



République Algérienne Démocratique et Populaire

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

Université Med Khider Biskra

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

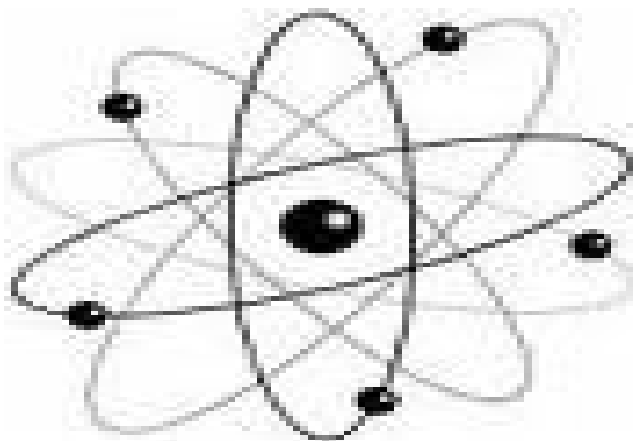


Département des Sciences de la Matière

Domaine : Sciences de la Matière

Filière : Chimie

Spécialité : Chimie pharmaceutique



*Mémoire de fin d'étude en Master*

*Intitulé :*

**Etude quantitative des relations structure/activité  
des propriétés cytotoxiques de dérivés de wogonine**

*Présenté par :*

**FELOUSSI Sara**

*Devant le jury :*

DAOUD Ismail

MELKEMI Nadjib

YOUCEF Oum Hani

*Président*

*Rapporteur*

*Examineur*



# *Dédicace*

*Je dédier ce modeste travail :*

- *À mes très chers parents qui ont aidés, soutenus et encouragés durant toutes ces années d'étude.*
- *Aux chandelles de ma maison mes chères sœurs « Asma, Maroua, Safaa, Rawane, Aya, Hiba ».*
- *À ma grande famille.*
- *À mes amis surtout : Merzakça, Garmia, Ahlam.*
- *À tous mes professeurs que j'ai connus durant mes études.*
- *À tous ceux qui me sont chers.*



# Remerciements

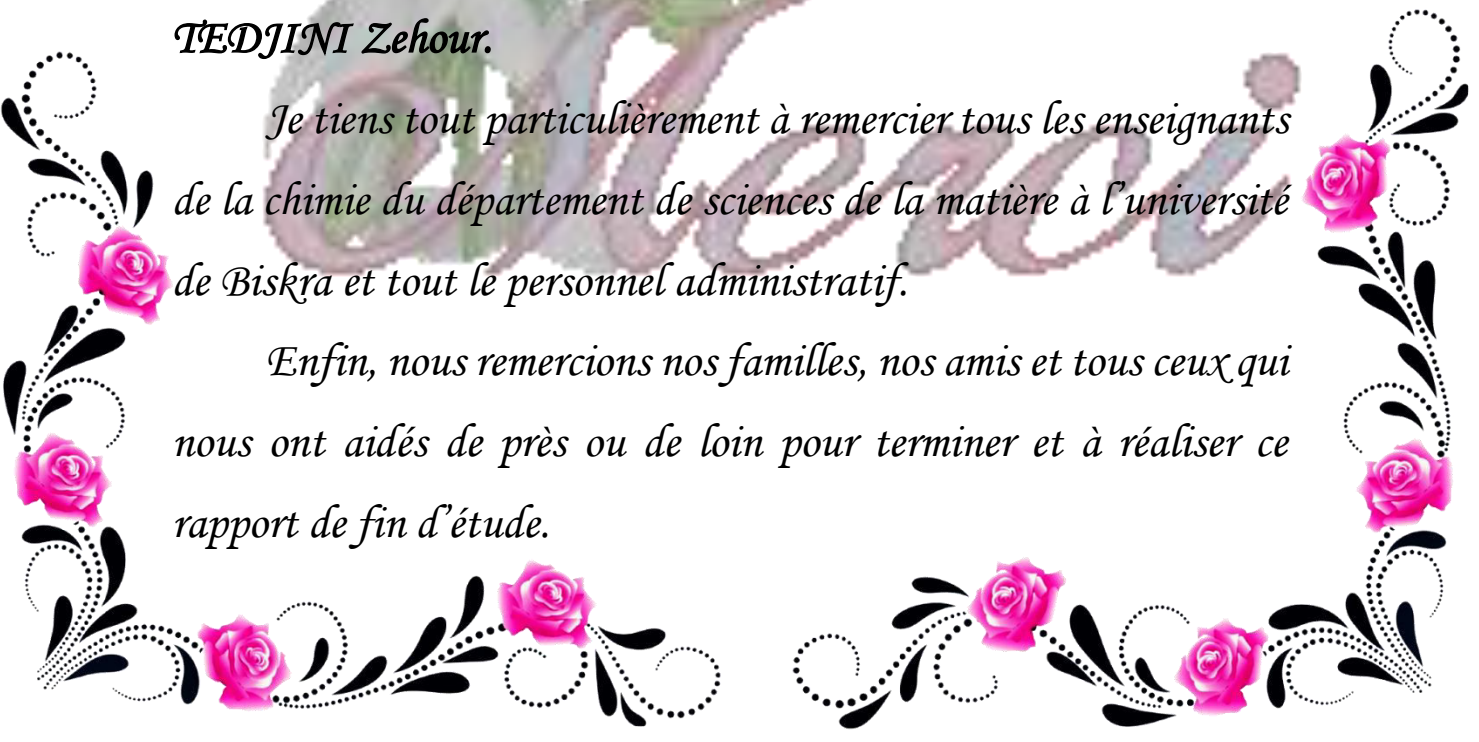
Tout d'abord, nous remercions « **Dieu** » le tout-puissant pour la volonté, la santé et la patience qu'il nous a données durant toutes ces années d'études.

