



Université Mohamed Khider de Biskra
Faculté des Sciences Exactes et des Sciences de la Nature et de la Vie
Département des Sciences de la Matière

MÉMOIRE DE MASTER

Domaine : Sciences de la Matière

Filière : Chimie

Spécialité : Chimie pharmaceutique

Réf. :

Présenté et soutenu par :

Cherroun Hana

Le : Mercredi 26 juin 2019

Étude quantitative de la relation structure/activité (QSAR) des dérivés pipéraziniques de la phénothiazine

Jury :

Mellaoui Malika	M.C.A	Université Med Khider de Biskra	Présidente
Almi Zineb	M.C.B	Université Med Khider de Biskra	Rapporteur
Sellami Mohamed	M.A.A	Université Med Khider de Biskra	Examineur

Année universitaire : 2018/2019



Dédicace

Je vous dédie ce travail en témoignage de mon profond amour. Puisse Dieu, le tout puissant, vous préserver et vous accorder santé, longue vie et bonheur.

A mes chers parents pour tous leurs sacrifices, leur amour, leur tendresse, leur soutien leurs prières tout au long de mes études.

Aux chandelles de ma maison mes chers frères : Sohaib ,rachid , fathalleh .

A ma grande famille.

A mes amis

A tous les proches que j'ai mentionnés et les autres que j'ai oubliés veuillez m'excuser.



Remerciements

En tout premier lieu, je remercie le bon dieu, tout puissant, de m'avoir donné la force pour survivre, ainsi que l'audace pour dépasser toutes les difficultés.

*Je tiens tout d'abord à remercier ma superviseuse, M^{me} **Almi Zineb** de université de Biskra pour d'avoir proposé, accepté et supervisé ce travail, et j'ai témoigné fidèlement de toute ma gratitude pour son esprit critique et sa direction qui ont grandement aidé mon travail, ainsi que pour la patience et la gentillesse qu'elle m'avait montrées. Je dois la remercier pour ses efforts qui m'ont permis d'aboutir dans mon travail de mémoire.*

*Je remercie également les membres de jury M^{me} **Mellaoui Malika** et M^r **Sellami Mohamed** pour l'honneur qu'ils m'ont fait en acceptant de juger mon travail.*

Enfin je veux dire merci à toute ma famille; mais surtout à ma mère, mes frères, qui me soutiennent pour leur aide, leur patience et leur encouragement pendant mon travail

Sommaire

Liste des tableaux	1
Liste des figures	2
Liste des principales abréviations	3
Introduction générale	5
<u>Chapitre I</u> : Généralités sur la maladie de cancer et de l'activité anticancéreuse des phénothiazines pipéraziniques sur les cellules MDR1.	
I.1.Généralité sur la maladie de cancer	9
I.1.1.Rappel historique	9
I.1.2.Définition.....	9
I.1.3. Classification de cancer	9
I.1.3.1. Tumeur	9
I.1.3.1.1. Définition	9
I.1.3.1.2.Types de tumeurs	9
I.1.3.2. Métastases.....	10
I.1.3.2.1. Définition.....	10
I.1.3.2.2. Génétique des métastases.....	10
I.1.4. Facteurs cancérogènes	12
I.1.5. La cellule cancéreuse	12
I.1.5.1. Caractères généraux de la cellule cancéreuse	12
I.1.5.1.1. Caractères résiduels	12
I.1.5.1.2. Caractères morphologiques acquis	12
I.1.5.1.3. Caractères dynamiques	13
I.1.5.1.4. Anomalies biochimiques	13
I.1.6. Chimiothérapie	13
I.1.6.1. Des agents anti-tumoraux	13
I.2. L'activité biologique anti -cancéreux	14
I.2. 1. Historique	14
I.2. 2. Les anticancéreux	14

I.2. 3. Mode d'action	14
I.2.4. Transporteurs d'efflux	15
I.2.5. Mécanisme d'inhibition de p-gp	16
I.3. Les phénothiazines	17
I.3.1. Définition	17
I.3.2. Les classes des phénothiazines	18
I.3.3. Pharmacologie et mode d'action des phénothiazines	20
I.3.4. Les effets indésirables des anticancéreux.....	21
Références bibliographiques.....	23

Chapitre II : Etude quantitative de la relation structure/activité QSAR.

II.1. Les bases de la chimie quantiques.....	26
II.2. Les méthodes de calcul.....	27
II.2.1. Les méthodes empiriques	27
II.2.2. Les méthodes semi-empiriques	27
II.2.3. Les méthodes quantiques	27
II.2.3.1. Théorie de la fonctionnelle de la densité(DFT)	28
II.2.3.2. Ab initio	28
II.2.4. QSAR (relation quantitative structure-activité)	29
II.3. Les paramètres utilisées dans l'étude QSAR.....	29
II.3.1. Les paramètres électroniques	29
II.3.2. Les paramètres d'hydrophobicité	29
II.3.3. Les paramètres stériques	29
II.4. Le model QSAR	33
II.4.1. Définition	33
II.4.2. Outils et Méthodologie de QSAR.....	33
II.4.2.1. Paramètres biologiques.....	33
II.4.2.2. Les descripteurs moléculaires	34
II.4.3. La méthode de la régression linéaire multiple (MLR).....	36
II.4.3.1. Description de la méthode	36
II.4.3.2. Test de la signification totale de la régression.....	37

II.4.3.3. Le choix du meilleur modèle	38
II.4.3.4. Validation du modèle LOO	39
II.4.4. Les applications de l'étude QSAR	40
II.5. Les règles de lipinsky.....	40
Références bibliographiques.....	42

Chapitre III: Résultats et discussions.

III.1. Introduction	46
III.2. Données expérimentales	46
III.3. Etude des paramètres physico-chimiques et électroniques des phénothiazines pipéraziniques.....	47
III.4. Etude quantitative de la relation structure-activité QSAR des phénothiazines pipéraziniques	52
III.4.1. Développement de modèle QSAR	52
III.4.2 Validation de modèle QSAR.....	53
III.4.3. Prédiction de l'activité.....	54
III.4.4. La matrice de corrélation.....	56
III.5. Les règles de Lipinsky	56
Référence bibliographiques	58
Conclusion générale	61
Annexe	

Liste des tableaux

Tableau I.1 : Les formules générale des dérivées de la phénothiazine «ChemDraw».....	19
Tableau I.2 : Les structures chimiques et l'activité biologique des phénothiazines pipéraziniques.....	20
Tableau II.1 : Les types des données biologiques utilisées dans l'analyse QSAR.....	34
Tableau III.1: L'activité biologique de la série des dérivés des phénothiazines pipéraziniques étudiée.....	47
Tableau III.2 : Les valeurs des paramètres physicochimiques des phénothiazines pipéraziniques.....	48
Tableau III.3 : Les valeurs des descripteurs électroniques utilisés dans l'étude QSAR.....	51
Tableau III.4: Les paramètres de validation croisée.....	53
Tableau III.5: L'activité anticancéreuse expérimentale, prédite et résiduelle des phénothiazines pipéraziniques.....	54
Tableau III.6 : La matrice de corrélation du modèle.....	56
Tableau III.7 : Les règles de Lipinsky des phénothiazines pipéraziniques.....	57

La liste des figures

Figure I. 1 : Le processus malin comparé au processus bénin.....	10
Figure I.2 : Les différents types de cancer.....	11
Figure I.3 : Schéma représentant les mécanismes de résistance à de multiples drogues des cellules cancéreuses.....	15
Figure I.4: pompe d'efflux de l'exemple de la P-gp.....	16
Figure I.5 : Schéma de base des phénothiazines «ChemDraw».....	17
Figure I.6: Formule générale des dérivés de la phénothiazine «ChemDraw».....	18
Figure I.7: Formule générale des phénothiazines pipéraziniques «ChemDraw».....	19
Figure I.8 : La structure des phénothiazines pipéraziniques «ChemDraw».....	20
Figure II.1: La liaison hydrogène	31
Figure II. 2: Polarité des différents milieux cellulaires.....	32
Figure II.3: Deux exemples sur les liaisons donneurs et accepteurs des hydrogènes.....	41
Figure III.1: Structure générale des dérivés de la phénothiazine «ChemDraw».....	46
Figure III.2: Les sites accepteurs dans le composé T116.....	49
Figure III.3: Les sites accepteurs dans le composé T174.....	49
Figure III.4: Schéma de la phénothiazine «ChemDraw».....	50
Figure III.5: La courbe de la régression linéaire des valeurs prédites en fonction des valeurs expérimentales pour le modèle QSAR.....	55
Figure III.6: Représentation graphique des valeurs résiduelles en fonction des valeurs expérimentales de pIC50.....	55

Liste des principales abréviations

MDR	<i>Multiple Drug Resistant</i>
P-gp	P-glycoprotéine
Ĥ	Opérateur hamiltonien
Ψ	fonction d'onde
PM3	Méthode paramétrique 3
DFT	La théorie fonctionnelle de la densité
DM	Moment dipolaire
HBD	Donneur de liaison hydrogène
HBA	Accepteur de liaison hydrogène
Log P	Coefficient de Partage
MR	Réfractivité molaire
V	Volume
HE	l'énergie d'hydratation
Pol	la polarisabilité
MW	la masse moléculaire
SAG	Surface grille
IC₅₀	Concentration inhibitrice de 50%.
QSAR	Relation quantitative structure-activité
LOO	Leave-one-out.
MLR	La régression linéaire multiple.
PRESS	somme des carrés résiduels
R_{2adj}	Qualité d'ajustement.
R_{2cv}	Coefficient de corrélation de validation croisée.
PE	L'erreur prédictive du coefficient de corré

Introduction générale

Introduction générale

Le cancer est une maladie répandue dans le monde, une maladie tue de nombreuses personnes et que l'on appelle également une croissance anormale des tissus cellulaires dans le corps, de sorte qu'elle se propage de manière incontrôlable.

Il faut noter qu'il affecte différentes parties du corps et ses symptômes varient en fonction du tissu ou de l'organe touché et peuvent se propager à tous les organes du corps par l'invasion des cellules en croissance entourant le tissu, entraînant l'apparition de la tumeur à différents endroits du corps et pouvant être éloignés du membre blessé.

La résistance aux cytostatiques est l'obstacle majeur au traitement de nombreux cancers. Les causes d'échecs des chimiothérapies sont multiples, la principale étant une résistance intrinsèque de la cellule tumorale, la reconnaissance d'une glycoprotéine de 170 kDa (la P-gp) à la surface de cellules de lignées tumorales rendues résistantes à des cytostatiques de monde d'action différents, a permis de décrire un phénotype de cellules « multi-résistantes » (MDR, pour *multi drug resistance*), cette P-gp, codée par le gène *mdr1* chez l'homme, a un rôle de pompe, chassant hors de la cellule les cytostatiques, sa responsabilité dans la résistance clinique des tumeurs humaines est suspectée, en raison de son expression souvent élevée dans des cancers d'emblée ou devenus chimio résistants, des essais thérapeutiques d'inhibition de cette P-gp sont actuellement en cours [1].

La chimie-informatique, également dénommée chemoinformatique ou chimio-informatique, est le domaine de la science qui consiste en l'application de l'informatique aux problèmes relatifs à la chimie. Elle a pour objectif de fournir des outils et des méthodes pour l'analyse et le traitement des données issues des différents domaines de la chimie [2].

Le développement important des moyens informatiques a permis à la chimie de s'enrichir d'outils informatiques spécialement dédiés à la représentation des molécules et de mieux appréhender les problèmes liés au domaine moléculaire : la modélisation moléculaire[3.4].

La modélisation moléculaire est une application des méthodes théoriques et des méthodes de calcul pour résoudre des problèmes impliquant la structure moléculaire et la réactivité chimique [5].

La modélisation QSAR (Quantitative Structure-Activity Relationship) constitue actuellement un outil fiable et performant pour combler les lacunes des études *in vivo* et *in vitro* et limiter le recours aux tests biologiques. C'est une méthode qui tente à identifier une

relation quantitative entre l'activité mesurée expérimentalement et des descripteurs théoriques ou empiriques liés à la structure moléculaire [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) [7].

La régression linéaire multiple (MLR) est une méthode statistique utilisée pour quantifier la relation entre les variables indépendantes obtenues par les calculs et une variable dépendante donnée par l'expérience. Cette méthode repose sur un ensemble des paramètres statistiques qui définissent la combinaison linéaire entre les variables indépendantes (descripteurs moléculaires) et l'activité biologique des molécules étudiées [8].

Notre travail met en évidence une étude quantitative des relations structure/activité des propriétés anticancéreuses des phénothiazines pipéraziniques. Cette étude permet de développer un modèle mathématique utilisé dans la prédiction.

Le présent travail est divisé en deux parties :

✓ La première partie servant de rappel bibliographique, comporte de deux chapitres contenant respectivement :

Chapitre I : Généralités sur la maladie de cancer et de l'activité anticancéreuse des phénothiazines pipéraziniques sur les cellules MDR1

Chapitre II : Généralités sur les méthodes de recherche utilisées dans l'étude quantitative de la relation structure/activité QSAR.

✓ La deuxième partie, dans laquelle nous avons analysé les résultats de nos calculs, elle est composée de chapitre III :

Chapitre III : les résultats obtenus par l'étude quantitative "QSAR" de l'activité anticancéreuse des phénothiazines pipéraziniques et leurs discussions, et aussi les résultats de règle de lipinsky.

Références bibliographiques

- [1]. J.P. Marie, Le phénomène de résistance multiple aux anti-cancéreux : les gènes MOR et la P-gp, médecine/sciences. 6, 443-8 (1990).
- [2]. R.Djani, Mémoire de master, Application de la méthode DFT dans l'étude quantitative des relations structure-activité des propriétés chimio-préventives de certains dérivés de 1,2-dithiole-3-thione, Université de Biskra, (2014).
- [3]. T. Clark, Hand book of computational chemistry, Edition, wiley, london, (1985).
- [4]. S. Belaidi, M. Omari, T. Lanez et A. Dibi, J. Soc. Alger. Chim., pp.14, 27, (2004).
- [5]. A.D.Mighell, V.L. Himes et J.R. Rodgers, Acta Cryst. A39, 737 (1983).
- [6]. Y. Yousfi, Mémoire de master, Etude QSAR de l'activité antioxydant d'une série de composés phénoliques, Université Tlemcen, (2017).
- [7]. C.G. Wermuth, The Practice of Medicinal Chemistry 3rd edition, Elsevier Ltd, (2008).
- [8]. N. Goudarzi, M. Goodarzi, QSAR prediction of HIV inhibition activity of styryl quinoline derivatives by genetical gorithm coupled with multiple linear regressions, medicinal chemistry research. 21, 437-443 (2012).

Chapitre I

Généralité sur la maladie de cancer et de l'activité anticancéreuse des phénothiazines pipéraziniques sur les cellules MDR1

I .1.Généralité sur la maladie de cancer

I. 1.1.Rappel historique

Hippocrate fut le premier à définir le cancer, qu'il appela alors «carcinome ».Pour Hippocrate, le «carcinome» était une tumeur envahissante conduisant à une mort inéluctable, au début du 18ème siècle eurent lieu les premières observations épidémiologiques relatives au cancer. En 1713, Ramazzini observa que le cancer du sein était particulièrement fréquent chez les religieuses puis Pott, en 1775, nota une prévalence élevée du cancer de la peau du scrotum chez les ramoneurs .Ces études ont permis d'établir la notion de facteurs de risques, qui sont des facteurs pour lesquels la relation de causalité avec le cancer est probable [1].

I.1.2.Définition

Le cancer est une maladie, provoquée soit par des changements sous-jacents de l'information génétique, se traduisant soit par des modifications du comportement cellulaire, soit par un comportement anormal de cellules souches.

Ces modifications peuvent se traduire par une prolifération cellulaire incontrôlée dans certains tissus conduisant à la formation d'une masse tissulaire dénommée tumeur. Les tumeurs peuvent être bénignes (non cancéreuses) ou malignes (cancéreuses). Les cellules formant des tumeurs bénignes demeurent localisées dans le tissu d'origine. Les cellules à l'origine des tumeurs malignes envahissent par définition les tissus voisins (invasion locorégionale) ou des organes à distance (métastases) [2].

I.1.3. Classification de cancer

I.1.3.1. Tumeur

I.1.3.1.1. Définition

C'est une prolifération cellulaire excessive aboutissant à une néoformation tissulaire ayant tendance à persister et à s'accroître indéfiniment, ce qui témoigne d'une certaine autonomie biologique.

L'expression de pseudo- tumeur est appliquée à des lésions qui, tout en offrant des ressemblances cliniques ou morphologiques avec ces tumeurs, n'en ont pas comportement biologique. Ce sont les lésions dystrophiques, malformatives ou inflammatoires [1].

I.1.3.1.2.Types de tumeurs

Les tumeurs peuvent être classées en deux catégories : les tumeurs bénignes et les tumeurs malignes ou cancers (Figure I.1).

- a) Les tumeurs bénignes sont entourées d'une capsule, qui les rend non invasives. Elles ont une croissance limitée et sont rarement mortelles.
- b) les tumeurs malignes, les cellules ont la capacité de donner des foyers tumoraux secondaires (métastases) situés à distance de la tumeur primaire. De plus, les tumeurs malignes manifestent souvent une plus grande agressivité locale en termes d'envahissement et de destruction vis-à-vis des tissus environnants [3].

