	70.48	0	84.684

Étude qualitative et quantitative des propriétés QSAR d'une série de dérivés de phtalazine

18	3	70.48	0	84.684
19	4	79.72	0	81.500
20	5	50.17	0	91.691
21	5	50.17	0	91.691
22	5	50.17	0	91.691
23	5	50.17	0	91.691
24	5	50.17	0	91.691
25	6	59.41	0	88.503

NRB : Le nombre des liaisons rotatives, **PSA** : La surface polaire

Interprétation des résultats

D'après les résultats de ce tableau on remarque que tous les valeurs de nombre de liaison rotative sont inférieures à 10 dans les composés étudiés. Alors, tous les composé sont flexible et capable de changé leur conformation.

Pour les résultats de la surface polaire on remarque que tous les composés de la série étudiée ont des valeurs inférieures À 140 ($A^{\circ 2}$) ce qui montre la bonne prédiction de la biodisponibilité orale et le transport à travers les membranes.

Alors, toutes les conditions sont vérifiées donc tous les composés de la série phtalazine étudiés sont en accord avec la règle de Veber.

On a aussi les valeurs de pourcentage d'absorption (%ABS), on peut observer que tous les composés présente une grande pourcentage d'absorption allant 71,047% du composé 3 à 91,691% du composé (20-21-22-23-24) Cela signifie que ces composés avoir une bonne perméabilité membranaire cellulaire.

III.5. Efficacité de ligand « LE »

L'efficacité du ligand est une mesure de l'énergie de liaison par atome d'un ligand à son partenaire de liaison, tel qu'un récepteur ou une enzyme. [17]

L'efficacité des ligands est utilisée dans les programmes de recherche de découverte de médicaments pour aider à restreindre l'attention aux composés de plomb avec des combinaisons optimales de propriétés physicochimiques et de propriétés pharmacologiques [18].

Mathématiquement, l'efficacité du ligand (LE) peut être définie comme le rapport entre l'énergie libre de Gibbs (ΔG) et le nombre d'atomes non hydrogène du composé :

Étude qualitative et quantitative des propriétés QSAR d'une série de dérivés de phtalazine

$$LE = (\Delta G) / NH$$

Où $\Delta G = -RT \ln K_i$ et NH est le nombre d'atomes autres que l'hydrogène. [19] Il peut être transformé en l'équation : [20]

$$LE = 1,4 (-\log IC_{50}) / NH$$

Tableau III.5 : l'efficacité de ligand de la série phtalazine.

Composés	pIC ₅₀	NH	LE
1	6.853	21	0.456
2	6.880	27	0.357
3	7.075	27	0.366
4	7.065	27	0.396
5	8.173	27	0.423
6	8.221	18	0.639
7	8.387	19	0.618
8	8.356	22	0.531
9	7.113	15	0.663
10	7.119	17	0.586
11	7.283	16	0.637
12	7.187	21	0.479
13	7.193	22	0.457
14	7.275	22	0.462
15	7.327	17	0.603
16	7.638	17	0.629
17	7.346	22	0.467
18	7.376	23	0.449
19	7.443	23	0.453

Étude qualitative et quantitative des propriétés QSAR d'une série de dérivés de phtalazine

20	8.251	27	0.427
21	8.200	28	0.410
22	8.161	27	0.423
23	8.337	27	0.432
24	8.187	27	0.424
25	8.267	28	0.413

Interprétation des résultats

Lorsqu'on observe les résultats de ce tableau on remarque que les valeurs de l'efficacité de ligand augmentent quand le nombre des atomes lourds diminue, alors pour obtenir une efficacité de ligand très forte nécessite nombre des atomes lourds faible on explique cette relation par la corrélation qui existe entre la taille de la molécule et leur propriétés physicochimique. Par contre l'efficacité de ligand varié proportionnellement avec les valeurs de pIC₅₀.

On étudié dans cette partie, l'efficacité de ligand pour pénaliser les grands composés sur les petits composés ayant une puissance similaire, car les molécules de plus grande taille ont des propriétés physicochimiques et des propriétés ADME faible [21,22]. D'après les résultats de tableau on a par exemple les composé 9, 10, 11, 15, 16 qui ont les petits nombres d'atome lourds et une grande efficacité de ligand suivante (0.663, 0.586, 0.637, 0.603, 0.629), et il y a aussi les deux composés 2, 3, 4 ayant une grande nombre d'atome lourd et une petite valeur de l'efficacité de ligand (0.357, 0.366, 0.396) ces deux composé peut avoir des mauvaises propriétés physicochimiques et ADME.

III.6.Efficacité lipophile de ligand (LLE)

Efficacité lipophile (LiPE), parfois appelée efficacité ligand-lipophilie (LLE) est un paramètre utilisé dans la conception de médicaments et la découverte de médicaments pour évaluer la qualité des composés de recherche, reliant la puissance et la lipophilicité dans une tentative d'estimer drug-likeness. [23] Pour un composé donné, LiPE est défini comme le pIC₅₀ (ou pEC₅₀) d'intérêt moins le LogP du composé.

$$LiPE = pIC_{50} - LogP$$

Étude qualitative et quantitative des propriétés QSAR d'une série de dérivés de phtalazine

Tableau III.6 : l'efficacité lipophile de ligand de la série phtalazine.

Composés	PIC50	LogP	LLE
1	6.853	-0.56	7.413
2	6.880	-0.12	7.000
3	7.075	0.11	6.965
4	7.065	-0.07	7.135
5	8.173	0.92	7.253
6	8.221	0.69	7.201
7	8.387	1.02	7.367
8	8.356	1.53	6.826
9	7.113	0.92	6.193
10	7.119	1.67	5.449
11	7.283	2.53	4.753
12	7.187	1.96	5.227
13	7.193	1.60	5.593
14	7.275	0.97	6.305
15	7.327	1.47	5.857
16	7.638	0.89	6.748
17	7.346	1.62	5.726
18	7.376	1.40	5.976
19	7.443	0.63	6.813
20	8.251	4.33	3.921
21	8.200	4.11	4.090
22	8.1611	3.73	4.431

Étude qualitative et quantitative des propriétés QSAR d'une série de dérivés de phtalazine

23	8.337	4.11	4.227
24	8.187	3.88	4.307
25	8.267	3.49	4.777

Interprétation des résultats

D'autre part nous avons étudié l'efficacité lipophile pour maximiser la puissance tout en maintenant la lipophilie aussi faible que possible, en raison de l'association entre une lipophilie élevée et plusieurs problèmes, y compris une faible solubilité, perméabilité membranaire, stabilité métabolique, ...etc.

Comme le lipophile est le principal facteur de la promiscuité des composés, LipE devraient être plus sélectifs pour les composés optimisés. Il est suggéré de cibler un LipE dans une gamme de 5-7 ou même plus [24]

D'après les résultats de tableau (III 7), les composés (2, 3, 8, 9, 10, 12, 13, 14, 15, 16, 17, 18, 19) qui sont encore au-dessus de la fourchette proposée de 5-7. C'est ce qui explique que log P est le paramètre le plus important pour l'activité inhibitrice contre le cancer de la peau.

III.7. Etude quantitative de la relation structure-activité

Pour élaborer un modèle QSPR/QSAR nous avons besoin d'une méthode d'analyse de données, cette méthode permet de quantifier la relation qui existe entre la propriété/activité et la structure en utilisant des descripteurs moléculaires.

Il existe plusieurs méthodes pour construire un modèle et analyser les données statistiques de ce dernier, certaines sont linéaires telles que la régression linéaire multiple (MLR), la régression aux moindres carrés partiels (PLS), ces méthodes sont disponibles dans des logiciels tels que, Excel, Origin Microcal, Minitab, Statistica, SPSS,...

La méthode utilisée dans nos études est la méthode de Régression Linéaire Multiple (MLR) implémentée sur SPSS. [25]

III.7.1. Descripteurs quantique/électronique

L'étude des propriétés quantiques des dérivés phtalazine (tableau III-7) a été réalisée par la méthode de QSAR, en utilisant le logiciel Gaussian09 pour calculer l'énergie total, les énergies des orbitales frontières, moment dipolaire et les charges de cette série des molécules.

III.7.1.1. Energies des orbitales HOMO et LUMO

Étude qualitative et quantitative des propriétés QSAR d'une série de dérivés de phtalazine

Deux orbitales moléculaires, appelées orbitales frontières, jouent un rôle particulier :

- HOMO (Highest Occupied Molecular Orbital) Traduit le caractère électro-donneur (nucléophile) de la molécule. Plus l'énergie de cette OM est élevée, plus la molécule cédera facilement des électrons.
- LUMO (Lowest Unoccupied Molecular Orbital) Traduit le caractère électro-accepteur (électrophile) de la molécule. Plus l'énergie de cette OM est faible, plus la molécule acceptera facilement des électrons.

Dans le cas de l'éthylène, l'orbitale π joue le rôle de HOMO et l'orbitale π^* le rôle de LUMO.