Je souhaite renouveler mes remerciements à mon encadreur **MELKEMI Nadjib** pour avoir proposé, accepté et encadré cette thèse, qui a fourni des conseils judicieux et qui a largement contribué à la réalisation et à l'achèvement de ce travail ; Je me dois également de le remercier pour ses efforts qui m'ont permis d'aboutir dans mon travail de mémoire.

Je remercie également les membres de jury **DAOUD Ismail** et **YOUCEF Oum Hani** pour l'honneur qu'ils m'ont fait en acceptant de juger notre travail. Et je remercie responsable d'informatique **TEDJINI Zehour**.

Je tiens tout particulièrement à remercier tous les enseignants de la chimie du département de sciences de la matière à l'université de Biskra et tout le personnel administratif.

Enfin, nous remercions nos familles, nos amis et tous ceux qui nous ont aidés de près ou de loin pour terminer et à réaliser ce rapport de fin d'étude.



# Sommaire

<b>Liste des tableaux</b> .....	1
<b>Liste des figures</b> .....	2
<b>Liste des principales abréviations</b> .....	3
<b>Introduction générale</b> .....	4
<b>I. Généralités sur le cancer et les dérivés cytotoxiques de wogonine</b> .....	7
I.1 Généralités sur le cancer .....	7
I.1.1 Rappels historiques .....	7
I.1.2 Définition .....	7
I.1.3 Physiopathologie .....	7
I.1.4 Métastases .....	9
I.1.4.1 Définition .....	9
I.1.4.2 Génétique des métastases .....	9
I.1.5 Tumeur .....	10
I.1.5.1 Définition .....	10
I.1.5.2 Types de tumeurs .....	10
I.1.6 Mécanisme de cancérisation .....	11
I.1.6.1 Les propriétés de la cellule cancéreuses .....	12
I.1.7 Classification des médicaments utilisés en chimiothérapie anticancéreuse .....	12
I.1.7.1 Les agents alkylants .....	12
I.1.7.2 Les agents intercalants .....	13
I.1.7.3 Les agents antimétabolites et les inhibiteurs de topoisomérase .....	13
I.2 Généralités sur les wogonines .....	14
I.2.1 Définition .....	14
I.2.2 La structure .....	14
I.2.3 Synthèse des dérivés cytotoxiques de wogonine .....	14
I.2.4 Mécanismes anticancéreux pour le wogonine « Perspectives moléculaires » .....	18
Références bibliographiques .....	20
<b>II. Aperçu théorique sur les méthodes de calcul utilisées (DFT, QSAR)</b> .....	22
II.1 La théorie de la fonctionnelle de la densité (DFT) .....	22