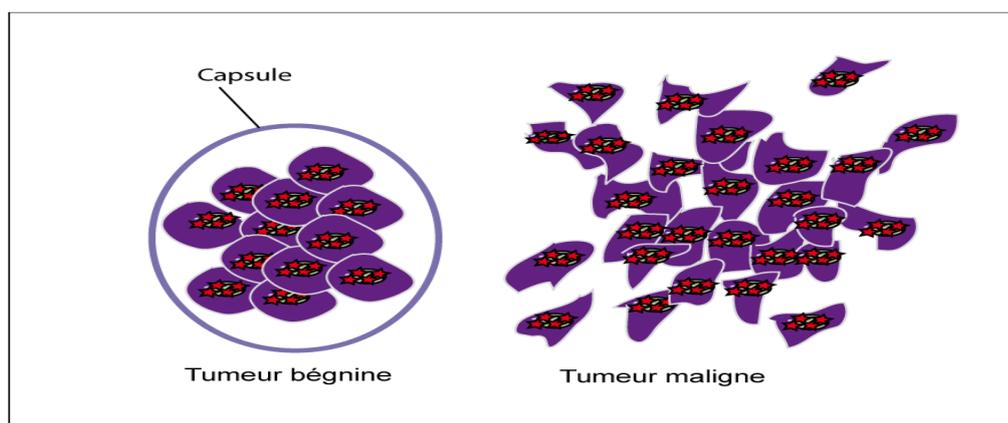


Figure I. 1 : Le processus malin comparé au processus bénin [3].

I.1.3.2. Métastases

I.1.3.2.1. Définition

Le terme métastase signifiant changement de place en grec fait référence à la croissance de tumeurs secondaires en des sites distants d'un néoplasme primitif [4] OMS. Ce terme différencie donc les lésions malignes des lésions bénignes et caractérise la dernière étape de la progression tumorale. La croissance métastatique est la principale cause d'échec des traitements et de décès des patients cancéreux. Ainsi 90% des patients atteints d'un cancer décèdent de leurs métastases [5]. Le terme métastase est réservé à la dissémination des cellules tumorales via le sang ou le système lymphatique mais la propagation dans le fluide cérébrospinal et le passage transcoelomique est aussi possible. La plupart des patients cancéreux (60 à 70%) présentent des métastases manifestes ou occultes lors du diagnostic, et le pronostic pour la majorité est sombre [4].

I.1.3.2.2. Génétique des métastases

Avec la publication de la séquence du génome humain, et différentes initiatives majeures telles que le cancer génome projet au Royaume-Uni et le cancer génome anatomie projet aux USA, la recherche de gènes régulés de manière positive, mutés ou perdus dans des

cancers métastatiques a gagné du terrain. Il est maintenant possible par capture laser et l'analyse en série de l'expression des gènes, d'isoler des cellules cancéreuses invasives et de comparer leur expression protéique ou génique à celle de cellules normales ou non invasives du même patient. Avant ceci, les techniques de transfection de chromosomes ou d'ADN de cellules métastatiques à des cellules non métastatiques (ou vice versa), ont permis d'identifier certains gènes spécifiquement liés aux métastases, bien que beaucoup d'autres ainsi identifiés soient aussi associés à la croissance tumorale. [2]

- On peut être classé les cancers en fonction du type de tissu affecté (Figure I.2) [3].
- ✓ Les carcinomes sont les cancers les plus courants. Ils proviennent des cellules qui recouvrent les surfaces internes et externes du corps c'est-à-dire les tissus épithéliaux comme le poumon et le côlon.
- ✓ Les sarcomes sont des cancers qui touchent les tissus conjonctifs c'est-à-dire les tissus de soutien de la structure de l'organisme comme les tissus osseux, cartilagineux, musculaires, adipeux et vasculaires. L'incidence de ce type de cancer est rare.
- ✓ Les lymphomes touchent les tissus hématopoïétiques qui regroupent les ganglions lymphatiques et les tissus du système immunitaire.
- ✓ les leucémies concernent les tissus de la moelle osseuse responsable de la production des globules blancs .L'incidence de ce type de cancer est rare.
- ✓ les leucémies concernent les tissus de la moelle osseuse responsable de la production des globules blancs.

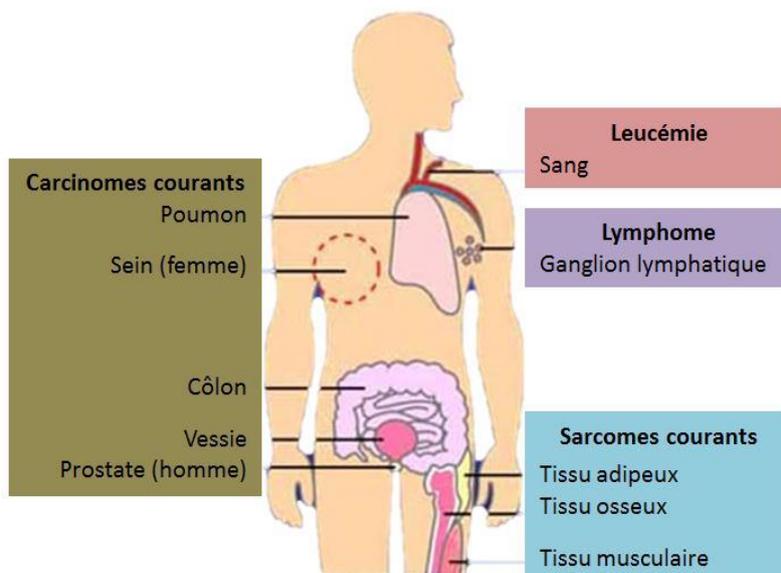


Figure I.2 : Les différents types de cancer [3].

I.1.4. Facteurs cancérogènes

Tous les agents carcinogènes se sont révélés des initiateurs de la cancérogenèse c'est-à-dire qu'ils sont capables de transformer une cellule normale en une cellule tumorale dormante. Cette action est très rapide, parfois instantanée est toujours irréversible. Des promoteurs assurent le passage des cellules tumorales dormantes aux tumeurs évolutives; leur action est lente et parfois irréversible [6].

I.1.5. La cellule cancéreuse

Le développement d'une tumeur maligne implique qu'elle n'est pas rejetée par les tissus environnants et par l'organisme en général. Des caractères particuliers, acquis des cellules malignes, permettent le développement du cancer.

I.1.5.1. Caractères généraux de la cellule cancéreuse

I.1.5.1.1. Caractères résiduels

La cellule cancéreuse garde à un certain degré, la fonction de sa différenciation et des caractéristiques de son tissu d'origine.

I.1.5.1.2. Caractères morphologiques acquis

Les anomalies morphologiques des cellules cancéreuses sont nombreuses, mais aucune n'est constante ou spécifique du cancer :

- ✓ Anomalies du noyau : souvent volumineux, irrégulier, parfois multiples. Les nucléoles sont visibles. Les anomalies chromosomiques sont fréquentes et plus souvent à type d'hyperploïdie.
- ✓ Anomalies cytoplasmiques : cytoplasme souvent très basophile avec augmentation du rapport nucléo cytoplasmique.
- ✓ Anomalies des mitoses : elles sont en général plus volumineuses que dans le tissu normal, avec désorientation du fuseau de division, voire division multipolaires.
- ✓ Anomalies de la membrane cytoplasmique : La seule définition à peu près universelle de la cellule cancéreuse est donnée par sa capacité de proliférer, sans être parfaitement contrôlée par les mécanismes normaux de régulation. Cette incapacité de recevoir les messages régulateurs émis par les autres cellules et les autres tissus incite évidemment à étudier les membranes des cellules cancéreuses, intermédiaires obligatoires des relations intercellulaires. Comme anomalies majeures, il y a perte de l'inhibition de contact, modification de l'adhésivité, modification des antigènes de surface.

I.1.5.1.3. Caractères dynamiques

Le taux de multiplication des cellules cancéreuses est anormal, élevé, autonome, anarchique et fragile (importante mort cellulaire par hypoxie) et indéfini. Le phénomène de tumorigénicité est observé après injection de la cellule cancéreuse dans un organisme immunotolérant (souris nude) [6].

I.1.5.1.4. Anomalies biochimiques

Les anomalies biochimiques des cellules cancéreuses sont fréquentes et nombreuses. Aucune des anomalies rapportées ne semble spécifique, la cause ou la condition de la cancérisation. L'activité de synthèse, de sécrétion, d'excrétion de la cellule cancéreuse est en général moindre que celle de la cellule normale du tissu d'origine (moindre différenciation). Par ailleurs, la dé-répression de certains gènes peut aboutir à la synthèse de substances en quantité et /ou en qualité anormale ou de substance de type embryonnaire [7.8].

I.1.6. Chimiothérapie

La chimiothérapie tient une place importante dans le traitement des cancers. Des milliers de substances sont proposées chaque année par les laboratoires de recherches. La plupart proviennent en réalité de modifications chimiques, plus ou moins simples, d'agents chimiothérapeutiques connus [6].

Ces dérivés ou encore homologues sont des agents dont le spectre anti-tumoral peut être différent de celui de la substance d'origine. Par ailleurs, certains dérivés sont développés dans le but de diminuer la toxicité liée à la substance sans chercher à augmenter son efficacité. Seul un petit nombre de ces substances satisfait à des critères précis d'activité anti-tumorale [9.10].

I.1.6.1. Des agents anti-tumoraux

Le cancer étant en gros provoqué par un développement anarchique des tissus dont les acides nucléiques sont considérés comme éléments formateurs, il était inévitable que l'on s'ingénie à freiner ce développement par des antagonistes de constituants de ces acides « créés » classiquement par analogie structurale [11].

Il est difficile d'élaborer une classification homogène du grand nombre de drogues anticancéreuses actuellement disponibles. En effet, il existe une très grande diversité des formules chimiques et les origines de ces substances, naturelles ou synthétiques, sont également très variées.

Classification schématique des drogues anticancéreuses selon leur principal mécanisme d'action cellulaire :

- Alkylants
- Intercalant: Modification (s) directe(s) de la molécule d'ADN
- Scindant
- Antimétabolites : Inhibition de la synthèse d'ADN.
- Poisons de fuseaux: Inhibition de la polymérisation de la tubuline [6].

I.2. L'activité biologique anti -cancéreux

I.2. 1. Historique

La découverte des anti-oncogènes, ou gènes suppresseurs de tumeurs a permis aux chercheurs de mieux comprendre le mécanisme génétique d'apparition des tumeurs, ceci permet d'entrevoir la possibilité d'une approche entièrement nouvelle des cancers [12].

I.2. 2. Les anticancéreux

Les anti-cancéreux (MDR) est définie comme étant une résistance croisée, étendue à résistance multiple aux des cytostatiques de famille et de mode d'action divers.

Le phénomène « MDR », a été mis en évidence par l'observation désignées tumorales exposées de façon continue et prolongée à des doses croissantes d'un cytostatique [13.14]. La constatation que ces lignées étaient devenues d'emblée résistante sa d'autres cytostatiques, auxquelles elles étaient exposées pour la première fois, a conduit Biedler et Riehm, dès 1970, à envisager un mécanisme commun de résistance. Les médicaments intéressés par le phénomène MDR contiennent généralement un noyau aromatique [15].

La sous famille MDR (*Multiple Drug Resistant*) représentée par au maximum 3 membres selon les espèces, chacun des gènes codant pour un isoforme de pompe transmembranaire appelé Glycoprotéine P (Gly-P). Ils sont capables d'expulser à l'extérieur de la cellule des xénobiotiques non modifiés [16].

I.2. 3. Mode d'action

Ci –dessus est dressée une liste non exhaustive des mécanisme impliqués dans le phénotype MDR .Gillet et Gottesmann ont élaboré un schéma récapitulant ces différents mécanismes (figure I.3) .Malheureusement , bien que les études in vitro aient permis d'identifier et de valider ces mécanismes , l'équipe du Dr Gottesman a mis en évidence la grande différence entre les gènes impliqués dans les modèles cellulaires utilisés et les situation cliniques réelles, insistant sur la nécessité de développer de nouveaux modèles précliniques se rapprochant au plus près des situations cliniques.

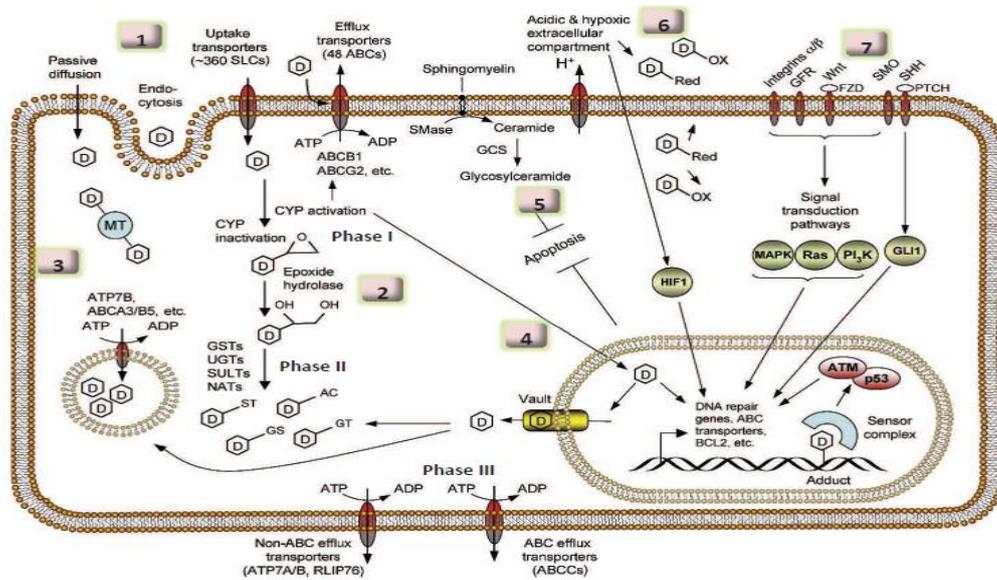


Figure I.3 : Schéma représentant les mécanismes de résistance à de multiples drogues des cellules cancéreuses [17].

De plus, même dans les modèles *in vitro*, le phénotype MDR, est souvent du à plusieurs mécanismes. Parfois, un mécanisme en engendrer un autre par une sorte de réaction en cascade que la cellule met en œuvre pour se défendre contre les agents chimiothérapeutiques. Par exemple, il a été démontré que la protéine p53 régule l'expression d'ABCB1 ; p53 sauvage entraînait une baisse de l'expression d'ABCB1 alors que la protéine mutante (exprimée par les cellules cancéreuses de phénotype MDR) induisait une surexpression du transporteur ABC. Cependant, ce phénomène est tissu-dépendant, ensuite, Tew et Townsend ont démontré que la résistance dues à la GST pouvait être issue de différents mécanismes. D'autre part, plusieurs études ont montré que la résistance à un même composé pouvait résulter de différents mécanismes, comme par exemple l'irinotécan, au Cis-platine ou encore à la mitoxantrone. En fin la complexité de l'étude de ce phénotype MDR réside également dans le fait que les gènes impliqués peuvent être différents d'un tissu, voire même d'un individu à l'autre. Cette observation indique l'importance de l'émergence de nouvelles thérapies ciblées et adaptées à chaque situation. [17].

I.2.4. Transporteurs d'efflux

Les protéines de la superfamille ABC utilisent l'hydrolyse de l'ATP pour transporter des substrats du milieu intracellulaire vers le milieu extracellulaire, le plus souvent contre un gradient de concentration. Le transporteur MDR1 ou P-glycoprotéine (P-gp), codé par le gène ABCB1, est physiologiquement impliqué dans l'efflux de très nombreux toxiques et médicaments [18]. Il est présent au pôle canaliculaire des hépatocytes, mais également au

niveau d'autres barrières clés de l'organisme (intestin, rein, barrière hémato-encéphalique)[19].

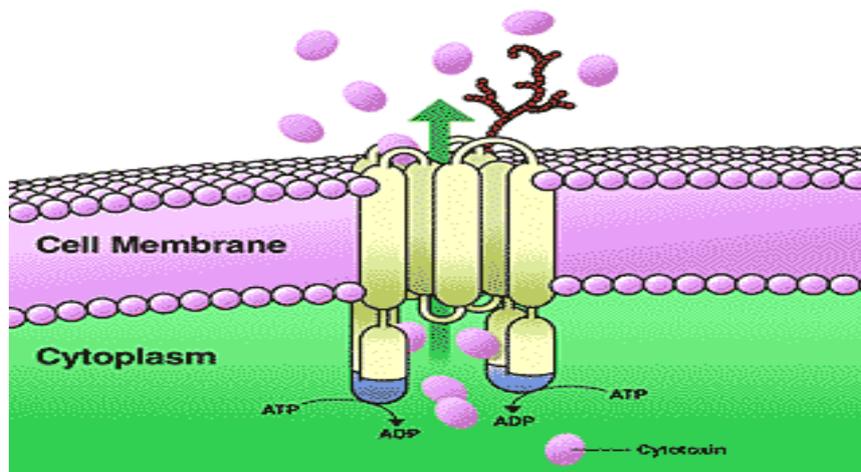


Figure I.4: pompe d'efflux de l'exemple de la P-gp

I.2.5. Mécanisme d'inhibition de p-gp

De multiples mécanismes moléculaires sont responsables de l'inhibition de la P-gp. Il est difficile de savoir lequel intervient lors d'une interaction entre un substrat et un inhibiteur donné. Cette complexité empêche de prédire les conséquences quantitatives ou qualitatives des interactions. La P-gp comprend plusieurs sites de fixation des substrats, qui interagissent avec les deux domaines fixant l'ATP pour former une unité fonctionnelle. Inhibition "substrat-dépendante" : Le modèle d'inhibition de la P-gp semble être "substrat-dépendant". Il peut être classé en deux grandes catégories [20].