[26]

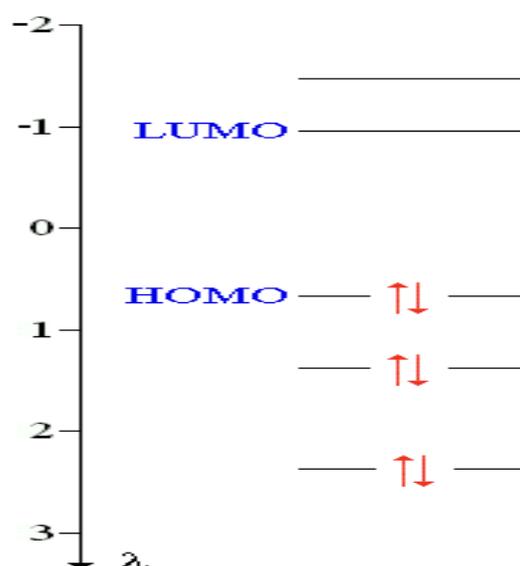


Figure III.2 : les orbitales frontières HOMO/LUMO

III.7.1.2. Moment dipolaire

La notion de moment dipolaire en physique et en chimie repose sur l'existence de dipôles électrostatiques. Il s'agit d'une répartition hétéroclite de charges électriques telles que le barycentre des charges positives ne coïncide pas avec celui des charges négatives. Le dipôle le plus simple est donc un couple de deux charges, de signes opposés, séparées par une distance non nulle. Un moment dipolaire est un vecteur noté généralement P ou D.

Dans le cas de deux charges identiques mais de signes opposés, donc valant respectivement $+q$ et $-q$, séparées par une distance d , le moment dipolaire est décrit par le vecteur $\mathbf{P}=\mathbf{q}\cdot\mathbf{d}$. Le vecteur d étant orienté de la charge négative à la charge positive. Plus généralement, un moment dipolaire est la somme de n vecteurs $\mathbf{P}_i=\mathbf{q}_i\mathbf{r}_i$ où q_i et r_i sont respectivement les valeurs de n charges q_i ayant n vecteurs positions r_i . [27]

Étude qualitative et quantitative des propriétés QSAR d'une série de dérivés de phtalazine

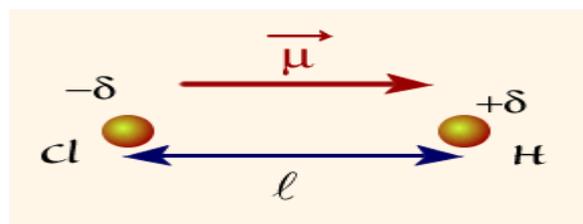


Figure III.3 : moment dipolaire entre deux atomes

III.7.1.3. Energie totale

Pour une molécule isolée à l'état fondamental, l'énergie totale calculée, notée E_T , mesurée en eV, peut être utilisée comme descripteur moléculaire quantique. Cette énergie approximative a été calculée pour une conformation optimisée de la géométrie la plus stable dont la structure d'énergie est minimale. [28]

L'étude des propriétés électroniques des dérivés de phtalazine a été réalisée par le logiciel Gaussian 09, la base utilisée au cours de calculs est : B3LYP/6-31++G(d,p).

Tableau III.7 : Les valeurs des descripteurs moléculaires utilisés dans l'analyse de régression.

composés	pIC50	HOMO	LUMO	MD	E_T	N1	N2
1	6.853	-0.276	-0.014	7.045	-1308.327	-0.291	-0.236
2	6.886	-0.280	-0.019	6.254	-1592.497	-0.289	-0.237
3	7.076	-0.269	-0.012	8.257	-1555.427	-0.292	-0.236
4T	7.065	-0.268	-0.015	9.116	-1571.475	-0.322	-0.230
5T	8.173	-0.269	-0.014	2.019	-1650.125	-0.293	-0.237
6T	8.221	-0.264	-0.018	6.670	-803.629	-0.331	-0.230
7	8.387	-0.268	-0.016	10.201	-796.636	-0.328	-0.228
8	8.356	-0.130	-0.127	4.352	-972.796	-0.303	-0.237
9	7.113	-0.155	-0.155	8.217	-1020.221	-0.274	-0.235
10	7.119	-0.191	-0.163	4.900	-1098.957	-0.307	-0.238
11	7.283	-0.191	-0.162	5.000	-1177.592	-0.307	-0.238
12	7.187	-0.191	-0.166	8.786	-1251.376	-0.316	-0.232
13	7.193	-0.132	-0.144	2.461	-1672.203	-0.302	-0.237
14	7.275	-0.208	-0.082	7.788	-1365.893	-0.265	-0.247
15	7.327	-0.266	-0.019	6.122	-1250.394	-0.296	-0.231
16	7.638	-0.263	-0.025	3.670	-1329.035	-0.312	-0.236
17	7.346	-0.248	-0.013	5.611	-1402.817	-0.317	-0.234
18	7.376	-0.263	-0.030	2.546	-1862.415	-0.314	-0.236
19	7.443	-0.259	-0.023	5.426	-1517.351	-0.317	-0.236
20	8.251	-0.251	-0.001	3.595	-1106.294	-0.309	-0.234
21T	8.200	-0.255	-0.005	2.411	-1565.889	-0.308	-0.234
22	8.161	-0.253	-0.005	3.928	-1205.534	-0.313	-0.234
23	8.337	-0.254	-0.004	3.862	-1565.889	-0.309	-0.234
24T	8.187	-0.247	-0.010	4.264	-1237.925	-0.310	-0.221
25T	8.267	-0.245	-0.099	5.088	-1260.146	-0.310	-0.233

T : test set de validation externe, **N1** : la charge de l'atome d'azote n°1, **N2** : la charge de l'atome d'azote n°2

III.7.2. Le modèle QSAR

Le succès de l'approche SAR pour la conception de médicaments ne dépend pas seulement des connaissances et l'expérience du groupe de conception de médicaments, mais aussi liée à la chance. QSAR est un tenter d'enlever le facteur chance de la conception de médicaments en établissant une relation dans le forme d'une équation mathématique entre l'activité biologique et le mesurable paramètres physicochimiques d'un médicament qui représente ses propriétés telles que la lipophilie, la forme et la distribution des électrons, qui ont des effets majeurs sur l'activité [29-30].

Les relations quantitatives structure-propriété (QSPR) sont également utilisées, en la propriété spécifique autre que l'activité biologique pIC50 [31] est considérée. Si une équation est formée, alors un chimiste médicinal pourrait déterminer à partir de l'équation la valeur du paramètre, et d'où les structures, qui optimiseraient l'activité. Ces prédictions permettent médicaments chimistes de faire un choix plus éclairé quant à ce que les analogues à concevoir et synthétiser. Évidemment, cela pourrait considérablement réduire le coût du développement de médicaments.

La relation entre les variables indépendantes (descripteurs physico-chimiques (Tableau III.2), descripteurs quantiques (Tableau III.7)) et les variables dépendantes (activités biologiques pIC50) a été déterminée statistiquement en utilisant l'analyse de régression. Dans ce travail, l'analyse de régression linéaire multiple MLR de descripteurs moléculaires a été réalisée en utilisant le logiciel statistique SPSS 20.

La corrélation entre les activités biologiques et les descripteurs physico-chimique et quantique exprimés par l'équation (V.1) suivante :

$$\begin{aligned} \text{Log (1/IC50)} &= 26,998-0,05 \text{ SAG}+0,029 \text{ V}+0,096 \text{ HE}-0,062 \text{ MR}+0,131 \text{ POL} \\ &- 1,377 \text{ HOMO}+0,088 \text{ MD}+0,001 \text{ E}_{\text{TOT}}+63,25 \text{ N2} \\ \mathbf{n} &= 19, \mathbf{R} = 0.948, \mathbf{R}^2 = 0.899, \mathbf{SE} = 0.233, \mathbf{F} = 8.919, \mathbf{Q} = 4.068 \end{aligned}$$

Étude qualitative et quantitative des propriétés QSAR d'une série de dérivés de phtalazine

Où : n est le nombre des composés, R est le coefficient de corrélation, R^2 est le coefficient de détermination, SE est l'erreur standard de l'estimation, F est la statistique Fischer et Q est la qualité de l'ajustement ou bien l'adaptation.

Interprétation du modèle

Le développement d'une série de descripteurs qui caractérisent les structures des 25 composés étudiés en vue de les relier à l'affinité de phtalazine, des outils d'analyse de données sont alors employés pour orienter le choix des descripteurs et mettre en place le modèle à proprement dit.

Le modèle QSAR doit considérer un $R^2_{cv} > 0.6$ pour qu'il soit valide. Les valeurs de $R = 0.948$ et de $R^2_{cv} = 0.899$, nous ont permis d'indiquer fermement la corrélation entre les différents paramètres (variables indépendantes) ayant une activité spécifique de protéine tyrosine kinase.

La valeur de F calculée pour le modèle QSAR généré dépasse la valeur F de tableau critique par grande marge comme on le souhaite pour une régression significative. De plus, la valeur de F a été jugée statistiquement significative au niveau de 95 % pour ce modèle. La valeur positive de facteur de qualité (Q) suggère son haut pouvoir prédictif, ainsi que le faible écart type du modèle QSAR démontre la précision du modèle.

D'après l'équation du modèle (III.1), les coefficients positives V , HE , POL , MD , E_T et N_2 montrent que toute augmentation dans les valeurs de ces paramètres entraîne une augmentation de l'activité biologique.

Dans l'équation(III.1), les coefficients négatifs de SAG , MR et $HOMO$ montrent que toute diminution dans les valeurs de ces paramètres entraîne une diminution de l'activité inhibitrice contre le cancer de la peau (mélanome) pour cette série des molécules.

Les matrices de corrélation entre les descripteurs obtenus par l'analyse MLR et l'activité biologique sont reportée dans le tableau III.7.