II.1.1	Aperçu historique.....	22
II.1.2	Principes de la théorie DFT « théorèmes de Kohn Hohenberg » .....	22
I.1.2.1	Première théorème (HK 1) .....	22
I.1.2.2	Deuxième théorème (HK 2) .....	23
II.1.3	Les équations de Kohn et Sham.....	23
II.1.4	Les différentes approximations.....	24
II.1.4.1	LDA ou approximation de la densité locale .....	24
II.1.4.2	GGA ou approximation des gradients généralisée .....	25
II.1.5	Succès et limites de la DFT .....	25
II.2	Etude quantitative des relations structure-activité (QSAR) .....	27
II.2.1	Introduction.....	27
II.2.2	Principe de QSAR.....	27
II.2.3	Application de QSAR .....	28
II.2.4	Outils et Méthodologie de QSAR.....	29
II.2.4.1	Paramètres biologiques.....	29
II.2.4.2	Descripteurs moléculaires.....	30
II.2.4.2.1	Les descripteurs physico-chimiques .....	30
II.2.4.2.2	Les descripteurs électroniques .....	30
II.2.4.2.3	Les descripteurs topologiques .....	30
II.2.4.3	Modèles statistiques.....	31
II.2.5	Régression linéaire multiple .....	31
II.2.5.1	Description de la méthode .....	31
II.2.5.2	Test de la signification totale de la régression.....	33
II.2.5.3	Validations des modèles .....	34
II.2.6	Domaine d'applicabilité.....	34
II.3	L'activité biologique.....	36
II.3.1	Découverte des médicaments.....	36
II.3.2	Les différentes cibles des médicaments .....	36
II.3.3	Les tests biologiques .....	36
II.3.3.a	Tests in vivo.....	36
II.3.3.b	Tests in vitro .....	36
II.3.4	Etude de la liaison ligand-récepteur.....	37
	Références bibliographiques.....	38
<b>III.</b>	<b>Résultats et discussion.....</b>	<b>41</b>

III.1 Etude quantitative des relations structure-activité (QSAR) des propriétés cytotoxiques d'une série de dérivés de wogonine.....	41
III.1.1 Introduction .....	41
III.1.2 Méthodologie .....	41
III.1.2.1 Base de données.....	41
III.1.2.2 Sélection des descripteurs et méthodes de calcul .....	45
III.1.3 Développement des modèles QSAR .....	50
III.1.4 Validation des modèles QSAR.....	53
III.1.4.1 Validation interne .....	53
III.1.4.2 Validation externe.....	54
III.1.5 Prédiction de l'activité .....	54
Références bibliographiques.....	58
<b>Conclusion générale</b> .....	<b>60</b>
<b>Annexe</b> .....	<b>61</b>

## *Liste des tableaux*

<b>Tableau II. 1 :</b> Types des données biologiques utilisées dans l'analyse QSAR .....	29
<b>Tableau III.1 :</b> Structures chimique et activité biologique expérimentale des molécules utilisées dans l'étude QSAR.....	42
<b>Tableau III.2 :</b> Valeurs des descripteurs physicochimiques utilisés dans l'étude QSAR .....	45
<b>Tableau III.3 :</b> Valeurs des descripteurs quantiques utilisés dans l'étude QSAR dans la phase gazeuse .....	47
<b>Tableau III.4 :</b> Valeurs des descripteurs quantiques utilisés dans l'étude QSAR dans la phase aqueuse .....	49
<b>Tableau III.5 :</b> matrice de corrélation du 1er modèle .....	52
<b>Tableau III.6 :</b> matrice de corrélation du 2ème modèle .....	52
<b>Tableau III.7 :</b> Paramètres de validation croisée .....	53
<b>Tableau III.8 :</b> Activité cytotoxique expérimentale, prédite et résiduelle des dérivés de wogonine .....	55

## *Liste des figures*

<b>Figure I.1</b> : Les Principales étapes de la progression tumorale.....	8
<b>Figure I.2</b> : L'apparence de la tumeur bénigne.....	10
<b>Figure I.3</b> : L'envahissement des cellules cancéreuses.....	11
<b>Figure I.4</b> : <i>Scutellaria baicalensis</i> .....	14
<b>Figure I.5</b> : Structure chimique de wogonine (5,7-dihydroxy-8-methoxyflavone).....	14
<b>Figure II.1</b> : Présentation de la méthodologie de QSAR .....	28
<b>Figure III.1</b> : Corrélation entre l'activité biologique expérimentale et prédite pour les modèles (1) et (2) .....	57
<b>Figure III.2</b> : Corrélation entre les valeurs résiduelles et l'expérimentales des modèles (1) et (2) .....	57