- Inhibition compétitive : Deux substrats agissent sur le même site de fixation et un seul peut se fixer à la fois, exemple: le vérapamil et la vinblastine sont deux substrats de la P-gp. Lorsqu'ils sont administrés simultanément, le vérapamil inhibe le transport de la vinblastine en se fixant à sa place sur le site de liaison du substrat.
- Inhibition non compétitive : le substrat et l'inhibiteur se fixent simultanément sur deux sites distincts et fonctionnellement indépendants de la P-gp. Il existe alors 3 cas différents :
 - ✓ Soit l'inhibiteur se fixe à la place de l'ATP sur son site de fixation,
 - ✓ Soit il bloque l'hydrolyse de l'ATP nécessaire au transport du substrat,
 - ✓ Soit le substrat et l'inhibiteur se fixent simultanément sur deux sites différents :

L'interaction entraîne un effet allostérique, l'inhibiteur bloque le changement de conformation de la P-gp nécessaire pour faire passer le substrat d'un côté à l'autre de la membrane cellulaire,

exemple : La ciclosporine A est à la fois substrat et inhibiteur de la P-gp. Elle peut se fixer sur le site de liaison du substrat pour être transportée. Elle peut également bloquer l'hydrolyse de l'ATP et inhiber ainsi le transport d'un autre substrat [21].

I.3. Les phénothiazines

I.3.1. Définition

Les phénothiazines sont des composés hétérocycliques qui ont des activités spécifiques contre des constituants de la membrane plasmique des eucaryotes. Bien que les premières phénothiazines ont été synthétisées à la fin du 19^{ème} siècle, la structure de base de la phénothiazine est identique à celle du bleu de méthylène coloré, qui a été préparé pour la première fois comme une tache en 1876 par le chimiste allemand Heinrich Caro comme dérivée d'une phénothiazine existant. Les substituants du cycle de la phénothiazine, qui augmentent la lipophilie, amplifient leur activité anti-proliférative dans les cellules cancéreuses. La présence de différents substituants dans le squelette de la phénothiazine, ainsi que la modification du système cyclique tricyclique, modifie leur effet biologique. En raison de leur activité biologique importante, de nombreux dérivés des phénothiazines ont été synthétisés et brevetés [22].

La phénothiazine elle-même consiste en un noyau tricyclique de 2 cycles benzéniques (phéno), reliés par un cycle central contenant un atome de soufre (thio, position 5) et un atome d'azote (azo, position 10) [23].

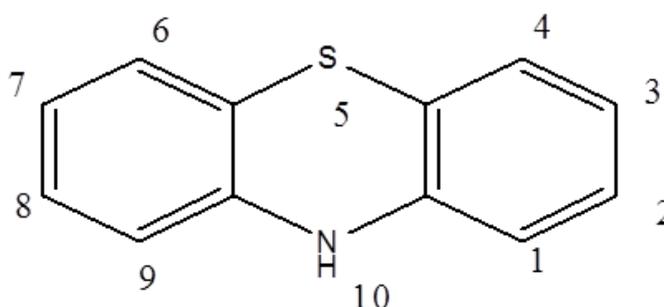


Figure I.5 : Schéma de base des phénothiazines «ChemDraw».

* Par contre, pour les dérivées de la phénothiazine, deux substituants se greffent sur cette structure de base:

–R :est un groupement qui varie suivant le type de phénothiazine : une chaîne latérale, ou un groupement cyclique qui contient l'atome d'azote N.

–X : est un reste de faible encombrement: Cl, OCH₃, COCH₃, CF₃, CN, SCH₃, ...etc

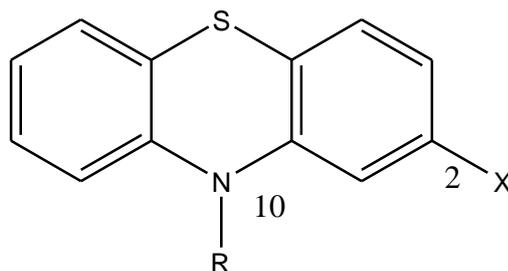


Figure I.6: Formule générale des dérivés de la phénothiazine «ChemDraw»

I.3.2. Les classes des phénothiazines

Ces composés sont classés par trois groupes selon la nature de la chaîne latérale, de l'intensité de leur activité et leur rapidité d'action [24].

- Phénothiazines à chaîne linéaire.
- Phénothiazines pipéridiniques.
- Phénothiazines pipéraziniques.

A-Phénothiazines à chaîne linéaire

Elles sont caractérisées par la présence d'une chaîne basique linéaire en (position R) sur l'azote et d'un substituant linéaire variable sur le sommet 2 (position X).

B- Phénothiazines pipéridiniques

Les dérivés de ce groupe présentent une chaîne à cycle pipéridinique, qui freine leur métabolisme et réduit leurs effets secondaires extrapyramidaux. Elles sont contenues en position 10 sur l'azote.

C- Phénothiazines pipéraziniques

Elles sont caractérisées en position 10 sur l'azote qui est ressemblé par des groupements variables selon des différents constituants. Ces composés constituent des groupes désinhibiteurs, Ils possèdent des propriétés anticancéreuses et donnent lieu à des activités anti-prolifératives et anti-MDR des phénothiazines et inhibition de la fonction de transport P-gp.

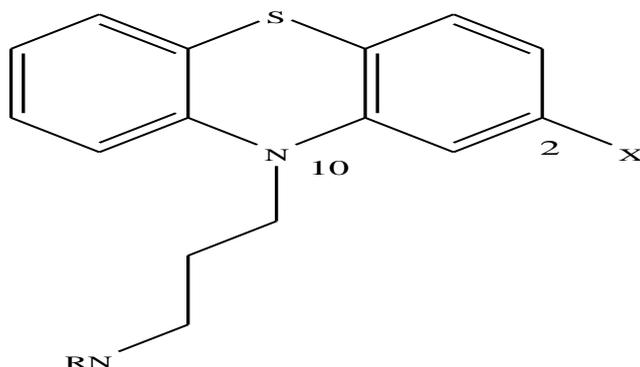


Figure I.7: Formule générale des phénothiazines pipéraziniques «ChemDraw».

Le tableau I-1 suivant présente les structures des phénothiazines en position 10 des différents constituants :

Tableau I.1: Les Formules générale des dérivées de la phénothiazine «ChemDraw» [25].

Phénothiazines à chaîne linéaire	Phénothiazines pipéridiniques	Phénothiazines pipéraziniques

Dans notre travail, on va étudier la structure de la phénothiazine pipérazinique des différents groupements de la structure suivant :

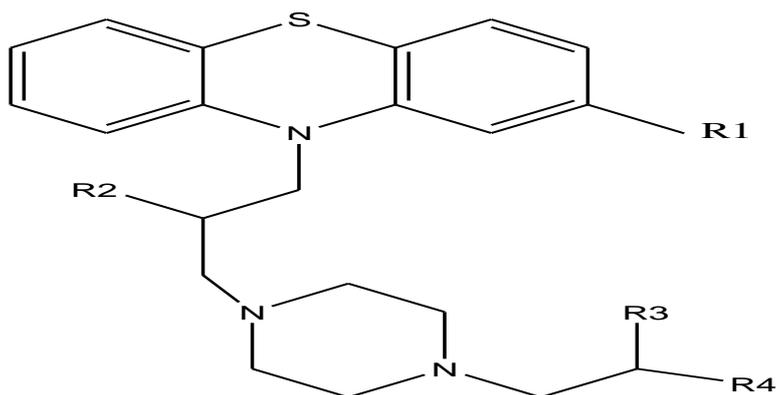


Figure I.8 : La structure des phénothiazines pipéraziniques «ChemDraw» [26].

Le tableau (I.9) suivant présente les structures chimiques et l'activité biologique des phénothiazines pipéraziniques.

Tableau I.2 : Les structures chimiques et l'activité biologique des phénothiazines pipéraziniques.

Phénothiazines Pipéraziniques	R1	R2	R3	R4	IC50 μ M [26]
T5	O	H	H		1.35
T114	OMe	O	H		3.13
T116	Cl	O	H		1.45
T145	Cl	O	Me		0.87
T158	OMe	O	Me		0.71
T172	Cl	Me	H		1.65
T174	Cl	Me	Me		1.59

I.3.3. Pharmacologie et mode d'action des phénothiazines

Les médicaments de la famille de phénothiazine présentent une large gamme des effets biologiques, avec leur principal neuroleptique action, autres activités biologiques importantes à leur effet chimioventif (activité anti-CaM, inhibition de l'activité PKC, diminution de la prolifération cellulaire et inhibition de la fonction de transport P-gp) [25]. Le mécanisme de

ces activités ont déjà été bien reconnus et liés à la structure chimique des composés de la famille de phénothiazine.

En oncologie, l'incidence des effets indésirables, notamment les effets graves, est particulièrement élevée, les anticancéreux, en raison de leur mécanisme d'action, vont, dans de nombreux cas, agir sur la cellule ou le tissu cancéreux mais aussi sur certaines cellules ou tissus normaux, la toxicité hématologique, digestive de nombreux anticancéreux est souvent indissociable de l'activité anticancéreuse, selon les molécules utilisées, certains organes-cibles seront particulièrement à risque (toxicité rénale du cis platine et de ses dérivés, toxicité cardiaque des anthracyclines, etc.)[27].

Les phénothiazines présentent un effet anti-prolifératif sur le néoplasme et cellules endommagées génotoxiquement. L'activité biologique importante des phénothiazines est leur capacité de renverser le RIM des cellules néoplasiques aux cytostatiques drogues. Les phénothiazines pourraient être incluses dans les groupes composés qui augmentent la sensibilité cellulaire aux cytostatiques médicaments (ils rétablissent la sensibilité des médicaments à la drogue cellules), principalement par une forte inhibition du Mécanisme dépendant du P-gp du MDR [26].

I.3.4. Les effets indésirables des anticancéreux

La plupart des médicaments anticancéreux ont un index thérapeutique très faible. Les chimiothérapies sont pour la plupart cytotoxiques. Quand un médicament anticancéreux agit, c'est qu'on est déjà dans une zone de concentration toxique : donc il y a souvent des effets indésirables (EI). Les EI aigus sont à distinguer des chroniques[28] :

1. Effets indésirables aigus

Index thérapeutique très faible : Effets toxiques apparaissent de quelques heures à quelques jours après l'administration et durent de quelques heures à huit semaines, en général réversibles.

- A. Toxicité hématologique
- B. Toxicité gastro-intestinale
- C. Toxicité rénale
- D. Toxicité vésicale
- E. Toxicité dermatologique
- F. Neurotoxicité

G. Réactions allergiques

H. Extravasation

2. Effets indésirables chroniques

Inconstamment et incomplètement réversible : irréversible

A. Myélotoxicité

B. Toxicité cardiaque

C. Toxicité hépatique

D. Neurotoxicité

E. Toxicité pulmonaire

Références bibliographiques

- [1]. S .Felloussi, Mémoire de master, Etude quantitative des relations structure/activité des propriétés cytotoxiques de dérivés de wogonine, Université Biskra, (2018).
- [2]. J. Kouame, Thèse de doctorat , Etudes des propriétés cytotoxiques et anti radicalaires d'extraits de feuilles et de galles de guiera senegalensis j. F. Gmel (combretaceae), Université d'ouagadougou, (2009).
- [3]. J. Seguin, Thèse de doctorat, Caractérisation de modèles de tumeurs murines et leurs applications en thérapie anti-angiogénique vectorisée, Université de pierre et marie de curie, (2012).
- [4]. OMS, Le cancer dans le monde. IARC : Press Lyon, (2005).
- [5]. P. Mehlen, A. Puisieux, Metastasis: a question of life or death. Nat Rev Cancer.6(6), 449-458 (2006).
- [6]. M. Benamar, Thèse de doctorat : Etude de l'activité antimittotique et anticancéreuse des alcaloides naturels ou synthétiques d'arisarum vulgare targ et de pancratium foetidum pom sur deux lignées cellulaires cancéreuses p815 et HEP, Université Mohammed Premier Oujda,(2009).
- [7]. L.Israel, Cancérologie. Medsi/McGraw.Hill, Paris, 1-137, (1983).
- [8]. V.Levy, M. Levy-Sussan, Cancérologie, Editions Estem et Med-Line, Paris, 1- 275, (1996).
- [9]. R.Abele, Les nouveaux médicaments antitumoraux. Le journal international de médecine. 7 (38), 131-136 (1986).
- [10]. M.Hubert-Habart, B. Kokel, G.Menechi, Réflexions sur la chimiothérapie des cancers. L'actualité chimique p.105-113,(1987).
- [11]. A.Lespagnol, Précis de pharmacie chimique usuelle fascicule 1. Edition : Technique et documentation PARIS, (1977).
- [12]. T.Yatabe, Un genre supprimeur de cancer. La recherche 20 (208) p.404-408, (1989).
- [13]. I.Pastan, MM.Gottesman .Multidrug resistance in humain cancer, N EnglJ Med .316 ,1388-93(1987).
- [14]. G.Bradley ,P.F. Juranka, V. Ling .Mechanis mog drug resistance Biochim Biophys Acta . 948 ,87-128 (1988) .
- [15]. J.P. Marie Le phénomène de résistance multiple aux anti-cancéreux : les gènes MOR et la P-gp, médecine/sciences. 6, 443-8 (1990).

- [16]. K.L.Buker ,Thèse de doctorat , Rôle de gène MDR1 dans la prédisposition aux maladies inflammatoires chroniques intestinales chez les colleys , École nationale vétérinaire d'Alfort ,(2008).
- [17]. C. Gauthier, Diplôme de doctorat , Indentification et mécanisme d'action de modulateurs sélectifs du transporteur ABCG2 responsable de la chimiorésistance de cellules cancéreuses, Université Claude Bernard Lyon 1, (2014).
- [18]. C.Fahrmayr ,MF.Fromm ,J.König .Hepatic OATP and OCT up take transporters: theirrole for drug-drug interactions and pharmacogenetic aspects. Drug Metab Rev. 42,380-401 (2010).
- [19]. Ch. Le Guellec, I. Benz, Rôle des transporteurs dans l'élimination des médicaments et leurs effets indésirables,La Lettre du Pharmacologue. 25 (4) ,132 (2011).
- [20]. H. J.LIN. Drug-drug interaction mediated by inhibition and induction of P glycoprotein. Adv. Drug. Deliv. Rev.55, 53-81 (2003).
- [21]. A.Grenoble, Thèse de doctorat, Glycoproteine –p et interactions medicamenteuses en pratique courante a l'officine,Université Joseph Fourier, (2004).
- [22]. G.Spengler, A.Csonka, J.Molnar, L.Amaral, The Anticancer Activity of the Old Neuroleptic Phenothiazine-type Drug Thioridazine, Anticancer Research. 36(11), 5701-5706 (2016).
- [23]. Z.Almi, Thèse de doctorat , Etude qualitative et quantitative des relations structures activités dans des hétérocycles à intérêt pharmaceutique, université biskra, (2015).
- [24]. Z.Almi , Y. Soualah , Mémoire de master , Etude de la relation structure activité des phénothiazines a chaine linéaire ,université biskra ,(2010).
- [25]. A. Jaszczyszyn, K. Gasiorowski, P. Swiatek, W. Malinka, K. Cieslik- Boczula, J. Petrus, and B. Czarnik-Matusiewicz, Pharmacological Reports .64 ,16 -23(2012).
- [26]. K.Attila, A. Attila, P.Satorhelyi, K.Neémeth, I. Kurucz, Biochemical and Biophysical Research communication .346, 45-50(2006).
- [27]. E. Apretna, F. Thiessard, G. Miremont-Salamé, F. Haramburu Centre de pharmacovigilance, EA 3676 MPSS, Université Victor Segalen, Bordeaux, France, Oncologie .6, 66-71 (2004).
- [28]. A.Chlo, E. Grouard, Pharmacologie des anticancéreuses, UE10 - Tissu sanguin Guérin-Dubourg, (2017).

Chapitre II

*Etude quantitative de la relation
structure /activité QSAR.*

II.1. Les bases de la chimie quantiques

Les fondements de la chimie quantique sont apparus et se sont beaucoup développés dans les années 1920 grâce à des scientifiques comme Bohr, Schrödinger, Born, Oppenheimer, Hartree ou encore Slater, en 1930, Hartree et Fock développèrent la méthode du champ autocohérent qui permit d'effectuer les premiers calculs ab initio sur des systèmes diatomiques, il fallut attendre les années 1950 pour que l'ordinateur fasse son apparition. En 1964, Hohenberg et Kohn ont défini un théorème, la théorie de la fonctionnelle de la densité (DFT), en 1970, Pople créa Gaussian qui est aujourd'hui le logiciel de modélisation de très loin le plus utilisé. Les méthodes semi-empiriques et l'utilisation des bases de fonction gaussiennes se sont particulièrement développées dans les années 1970 et 1980, avec la forte augmentation de la puissance de calcul, la modélisation s'est invitée dans nos ordinateurs à partir des années 1990. En 1993 apparaît la méthode B3LYP, méthode hybride qui permet d'effectuer des calculs DFT. en 1998, le prix Nobel de chimie a été décerné à John. A. Pople et Walter Kohn pour leurs travaux dans le domaine de la chimie informatique et la modélisation moléculaire (chimie quantique) [1.2].