Tableau III.8 : Matrice de la corrélation du modèle.

	pIC50	SAG	V	HE	MR	POL	HOMO	MD	E _T	N ₂
pIC50	1									
SAG	0.087	1								
V	0.142	0.986	1							
HE	0.468	0.059	0.090	1						
MR	0.313	0.455	0.583	0.329	1					

Étude qualitative et quantitative des propriétés QSAR d'une série de dérivés de phtalazine

POL	0.347	0.409	0.538	0.430	0.979	1				
HO	-0.039	-0.089	-0.047	-0.046	-0.003	0.074	1			
MO										
MD	-0.224	-0.197	-0.252	-0.421	-0.291	-0.388	-0.073	1		
E _T	0.399	-0.252	-0.322	-0.116	-0.476	-0.494	-0.070	0.425	1	
N ₂	0.343	-0.075	-0.018	-0.131	0.134	0.131	0.122	-0.454	-0.084	1

Le tableau III.12 montre que le descripteur **HE** est le plus important dans la corrélation entre les descripteurs sélectionnés et l'activité biologique

III.7.2.1. Validation de modèle

Pour la validation de modèle on utilisant la méthode de validation croisée, est une méthode d'estimation de fiabilité d'un modèle fondé sur une technique d'échantillonnage. En fait, il y a au moins trois techniques de validation croisée : (test set validation) ou (hold out method), (k-fold cross-validation) et (leave-one-out cross-validation **LOO**) cette dernière validé notre model par calcule de ses paramètres statistiques : Press, SSY, Press/SSY, Spress, R^2_{cv} , R^2_{adj} et PE.

Tableau III.9 : paramètres de validation croisée.

Modèle	PRESS	SSY	PRESS/SSY	Spress	R^2_{cv}	R^2_{adj}	6PE
1	0.491	4.872	0.100	0.160	0.899	0.798	0.093

Le **PRESS** (Prédictive somme des carrés résiduelle) est un paramètre important. Sa valeur inférieure à **SSY** (somme des carrés des distances) indique que le modèle proposé à un bon pouvoir prédictif et c'est mieux que le hasard. Tell que dans le (**Tableau III.9**) la valeur de **PRESS** est **0.491** et la valeur de **SSY** est **4.872**.

Dans notre cas, (**Tableau III.9**) pour le modèle proposé **PRESS** < **SSY** indiquant qu'il est un bon pouvoir prédictif et sont mieux que le hasard. Pour un modèle bon **PRESS / SSY** devrait être inférieur à **0.4** et sa valeur dans notre modèle est **0.100** [32].

Autre paramètres de validation croisée important est le coefficient de corrélation prédictive R^2_{cv} , la valeur la plus élevée qui a donné l'excellent modèle, et dans notre étude (**Tableau IV.3**) $R^2_{cv} = 0.899$, Pour R^2_{adj} est **0.798** aussi la valeur plus élevée exprime un meilleur modèle.

Étude qualitative et quantitative des propriétés QSAR d'une série de dérivés de phtalazine

L'erreur de prédiction du coefficient de corrélation (**PE**) est un autre paramètre utilisé pour déterminer le pouvoir prédictif des modèles proposés. Nous avons calculé la valeur de **6PE** de ce modèle proposé et il est présenté dans le tableau. Pour le modèle la condition $R > 6PE$ est satisfaite et donc ils ne peuvent être considérés comme ayant un bon pouvoir prédictif [29].

III.7.2.2. Validation externe

Pour s'assurer de la bonne capacité prédictive effective du modèle, nous avons opéré par validation externe sur l'ensemble de 6 composés choisis aléatoirement (4-5-6-21-24-25) et qui ne font pas partie de l'ensemble d'essai. Une validation rigoureuse du modèle se traduit par une proportion importante de prédictions exactes données sur l'ensemble de la validation. La performance du modèle est alors mesurée par le coefficient de régression R^2 .

Le coefficient qui décrit la validation est donné par l'équation ci-dessous :

$$R^2_{\text{Pred}} = 1 - \frac{\sum_{i=1}^{\text{test}} (y_{\text{exp}} - y_{\text{pred}})^2}{\sum_{i=1}^{\text{test}} (y_{\text{exp}} - \bar{y}_{\text{tr}})^2}$$

$R^2_{\text{pred}} = 0,715 > 0,6$. Alors la validation externe est appliquée avec succès.

Tableau III.10 : Valeurs observées, prédites et les résidus ordinaires de pIC₅₀ par MLR pour l'ensemble de validation.

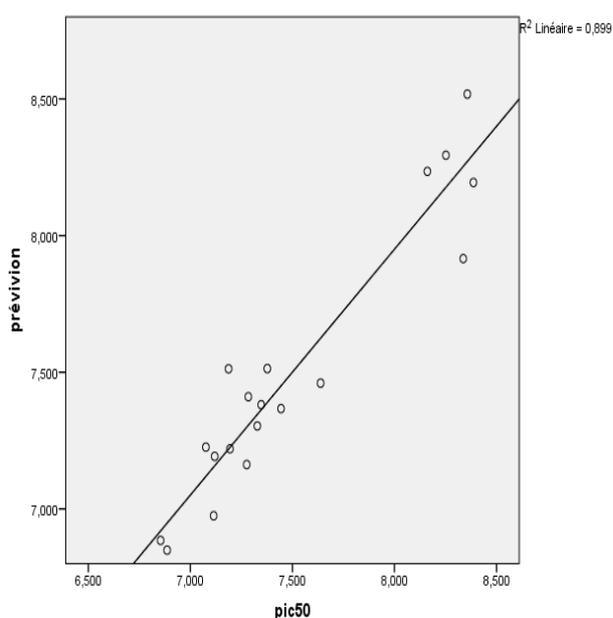
Composés	pIC ₅₀ Exp	pIC ₅₀ Préd	pIC ₅₀ Rés
1	6.854	7.205	-0.030
2	6.886	7.256	0.036
3	7.076	7.616	-0.150
4 VE	7.065	7.070	-0.005
5 VE	8.173	8.164	0.009
6 VE	8.221	8.210	0.011
7	8.387	8.347	0.192
8	8.357	8.724	-0.161
9	7.114	7.225	0.138
10	7.119	7.447	-0.073
11	7.284	7.679	-0.126
12	7.187	7.682	-0.325
13	7.194	7.713	-0.026
14	7.276	7.542	0.113
15	7.328	7.601	0.024
16	7.638	7.803	0.178

Étude qualitative et quantitative des propriétés QSAR d'une série de dérivés de phtalazine

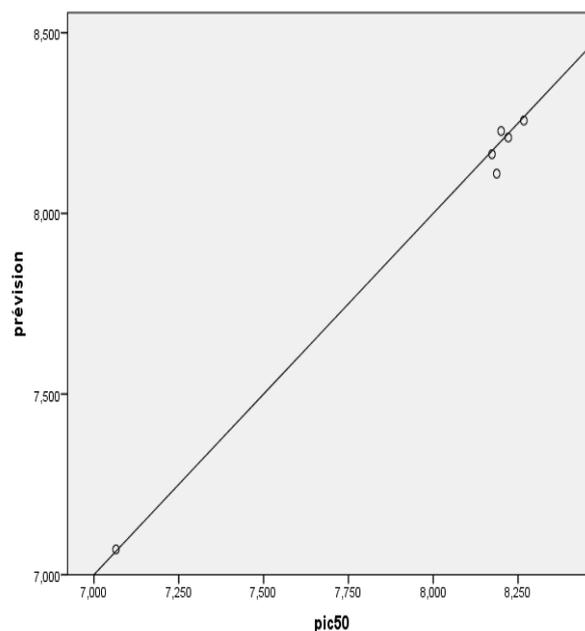
17	7.347	7.713	-0.034
18	7.377	8.012	-0.136
19	7.444	7.753	0.076
20	8.252	8.543	-0.042
21 VE	8.200	8.228	-0.028
22	8.161	8.518	-0.073
23	8.337	8.326	0.421
24 VE	8.187	8,110	0.077
25 VE	8.267	8.257	0.010

VE : validation externe, Exp : expérimental, Préd : prédictif, Rés : résiduel.

La figure III-4 ci-dessous montre deux courbes de régression linéaire des valeurs prédites par rapport aux valeurs expérimentales de l'activité biologique de phtalazines. La première courbe a) présente le modèle représenté pour être plus pratique avec $R^2 = 0,899$. Il indique que le modèle peut être appliqué avec succès pour prédire l'activité inhibitrice anti-cancer dans cette série de molécules, la deuxième courbe b) présente la validation externe pour 6 molécules peut être appliqué avec succès.



a)



b)

Figure III.4 : Les courbes des valeurs prédictives en fonction des valeurs expérimentales de pIC50.

D'après l'étude de la présence d'une erreur systématique dans l'élaboration de modèle QSAR courbe (a) l'ensemble d'entraînement (training set) et la courbe (b) l'ensemble d'essai

Étude qualitative et quantitative des propriétés QSAR d'une série de dérivés de phtalazine

(test set), les résidus des valeurs prévus de l'activité biologique pIC_{50} ont été tracés en fonction des valeurs expérimentales, comme le montre dans la figure III.4.

La propagation des résidus de côté de zéro indique qu'il n'y a aucune erreur systématique. Il indique que ce modèle peut être appliqué avec succès pour prédire l'activité inhibitrice contre le cancer de la peau (mélanome) pour cette série des molécules [32].