## *Liste des principales abréviations*

<b>2D</b>	Deux dimensions
<b>3D</b>	Trois dimensions
<b>ADN</b>	Acide déoxyribo nucléique
<b>B3LYP</b>	Becke 3-parameter lee-yang-parr
<b>DM</b>	Dipole moment
<b>DFT</b>	Density functional theory
<b>GGA</b>	Generalized gradient approximation
<b>HF</b>	Hartree-fock
<b>HOMO</b>	Highest occupied molecular orbital.
<b>IC<sub>50</sub></b>	Inhibitory Concentration of 50%.
<b>QSAR</b>	Quantitative structure-activity relationships
<b>QSPR</b>	Quantitative structure-property relationships
<b>LDA</b>	Local density approximation
<b>LUMO</b>	Lowest unoccupied molecular orbital.
<b>LOO</b>	Leave-one-out.
<b>MLR</b>	Multiple linear regression.
<b>PRESS</b>	Predicted residual error some square
<b>SAG</b>	Surface area grid.

*Introduction*

*générale*

### Introduction générale

Le cancer a été la deuxième cause commune de la mort humaine dans le monde développé. Plusieurs facteurs sont responsables de cette maladie, parmi lesquelles on trouve l'utilisation des produits chimiques dans plusieurs domaines industrielles ou agricoles. La majorité de ces produits chimiques ont des propriétés carcinogènes ou toxiques.

Les produits naturels ont longtemps été une source des médicaments antitumoraux. Parmi ces produits, on trouve les flavonoïdes qui sont un groupe de composés polyphénoliques présents dans les plantes et ils montrent diverses activités biologiques telle que l'activité anti-inflammatoire, neuroprotection et antitumorale. Des études épidémiologiques ont montré que les flavones dérivées de *Scutellaria* possédaient des activités cytostatiques et cytotoxiques contre de nombreuses lignées cellulaires cancéreuses humaines avec peu ou pas de toxicité pour les cellules normales. Comme l'un des flavonoïdes représentatifs, le produit naturel wogonine (5,7-dihydroxy-8-méthoxyflavone) et ses dérivés ont récemment été décrits comme possédant des activités inhibitrices contre le cancer. Des grands efforts ont été consacrés à explorer le mécanisme de son activité antitumorale et de nombreux résultats suggèrent qu'il possédait des activités antitumorales puissantes *in vitro* et *in vivo* [1].

La chimie informatique est maintenant devenue un outil très important pour l'industrie et le milieu académique, avec des méthodes théoriques de calcul de plus en plus sophistiquées.

La recherche et la synthèse de nouveaux composés chimiques et biochimiques sont aujourd'hui, souvent associées à une étude par modélisation moléculaire.

La modélisation moléculaire implique l'utilisation des méthodes de calcul théoriques (mécanique moléculaire, dynamique moléculaire, mécanique quantique *ab-initio* ou semi-empirique,...) permettant de déterminer la représentation graphique de la géométrie ou de la configuration des atomes d'une molécule et d'évaluer les propriétés physicochimiques de la molécule étudié [2]. 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 ou l'activité biologique [3].

La relation quantitative structure-activité (QSAR) est l'un des outils les plus pratiques en physico-chimie computationnelle. L'idée fondamentale de QSAR consiste en la possibilité d'une relation entre un ensemble de descripteurs, et une réponse moléculaire. QSAR peut être considérée comme une règle calculée par ordinateur qui décrit quantitativement l'activité

biologique en termes de descripteurs chimiques, il a été fréquemment utilisé pour prédire les activités biologiques des nouveaux composés [4].

La prédiction de L'activité biologique est basée sur des calculs mathématique et statistiques, c'est pour cette raison on met le choix d'une méthode statistique, appelée la régression multilinéaire (RML).

La régression linéaire multiple (MLR) est une méthode statistique utilisée pour quantifier la relation entre les variables indépendantes obtenus par les calculs et une variable dépendante données 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 [5].