La chimie quantique consiste en l'utilisation des méthodes basées sur la résolution de l'équation de Schrödinger indépendante de temps (état stationnaire) .En résolvant l'équation aux valeurs propres et vecteurs propres $\hat{H}\Psi=E\Psi$, ou H est l'Hamiltonien non relativiste, E l'énergie totale et Ψ la fonction d'onde du système , il sera alors possible de déterminer toutes les informations du système étudié .Il n'est cependant pas possible de résoudre rigoureusement une telle équation dans la théorie quantique proposée dès les années 20 afin de pouvoir résoudre cette équation de façon approchée [3.4].

Bien que l'étude des relations structure-activité ait commencé à la fin du 19^{ème} siècle, ce n'est qu'au début des années 60 que les travaux de Corwin Hansch ont proposé un modèle mathématique pour corréler 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 [5].

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) [6].

II.2. Les méthodes de calcul

II.2.1. Les méthodes empiriques

Parmi les méthodes empiriques " non quantique", la mécanique moléculaire et la dynamique moléculaire classique sont les plus utilisées. La première est une méthode d'analyse conformationnelle basée sur l'utilisation de champs de forces empiriques et la minimisation d'énergie. Elle permet de décrire l'énergie en fonction des coordonnées atomiques et de chercher des minima de l'énergie correspondant à des conformères stables [7,8].

Toutefois, elle ne permet pas de décrire l'énergie en fonction du temps et de la température (entropie, énergie libre) et de franchir des barrières énergétiques. La dynamique moléculaire classique a pour but de calculer les mouvements des molécules, le plus souvent à partir des énergies de la mécanique moléculaire, en appliquant les lois de la mécanique classique. Elle permet de simuler l'évolution des systèmes dans le temps[9].

II.2.2. Les méthodes semi-empiriques

Les méthodes semi-empiriques sont exclusivement basées sur l'expérience. Elles sont dérivées des méthodes ab-initio dont elles reprennent le principe en ne considérant que les électrons de la couche de valence de chaque atome et en négligeant certaines intégrales rendant compte du recouvrement des orbitales entre elles. Ces approximations sont alors compensées par la mise en jeu des grandeurs physiques expérimentales, les méthodes semi empiriques sont généralement utilisées pour traiter de grosses molécules, afin d'obtenir une première description de ses propriétés moléculaires (structure, réactivité). On distingue plusieurs variantes : **CNDO**, **INDO**, **MINDO/3**, **MNDO**, **AM 1**, **PM 3**, **SAM 1** [10].

PM 3 : (Parametric Method 3) Proposée par Stewart en 1989, présente beaucoup de points en commun avec AM1, D'ailleurs il existe toujours un débat concernant les mérites relatifs de paramétrisation de chacune d'elles [11].

II.2.3. Les méthodes quantiques

Les méthodes quantiques pures basées sur la mécanique quantique, la mécanique quantique est une technique mathématique rigoureuse basée sur l'équation de Schrödinger, la solution de cette équation permet d'obtenir des informations précises sur les propriétés géométriques et électroniques de la molécule [12].

Ces méthodes sont des techniques de résolution de l'équation de Schrödinger de systèmes à plusieurs électrons elles utilisent des données ajustées sur des résultats expérimentaux afin de simplifier les calculs, les méthodes semi-empiriques contiennent une base minimale par défaut (STO-3G) par contre les méthodes quantiques pures utilisent des bases et des corrélations différentes ça dépend le type de calcul. Le logiciel HyperChem fournit ces méthodes semi-empirique (CNDO, INDO, MINDO3, PM3, RM1..) et quantique (Ab-initio et DFT) [13].

II.2.3.1. Théorie de la fonctionnelle de la densité (DFT)

Les bases de la théorie de la fonctionnelle de la densité (DFT) ont été élaborées en 1927 par Thomas et Fermi qui calculèrent l'énergie d'un atome en représentant son énergie cinétique en fonction de la densité électronique [14]. En 1928, Dirac introduit le terme d'échange prédit par Hartree mais il n'y a toujours aucune prise en compte de la corrélation électronique qui fût finalement ajoutée par Wigner. Dans ce modèle, les n électrons dépendants de $3n$ coordonnées d'espace sont remplacés par leur densité $\rho(\mathbf{r})$ qui ne dépend plus que de 3 variables.

L'état fondamental de ce système est décrit par la fonction d'onde $\Psi_0(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_n)$ qui correspond à une unique densité électronique $\rho(\mathbf{r})$. Cette fonction d'onde, et l'énergie E_0 qui lui est associée, sont déterminées par la minimisation de l'énergie totale du système. Le potentiel externe $v_{\text{ext}}(\mathbf{r})$ créé par les N noyaux du système est alors complètement déterminé et fixe donc l'hamiltonien. Ainsi, le nombre d'électron n et le potentiel $v_{\text{ext}}(\mathbf{r})$, définissent toutes les propriétés de l'état fondamental [15].

II.2.3.2. Ab-initio

Les méthodes Ab initio sont caractérisées par l'introduction d'une base arbitraire pour étendre les orbitales moléculaires et alors le calcul explicite toutes les intégrales exigées qui impliquent cette base [16].

Les méthodes *ab initio* se divisent en deux sous familles : les méthodes Hartree - Fock (HF, RHF, UHF, ROHF) (Hartree, 1928 ; Fock, 1930), et les méthodes post Hartree-Fock, (MPn, CAS, ...) (Moller, 1934). La principale différence entre ces deux méthodes est que les interactions électroniques sont négligées dans les méthodes HF et réintroduites dans les méthodes post HF. Ces méthodes ne peuvent être appliquées qu'à des systèmes de quelques dizaines d'atomes pour les méthodes HF et d'une dizaine d'atomes seulement pour les méthodes post HF [17].

II.2.4. QSAR (relation quantitative structure-activité)

Le modèle QSAR (*Quantitative Structure-Activity Relationships*) est une corrélation basée sur un calcul statistique, de paramètres physico-chimiques avec l'activité biologique [18].

La relation Quantitative Structure et Activité (QSAR) est un outil qui permet d'expliquer l'interaction des composés chimiques avec les sujets actuels, le principe des méthodes QSAR est comme leur nom l'indique, de mettre en place une relation mathématique reliant de manière quantitative des propriétés moléculaires, appelées descripteurs, avec une observable macroscopique (activité biologique, toxicité physico-chimique, etc.), pour une série de molécules biologiquement actives à l'aide des méthodes d'analyse de données "une équation de corrélation" [19].

II.3. Les paramètres utilisés dans l'étude QSAR

Les 3 types de paramètres sont les paramètres électroniques et Les paramètres d'hydrophobicité et Les paramètres stériques :

II.3.1. Les paramètres électroniques

Les modèles QSAR comportent des descripteurs de chimie quantique comprenant des renseignements sur la nature des forces inter-moléculaires impliquées dans la réponse biologique.

Les descripteurs électroniques les plus utilisés, tels que les charges atomiques nettes, les énergies des orbitales moléculaires les plus occupées / les orbitales moléculaires inoccupées (HOMO-LUMO), les densités électroniques des orbitales frontières...etc, montre une bonne corrélation avec diverses activités biologiques [20].

II.3.2. Les paramètres d'hydrophobicité

Les interactions hydrophobes sont très importantes dans les domaines de la chimie. Il s'agit notamment des interactions de ligand-enzyme, l'assemblage de lipides dans les membranes biologiques, l'agrégation des tensioactifs et coagulation. L'hydrophobicité de solutés peut être facilement déterminée en mesurant les coefficients de partage P.

Le coefficient de partage P est défini comme le rapport de la concentration du soluté dans l'octanol à sa concentration dans l'eau [20].

II.3.3. Les paramètres stériques

Les paramètres stériques ont une importance dans les interactions ligand-récepteur ainsi que dans les phénomènes de transport dans les systèmes cellulaires.

L'un des paramètres stériques le plus utilisé est la réfraction molaire (MR). Bien qu'elle soit généralement considérée comme une mesure de masse brute globale, elle intégrera un élément de polarisabilité qui peut décrire la cohésion [20].

✓ Les différentes propriétés fournies par l'Hyperchem sont :

➤ **Surface moléculaire(SAG)**

C'est une enveloppe entourant les atomes localisés à la périphérie, qui explique la surface de contact Ligand-Récepteur appelé l'affinité de liaison L-R [21.22].

La surface grille (SAG) est calculée par l'HyperChem, c'est une méthode de grille où une méthode plus rapide plus approximative; quel que soit la zone accessible au solvant ou la surface de Van der Waals [23]. Dans cette théorie, chaque atome de la molécule est représenté par une sphère. La surface extérieure de toutes les sphères atomiques définit la surface de Van der Waals [20].

➤ **Volume moléculaire (V)**

C'est un volume occupé par un corps, c'est un volume spécifique (à température et pression normales). Son calcul est très similaire à celui de la surface.

Le volume est défini par la relation suivante [24] :

$$V = \frac{W}{d} \quad (\text{II.1})$$

Où W: est la masse moléculaire et d: est la densité.

➤ **Energie d'hydratation (HE)**

L'hydratation est la formation d'une solution implique l'interaction du soluté avec des molécules de solvant, différents liquides peuvent être utilisés comme solvants, mais l'eau est le solvant le plus couramment utilisé.

La liaison hydrogène joue un rôle primordial dans la solubilité des molécules médicamenteuses et leurs interactions avec les récepteurs biologiques [25].

Dans les milieux, les molécules polaires s'entourent pas des molécules d'eau, ce qui fait apparaître des liaisons hydrogènes entre eux; évidemment, les sites donneurs de proton interagissent avec l'atome d'oxygène de l'eau et les sites accepteurs de proton avec l'atome d'hydrogène[20].

L'oxygène (O) est appelé l'accepteur (accepteur de proton H+) et l'azote (N) est appelé le donneur (donneur de proton H+) présenté dans la figure(II.1).

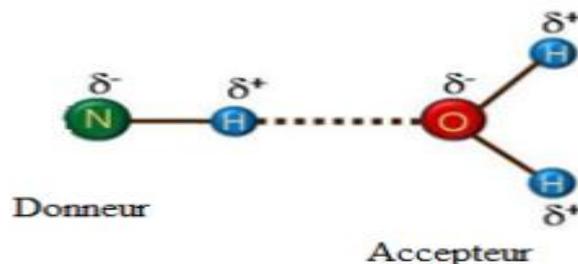


Figure II.1 : La liaison hydrogène[20].

L'énergie d'hydratation est un facteur déterminant de la stabilité des différentes conformations moléculaires dans les solutions aqueuses [26]. Le calcul de l'énergie d'hydratation est basé sur la surface exposée qui dépend du type d'atomes des groupements moléculaires qui peuvent être donneurs des liaisons hydrogènes tels que : O-H, N-H, P-H... ou bien des groupements accepteurs qui portent des doublets libres tels que :O, N, S,P...[20]

➤ **Coefficient de partition (Log P)**

La lipophilie est une propriété qui a un effet important sur la solubilité, l'absorption, la distribution, la métabolisation et l'excrétion des médicaments. Hansch et Leo ont estimé quelles molécules qui portent une grande lipophilie seront conservées à l'intérieur des lipides membranaires [26].

La meilleure méthode pour estimer la capacité d'un composé à se dissoudre dans les milieux aqueux du cytoplasme ainsi que dans le milieu apolaire de la membrane cellulaire est la mesure de la lipophile (figure II. 2).

Le coefficient de partage P se calcule de la manière suivante:

$$\text{Log P} = \frac{[\text{concentration du médicament}]_{\text{Octanol}}}{[\text{concentration du médicament}]_{\text{Eau}}} \quad (\text{II.2})$$

Une valeur positive pour log P, ($P > 1$), indique que le composé est lipophile.

Une valeur négative pour log P, ($P < 1$) indique que le composé est hydrophile.

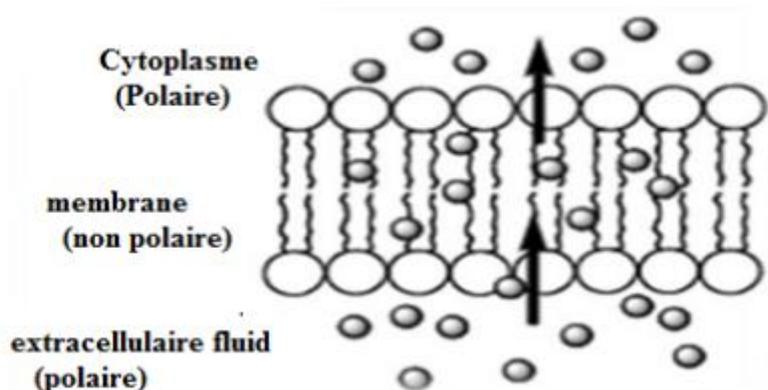


Figure II. 2: Polarité des différents milieux cellulaires [20].

➤ **Réfractivité molaire (MR)**

C'est une réfractivité ramenée à une quantité de matière. La réfractivité est un cas particulier de volume moléculaire, c'est une réfractivité ramenée à une quantité de matière, ainsi qu'elle est généralement désignée comme une simple mesure du volume occupé soit par un atome individuel ou par un groupe d'atomes [27].

En chimie, la réfractivité moléculaire est un critère important pour mesurer le facteur stérique, elle est importante dans le cas où le substituant possède des électrons π ou des doublets libres [26].

La réfractivité moléculaire est définie par la relation suivante :

$$MR = V \frac{n^2-1}{n^2+1} \quad (\text{II.3})$$

Où : V : volume moléculaire et n: est l'indice de réfraction.

➤ **Polarisabilité (Pol)**

On appelle polarisabilité d'une molécule, la facilité avec laquelle son nuage électronique se déforme par un champ électrique externe. La polarisabilité d'une molécule peut améliorer sa solubilité aqueuse qui joue un rôle très important dans la modélisation de nombreuses propriétés moléculaires et des activités biologiques [28].

Une molécule très polarisée peut donner des attractions fortes avec d'autres molécules. La polarisabilité d'une molécule peut également améliorer sa solubilité aqueuse [29].

La polarisabilité moléculaire est définie par la relation suivante :

$$P(e) = \epsilon_0 \alpha E \quad (\text{II.4})$$

Où : P(e) : Coefficient de polarisabilité. ϵ_0 : Constante diélectrique.

α : Moment dipolaire électrique induit.

II.4. Le model QSAR

II.4.1.Définition

Le modèle QSAR est une relation mathématique (souvent une corrélation statistique), qui s'appuie sur la présence ou de l'absence d'une propriété liée à la structure, ou l'activité d'une autre propriété ou une activité d'intérêt.

La relation mathématique d'un modèle QSAR est :

$$\text{Activité biologique} = f(\text{Propriété physico-chimique ou structurelles}) \quad (\text{II.5})$$

Pour déterminer un modèle QSAR nécessite les trois éléments suivants :

Un ensemble des données qui fournit l'activité biologique (habituellement mesurée expérimentalement) pour un groupe de produits chimiques. Ces groupes de produits chimiques est généralement définie par des critères de sélection, des critères structurels ou des données des propriétés liées à la structure définis pour le même groupe des produits chimiques (par exemple, les variables indépendantes). Un moyen de trait de ces deux ensembles des données .Méthodes pour la structure relative à la gamme d'activité de la régression linéaire simple, à travers des méthodes plus complexes tels que la méthode de moindre carrés à des techniques les plus complexes d'apprentissage machine tels que les réseaux de neurones [30].

II.4.2.Outils et Méthodologie de QSAR

II.4.2.1.Paramètres biologiques

Les données biologiques sont habituellement exprimées sur une échelle logarithmique en raison de la relation linéaire entre la réponse et le logarithme de dose dans la région centrale de la courbe de log dose-réponse. Les logarithmes inverses de l'activité ($\log 1/C$) sont également utilisés pour obtenir des valeurs mathématiques plus élevées lorsque les structures sont biologiquement très efficaces. Des exemples de données biochimiques ou biologiques, utilisés dans l'analyse de **QSAR**, sont décrits dans le tableau(II.1)[31].

Tableau II.1 : Les types des données biologiques utilisées dans l'analyse QSAR.

Source d'activité	Paramètres biologiques
1. Récepteurs isolés	
Constante de vitesse	Log k
Constante de Michaelis-Menten	Log $1/K_m$
Constante d'inhibition	Log $1/K_i$
2. Systèmes cellulaires	
Constante d'inhibition	Log $1/IC_{50}$
Résistance croisée	Log CR
Données biologiques <i>in vitro</i>	Log I/C
Mutation de gene	Log TA_{98}
3. Systèmes <i>in vivo</i>	
Facteur de bioconcentration	Log BCF
Vitesses de la réaction <i>in vivo</i>	Log I (induction)
Vitesses pharmacodynamiques	Log T (clairance totale)

II.4.2.2. Les descripteurs moléculaires

Les descripteurs moléculaires sont généralement classés en trois catégories ; les descripteurs physicochimiques, topologiques et électroniques. Ces descripteurs sont caractéristiques de la structure bidimensionnelle ou tridimensionnelle de la molécule, l'étude QSAR-2D permet de rationaliser une propriété avec des descripteurs calculés à partir de la formule développée avec une équation de régression linéaire, les études QSAR- 3D, décrites pour la première fois par Cramer et ses collaborateurs [32] permettent d'établir une relation entre une propriété et des descripteurs calculés à partir d'une conformation spatiale de la molécule définie.