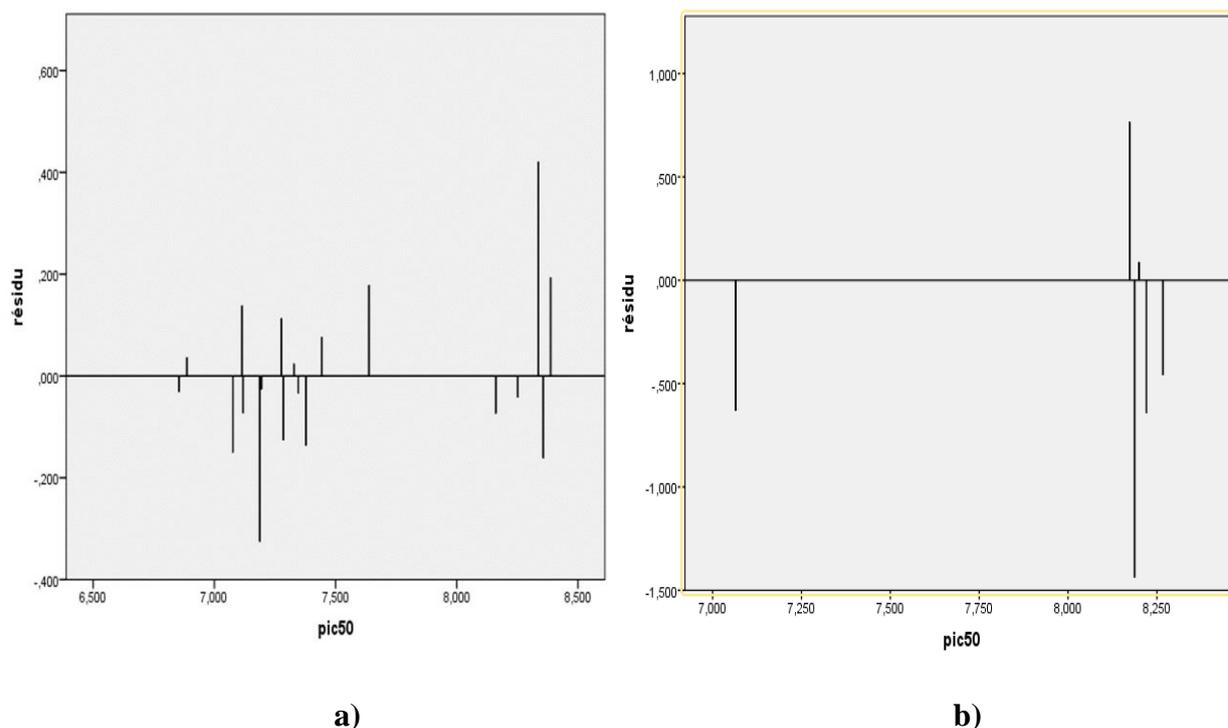


Figure III.5 : La courbe des valeurs résiduelles par rapport à l'observé expérimentalement.

La figure III.5 montre deux courbes de régression linéaire des valeurs résidues par rapport aux valeurs expérimentales de l'activité biologique de phtalazines. La première courbe (a) présente l'ensemble d'entraînement (training set) et la deuxième courbe (b) présente l'ensemble d'essai (test set) ces deux courbes sont asymétriques alors sont appliqués avec succès.

III.8. Conclusion

L'application de ces études des relations structure-activité sur une nouvelle série de phtalazine qui peuvent effectivement traiter la mélanome avec des effets secondaires mineurs en évaluant les corrélations entre les différents descripteurs.

La polarisabilité et la réfractivité molaire augmentent relativement avec la taille et le poids moléculaire des dérivés de phtalazine.

Les valeurs importantes de la masse, le volume, la surface, la polarisabilité et la réfractivité molaire sont respectivement 406.46 UMA, 1012.88 Å³, 587.02 Å², 46.42 Å³ et 132.54 Å³.

Le composé 9 est le petit dérivé dans la série étudiée, il possède des valeurs faibles de la polarisabilité (24.33 Å³), de la réfractivité molaire (66.02 Å³) et énergie d'hydratation en valeur absolue importante (21.90 kcal/mole), ainsi que, la présence des groupes hydrophiles augmentent l'énergie d'hydratation dans le composé 9 qui possède trois sites donneurs des protons (2 NH, 1NH₂) et un site accepteur de proton (S).

Les composés 1, 2 et 4 présentent des petits coefficients de lipophilie (-0.56), (-0.12) et (-0.07), respectivement. Ces molécules sont les produits les plus hydrophiles. Ces composés donnent une bonne solubilité aqueuse, mais une faible absorption et pénétration, en raison de leur mauvaise perméabilité par diffusion passive. Ces composés ont une tendance élevée d'éliminer par les reins.

Pour les résultats liés de l'efficacité de ligand, on a trouvé une grande valeur de LE est de (0.663) pour le composé 9 avec N_H égale 15, donc cette valeur de LE explique que ce composé a tendance à avoir les meilleures propriétés physico-chimiques et ADME.

Par comparaison, des résultats de calcul et les critères de la règle de Lipinski on a remarqué que tous les composés de la série confirment aux ces critères, ils n'ont pas des problèmes d'absorption ou de perméabilité et ils sont acceptables d'être administrés par voie orale. Notre travail est ensuite porté l'étude la règle de Veber explique que tous les composés de la série ayant le nombre des liaisons inférieure à 10, et aussi que tous les composés de notre série avec des valeurs de la surface polaire sont en accord avec la règle 2^{ème} de Veber, c'est-à-dire tous les valeurs de la surface polaire inférieure à 140 Å², et ce résultat qui montre une

Étude qualitative et quantitative des propriétés QSAR d'une série de dérivés de phtalazine

bonne corrélation avec le transport moléculaire passive à travers les membranes. Donc, il permet l'estimation des propriétés de transport des médicaments.

L'étude de l'efficacité lipophile dans la série des de dérivés de phtalazine, à partir de nos résultats on a trouvés les composés (2,3, 8,9,10,12;13,14,15,16,17,18,19) qui se situé dans la gamme suggérée de 5-7, ce qui indique que ce composé été optimisé avec succès.

L'analyse QSAR a été effectuée pour déterminer la relation quantitative entre la structure moléculaire des composés avec leur activité. Le modèle QSAR indique que les descripteurs physico-chimiques et descripteurs quantiques ont des relations significatives avec la bioactivité observée. Nous avons observé une similitude élevée entre les valeurs expérimentales et les valeurs prédictives de l'activité, qui indique une qualité du modèle QSAR magnifique.

Référence

- [1] Debnath, A. K., Quantitative structure-activity relationship (QSAR) paradigm--Hansch era to new millennium. Mini reviews in medicinal chemistry, 2001, 1(2), 187-195.
- [2] Waterbeemd H., Rose S., "Quantitative approaches to structure-activity relationships", in Book "Quantitative approaches to structure-activity relationships". Elsevier, 2003, 351-367.
- [3] Cohen Y., Pharmacologie moléculaire , (1978).
- [4] Belaidi, S., Melkemi, N., & Bouzidi, D. Molecular geometry and structure-property relationships for 1, 2-dithiole-3-thione derivatives. Int J Chem Res, (2012), 4(2), 134-139.
- [5] Site internet : aterloo.ca/science.
- [6] site internet : formules-physique.com
- [7] Galvez M., BIBLIOGRAPHIE DE LA LITTÉRATURE FRANÇAISE (XVI e-XXI e SIÈCLES). Année 2013. Revue d'Histoire littéraire de la France, (2014), 11-761.
- [8]. Witan H., Kern A., Koziollek-Drechsler I., Wade R., Behl C., & Clement A. M. Heterodimer formation of wild-type and amyotrophic lateral sclerosis-causing mutant Cu/Zn-superoxide dismutase induces toxicity independent of protein aggregation. Human molecular genetics, (2008), 17(10), 1373-1385.
- [9] Abraham M. H., Chadha H. S., Leitao R. A., Mitchell R. C., Lambert W. J., Kaliszan R., ... & Haber P. Determination of solute lipophilicity, as log P (octanol) and log P (alkane) using poly (styrene–divinylbenzene) and immobilised artificial membrane stationary phases in reversed-phase high-performance liquid chromatography. Journal of Chromatography A, (1997), 766(1-2), 35-47.
- [10] Abraham M. H., Gola J. M., Kumarsingh R., Cometto-Muniz J. E., & Cain W. S. Connection between chromatographic data and biological data. Journal of Chromatography B: Biomedical Sciences and Applications, (2000), 745(1), 103-115.
- [11] Lipinski C. A., Lombardo F., Dominy B. W., & Feeney P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Advanced drug delivery reviews, (2001), 46(1-3), 3-26.
- [12]. Rouahna Y., Belaidi S., Harkati D., & Kerassa A. Qualitative QSAR Studies, Electronic Structure, Drug Likeness of 1, 2-Dithiole-3-One Derivatives. Journal of Computational and Theoretical Nanoscience, (2015),12(11), 4233-4241.