Le principal objectif de ce travail est l'étude quantitative des relations structure/activité des propriétés cytotoxiques de certains dérivés de Wogonine. Cette étude permet de développer des modèles mathématiques utilisés dans la prédiction des propriétés bioactives de nouveaux dérivés non synthétisés de Wogonine. Notre travail comporte trois chapitres, Le premier chapitre représente des rappels généraux sur le cancer et les dérivés cytotoxiques de wogonine. Dans le deuxième chapitre nous présenterons des généralités sur les méthodes de recherche utilisés à savoir : DFT, QSAR, .... Dans le troisième chapitre nous présentons les résultats obtenus par l'étude QSAR de l'activité anticancéreuse des dérivés de wogonine et leurs discussions.

### *Références bibliographiques*

- [1]. J. Bian, T. Li, T. Weng, J. Wang, Y. Chen, Z. Li, Synthesis, evaluation and quantitative structure–activity relationship (QSAR) analysis of Wogonin derivatives as cytotoxic agents, *Bioorganic & Medicinal Chemistry Letters*, (2016).
- [2]. S. Belaidi, M. Omari, T. Lanez, et A. Dibi, *J. Soc. Alger. Chim.*, 14, 27, (2004).
- [3]. H. Chermette, *Spectra* 2000. 22, 15, (1993).
- [4]. Camille Georges Wermuth, *The Practice of Medicinal Chemistry* 3rd edition, Elsevier Ltd, (2008).
- [5]. N. Goudarzi, M. Goodarzi, QSAR prediction of HIV inhibition activity of styrylquinoline derivatives by genetic algorithm coupled with multiple linear regressions, *medicinal chemistry research* vol 21, pp. 437-443, (2012).

# *Chapitre I*

---

*Généralités sur le cancer et  
les dérivés cytotoxiques de  
wogonine*

## **I.1 Généralités sur le cancer**

### **I.1.1 Rappels historiques**

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 correspond à la prolifération non contrôlée de cellules dites malignes, en raison de certains caractères anormaux. La prolifération peut rester localisée (tumeur) ou se propager dans d'autres sites, soit par voie lymphatique (envahissement ganglionnaire), soit par voie sanguine (métastases). Le cancer est une maladie de l'ADN. Certains gènes présentent des anomalies existant déjà dans l'œuf initial pour les rares cancers héréditaires et apparaissant sur certaines cellules au cours de la vie pour les nombreux cancers acquis.

Ceci signifie que des gènes de structure normale au départ se sont modifiés. Le cancer est dû à des anomalies génétiques touchant une cellule. Au moins deux modifications et souvent davantage, en général quatre, sont nécessaires pour que la cellule devienne maligne.

La cellule perd sa forme spécifique, ne réagit plus aux signaux extérieurs, en particulier aux signaux d'inhibition de la croissance. Elle se multiplie exagérément, ainsi que ses descendantes. Une seule cellule maligne peut être à la base d'un cancer [2].

### **I.1.3 Physiopathologie**

Le tissu cancéreux comporte deux composantes : les cellules cancéreuses et le stroma. Le stroma est un tissu non tumoral provenant de l'hôte. Il est le tissu nourricier et de soutien des cellules tumorales.

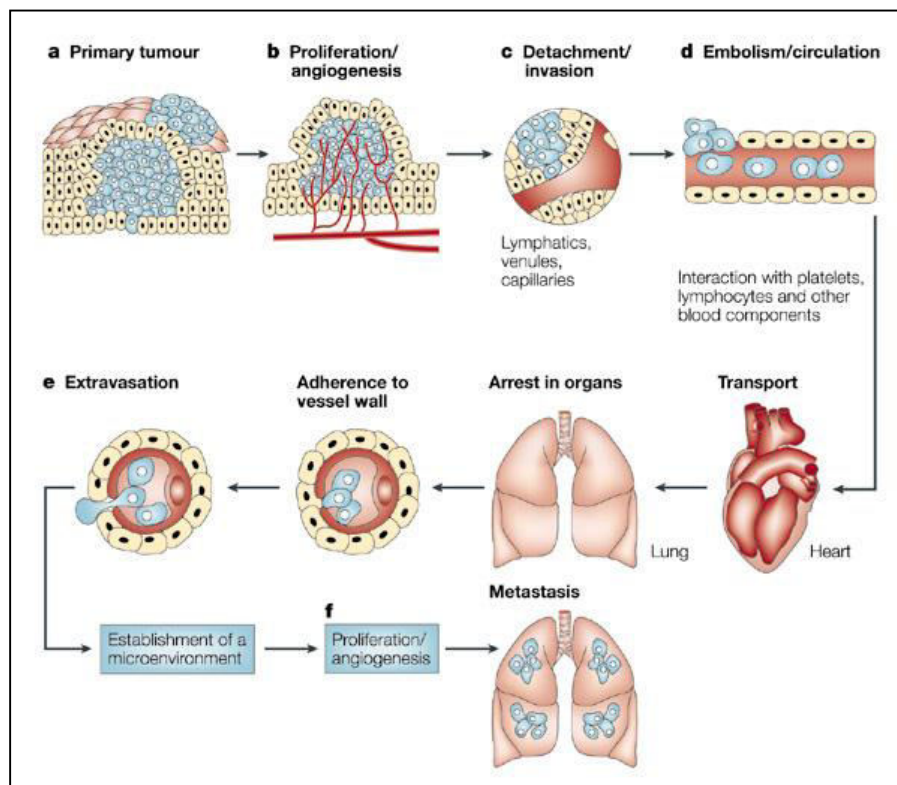
Deux conséquences directes de la migration des cellules cancéreuses sont l'angiogenèse, nécessitant un recrutement de cellules endothéliales et de vaisseaux sanguins pour irriguer la tumeur, et la formation de métastases [1].