➤ Les descripteurs physico-chimiques

Les descripteurs physico-chimiques caractérisent généralement la structure - bidimensionnelle de la molécule.

Parmi ces descripteurs, certains reflètent la composition moléculaire du composé, soit le nombre et le type d'atomes et de liaisons présents dans la molécule, son nombre de cycle, son poids moléculaire... : ce sont les descripteurs constitutionnels [33]. Certains représentent la surface accessible au solvant (nommée « Connolly Accessible Surface »), le volume des solvants couverts par cette surface (« Connolly Solvent- Excluded Volume ») [34], le caractère hydrophile ou lipophile de la molécule généralement évalué à partir du coefficient de partage octanol/eau représenté par le logP [35].

Ces descripteurs ne fournissent pas assez d'information sur la structure des molécules pour l'élaboration des modèles prédictifs plus complexes ; il est nécessaire d'ajouter d'autres types de descripteurs.

➤ Les descripteurs électroniques

Ces descripteurs caractérisent la distribution de charge des molécules (polarité des molécules) mais aussi les paramètres de la chimie quantique : le moment dipolaire, les énergies HOMO (Orbitale moléculaire occupée de plus haute énergie) et LUMO (Orbitale moléculaire non-occupée de plus basse énergie) [33].

➤ Les descripteurs topologiques

Ces descripteurs décrivent les connectivités atomiques dans la molécule. Ils sont pour la plupart représentés par des indices comme ceux de Balaban, Wiener... et sont issus de la théorie des graphes. Cette théorie a été développée par Euler en 1736 avec les sept ponts de Königsberg. Un graphe est un ensemble de point, certains reliés par des lignes ; il permet de représenter la topologie des molécules sans se soucier de la géométrie spatiale exacte de cette dernière [36].

★ Les descripteurs utilisés dans ce modèle QSAR sont :

- Surface moléculaire (**SAG**)
- Volume moléculaire (**V**)
- Energie d'hydratation (**HE**)
- Log P
- Réfractivité molaire (**MR**)
- Polarisabilité (**Pol**)

- Moment dipolaire (**DM**)
- La liaison hydrogène donneurs (**HBD**)
- La liaison hydrogène accepteurs (**HBA**)
- Les charges des atomes de noyau de base (**q**)

II.4.3. La méthode de la régression linéaire multiple (MLR)

La régression multi-linéaire (MLR) est une méthode statistique pour étudier la relation entre une variable dépendante et deux variables indépendantes ou plus. Dans cette méthode, une variable dépendante Y est décrite en termes d'une série de variables explicatives $X_1 \dots X_n$, comme indiqué dans l'équation :

$$Y = Y_0 + a_1X_1 + a_2X_2 + \dots + a_nX_n \quad (\text{II.6})$$

On suppose que toutes les variables explicatives sont indépendantes les unes des autres [37].

II.4.3.1. Description de la méthode

La régression multilinéaire est la méthode la plus simple et la plus employée pour le développement de modèles prédictifs [38]. Elle repose sur l'hypothèse qu'il existe une relation linéaire entre une variable dépendante Y (ici, la propriété) et une série de p variables indépendantes X_i (ici, les descripteurs). L'objectif est d'obtenir une équation de la forme suivante :

$$Y = \beta_0 + \beta_1X_1 + \beta_2X_2 + \beta_3X_3 + \dots + \beta_pX_p + \varepsilon \quad (\text{II.7})$$

Où β_i sont les coefficients de la régression.

La détermination de l'équation (II.7) se fait alors à partir d'une base de données de n échantillons pour laquelle a la fois les variables dépendantes et la variable indépendante sont connues. Il s'agit donc de considérer un système de p équations.

$$\begin{aligned} Y_1 &= \beta_0 + \beta_1X_{1,1} + \beta_2X_{2,1} + \dots + \beta_pX_{n,1} + \varepsilon_1 \\ Y_2 &= \beta_0 + \beta_1X_{1,2} + \beta_2X_{2,2} + \dots + \beta_pX_{n,2} + \varepsilon_2 \\ &\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ &\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ &\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ Y_n &= \beta_0 + \beta_1X_{1,p} + \beta_2X_{2,p} + \dots + \beta_pX_{n,p} + \varepsilon_n \end{aligned}$$

Ce système d'équations peut être écrit sous la forme matricielle suivante :

$$\begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{pmatrix} = \begin{pmatrix} 1 & X_{1,1} & X_{2,2} & \dots & X_{1,p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{n,1} & X_{n,2} & \dots & X_{n,p} \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}$$

La matrice $X(n,p)$ comporte dans sa première colonne un vecteur constitué uniquement de 1. Ce vecteur correspond à la constante X_0 . La matrice X est alors de dimension $(n, p+1)$. Le modèle s'écrit alors sous forme

$$Y = X\beta + \varepsilon \tag{II.8}$$

La méthode consiste alors à choisir les coefficients du vecteur β en faisant en sorte de minimiser la somme des carrés 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 est alors :

$$\hat{Y} = \beta_0 + \beta_1 X_{1,i} + \dots + \beta_p X_{p,i} \tag{II.9}$$

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

$$\varepsilon = Y_i - \hat{Y}_i \tag{II.10}$$

Il s'agit alors de trouver les coefficients $\hat{\beta}_i$ afin de minimiser la somme des carrés de ces résidus (RSS) pour l'intégralité de la base de données [39].

II.4.3.2. Test de la signification totale de la régression

a. Coefficient de détermination (R^2)

Coefficient de détermination R^2 est une mesure dans laquelle mesure la droite de régression représente-t-elle les données sur un diagramme de dispersion. R^2 est une mesure de l'ajustement du modèle de régression. La R^2 peut être utilisé pour déterminer la relation linéaire entre les valeurs mesurées et estimées. R^2 va de 0 à 1.

Le coefficient de détermination R^2 est le rapport de la somme expliquée des carrés à la somme totale des carrés [37].

$$R^2 = \frac{ESS}{TSS} = \frac{TSS-RSS}{TSS} = 1 - \frac{RSS}{TSS} \quad (\text{II.11})$$

TSS est la somme des carrés totaux : $TSS = \sum(Y_{obs} - \bar{Y})^2$

ESS est la somme des carrés expliqués : $ESS = \sum(Y_{cal} - \bar{Y})^2$

RSS est la somme des carrés résiduels : $RSS = \sum(Y_{obs} - Y_{cal})^2$

b. Coefficient de corrélation (R)

La quantité R , appelée coefficient de corrélation (R), est un coefficient de corrélation entre les valeurs observées et prédites des variables dépendantes Y , où $0 < R < 1$.

$$R = (ESS/TSS)^{0.5} = (1 - RSS/TSS)^{0.5} \quad (\text{II.12})$$

c. Le Test de Fisher-Snedecor (F)

Le Test de Fisher, reflète la ration de la variance expliquée par le modèle et la variance en raison de l'erreur dans le modèle, les valeurs élevées du test F indiquent la signification de l'équation.

$$F = [ESS/(K)]/[RSS/(n - K - 1)] \quad (\text{II.13})$$

Où n , et $(k-1)$ sont des degrés de liberté associés à ESS et RSS respectivement [37].

d. Ecart type (s)

L'écart type (s) est un autre paramètre habituellement rapporté ; il indique dans quelle mesure la fonction de régression prédite les données observées, ce paramètre est donnée par :

$$S = \sqrt{\frac{RSS}{n-p-1}} \quad (\text{II.14})$$

Où p est le nombre de variables indépendantes.

e. Coefficient de Prédiction (Q^2)

Si R^2 est une mesure de la justesse de la corrélation, Q^2 est une mesure de la justesse de la prédiction. Il mesure la capacité prédictive d'un modèle et donné par [39]:

$$Q^2 = 1 - \frac{PRESS}{SSY} \quad (\text{II.15})$$

II.4.3.3. Le choix du meilleur modèle

Le meilleur modèle QSAR est évalué par la valeur de la probabilité conditionnelle sig (< 0.05) , la valeur de R coefficient de corrélation entre les valeurs observées et les prévisions de la variable dépendante, sa valeur variée entre 0 et 1.

La valeur de F est le rapport de deux carrés des moyennes. Quand la valeur de F est importante l'hypothèse nulle peut être rejetée, en d'autres termes F est un test statistique utilisé

pour tester l'hypothèse nulle que l'introduction d'une variable supplémentaire n'entraîne pas un accroissement significatif du R^2 .

La valeur de **SE** (standard error of the estimate) est une mesure de variation de la valeur d'un test statistique d'un échantillon à l'autre. Il s'agit de l'écart-type de la distribution de l'échantillon pour une statistique.

La valeur **Q** : facteur de qualité d'ajustement, suggère la puissance de prédiction. Quand nous aurons le meilleur modèle avec les valeurs de F la plus grande et SE la plus petite, l'équation de régression est déterminée par $Y = B_0 + \sum B_n X_n$ [20].

II.4.3.4. Validation du modèle LOO

Pour tester la validité de la puissance de prédiction d'un modèle MRL, on utilise la technique de validation croisée LOO (leave one out) qui s'applique pour déterminer le nombre de composantes principales qui produisent des modèles prédictifs de façon optimale, cette technique est basée sur le calcul des paramètres statistiques suivants : PRESS, SSY, Spress, R^2_{cv} , R^2_{adj} et PE.

Ces paramètres statistiques sont calculés à partir des relations suivantes [13] :

$$PRESS = \sum (Y_{obs} - Y_{calc})^2 \quad \text{(II.16)}$$

$$SSY = \sum (Y_{obs} - Y_{mean})^2 \quad \text{(II.17)}$$

$$SPRESS = \sqrt{\frac{PRESS}{n}} \quad \text{(II.18)}$$

$$r^2_{cv} = 1 - \frac{PRESS}{SSY} \quad \text{(II.19)}$$

$$r^2_{adj} = \left([1 - (r^2)] \left(\frac{n-1}{n-p-1} \right) \right) \quad \text{(II.20)}$$

$$PE = 0.6745 (1 - r^2) / \sqrt{n} \quad \text{(II.21)}$$

Avec :

PRESS : somme des carrés des écarts entre les valeurs observées et les valeurs prévues à partir du modèle.

SSY : somme des carrés des distances des valeurs observées pour une variable par rapport à la moyenne de cette variable, la somme des carrés permet de mesurer la variation totale dans une variable.

Spress : la capacité prédictive des modèles est évaluée par l'erreur de racine carrée moyenne.

r^2_{adj} : Estimation de la qualité d'ajustement du modèle à la population.

r^2_{cv} : est la variation de la statistique du R^2 obtenue en ajoutant ou en enlevant une variable indépendante.

PE : L'erreur de prédiction du coefficient de corrélation est utilisée pour déterminer le pouvoir prédictif des modèles proposés.

II.4.4. Les applications de l'étude QSAR

Il existe un grand nombre d'applications de ces modèles tels que [40] :

- L'optimisation de l'activité pharmacologique.
- La conception rationnelle de nombreux autres produits tels que des agents tensioactifs, des parfums, des colorants et des produits chimiques fins.
- L'identification des composés dangereux dans les premiers stades de développement de produits ou la projection des stocks de composés existants.
- La prédiction de la toxicité et les effets secondaires de nouveaux composés.
- La prédiction de la toxicité pour les espèces environnementales.
- La sélection des composés ayant des propriétés pharmacocinétiques optimales, que ce soit la stabilité ou la disponibilité dans les systèmes biologiques.
- La prédiction d'une variété de propriétés physico-chimiques des molécules.
- La prédiction du devenir des molécules qui sont libérées dans l'environnement.
- La prédiction des effets conjugués de molécules, que ce soit dans des mélanges ou des formulations.

II.5. Les règles de lipinsky

La règle de Lipinsky est la plus utilisée pour caractériser les composés « drug-like ». Nous rappelons que cette règle a pour but d'identifier les composés posant des problèmes d'absorption et de perméabilité par voie orale, et qu'elle a été établie à partir d'une liste de composés ayant passé avec succès les tests cliniques de phase II.

Nous avons choisi d'utiliser des règles « drug-like » basées sur des limites physicochimiques. Cela donne un filtre « drug-like » plus facilement interprétable. Cependant, pour éviter les inconvénients des limites fixes, nous avons choisi d'utiliser des limites progressives pour notre score.

Avant toute chose, on ne considèrera que les composés ne contenant pas d'autres atomes que C, O, N, S, P, F, Cl, Br, I, Na, K, Mg, Ca, ou Li comme pouvant être « drug-like ». Nous avons choisi dans un premier temps de définir des règles utilisant des limites fixes, en se basant sur une autre étude du laboratoire qui s'appuie sur des règles publiées pour évaluer les proportions de composés « drug-like » dans les bases commerciales [41].

Les règles de lipinsky sont :

- masse moléculaire ≤ 500 Da
- $\text{Log P} \leq 5$
- accepteurs de liaisons H ≤ 10
- donneurs de liaisons H ≤ 5

Les définitions suivantes seront utilisées :

- accepteurs de liaisons H : atomes d'azote, d'oxygène, de phosphore et de soufre, sauf dans les cas suivants : oxygène et soufre aromatiques, azote aromatique connecté à trois autres atomes, azote de valence 5, soufre de valence 6 ou 7.
- donneurs de liaisons H : hétéroatomes avec au moins un atome d'hydrogène et sans charge négative.

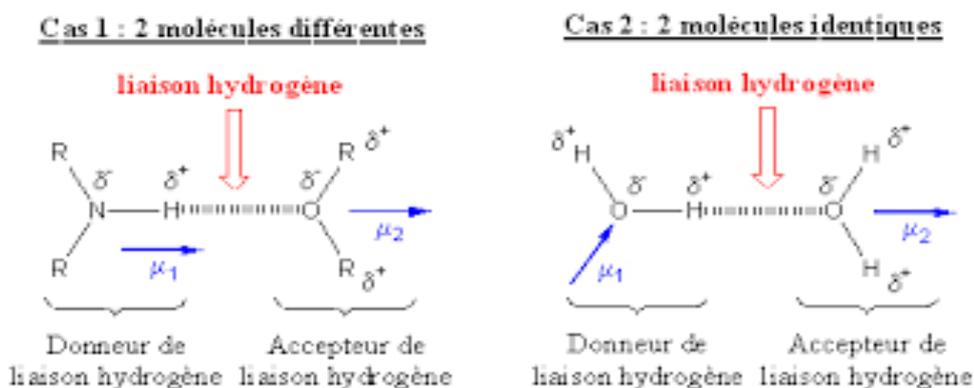


Figure II.3 : Deux exemples sur les liaisons donneurs et accepteurs des hydrogènes.

Références bibliographiques

- [1]. G. Ohanessian, Applications de la Chimie Quantique en Chimie-Cours, École Polytechnique, Orsay, (2005).
- [2]. R. Pollet, Méthodes de la Chimie Quantique, Cours de Master de Chimie–2^{ème}année, Université Paris-Sud 11 et ENS Cachan, (2006).
- [3]. J.P.Doucet , J. Weber , Computer –aided molecular design : Teory and applications , AcademicPress, London ,266(1996).
- [4]. I.N.Levine ,Quantum Chemistry ,5th edition, 455(1999).
- [5]. A. K. Debnath, Quantitative Structure-Activity Relationship (QSAR) Paradigm-HanschEra to New Millenium. Mini Reviews in Medicinal Chemistry. 1(2), 187-195(2001).
- [6]. H.N.G.Phuong, Thèse de doctorat, Synthèse et étude des relations structure - activité quantitatives (QSAR/2D) d'analogues benzo [c] phénathridiniques, Université D'angers, (2007).
- [7]. A. R . Leach, Molecular Modelling. Principles and Applications, Longman, (1996).
- [8]. P.W. Atkins, Chimie Physique, De Boeck Université, (2000).
- [9]. T. Saoud, Mémoire de magister, Calcul computationnel des modes vibrationnels d'un dérivé de l'adénine :Adéniniumhémi sulfate hydrate, Université Constantine ,(2015).
- [10]. R. Bounehas, Mémoire de master, Approche qualitative de la relation structure activité dans des dérivés bioactives de 1, 2,3triazine, Université Biskra, (2014).
- [11]. D.Harkati,Thèse de doctorat , Etude de la structure et des propriétés physico-chimiques associées, de quelques molécules bioactives à intérêt pharmaceutique, (2015).
- [12]. Y.Aouadi ,F.N. Daha , Mémoire de master ,Étude de la relation structure–activité dans des oxazoles antibiotiques ,Université D'el-oued ,(2014).
- [13]. T.Salah, Mémoire de master, Développement des modèles QSAR pour la prédiction des activités inhibitrices antitrypanosomiennes des dérivées cryptolepine, Université Biskra, (2013).
- [14]. P.A.M. Dirac, Proc. Roy. Soc. London A117 , 610-624(1928).
- [15]. J. A. Pople, D. P. Santry, G. A. Segal, J. Chem. Phys .S129,43, (1965).
- [16]. M.Mellaoui,Thèse de doctorat, Analyse des relations structures propriétés pour des dérivés de la céphalosporine et de l'oxazole, Université Biskra, (2014).
- [17]. C. Piron, Mécanique quantique: Bases et applications, PPUR presses polytechniques,(1998).