Étude qualitative et quantitative des propriétés QSAR d'une série de dérivés de phtalazine

- [13] Veber D. F., Johnson S. R., Cheng H. Y., Smith B. R., Ward K. W., & Kopple K. D. Molecular properties that influence the oral bioavailability of drug candidates. *Journal of medicinal chemistry*, (2002), 45(12), 2615-2623.
- [14] Dugas H., *Principes de base en modélisation moléculaire, Aspects théoriques et pratiques*, Hermann Dugas, quatrième édition, librairie de l'université de Montréal, 1996.
- [15] Remko M., Swart M., & Bickelhaupt F. M. Theoretical study of structure, pKa, lipophilicity, solubility, absorption, and polar surface area of some centrally acting antihypertensives. *Bioorganic & medicinal chemistry*, (2006), 14(6), 1715-1728.
- [16] E. P., R. B., and S. P., *J. MedChem.* (2000), 43,3714.
- [17] Kuntz I. D., Chen K., Sharp K. A., & Kollman P. A. The maximal affinity of ligands. *Proceedings of the National Academy of Sciences*, (1999), 96(18), 9997-10002.
- [18] Abad-Zapatero C., & Metz J. T. Ligand efficiency indices as guideposts for drug discovery. *Drug discovery today*, (2005), 10(7), 464-469.
- [19] Hopkins A. L., Groom C. R., & Alex A. *Ligand efficiency: a useful metric for lead selection. Drug discovery today*, (2004), 9(10), 430-431.
- [20] Shultz M. D. *Setting expectations in molecular optimizations: strengths and limitations of commonly used composite parameters. Bioorganic & medicinal chemistry letters*, (2013), 23(21), 5980-5991.
- [21] Ryckmans T., Edwards M. P., Horne V. A., Correia A. M., Owen D. R., Thompson L. R., ... & Young, T. Rapid assessment of a novel series of selective CB 2 agonists using parallel synthesis protocols: a lipophilic efficiency (LipE) analysis. *Bioorganic & medicinal chemistry letters*, (2009), 19(15), 4406-4409.
- [22] Edwards, M. P., & Price, D. A. Role of physicochemical properties and ligand lipophilicity efficiency in addressing drug safety risks. In *Annual Reports in Medicinal Chemistry*. Academic Press. (2010), Vol. 45, pp. 380-391.
- [23] Leeson P. D., & Springthorpe B. The influence of drug-like concepts on decision-making in medicinal chemistry. *Nature Reviews Drug Discovery*, (2007), 6(11), 881.
- [24] Pathak J., Narayan V., Sinha L., & Prasad O. Theoretical Raman and FTIR vibrational analysis of 2-phenyl-1H-indene-1, 3 (2H)-dione by ab initio method. *J Atomic Mol Sci*, (2012), 3(2), 95-105.

Étude qualitative et quantitative des propriétés QSAR d'une série de dérivés de phtalazine

[25] Zhang L., Zhu H., Oprea T. I., Golbraikh A., & Tropsha A. QSAR modeling of the blood–brain barrier permeability for diverse organic compounds. *Pharmaceutical research*, (2008), 25(8), 1902.

[26] Anh, N. T. *Orbitales frontières*. CNRS Éditions, Paris, (1995), 123.

[27] Site internet : www.futura-sciences.com

[28] Fukui K. *Theory of orientation and stereoselection, reactivity and structure, concepts in organic chemistry*, (1975).

[29] Srivastava A. K., & Shukla N. Quantitative structure activity relationship (QSAR) studies on a series of imidazole derivatives as novel ORL1 receptor antagonists. *Journal of Saudi Chemical Society*, (2013),17(3), 321-328.

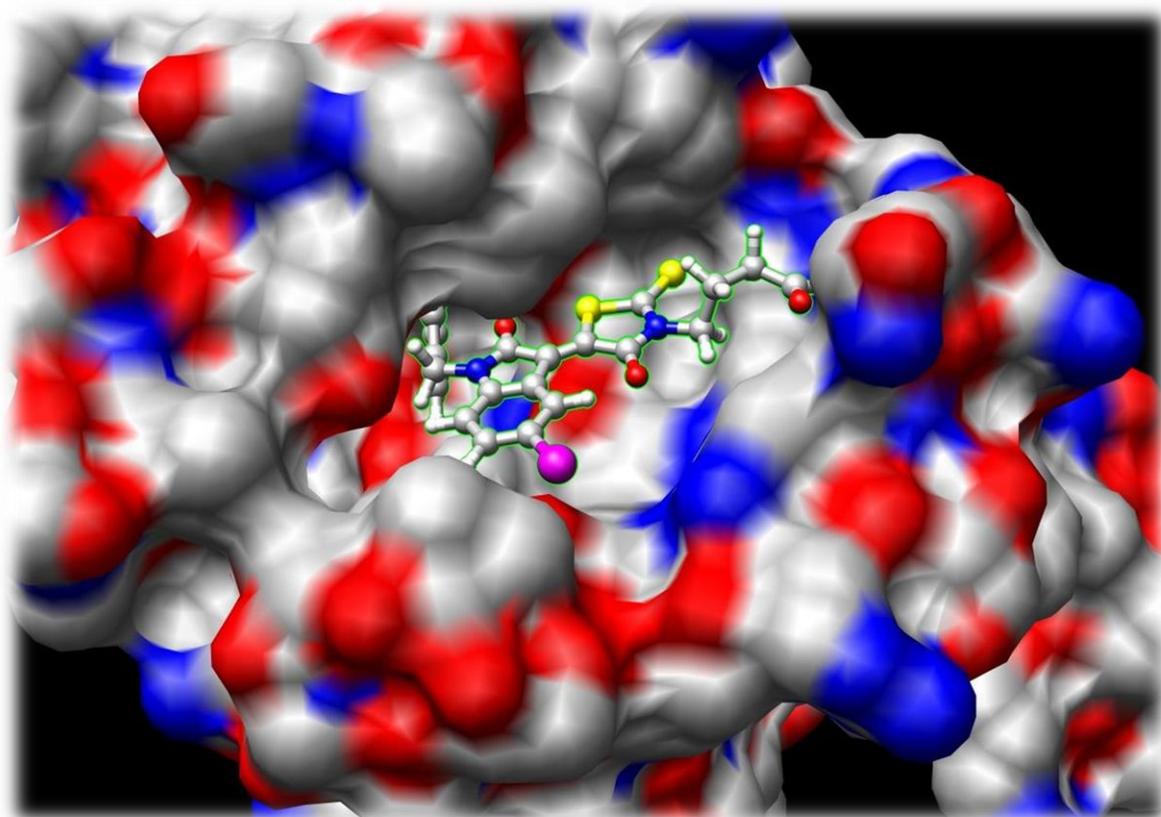
[30] Amin K. M., Barsoum F. F., Awadallah F. M., & Mohamed N. E. Identification of new potent phthalazine derivatives with VEGFR-2 and EGFR kinase inhibitory activity. *European journal of medicinal chemistry*, (2016), 123, 191-201.

[31] Bouchlaleg L., Belaidi S., Salah T., & Alafeefy A. M. Quantitative Structure Activity Relationship Study for Development of Plasmin Inhibitors Controlled by the Spacer Hydantoin. *Journal of Computational and Theoretical Nanoscience*, (2015), 12(10), 3949-3955.

[32] Allinger N., Q. C. P. E., (1983), 3, 32.

Chapitre IV

Etude par Doking Moléculaire des Interactions entre L'enzyme 1M17 et des Dérivés de Phtalazine



Etude par Doking Moléculaire des Interactions entre L'enzyme 1M17 et des Dérivés de Phtalazine

IV.1.Introduction

Historiquement, les premiers outils de docking obéissaient au principe dit « clef-serrure » ou « lock-and-key » selon lequel le ligand qui représente la clef, est complémentaire au niveau géométrique du site actif du récepteur, qui représente la serrure. Considérant donc le ligand et le récepteur comme deux entités rigides, la recherche de la pose optimale se limite au positionnement en énumérant toutes les rotations et les translations possibles pour le ligand à l'intérieur du site d'interaction. Ensuite toutes les poses incompatibles avec le site actif sont éliminées par une image négative du site actif. Enfin, les poses sélectionnées se voient attribuées un score et les meilleures sont optimisées [1].

Etude par Doking Moléculaire des Interactions entre L'enzyme 1M17 et des Dérivés de Phtalazine

IV.2. Les étapes de Docking moléculaire

Les outils de modélisation moléculaire sont particulièrement performants pour l'étude au niveau atomique d'interactions entre deux entités moléculaires (l'enzyme et l'inhibiteur pour former un complexe stable) et ce type d'étude fait appel le plus souvent aux techniques de « Docking Molculaire » [2].

IV.2.1. Structure de la macromolécule (l'enzyme)

L'enzyme utilisée dans cette partie est la tyrosine kinase, cette dernière a été téléchargée à partir d'une base de données RCSB Protein Data Bank [3] avec le code d'accès est 1M17.

La structure 3D du 1M17 a été obtenue par diffraction en rayon X avec une résolution de 2.6 Å, constituée par une chaîne A. On remarque que le 1M17 Co-cristallisé **figure IV.1** avec le ligand de référence 4-anilinoquinazoline ($C_{22} H_{23} N_3 O_4$) qui est constitué de 52 atomes.