L'invasion tumorale implique plusieurs mécanismes hautement coordonnés, à savoir :

- ❖ L'adhésion de la cellule aux composants de la matrice extracellulaire, impliquant des modifications de la composition moléculaire de cette matrice ;
- ❖ La motilité cellulaire, impliquant la réorganisation du cytosquelette ;
- ❖ L'invasion cellulaire, impliquant la dégradation de la matrice extracellulaire par des enzymes protéolytiques sécrétés par la tumeur.

La progression tumorale dépend du pouvoir prolifératif et du pouvoir métastasant de la tumeur [3]. Les principales voies de migration sont sanguines ou lymphatique. Quel que soit la voie, les cellules cancéreuses qui quittent le foyer tumoral initial doivent franchir des étapes successives : le détachement de la cellule de la tumeur primitive, l'invasion de la matrice extracellulaire, le passage dans la circulation par intravasation, la survie dans la circulation, l'extravasation, la survie et la prolifération dans un site étranger [4].

Pour se détacher de la tumeur primaire et envahir le tissu adjacent, la cellule cancéreuse dégrade la membrane basale et la matrice extracellulaire.



**Figure I.1** : Les Principales étapes de la progression tumorale.

Cette figure décrit les différentes étapes de la progression d'une tumeur.

- a. Formation et croissance de la tumeur primaire.
- b. Hyper vascularisation de la tumeur par la formation d'un réseau capillaire.



- c. Détachement de cellules de la tumeur primaire et envahissement de la membrane basale.
- d. Ces cellules détachées seront transportées par la circulation, (et pour survivre dans la circulation générale devront résister aux défenses immunitaires de l'organisme) pourront s'arrêter dans un lit capillaire et adhérer à la membrane basale subendothéliale.
- e. Ces cellules pénètrent dans le parenchyme de l'organe atteint.
- f. La prolifération dans le parenchyme de l'organe complète le processus métastatique.

Pour continuer à croître, les micro-métastases développent un réseau vasculaire et déjouent les défenses de l'hôte. Enfin, les cellules tumorales vont envahir les vaisseaux sanguins et produire d'autres métastases [5].

## **I.1.4 Métastases**

### **I.1.4.1 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 [6]. 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 [7]. 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[6].

### **I.1.4.2 Génétique des métastases**

Avec la publication de la séquence du génome humain, et différentes initiatives majeures telles que le cancer génome project au Royaume-Uni et le cancer génome anatomie project 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. Cependant, la capacité de prédire la formation et la localisation

des métastases, la recherche de thérapies ciblées efficaces contre les pathologies métastatiques demeurent des défis majeurs en oncologie expérimentale [1].

### I.1.5 Tumeur

Un cancer résulte de la prolifération anarchique de n'importe quel type de cellule composant un organisme ; aussi, observe-t-on plus d'une centaine de types de cancers, parfois très différents les uns des autres dans leur comportement et leur sensibilité aux traitements [8].