- [18]. V. N. Viswanadhan, M. R. Reddy, R.J. Bacquet et M.D. Erion, *Journal of computational chemistry*. 14, 1019(1993).
- [19]. G.Fayet, .Thèse de doctorat ,Développement de modèles QSAR pour la prédiction des propriétés d'explosibilité des composés nitro- aromatiques, Université pierre et marie curie, Paris VI, (2010).
- [20]. K.Dermeche, Thèse de doctorat, Etude in silico de la thalidomide : Apport de la modélisation moléculaire, université d'Oran, (2017).
- [21]. Kubinyi, Hugo, G. Folkers, and Yvonne C. Martin, eds. *3D QSAR in Drug Design: Volume 2: Ligand-Protein Interactions and Molecular Similarity*. Vol. 2. Springer Science & Business Media, (1998).
- [22]. A. Thakur ,M.Thakur ,Totale QSAR study on Phenyl Acridine in Reference to DNA Binding Affinity , 3rd Internationale conference on Medical , Biological and Pharmaceutical Science ,(2003).
- [23]. HyperChem .Molecular Modeling System .Hypercube, Inc. ...,1115NW, 4th Street, Garines ville, FL32601 ; USA, (2008).
- [24]. Y. Cohen, *Pharmacologie moléculaire*, Collection de biologie moléculaire Masson, Paris, 187, (1978).
- [25]. E.H.Kerns,L.Di ,Drug-like Properties : Concepts ,Strusture Design and Methods :from ADME to toxicity Optimization ,Academic , Burlington.(2008).
- [26]. S .Belaidi, N.Melkemi and D.Bouzidi,Molecular geometry and structure –Property Relationchips for 1,2-dithiole-3-thione derivatives ,*Int J Chem Res*. 4(2) ,134 - 139(2012).
- [27]. G.L. Patrick, H.D.Winter, W.langenaeker ,J.P.Tollenaere . *Computational Medicinal Chemitry for Drug Discovery* ,Marcel Dekker ,New York ,(1995).
- [28]. J.Wang,X.Xie ,T.Hou .FastApproaches for Molecular Polarizability calculations .*Journal of Physical Chemistry. A* 111 ,4443(2007).
- [29]. Z.Qaneinasah. *Density Functional TheoryStudy of Henry's Law Constant on the Benzimidazole derivatives* .*Journal of Chemical and Pharmaceutical Research*. 3 ,561(2011).
- [30]. O.AFoufou, Mémoire de master, Etude quantitative des relations structures activité dans les dérivés de 4-anilinoquinazolines, agents anticancéreux, Université Biskra, (2014).

- [31]. C. D. Selassie; "History of quantitative Structure-Activity relationships"; in Book "History of quantitative Structure-Activity relationships"; D. J. Abraham; Editor; John Wiley& Sons, Inc.: California ,(2003).
- [32]. R.D. Cramer. D.E .Patterson, and J.D. Bunce .Comparative molecular field analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins. J.Am.Chem. Soc. 110, 5959-5967 (1988).
- [33]. R.Bosque, J.Sales, E.Bosch, M.Rosès, M.C.Garcia-Alvarez-Coque, and J.R.Torres Lapasio. A QSPR study of the p solute polarity parameter to estimate retention in HPLC. J. Chem. Inf. Comput. Sci. 43, 1240-1247(2003).
- [34]. M.L. Connolly, Computation of molecular volume. J. Am. Chem. Soc. 107,118-1124(1985).
- [35]. V.N. Viswanadhan, A.K.Ghose, G.R. Revankar, and R.K. Robins. Atomic physicochemical parameters for three dimensional structure directed quantitative structure-activity relationships. 4. Additional parameters for hydrophobic and dispersive interactions and their application for an automated superposition of certain naturally occurring nucleoside antibiotics. J.Chem. Inf. Comput. 29, 163- 172(1989).
- [36]. H.P. Schultz, Topological organic chemistry. 1. Graph theory and topological indices of alkanes. J. Chem. Inf. Comput. Sci. 29, 227-228(1989).
- [37]. Z. Almi, Thèse de doctorat, Etude qualitative et quantitative des relations structures-activités dans des hétérocycles à intérêt pharmaceutique, Université Mohamed Khider-Biskra, (2015).
- [38]. I. Charif,Thèse de doctorat, Elaboration des corrélations quantitatives structure-activité des acides carbonés. Etude théorique des effets de solvants sur les équilibres libres cétoénoliques des composés b-dicarbonylés cycliques, Université de Tlemcen, (2012).
- [39]. F. Soualmia, Thèse de doctorat, Analyse multivariée des relations structures –activités dans des nouveaux composés organiques à utilisation médicinales, Université d’Oran, (2018).
- [40]. J. Ileszczynski, Challenges and advances in computational chemistry and physics, department of chemistry, jackson state university, U.S.A, (2010).
- [41]. M. Aurélien, Thèse de doctorat, Création et utilisation de chimiothèques optimisées pour la recherche « in silico » de nouveaux composés bioactifs, Université D’Orléans, (2006).

Chapitre III

Résultats et discussions

III.1.Introduction

Les méthodes QSAR représentent un domaine de recherche très important de la chimie-physique. L'objectif est d'établir une relation ou une équation mathématique reliant des descripteurs structurels d'une série de composés étudiés aux activités biologiques. Ces méthodes ont permis de nouvelles avancées vers la compréhension des données biologiques en utilisant des descripteurs simples et cibles [1].

La régression linéaire multiple est parmi les méthodes statistiques la plus utilisée, du fait de la simplicité de ses modèles linéaires et de la facilité de l'interprétation de ses résultats[2].Cependant, cette méthode présente des limitations, car elle dépend des lois et des hypothèses statistiques, notamment dans le cas où les données ne sont pas distribuées normalement [3].

L'objectif de ce travail est d'étudier les propriétés physico-chimiques et électroniques des phénothiazines pipéraziniques qui influent sur l'activité biologique et d'élaborer un modèle QSAR en utilisant les descripteurs qui sont étudiées. En fin, on a utilisé les règles de lipinsky « règle de 5 » pour identifier et sélectionner les composés posant des problèmes d'absorption et de perméabilité.

III.2.Données expérimentales

Une série des phénothiazines pipéraziniques (figure III.1) a été utilisée pour développer un modèle QSAR. Les structures et l'activité biologique expérimentale des phénothiazines pipéraziniques sont rapportées dans le tableau III.1. L'activité d'inhibition de des cellules MDR1 (transporteur P-gp) a été exprimée par PIC_{50} (μM) qui a été rapportée à partir des travaux de :Attila Konya, Attila Andor, Péter Satorhelyi, KlaraNémeth et Istvan Kurucz[4]. Les résultats biologiques sont exploités par la méthode de régression linéaire multiple ($MLR = multiple\ linear\ regression$) dans le but d'établir un modèle QSAR.

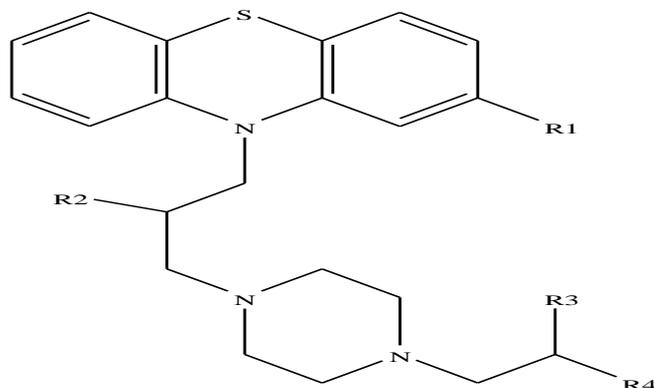
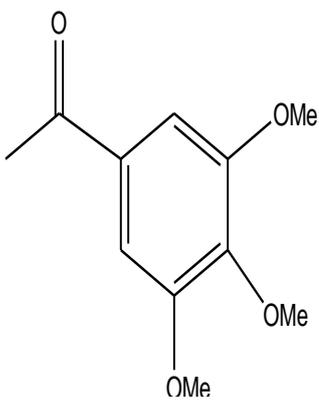


Figure III.1 : Structure générale des dérivés de la phénothiazine «ChemDraw».

Tableau III.1: L'activité biologique de la série des phénothiazines pipéraziniques étudiée.

Les composés	R1	R2	R3	R4	pIC ₅₀ μM(exp)[4]
T5	O	H	H		-0,130
T114	OMe	O	H		-0,328
T116	Cl	O	H		-0,161
T145	Cl	O	Me		0,060
T158	OMe	O	Me		0,148
T172	Cl	Me	H		-0,217
T174	Cl	Me	Me		-0,201

III.3. Etude des paramètres physico-chimiques et électroniques des phénothiazines pipéraziniques

Un descripteur moléculaire peut être considéré comme la conséquence d'un processus logique et mathématique, appliqué à l'information chimique codifiée à travers la représentation d'une molécule [5].

Le choix des descripteurs dépend des outils dont on dispose, de la nature des composés décrits et de la propriété ciblée [6].

L'étude des paramètres physico-chimiques et électroniques des phénothiazines pipéraziniques a été réalisée par la méthode de QSAR, en utilisant le logiciel Hyperchem (8.0.3)/PM3. [7].

➤ Le module "propriétés QSAR" de l'HyperChem (8.0.3) a été utilisé pour calculer les propriétés mises en jeu sont représentées dans le tableau (III.2) :

- * La surface grille (**SAG**)
- * le volume moléculaire (**V**)
- * l'énergie d'hydratation (**HE**)
- * la lipophilie (**Log P**)
- * la réfractivité molaire (**MR**)
- * la polarisabilité (**Pol**)

* la masse moléculaire (MW)

Tableau III.2 : Les valeurs des paramètres physicochimiques des phénothiazines pipéraziniques.

Les composés	SAG (Å ²)	V (Å ³)	HE (Kcal/mol)	LogP	MR (Å ³)	Pol (Å ³)	MW (amu)
T5	790.67	1532.61	-6.53	-3.13	178.04	63.68	598.16
T114	744.19	1469.83	-8.03	-2.98	173.50	61.83	577.70
T116	778.65	1540.54	-9.61	-3.97	179.88	64.31	607.72
T145	786.00	1551.07	-7.24	-2.79	182.64	65.60	626.17
T158	787.57	1574.39	-8.80	-3.56	184.29	66.14	621.75
T172	800.08	1567.60	-6.18	-2.62	182.45	65.51	612.18
T174	810.92	1593.38	-5.56	-2.21	186.87	67.35	626.21

✓ D'après les résultats de ce tableau on remarque que :

La polarisabilité la réfractivité molaire augmentent relativement avec la taille et le poids moléculaire des phénothiazines pipéraziniques étudiées (Tableau III.2). Ces résultats sont en accord avec le critère de Lorentz-Lorenz qui donne une relation entre polarisabilité, la réfractivité molaire et le volume [8], cette relation montre que la réfractivité molaire et la polarisabilité sont proportionnelles avec le volume et la masse moléculaire. Par exemple, les composés T158 et T174 sont des grandes molécules « qui ont des valeurs élevées de volume V et de masse moléculaire MW », portent des groupements l'un de groupement méthoxy en position R1 et un groupement carbonyle en position R2 et un groupement méthyl en position R3, mais le composé T174 contient un atome de chlorure en position R1 et un groupement méthyl en position R2 et un atome d'hydrogène en position R3 ont des valeurs importantes de la polarisabilité (66.14 Å³), (67.35 Å³) et de la réfractivité molaire (184.29 Å³), (186.87 Å³) respectivement. Contrairement, le composé T114 est une petite molécule dans la série

étudiée des phénothiazines pipéraziniques, elle a une petite valeur de polarisabilité (61.83\AA^3) et de réfractivité molaire (173.50\AA^3).

On remarque aussi que la surface et le volume sont augmentées respectivement comme l'exemple de la molécule T5 et T174, on trouve les valeurs de la surface (790.67\AA^2 , 810.92\AA^2) et les valeurs de volume (1532.61\AA^3 , 1593.38\AA^3).

La présence des groupes hydrophobes dans les structures des phénothiazines pipéraziniques provoquent la diminution de l'énergie d'hydratation ; ainsi que, la présence des groupes hydrophiles augmentent l'énergie d'hydratation (tableau III.2). Le composé T116 possède dix sites accepteurs de proton (3N, 7O). Par contre Le composé T174 possède huit sites accepteurs (3N, 5O).

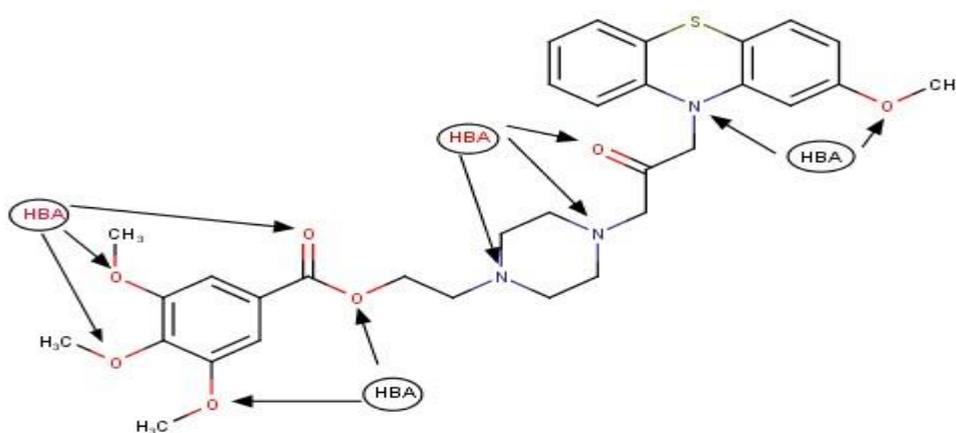


Figure III.2 : Les sites accepteurs dans le composé T116.

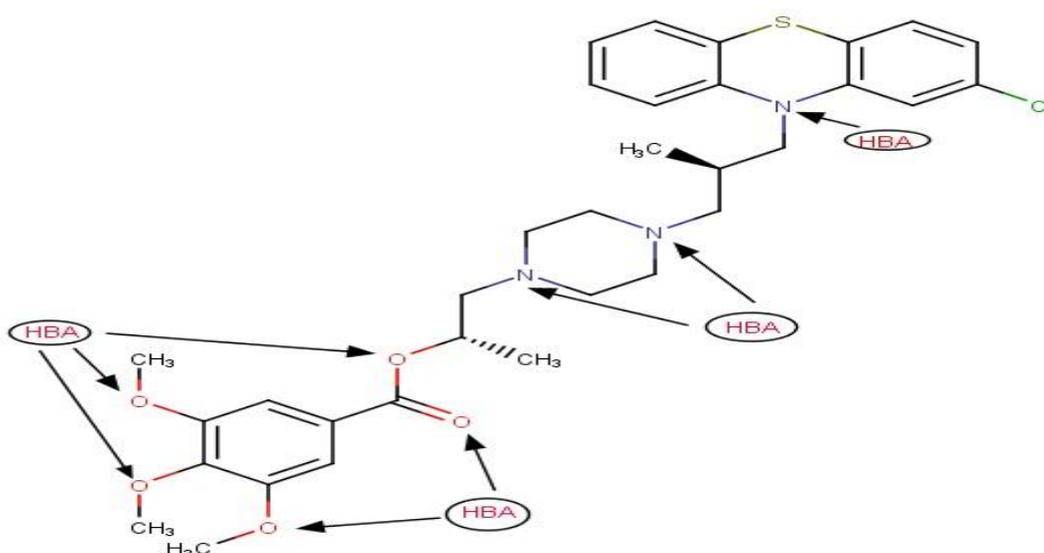


Figure III.3 : Les sites accepteurs dans le composé T174.

L'énergie d'hydratation en valeur absolue, la plus importante (9.61 kcal/mol) a été observée pour le composé T116. La plus faible valeur d'énergie d'hydratation (5.56kcal/mole) a été observée pour le composé T174 (tableau III.2).

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 sont généralement réversibles, on les observe avec les différents récepteurs biologiques [9].

La lipophilie 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 démontré que les molécules très lipophiles se répartissent et se conservent 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 ($0 < \log P < 3$). En effet, Pour un log P trop élevé, le médicament présente une faible solubilité aqueuse. Dans le cas d'un log P très faible, le médicament présente une difficulté de se pénétrer dans les bicouches lipidiques des membranes cellulaires [8].

Les résultats obtenus par le calcul de Log P des phénothiazines pipéraziniques montrent que tous les composés de cette série présentent des faibles coefficients de partition ($\log P < 0$) (-3.97), (3.56), (-3.13), (-2.98), (-2.79), (-2.62) et (-2.21), 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 [11].

➤ Les molécules optimisées ont été utilisées pour calculer un certain nombre des descripteurs électroniques tel que : le moment dipolaire (DM) et les charges atomiques nettes (qC1, qC2, qC3, qC4, qC6, qC7, qC8, qC9, qS5, qN10) dans le tableau(III.3).