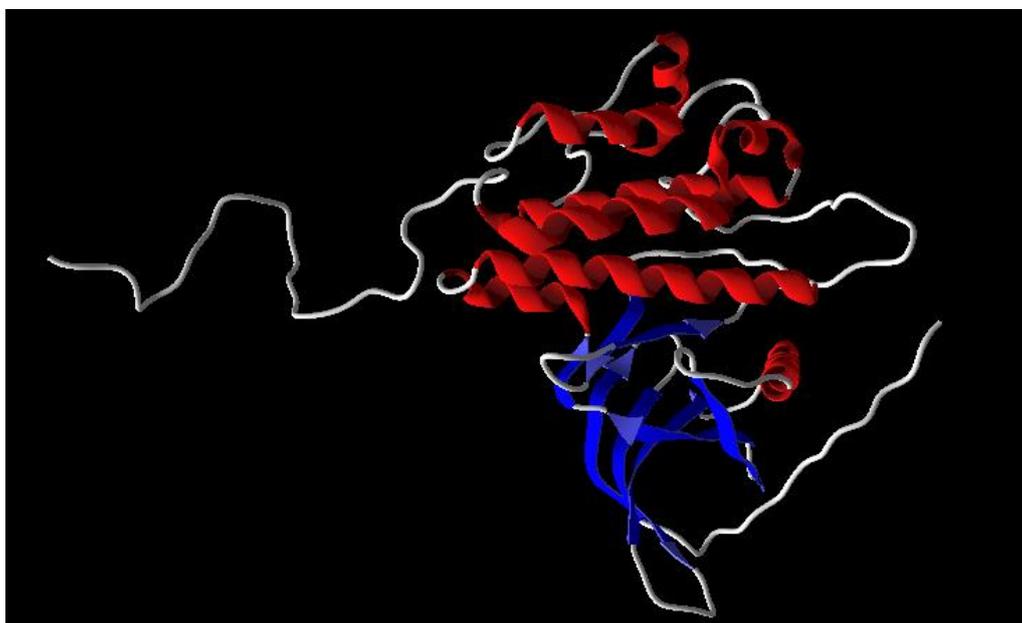


Figure IV.1 : Structure de récepteur du facteur de croissance épidermique tyrosine kinase 1M17.

IV.2.2. Structure des ligands

L'ensemble de 25 molécules (Tableau III.1) obtenues dans la première partie de travail sont utilisées dans le docking moléculaire. L'optimisation des géométries des ligands a été

Etude par Doking Moléculaire des Interactions entre L'enzyme 1M17 et des Dérivés de Phtalazine

effectuée (Tableau III.2) à l'aide du logiciel HyperChem (Version 8.0.8), par les méthodes de la mécanique moléculaire en utilisant le champ de force MM+ et semi empirique PM3.

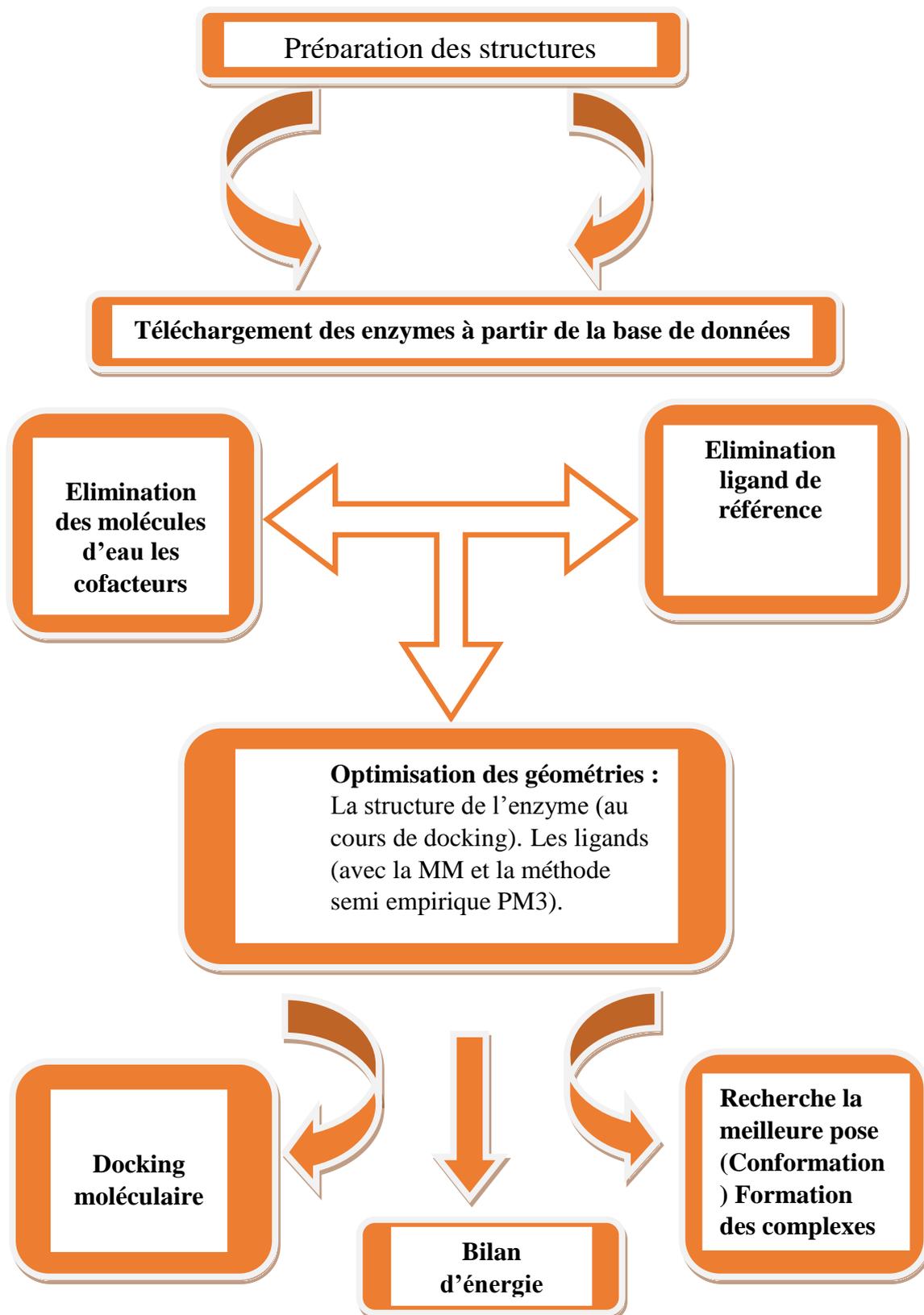


Figure IV.2 : protocole de calcul.

Etude par Doking Moléculaire des Interactions entre L'enzyme 1M17 et des Dérivés de Phtalazine

IV.3.Résultats de docking moléculaire

Nous avons utilisé la méthode de Docking à l'aide de logiciel Molegro Virtual Docker 5.5. Une fois que le complexe ligand-récepteur est formé, celui-ci va adapter la conformation la plus stable, c'est-à-dire avec le niveau énergétique le plus faible. Une fois le complexe formé nous allons rechercher la conformation la plus stable, et procéder à l'analyse des résultats donnés.

IV.3.1.Détection des cavités

Nous avons détectés les cavités dans la macromolécule (l'enzyme) à l'aide de logiciel MVD. Le tableau ci-dessous représenté le volume et la surface des 5 cavités détectées. On observe que le ligand de référence fait des interactions avec la première cavité, alors Nous choisissons cette cavité pour faire l'étude, cette cavité possède un volume plus important que les autres cavités détectés.

Tableau IV.1 : Volume et surface des cavités détectées.

Cavités	Volume (Å ³)	Surface (Å ²)
Cavité 1	279.552	903.68
Cavité 2	25.6	112.64
Cavité 3	20.992	88.32
Cavité 4	18.944	83.2
Cavité 5	13.824	56.32

Les valeurs de volume et la surface de la cavité 1 sont les plus élevée le volume est égale à 279.552(Å³) et la surface égale à 903.68(Å²) si pour ça en va choisi cette cavité.

Etude par Doking Moléculaire des Interactions entre L'enzyme 1M17 et des Dérivés de Phtalazine

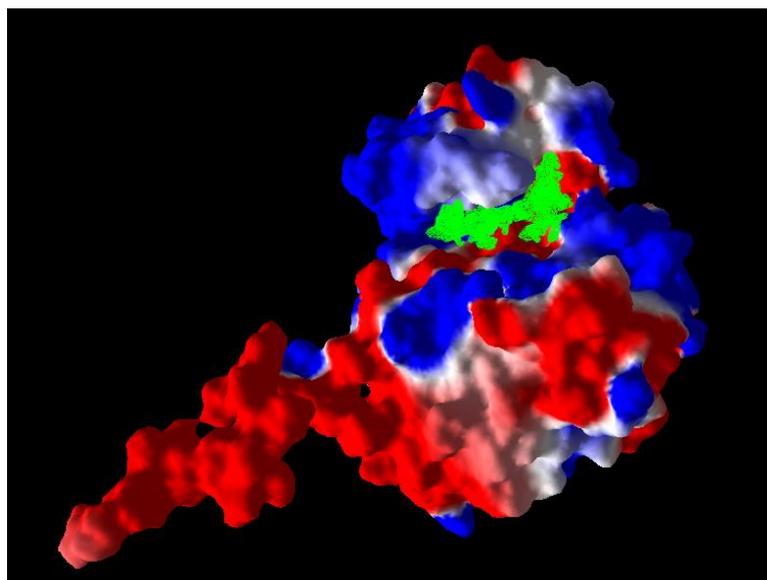


Figure V.3 : Interface graphique de la cavité 1 indiquée par Mol Dock.

IV.3.2.Énergie d'interaction

Nous mettons d'après le tableau ci-dessous, les différentes énergies d'interactions entre l'enzyme 1M17 et les dérivés de phtalazine sont calculées au cours de logiciel Molegro Vertual Docker.