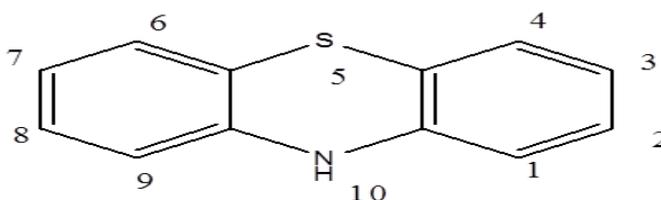


Figure III.4:Schéma de la phénothiazine «ChemDraw».

Tableau III.3 : Les valeurs des descripteurs électroniques utilisés dans l'étude QSAR.

Les composés	DM(μ)	qC1	qC2	qC3	qC4	qC6
T5	4.073	-0,141	-0,096	-0,125	-0,038	-0,055
T114	4.911	-0,144	-0,069	-0,122	-0,054	-0,066
T116	4.073	-0,227	0,119	-0,157	-0,022	-0,073
T145	4.201	-0,161	-0,093	-0,126	-0,036	-0,062
T158	4.873	-0,226	0,118	-0,158	-0,022	-0,071
T172	4.693	-0,143	-0,096	-0,126	-0,039	-0,055
T174	5.529	-0,121	-0,104	-0,120	-0,043	-0,059

La suite de tableau (III.3)

Les composés	qC7	qC8	qC9	qS5	qN10
T5	-0,117	-0,072	-0,120	0,190	0,095
T114	-0,123	-0,065	-0,143	0,192	0,123
T116	-0,111	-0,076	-0,123	0,190	0,108
T145	-0,116	-0,057	-0,141	0,193	0,139
T158	-0,111	-0,071	-0,120	0,190	0,106
T172	-0,120	-0,075	-0,120	0,190	0,097
T174	-0,114	-0,080	-0,103	0,188	0,080

✓ Interprétation des résultats :

On remarque pour le composé T174 présente une valeur maximale du moment dipolaire 5.529(μ), il provient de la présence d'un effet de résonance résultant d'effet mésomère.

Dans le composé T145, le soufre et l'azote présentent les charges positives maximales respectivement (0,193), (0.139), donc ces derniers favorisent une attaque nucléophile.

III.4. Etude quantitative de la relation structure-activité QSAR des phénothiazines pipéraziniques

III.4.1. Développement de modèle QSAR

Les différentes phénothiazines pipéraziniques (tableau III.1) ont été évalués par leur activité inhibition de MDR1 .En vue de déterminer le rôle des propriétés physico-chimiques et électroniques sur l'activité biologiques, nous proposons une étude par un model QSAR.

Le développement d'un modèle QSAR nécessite un ensemble diversifié de données et ce qui fait un grand nombre de descripteurs doivent être pris en considération. Les descripteurs sont des valeurs numériques codant pour différentes caractéristiques structurales des molécules et la sélection d'un ensemble de descripteurs appropriés à partir d'un grand nombre d'entre eux nécessite une méthode qui est capable de discriminer entre les paramètres, cette analyse a été effectuée sur les descripteurs à l'aide du logiciel statistique SPSS19 [12].

Une analyse de régression linéaire multiple des descripteurs moléculaires a été réalisée à l'aide de la stratégie par étapes de SPSS version 19 pour Windows [13].

La corrélation entre l'activité biologique et les descripteurs utilisés « HE, MR, Pol, SAG, V, log P, MW, DM, HBD, HBA, et les charges atomiques nettes : qC1, qC2, qC3, qC4, qC6, qC7, qC8, qC9, qS5, qN10 », représente le meilleur modèle de régression multilinéaire, est exprimée par la relation mathématique suivant :

$$PIC_{50} = 1.797 + 36.380 qC7 + 35.727 qC8 + 0.012 V - 0.022MW$$

$$N = 7 \quad R = 0.998 \quad S = 0.016 \quad F = 166.060 \quad Q = 62.375$$

N : indique le nombre de composés.

R : coefficient de corrélation.

S : erreur standard de l'estimation.

F : statistique Fischer.

Q : la qualité d'ajustement ou bien l'adaptation.

Les valeurs de la fraction de la variance peuvent varier entre **0** et **1**. Le modèle **QSAR** doit considérer un **R² > 0.6** pour qu'il soit valide[14]. Par exemple, dans notre modèle QSAR obtenu **R² = 0.997** nous ont permis d'indiquer fermement la corrélation entre les différentes paramètres (variables indépendants) ayant une activité spécifique anticancéreuse.

La valeur **F** est jugée la signification statistique au niveau de 95%, pour la valeur de **F** calculée est supérieure par rapport aux valeur lue dans la table du Fischer ($F_{cal} > F_{tab}$).

La valeur positive de facteur de qualité **Q** pour ce modèle suggère son haut pouvoir prédictif [15].

Dans le modèle QSAR, le coefficient négatif de **MW** montre que toute augmentation dans la valeur de ce paramètre entraine une augmentation de l'activité biologique. Donc les molécules lourds qui possèdent des grandes masses moléculaires **MW** ont une forte activité anticancéreuse.

Dans l'équation, le coefficient positif de **V** montre que toute augmentation dans les valeurs de ce paramètre entraine une diminution de l'activité anti-cancéreuse.

III.4.2 Validation de modèle QSAR

Afin de tester la validité de pouvoir prédictif de modèle QSAR sélectionné, la technique de validation croisée *LOO* (*leave one out*) a été utilisée. Le modèle développé a été validé par le calcul des paramètres statistiques suivants :

La somme des carrés résiduels (**PRESS**), la somme des carrés totaux (**SSY**), les erreurs de prédiction (**SPRESS**), coefficient de corrélation de validation croisée (**R²_{cv}**), qualité d'ajustement (**R²_{adj}**) et l'erreur de prédiction (**PE**) (Tableau III.4).

Tableau III.4: Les paramètres de validation croisée.

Modèle	PRESS	SSY	PRESS/SSY	SPRESS	R ² _{cv}	R ² _{adj}	6PE
P _{IC50}	0.000	0.165	0.000	0.000	0.997	0.991	0.009

- **PRESS** est un paramètre important de validation croisée car il s'agit d'une bonne approximation de l'erreur prédictive réelle du modèle. Sa valeur étant inférieure à TSS indique que ce modèle prédit mieux que le hasard et peut être considéré statistiquement significatif. La petite valeur de **PRESS** signifie la meilleure prévisibilité du modèle [16]. D'après les résultats présentés dans le tableau (III.4), le modèle est statistiquement significatif.
- En outre, pour un modèle raisonnable de QSAR, le rapport **PRESS/SSY** doit être inférieur à 0,4 [14]. La valeur présentée dans le tableau (III.4) indique que le modèle développé a un rapport de 0.000 qui est inférieure à 0.4.

- Les valeurs élevées du R^2_{cv} et R^2_{adj} sont des critères essentiels pour une meilleure qualification des modèles QSAR [17]. Dans notre résultats les valeurs de R^2_{cv} et R^2_{adj} sont : $R^2_{cv} = 0,997$ et $R^2_{adj} = 0,991$.
- **SPRESS** (incertitude de prédiction) est un bon paramètre à utiliser pour décider de l'incertitude dans la prédiction. Plus la valeur de ce paramètre est basse, la capacité prédictive du modèle sera mieux. Le modèle QSAR développé présente une faible valeur de Spress ($< 0,200$) ce qui indique que ce modèle présente des petites valeurs résiduelles entre l'activité biologique observée et prédite [18]. Dans notre cas, ce paramètre porte une petite valeur 0,000 ce qui explique que l'habileté de prédiction est la meilleure pour ce modèle.
- L'erreur prédictive du coefficient de corrélation (**PE**) est un autre paramètre utilisé pour évaluer la puissance prédictive de modèle proposé. Nous avons calculé la valeur de 6PE de ce modèle proposé, qui est présentée dans le tableau (III.4). Pour le modèle la condition $r > 6PE$ est satisfaite, ce qui indique que le modèle possède un bon pouvoir prédictif [19].

III.4.3. Prédiction de l'activité

Les activités anticancéreuses expérimentales, prédites et résiduelles des phénothiazines pipéraziniques sont représentés dans le tableau (III.5).

Tableau III.5 : L'activité anticancéreuse expérimentale, prédite et résiduelle des phénothiazines pipéraziniques.

Les composés	pIC ₅₀ expérimentale	pIC ₅₀ prédite	pIC ₅₀ résiduelle
T5	-0,130	-0,113	-0,016
T114	-0,328	-0,334	0,006
T116	-0,161	-0,161	0,000
T145	0,060	0,064	-0,004
T158	0,148	0,137	0,010
T172	-0,217	-0,225	0,008
T174	-0,201	-0,196	-0,004

La figure (III.5) ci-dessous montre la courbe de régression linéaire des valeurs prédites par rapport aux valeurs expérimentales de l'activité biologique anticancéreuse. Le courbe de modèle représente pour être plus pratique avec $R^2=0.997$. Il indique que le modèle peut être appliqué avec succès pour prédire l'activité inhibitrice de MDR1 (anticancéreuse) dans cette série des phénothiazines pipéraziniques.

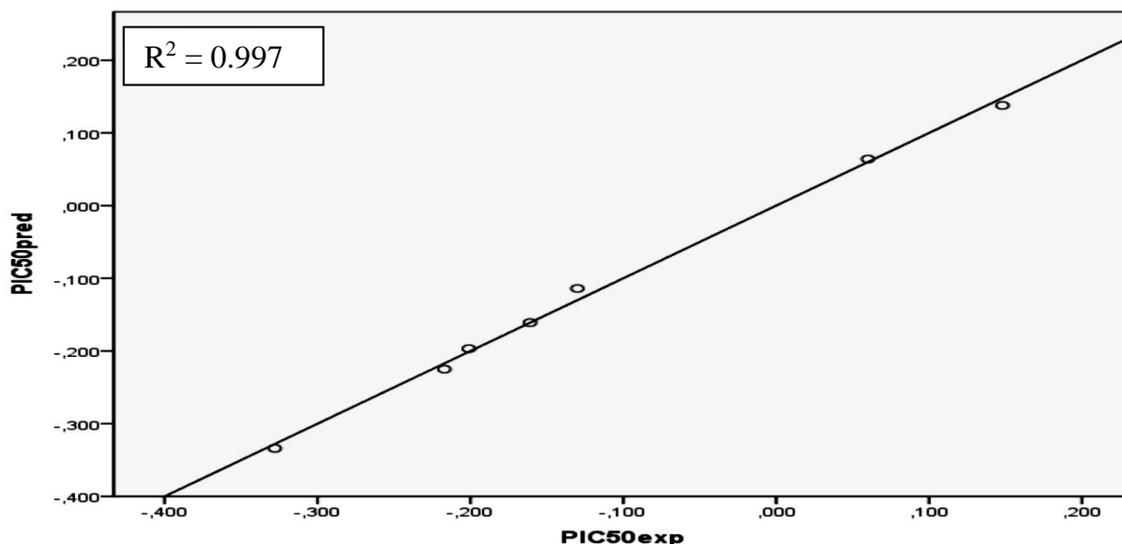


Figure III.5:La courbe de la régression linéaire des valeurs prédites en fonction des valeurs expérimentales pour le modèle QSAR.

Pour rechercher la présence d'une erreur systématique dans l'élaboration de modèle QSAR, les valeurs résiduelles de l'activité biologique ont été tracées en fonction des valeurs expérimentales comme montre la Figure (III.6).

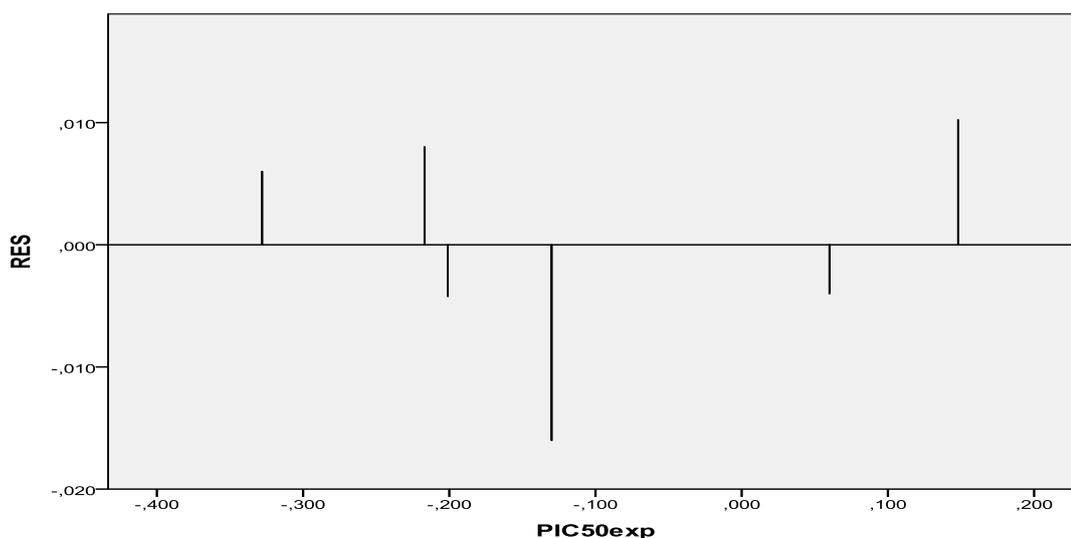


Figure III.6 : Représentation graphique des valeurs résiduelles en fonction des valeurs expérimentales de pIC50.

La propagation des résidus des deux côtés de zéro indique qu'il n'y a aucune erreur systémique. Il indique que ce modèle peut être appliqué avec succès pour cette série des molécules pour prédire l'activité inhibitrice de MDR1 des phénothiazines pipéraziniques.

III.4.4. La matrice de corrélation

La matrice de corrélation entre les descripteurs obtenus par l'analyse de la méthode MLR et l'activité biologique est reportée dans le tableau (III.6).

Tableau III.6 : La matrice de corrélation du modèle.

	pIC50	qC7	qC8	V	MW
pIC50	1				
qC7	0,660	1			
qC8	0,322	-0,317	1		
V	0,473	0,604	-0,490	1	
MW	0,655	0,643	-0,150	0,918	1

Le tableau (III.6) présente que les descripteurs obtenus sont corrélés entre eux et avec l'activité biologique, donc le descripteur **qC7** est le plus important dans la corrélation entre les descripteurs sélectionnés et l'activité anticancéreuse des phénothiazines pipéraziniques.

III.5. Les règles de Lipinsky

Lipinsky a défini un ensemble des règles permettant d'estimer la biodisponibilité d'un composé par voie orale à partir de sa structure bidimensionnelle (2D). Ces règles concernant les propriétés physico-chimiques ont été définies après l'analyse de 2245 médicaments commercialisés ou en phase finale de développement [20] :

- Le poids moléculaire du composé ne doit pas être supérieur à 500 (uma).
- Le logarithme décimal du coefficient de partage eau / 1-octanol, noté Log P, doit être inférieur à 5.
- Le nombre de donneurs de liaisons hydrogène doit être inférieur à 5.
- Le nombre d'accepteurs de liaisons hydrogène doit être inférieur à 10.

Nous avons utilisé les règles de Lipinsky pour identifier les composés qui posent des problèmes d'absorption et de perméabilité si ces composés ne valident pas au moins deux de

ses règles et on a utilisé le molispiration[21] pour calculer la liaison hydrogène donneurs (HBD) et la liaison hydrogène accepteurs (HBA) (tableau III.7).

Tableau III.7 : Les règles de Lipinsky des phénothiazines pipéraziniques.

Les composés	Log P	MW (uma)	HBD	HBA	N° de violation
T5	-3.13	598.16	0	8	1
T114	-2.98	577.70	0	9	1
T116	-3.97	607.72	0	10	1
T145	-2.79	626.17	0	9	1
T158	-3.56	621.75	0	10	1
T172	-2.62	612.18	0	8	1
T174	-2.21	626.21	0	8	1

Nous pouvons voir à travers le tableau (III.7) que tous les composés ont des valeurs moins de 5 pour log P, correspondant de -2.21 à -3.97, donc ces composés sont mieux solubilisés dans des solutions aqueuses. Nous pouvons également constaté dans le tableau (III.7) que tous ces composés ont des nombres des liaisons accepteurs d'hydrogène (HBA) inférieurs 10 (O, S, N) et aussi ils ont pas des liaisons donneurs d'hydrogène (HBD). Pour les poids moléculaires, tous ces composés ont des valeurs supérieures à 500 (uma), donc ils traversent difficilement les membranes cellulaires, la valeur maximale de MW est de 626.21 (uma) pour le composé T174.

Par conséquence, tous les composés étudiés dans ce travail ne répondent pas aux règles de Lipinsky "règles des cinq", suggérant que ces composés théoriquement auraient des problèmes avec la biodisponibilité orale.

Références bibliographiques

- [1]. Y.Yousfi, Mémoire de master, Etude QSAR de l'activité antioxydant d'une série de composés phénoliques, Université Tlemcen, (2017).
- [2]. H. Nocairi, M. Hanafi, E.M. Qannari. *Rev. Statistique Appliquee*, LIII(2), 29-41, (2005).
- [3]. P.Desmet. *Recherche et application en Marketing*, 11(2), 17-28 (1996).
- [4]. K.Attila, A.Attila ,P.Satorhelyi, K.Neémeth, I. Kurucz, *Biochemical and Biophysical Research communication* 346 ,45–50(2006).
- [5]. H. Gonzalez-Diaz, E. Olazabal, L. Santana, E. Uriarte, Y. Gonzalez-Diaz, N.Castanedo, *Bioorg. Med. Chem.*, 15, 962 (2007).
- [6]. S .Felloussi , Mémoire de master , Etude quantitative des relations structure/activité des propriétés cytotoxiques de dérivés de wogonine, Université Biskra , (2018) .
- [7]. HyperChem (Molecular Modeling System) Hypercube, Inc., 1115 NW, 4th Street, Gainesville, FL 32601, USA, (2008).
- [8]. H. Patrick, W. Langenacker, J. Tollenaere. *Computational Medicinal Chemistry for Drug Discovery*, marceldekker, new york (1995).
- [9]. L.B. Kier. *Molecular Orbital Theory in Drug Research*, Academic Press, New York (1981).
- [10]. H. Pajouhesh, G.R. Lenz. *Medicinal chemical properties of successful central nervous system drugs*. *J. Am. Soc. Exp. NeuroThera.*, 2, 541-553 (2005).
- [11]. E.H. Kerns, L. Di. *Drug-like Properties: Concepts, Structure Design and Methods: from ADME to Toxicity Optimization*, Academic Press, USA (2008).
- [12]. H.A.Carlson. *Curr. Op. Chem. Biol. S*, 6, 447-452, (2002).
- [13]. SPSS 19 software packages, SPSS Inc., 444 North Michigan Avenue, Suite 3000, Chicago, Illinois, 60611, USA.
- [14]. K.Ajeet, Bijander, *Int. J. Pharm. & Pharm. Sci.*, 4, 445 (2012).
- [15]. K.Dermeche , Thèse de doctorat , Etude in silico de la thalidomide : Apport de la modélisation moléculaire , université Oran , (2017).
- [16]. S.Wold, L. Eriksson. *Statistical validation of QSAR results. Validation tools .Methods And Principles in Medicinal Chemistry*, 2, 309-18 (1995).
- [17]. S. O.Podunavac-Kuzmanović, D.D. Cvetković, D.J. Barna, *Int. J. Mol. Sci.*, 10, 1670 (2009).

- [18]. S.Medjahid, Thèse de doctorat, Etude par des méthodes QSAR de plusieurs séries molécules hétérocycliques à intérêt pharmaceutique, Université d'Oran, (2018).
- [19]. A.K. Srivastava, N. Shukla. Quantitative structure activityc relationship (QSAR) studies on a seriesof imidazole derivatives as novel ORL1 receptor antagonists. J. SaudiChem. Soc., 17, 321-328 (2013).
- [20]. F. Soualmia, Thèse de doctorat, Analyse multivariée des relations structures –activités dans des nouveaux composés organiques à utilisation médicinales, Université d'Oran, (2018).
- [21]. <https://www.molinspiration.com/>

Conclusion générale

Conclusion générale

Dans ce travail nous avons réalisé une étude quantitative de la relation structure-activité (QSAR) sur une série des phénothiazines pipéraziniques. On a utilisé plusieurs descripteurs dans le développement de modèle QSAR tel que : HE, MR, Pol, SAG, V, log P, MW, DM, HBD, HBA, et les charges atomiques : qC1, qC2, qC3, qC4, qC6, qC7, qC8, qC9, qS5, qN10.

Le modèle QSAR a été établie en utilisant la méthode de régression multilinéaire (MLR), le modèle obtenu peuvent être utilisé pour prédire l'activité anticancéreuse.

Le pouvoir prédictif du modèle QSAR obtenu a été confirmé par la méthode de validation croisée LOO (Leave-one-out). Les résultats obtenus de validation externe confirme la fiabilité de modèle prédictif obtenue.

On a observé une forte corrélation entre les valeurs expérimentales et prédites de l'activité anti-cancéreuse, ce qui indique la validité et la qualité de modèle QSAR obtenu.

En fin, on a étudié les critères de règle de lipinsky "règle de 5" pour cette série des phénothiazines pipéraziniques, on a trouvé que ces critères sont non vérifiées : "tous les composés étudiés dans ce travail ne répondent pas aux règles de Lipinsky "règles des cinq" compte tenu des valeurs considérables de masse moléculaire "MW" de chaque composé de cette série, suggérant que ces composés théoriquement auraient des problèmes avec la biodisponibilité orale".

Table du F (de fichier)

F limite à p 0,05 (comparaison de 2 variances)

degré de liberté du numérateur

degré de liberté du dénominateur

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	20	30	40	50
1	161	199	216	225	230	234	237	239	241	242	243	244	245	245	246	248	250	251	252
2	18,5	19,0	19,2	19,2	19,3	19,3	19,4	19,4	19,4	19,4	19,4	19,4	19,4	19,4	19,4	19,4	19,5	19,5	19,5
3	10,1	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81	8,79	8,76	8,74	8,73	8,71	8,70	8,66	8,62	8,59	8,56
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96	5,94	5,91	5,89	5,87	5,86	5,80	5,75	5,72	5,70
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77	4,74	4,70	4,68	4,66	4,64	4,62	4,56	4,50	4,46	4,44
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06	4,03	4,00	3,98	3,96	3,94	3,87	3,81	3,77	3,75
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,64	3,60	3,57	3,55	3,53	3,51	3,44	3,38	3,34	3,32
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,35	3,31	3,28	3,26	3,24	3,22	3,15	3,08	3,04	3,02
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,14	3,10	3,07	3,05	3,03	3,01	2,94	2,86	2,83	2,80
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,98	2,94	2,91	2,89	2,86	2,85	2,77	2,70	2,66	2,64
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,85	2,82	2,79	2,76	2,74	2,72	2,65	2,57	2,53	2,51
12	4,75	3,89	3,49	3,26	3,11	3,00	2,91	2,85	2,80	2,75	2,72	2,69	2,66	2,64	2,62	2,54	2,47	2,43	2,40
13	4,67	3,81	3,41	3,18	3,03	2,92	2,83	2,77	2,71	2,67	2,63	2,60	2,58	2,55	2,53	2,46	2,38	2,34	2,31
14	4,60	3,74	3,34	3,11	2,96	2,85	2,76	2,70	2,65	2,60	2,57	2,53	2,51	2,48	2,46	2,39	2,31	2,27	2,24
15	4,54	3,68	3,29	3,06	2,90	2,79	2,71	2,64	2,59	2,54	2,51	2,48	2,45	2,42	2,40	2,33	2,25	2,20	2,16
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49	2,46	2,42	2,40	2,37	2,35	2,28	2,19	2,15	2,12
17	4,45	3,59	3,20	2,96	2,81	2,70	2,61	2,55	2,49	2,45	2,41	2,38	2,35	2,33	2,31	2,23	2,15	2,10	2,06
18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41	2,37	2,34	2,31	2,29	2,27	2,19	2,11	2,06	2,04
19	4,38	3,52	3,13	2,90	2,74	2,63	2,54	2,48	2,42	2,38	2,34	2,31	2,28	2,26	2,23	2,16	2,07	2,03	2,00
20	4,35	3,49	3,10	2,87	2,71	2,60	2,51	2,45	2,39	2,35	2,31	2,28	2,25	2,22	2,20	2,12	2,04	1,99	1,97
21	4,32	3,47	3,07	2,84	2,68	2,57	2,49	2,42	2,37	2,32	2,28	2,25	2,22	2,20	2,18	2,10	2,01	1,96	1,94
22	4,30	3,44	3,05	2,82	2,66	2,55	2,46	2,40	2,34	2,30	2,26	2,23	2,20	2,17	2,15	2,07	1,98	1,94	1,91
23	4,28	3,42	3,03	2,80	2,64	2,53	2,44	2,37	2,32	2,27	2,24	2,20	2,18	2,15	2,13	2,05	1,96	1,91	1,88
24	4,26	3,40	3,01	2,78	2,62	2,51	2,42	2,36	2,30	2,25	2,22	2,18	2,15	2,13	2,11	2,03	1,94	1,89	1,86
25	4,24	3,39	2,99	2,76	2,60	2,49	2,40	2,34	2,28	2,24	2,20	2,16	2,14	2,11	2,09	2,01	1,92	1,87	1,84
26	4,23	3,37	2,98	2,74	2,59	2,47	2,39	2,32	2,27	2,22	2,18	2,15	2,12	2,09	2,07	1,99	1,90	1,85	1,82
27	4,21	3,35	2,96	2,73	2,57	2,46	2,37	2,31	2,25	2,20	2,17	2,13	2,10	2,08	2,06	1,97	1,88	1,84	1,81
28	4,20	3,34	2,95	2,71	2,56	2,45	2,36	2,29	2,24	2,19	2,15	2,12	2,09	2,06	2,04	1,96	1,87	1,82	1,79
29	4,18	3,33	2,93	2,70	2,55	2,43	2,35	2,28	2,22	2,18	2,14	2,10	2,08	2,05	2,03	1,94	1,85	1,81	1,77
30	4,17	3,32	2,92	2,69	2,53	2,42	2,33	2,27	2,21	2,16	2,13	2,09	2,06	2,04	2,01	1,93	1,84	1,79	1,76
31	4,16	3,30	2,91	2,68	2,52	2,41	2,32	2,25	2,20	2,15	2,11	2,08	2,05	2,03	2,00	1,92	1,83	1,78	1,75
32	4,15	3,29	2,90	2,67	2,51	2,40	2,31	2,24	2,19	2,14	2,10	2,07	2,04	2,01	1,99	1,91	1,82	1,77	1,74
33	4,14	3,28	2,89	2,66	2,50	2,39	2,30	2,23	2,18	2,13	2,09	2,06	2,03	2,00	1,98	1,90	1,81	1,76	1,72
34	4,13	3,28	2,88	2,65	2,49	2,38	2,29	2,23	2,17	2,12	2,08	2,05	2,02	1,99	1,97	1,89	1,80	1,75	1,71
35	4,12	3,27	2,87	2,64	2,49	2,37	2,29	2,22	2,16	2,11	2,07	2,04	2,01	1,99	1,96	1,88	1,79	1,74	1,70
36	4,11	3,26	2,87	2,63	2,48	2,36	2,28	2,21	2,15	2,11	2,07	2,03	2,00	1,98	1,95	1,87	1,78	1,73	1,69
37	4,11	3,25	2,86	2,63	2,47	2,36	2,27	2,20	2,14	2,10	2,06	2,02	2,00	1,97	1,95	1,86	1,77	1,72	1,68
38	4,10	3,24	2,85	2,62	2,46	2,35	2,26	2,19	2,14	2,09	2,05	2,02	1,99	1,96	1,94	1,85	1,76	1,71	1,67
39	4,09	3,24	2,85	2,61	2,46	2,34	2,26	2,19	2,13	2,08	2,04	2,01	1,98	1,95	1,93	1,85	1,75	1,70	1,67
40	4,08	3,23	2,84	2,61	2,45	2,34	2,25	2,18	2,12	2,08	2,04	2,00	1,97	1,95	1,92	1,84	1,74	1,69	1,66
45	4,06	3,20	2,81	2,58	2,42	2,31	2,22	2,15	2,10	2,05	2,01	1,97	1,94	1,92	1,89	1,81	1,71	1,66	1,63
50	4,03	3,18	2,79	2,56	2,40	2,29	2,20	2,13	2,07	2,03	1,99	1,95	1,92	1,89	1,87	1,78	1,69	1,63	1,60
55	4,02	3,16	2,77	2,54	2,38	2,27	2,18	2,11	2,06	2,01	1,97	1,93	1,90	1,88	1,85	1,76	1,67	1,61	1,58
60	4,00	3,15	2,76	2,53	2,37	2,25	2,17	2,10	2,04	1,99	1,95	1,92	1,89	1,86	1,84	1,75	1,65	1,59	1,56
65	3,99	3,14	2,75	2,51	2,36	2,24	2,15	2,08	2,03	1,98	1,94	1,90	1,87	1,85	1,82	1,73	1,63	1,58	1,54
70	3,98	3,13	2,74	2,50	2,35	2,23	2,14	2,07	2,02	1,97	1,93	1,89	1,86	1,84	1,81	1,72	1,62	1,57	1,53
75	3,97	3,12	2,73	2,49	2,34	2,22	2,13	2,06	2,01	1,96	1,92	1,88	1,85	1,83	1,80	1,71	1,61	1,55	1,52
80	3,96	3,11	2,72	2,49	2,33	2,21	2,13	2,06	2,00	1,95	1,91	1,88	1,84	1,82	1,79	1,70	1,60	1,54	1,51
85	3,95	3,10	2,71	2,48	2,32	2,21	2,12	2,05	1,99	1,94	1,90	1,87	1,84	1,81	1,79	1,70	1,59	1,54	1,50
90	3,95	3,10	2,71	2,47	2,32	2,20	2,11	2,04	1,99	1,94	1,90	1,86	1,83	1,80	1,78	1,69	1,59	1,53	1,49
95	3,94	3,09	2,70	2,47	2,31	2,20	2,11	2,04	1,98	1,93	1,89	1,86	1,82	1,80	1,77	1,68	1,58	1,52	1,48
100	3,94	3,09	2,70	2,46	2,31	2,19	2,10	2,03	1,97	1,93	1,89	1,85	1,82	1,79	1,77	1,68	1,57	1,52	1,48
110	3,93	3,08	2,69	2,45	2,30	2,18	2,09	2,02	1,97	1,92	1,88	1,84	1,81	1,78	1,76	1,67	1,56	1,50	1,47
120	3,92	3,07	2,68	2,45	2,29	2,18	2,09	2,02	1,96	1,91	1,87	1,83	1,80	1,78	1,75	1,66	1,55	1,50	1,46
130	3,91	3,07	2,67	2,44	2,28	2,17	2,08	2,01	1,95	1,90	1,86	1,83	1,80	1,77	1,74	1,65	1,55	1,49	1,45
140	3,91	3,06	2,67	2,44	2,28	2,16	2,08	2,01	1,95	1,90	1,86	1,82	1,79	1,76	1,74	1,65	1,54	1,48	1,44
150	3,90	3,06	2,66	2,43	2,27	2,16	2,07	2,00	1,94	1,89	1,85	1,82	1,79	1,76	1,73	1,64	1,54	1,48	1,44
200	3,89	3,04	2,65	2,42	2,26	2,14	2,06	1,98	1,93	1,88	1,84	1,80	1,77	1,74	1,72	1,62	1,52	1,46	1,41
250	3,88	3,03	2,64	2,41	2,25	2,13	2,05	1,98	1,92	1,87	1,83	1,79	1,76	1,73	1,71	1,61	1,50	1,44	1,40
300	3,87	3,03	2,63	2,40	2,24	2,13	2,04	1,97	1,91	1,86	1,82	1,78	1,75	1,72	1,70	1,61	1,50	1,43	1,39
400	3,86	3,02	2,63	2,39	2,24	2,12	2,03	1,96	1,90	1,85	1,81	1,78	1,74	1,72	1,69	1,60	1,49	1,42	1,38
500	3,86	3,01	2,62	2,39	2,23	2,12	2,03	1,96	1,90	1,85	1,81	1,77	1,74	1,71	1,69	1,59	1,48	1,42	1,38
600	3,86	3,01	2,62	2,39	2,23	2,11	2,02	1,95	1,90	1,85	1,80	1,77	1,74	1,71	1,68	1,59	1,48	1,41	1,37
700	3,85	3,01	2,62	2,38	2,23	2,11	2,02	1,95	1,89	1,84	1,80	1,77	1,73	1,71	1,68	1,59	1,48	1,41	1,37
1000	3,85	3,00	2,61	2,38	2,22	2,11	2,02	1,95	1,89	1,84	1,80	1,76	1,73	1,70	1,68	1,58	1,47	1,41	1,36

Résumé

Le présent travail comporte une étude QSAR sur une série de dérivés des phénothiazines pipéraziniques. Les composés utilisés sont des inhibiteurs de MDR1, ce qui pourrait, au moins partiellement expliquer leur activité anticancéreuse en utilisant les descripteurs suivants : HE, MR, Pol, SAG, V, log P, MW, HBD, HBA, DM, et les charges atomiques. La régression multilinéaire (RML) a été utilisée pour élaborer la relation entre les descripteurs moléculaires et l'activité biologique des phénothiazines pipéraziniques.

La prédictivité du modèle QSAR a été confirmée par la méthode de validation croisée LOO, une forte corrélation a été observée entre les valeurs de l'activité expérimentale et prédite pour l'activité biologique anticancéreuse, ce qui indique la validité de modèle QSAR obtenu.

Mots Clés : phénothiazines pipéraziniques, Activité anti-cancéreuse, QSAR, MLR, LOO.

Abstract :

The present work involves a QSAR study on a series of derivatives of piperazinic phenothiazines. The compounds used are inhibitors of MDR1, which could at least partially explain their anticancer activity using the following descriptors: HE, MR, Pol, SAG, V, log P, MW, HBD, HBA, DM, and atomic charges. Multilinear regression (MLR) was used to develop the relationship between molecular descriptors and the biological activity of piperazinic phenothiazines.

The predictivity of the QSAR model was confirmed by the LOO cross validation method, a strong correlation was observed between the values of the experimental activity and predicted for the anticancer biological activity, which indicates the validity of the obtained QSAR model .

Keywords : piperazinic phenothiazines, Anticancerous activity, QSAR, MLR, LOO.