TableauIV.2 : Résultats de docking moléculaire des dérivés de phtalazine avec la cavité 1 de l'enzyme1M17.

Composés	E _T	E-inter stérique	HBond	E inter (total)
01	-96.777	14.693	-6.745	-111.67
02	-125.807	1.283	-6.816	-130.052
03	-113.236	11.154	-4.884	-125.424
04	-113.49	11.761	-8.061	-131.948
05	-118.86	11.83	-5.946	-133.329
06	-81.965	16.113	-1.738	-98.079
07	-87.534	15.449	-4.662	-104.805
08	-92.563	22.095	-8.889	-117.724
09	-83.984	7.570	-3.000	-94.895
10	-95.041	6.760	-3.956	-105.14
11	-91.260	-2.515	-4.056	-98.894
12	-104.033	4.527	-4.949	-112.088
13	-104.494	10.815	-3.861	-120.65
14	-116.314	-3.271	-4.460	-134.048
15	-93.571	6.228	-9.020	-101.393
16	-98.729	10.010	-7.307	-109.863
17	-102.755	16.107	-5.107	-118.862

Etude par Doking Moléculaire des Interactions entre L'enzyme 1M17 et des Dérivés de Phtalazine

18	-103.043	16.261	-3.952	-123.195
19	-104.609	16.237	-7.562	-122.473
20	-117.658	11.005	-4.279	-129.092
21	-112.659	10.720	-2.799	-135.224
22	-122.984	2.682	-3.927	-130.811
23	-121.643	3.592	-3.127	-131.015
24	-126.719	-1.010	-2.428	-134.264
25	-129.746	-0.386	-6.243	-130.5
AQ4_999	-122.086	-6.844	-2.5	-125.381

ET : énergie total ; **E-inter stérique** : énergie d'interaction stérique ; **HBond** : hydrogène bond ; **E-inter (total)** : énergie d'interaction total.

A partir des résultats obtenus dans ce tableau en basant sur la comparaison entre les différentes énergies de ligand de références AQ4_999 et les différentes énergies de la série de phtalazine et on doit se baser beaucoup plus sur l'énergie total.

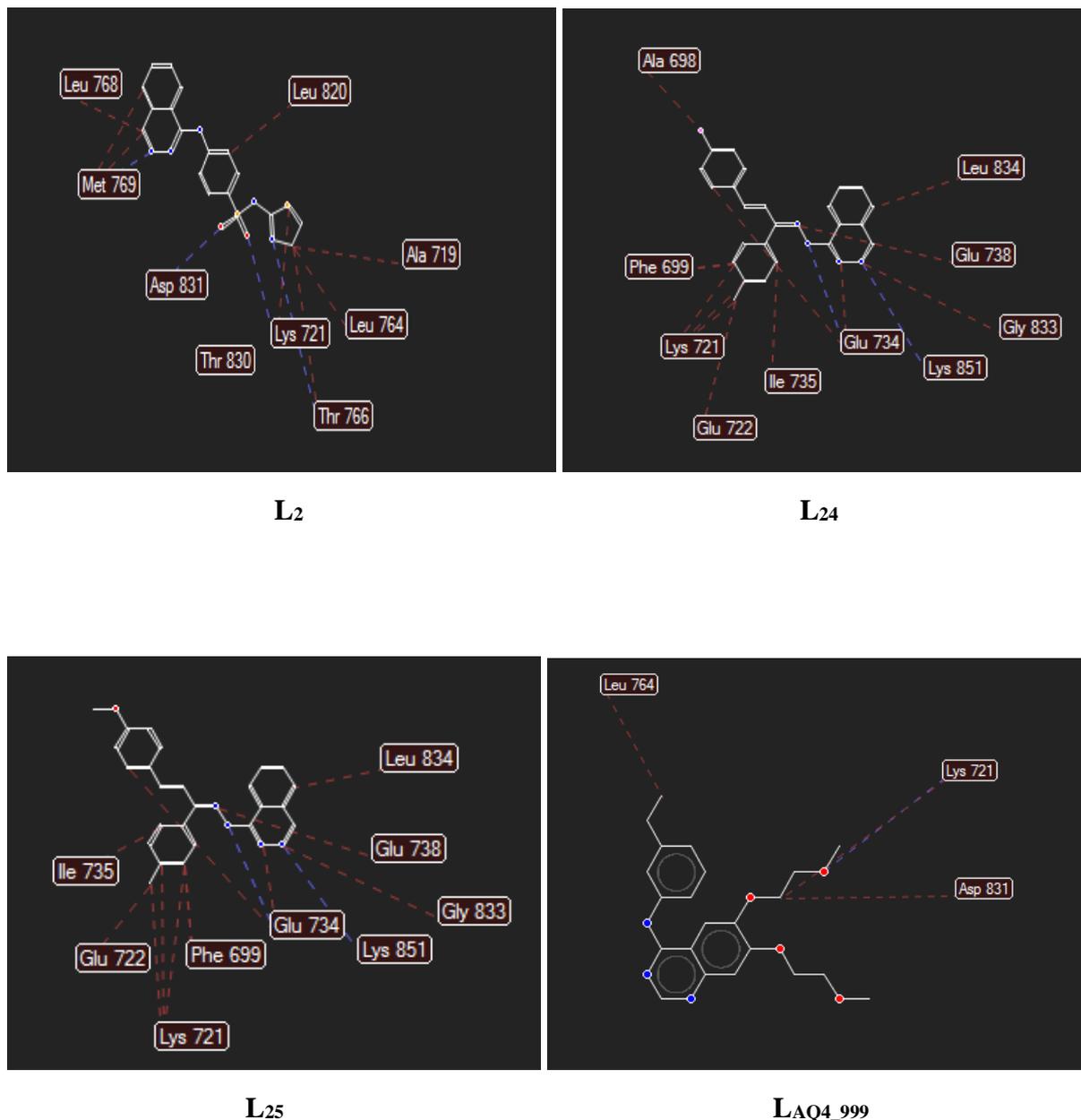
On remarque que la valeur de l'énergie total de AQ4_999 est égale à -122.086, alors que les résultats de ces recherches ont aboutis à trois molécules 2-24-25 qui possèdent des énergies total plus basse et équivalente à l'énergie total de ligand de référence que les autres complexes, respectivement -125.807,-126.719,-129.746.

Tableau IV.3 : Meilleur complexes formés par L₂, L₂₄, L₂₅.

Ligand	ET	E-inter (stérique)	H-bond	E-inter (total)
2	-125.807	1.283	-6.816	-130.052
24	-126.719	-1.010	-2.428	-134.264
25	-129.746	-0.386	-6.243	-130.5
AQ4_999	-122.086	-6.844	-2.5	-125.381

Les figures suivantes représentent les différentes interactions entre les résidus du site actif et les ligands L₂, L₂₄, L₂₅, LAQ4_999.

Etude par Doking Moléculaire des Interactions entre L'enzyme 1M17 et des Dérivés de Phtalazine



----- : Interactions stériques

----- : Interactions des liaisons hydrogène

Figure IV.4 : Les Interaction hydrogène et stérique entre les ligands et l'enzyme 1M17.

D'après la figure IV.5, les interactions striques et les interactions des liaisons hydrogène sont misent en jeu entre les résidus du site actif et les ligands qui on a choisi permettent d'obtenir l'énergie la plus basse et par la suite formation du complexe le plus stable.

Etude par Doking Moléculaire des Interactions entre L'enzyme 1M17 et des Dérivés de Phtalazine

Les distances mesurées entre les ligands et les acides aminés de la cavité 1 sont cités dans le tableau ci-dessous.

Tableau IV.4 : Distance (stérique et Hydrogènes) mesuré en Å entre les résidus du site actif et les atomes de ligand L₂.

Résidus	L ₂ -Atomes	Energie stérique	Résidus	L ₂ -Atomes	Liaison-H
Met769	C	2,96	Met769	N	3,14
Leu768	C	3,17	Thr830	O	3,10
Met769	C	2,70	Lys721	O	2,60
Leu820	C	3,04	Thr766	N	3.18
Lys721	S	3,08			
Thr766	C	2,78			
Leu764	C	3,19			
Ala719	C	3,16			

TableauIV.5 : Distance (stérique et Hydrogènes) mesuré en Å entre les résidus du site actif et les atomes de ligand L₂₄.

Résidus	L ₂₄ -Atomes	Interaction stérique
Ala698	F	3,11
Glu734	C	3,03
Phe699	C	3,08
Lys721	C	3,04
Lys721	C	3,18
Glu 722		2.47
Lys721	C	3 ,18
Ile735	C	2,98

Etude par Doking Moléculaire des Interactions entre L'enzyme 1M17 et des Dérivés de Phtalazine

Résidus	L ₂₄ -Atomes	Liaison-H
Glu738	N	2,58
Glu734	N	3,15
Gly833	N	3,12
Leu834	C	3,01

Tableau IV.6 : Distance (stérique et Hydrogènes) mesuré en Å entre les résidus du site actif et les atomes de ligand L₂₅.

Résidus	L ₂₅ -Atomes	Interaction stérique	Résidus	L ₂₅ -Atomes	Liaison-H
Glu734	C	3,04	Glu734	N	3.19
Ile735	C	3,00	Lys851	N	3.35
Glu722	C	2,51			
Lys721	C	3,13			
Lys721	C	3,15			
Lys721	C	3,19			
Phe699	C	2,91			
Glu738	N	2,60			
Glu734	N	3,10			
Lys851	N	3,10			
Leu834	C	3,01			

Interprétation des résultats

D'après Anne Imberty et col, les intervalles des valeurs des interactions ayant des distances de la liaison d'hydrogène sont comme suite [4] :

-[2.5Å-3.1Å] sont considérés comme fortes complexes.

-[3.1Å -3.55Å] sont supposées comme moyennes complexes.

-supérieures à 3.55Å sont considérées comme faibles complexes.

La formation d'un complexe stable dépend de la fixation du ligand dans le site active. La figure IV.5 et les tableaux (IV.4, IV.5, IV.6), les distances entre les résidus du site actif et les L₂, L₂₅, L₂₅ varient entre 2.60Å et 3.35Å expliquent l'existence de la formation des complexes stables, et par la suite, une meilleure fixation de ces ligands au niveau du site actif.

Etude par Doking Moléculaire des Interactions entre L'enzyme 1M17 et des Dérivés de Phtalazine

IV.4.Conclusion

La méthode de docking a été effectuée pour l'objectif de prédire la capacité d'une molécule à se lier au site actif d'une protéine, grâce à une simulation de la conformation et de l'orientation qu'elle adopte lors de sa liaison au récepteur [5-6] a été réalisée. On a utilisé différents énergies E_T , E stérique, E inter-stérique, H-Bond, peuvent être utilisés avec succès pour l'identification de ligands de la cible étudiée, avoir les différentes conformations moléculaires et leurs poses.

Après la comparaison entre les différentes énergies de la série phtalazine et les différentes énergies de ligand de référence, on a trouvé trois molécules nous pouvons les considérer comme des médicaments.

Etude par Doking Moléculaire des Interactions entre L'enzyme 1M17 et des Dérivés de Phtalazine

Référence

- [1] Yuriev E., Agostino M., & Ramsland P. A. (2011). Challenges and advances in computational docking: in review. *Journal of Molecular Recognition*, 2009, 24(2), 149-164.
- [2] Bower M. J., Cohen F. E., & Dunbrack Jr R. L. Prediction of protein side-chain rotamers from a backbone-dependent rotamer library: a new homology modeling tool. *Journal of molecular biology*, 1997, 267(5), 1268-1282.
- [3] Rehman S., Ijaz B., Fatima N., Muhammad S. A., & Riazuddin S. Therapeutic potential of *Taraxacum officinale* against HCV NS5B polymerase: In-vitro and In silico study. *Biomedicine & Pharmacotherapy*, 2016, 83, 881-891.
- [4] Imberty A., Hardman K. D., Carver J. P., & Perez S. Molecular modelling of protein-carbohydrate interactions. Docking of monosaccharides in the binding site of concanavalin A. *Glycobiology*, 1991, 1(6), 631-642.
- [5] Lyne P. D. Structure-based virtual screening: an overview. *Drug discovery today*, 2002, 7(20), 1047-1055.
- [6] Tuccinardi T. Docking-based virtual screening: recent developments. *Combinatorial chemistry & high throughput screening*, 2009, 12(3), 303-314.

Conclusion Générale



CONCLUSION GENERALE

Après ce travail notre application de différentes méthodes de la chimie informatique pour l'évaluation de l'activité biologique et les caractérisations physicochimiques des dérivés de phtalazine pour choisi les molécules qui nous pouvons les considérer comme des médicaments contre le cancer de la peau.

Notre travail contient plusieurs méthodes de modélisation moléculaire ont été utilisées sur les dérivés de phtalazine sont :

- Une étude quantitative et qualitative des relations structure-activité/propriétés sur une série de vingt-cinq molécules, qui on a développé un modèle QSAR avec des tests réussis.
- Une étude par doking moléculaire sur la série de phtalazine pour connaitre les molécules qui forme un complexe stable avec l'enzyme 1M17.

Cette première étude a mené à l'obtention d'un modèle prédictif de l'activité inhibitrice anticancéreux contre le cancer de la peau (mélanome). L'obtention d'un modèle a été établie par la détermination des paramètres statistiques approprié sen utilisant la méthode de régression multilinéaire (MLR), ce modèle obtenus peuvent être utilisé pour prédire l'activité cytotoxique de phtalazine, et permet de déterminer les paramètres qui sont influes sur l'activité cytotoxique des dérivés de phtalazine.

Le pouvoir prédictif de modèle obtenu a été confirmé par la méthode de validation croisée LOO et validation externe. Une forte corrélation a été observée entre les valeurs expérimentales et prédites des activités biologiques cytotoxiques, ce qui indique la validité et la qualité de modèle QSAR obtenu.

Dans le deuxième volet de ce travail, nous avons réalisé une étude de doking moléculaire avec le programme Molegro Virtual Docker ces études sont des interactions entre deux entités moléculaires l'enzyme et l'inhibiteur pour former un complexe stable.

Enfin, d'après les résultats de ce travail on a trois molécules qui sont considérés comme des médicaments 2, 24,25 parce qu'ils possèdent des énergies totales plus basses et équivalentes à l'énergie totale de ligand de référence.

Résumé

L'outil informatique est actuellement indispensable dans la recherche en biologie pour traiter le flot de données produites et optimiser ses avancées. Le "docking" moléculaire est, en effet, une des méthodes communément utilisée en pharmacochimie pour découvrir et mettre au point de nouveaux médicaments par criblage de milliers de composés pour une protéine cible donnée.

Les méthodes de modélisation moléculaire utilisées dans notre travail sont : MM⁺, PM3, (DFT /B3LYP/6-31++G (d,p)). Ces méthodes ont été utilisées pour déterminer les paramètres physicochimiques et quantiques associés aux molécules étudiées.

Une étude QSAR a été effectuée sur 25 dérivées de phthalazine. La régression linéaire multiple (MLR) a été utilisée pour quantifier la relation entre les descripteurs moléculaires et l'activité anticancéreuse (cancer de la peau) des dérivées de phthalazine. Le pouvoir prédictif de modèle obtenu a été confirmé par les deux méthodes de validation croisée LOO et validation externe. Une forte corrélation a été observée entre les valeurs expérimentales et prédites des activités biologiques, ce qui indique la validité et la qualité des modèles QSAR obtenus.

Enfin, une étude par docking moléculaire a été effectuée également pour une série bioactive de phthalazine, après le calcul on recherche la meilleure pose et la formation des complexes plus stable, les meilleures molécules obtenues sont des médicaments anticancéreux.

Mots-clés : Dérivées de Phtalazine, cancer de la peau, DFT, QSAR ,MLR ,Doking.

Abstract

The computer tool is currently requisite in research in biology to treat the stream of data produced and to optimize its progress. The " docking " is, indeed, one of the methods commonly used in pharmacochimistry to discover and to finalize of new medicines by screening of thousands of compounds for a protein target.

The molecular modeling methods used in our work are: PM3, (DFT / B3LYP / 6-31 ++ G (d, p)). These methods were used to determine the physicochemical and quantique parameters associated with the studied molecules.

A QSAR study was performed on 25 phthalazine derivatives. Multiple linear regression (MLR) was used to quantify the relationship between molecular descriptors and the activity anticancer (skin cancer) of phthalazine derivatives. The model predictive power obtained was confirmed by the two methods of LOO cross-validation and external validation. A strong correlation was observed between the experimental and predicted values of biological activities, indicating the validity and quality of the QSAR models obtained.

Finally, a molecular doking study was also carried out for a bioactive series of phthalazine, after the calculation we search the best pose (Conformation) and the formation of the complexes more stable, the best molecules obtained are anticancer drugs.

Keywords: Phtalazine derivatives, skin cancer, DFT, QSAR ,MLR, Doking.

المخلص

أداة الكمبيوتر ضرورية حاليا في بحوث البيولوجيا لمعالجة تدفق البيانات المنتجة وتحسين تقدمها. في الواقع ، بعد "الالتحام" الجزيئي أحد الأساليب الشائعة في الكيمياء الصيدلانية لاكتشاف وتطوير أدوية جديدة عن طريق فحص آلاف المركبات لبروتين مستهدف معين.

طرق النمذجة الجزيئية المستخدمة في عملنا MM⁺, PM3, (DFT / B3LYP / 6-31 ++ G (d,p))، تم استخدام هذه الطرق لتحديد المعلمات الفيزيوكيميائية والكوانتية المرتبطة بالجزيئات المدروسة .

تم إجراء دراسة QSAR على 25 مشتق من فتالازين تم استخدام الانحدار الخطي المتعدد (MLR) لتحديد العلاقة بين الواصفات الجزيئية والنشاط السام للخلايا لمشتقات الفتالين. تم تأكيد القدرة التنبؤية النموذجية التي تم الحصول عليها من خلال طريقتين للتدقيق المتبادل LOO والتحقق من الصحة الخارجية. وقد لوحظ ارتباط قوي بين القيم التجريبية والمتوقعة للأنشطة البيولوجية ، مما يشير إلى صحة وجودة نماذج QSAR التي تم الحصول عليها.

وأخيرا، تم إجراء الدراسة بواسطة Doking الجزيئية أيضا عن سلسلة من فتالازين النشطة بيولوجيا، بعد حساب نبحت أفضل تكيف وتشكيل المعقدات أكثر استقرارا، وأفضل الجزيئات التي تم الحصول عليها هي الأدوية المضاد للسرطان.

الكلمات الأساسية : فتالازين سرطان الجلد، DFT, QSAR , MLR, Doking.