#### I.1.5.1 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 [9].

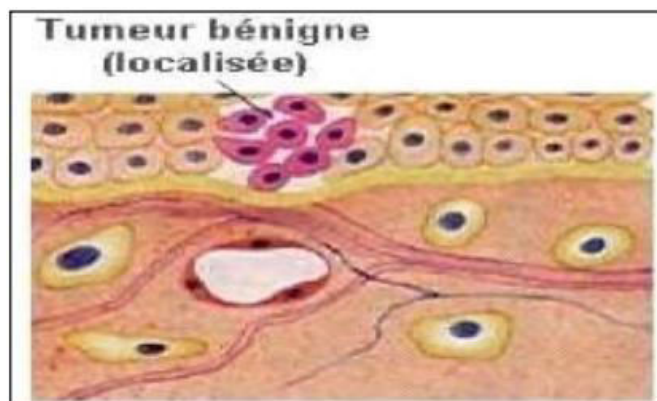
#### I.1.5.2 Types de tumeurs

La première classification c'est qu'on classe les tumeurs tant bénignes que malignes.

##### a) Les tumeurs bénignes

Une tumeur bénigne est une masse de cellules transformées logées à l'intérieur d'un tissu (**figure I.2**) [8]. Les tumeurs bénignes se développent localement et restent cantonnées au tissu dans lequel elles ont pris naissance. Leur croissance est lente. Toutefois, elles peuvent atteindre un volume et un poids importants. Elles ne récidivent pas après ablation chirurgicale, à condition que l'exérèse soit complète. Ces tumeurs ne métastasent jamais [10].

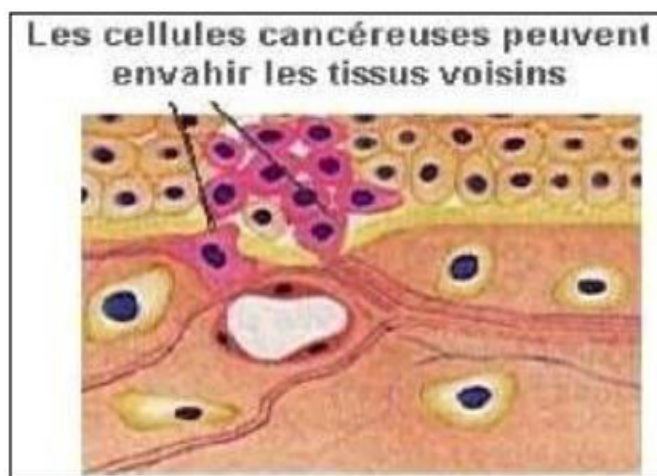
On note que les tumeurs bénignes ne posent de problème médicale sérieux que quand leur volume interfère avec le fonctionnement normal d'un organe ou quand elles sécrètent des substances biologiques actives, telles des hormones [8].



**Figure I.2** : L'apparence de la tumeur bénigne.

## b) Les tumeurs malignes

Les tumeurs malignes sont des cancers résultent d'une croissance illimitée et autonome d'un clone cellulaire anormale. Cette prolifération cellulaire aboutit à la formation d'une masse tumorale ou néoplasie maligne [8]. Les tumeurs malignes ont habituellement une croissance rapide. Elles donnent naissance à une dissémination tumorale à distance (surtout par voie lymphatique et sanguine) avec éclosion et développement de tumeur secondaires dans d'autres viscères (les métastases) (**figure I.3**). Les tumeurs malignes ont tendance à récidiver après éradication locale. L'évolution, en dehors des traitements se fait spontanément vers la mort [7].



**Figure I.3 :** L'envahissement des cellules cancéreuses.

La deuxième classification c'est selon le type de cellules dont elles proviennent, à savoir en carcinomes, en sarcomes et en leucémies. Les carcinomes, qui comptent à peu près 90% des tumeurs humaines, se composent de cellules épithéliales malignes. Les sarcomes, rares chez l'Homme, sont des tumeurs solides des tissus conjonctifs, tels les muscles, l'os, le cartilage et le tissu fibreux. Les leucémies et les lymphomes, qui présentent environ 8% des tumeurs malignes, sont issus respectivement de cellules destinées au courant sanguin et de la cellule du système immunitaire.

La troisième classification c'est selon l'organe dont elles sont issues, par exemple cancer du poumon, cancer du sein ...etc. [8]

### I.1.6 Mécanisme de cancérisation

Le cancer est une maladie génétique des cellules caractérisé par la présence dans l'organisme d'une tumeur maligne. Les tumeurs malignes résultent d'une croissance illimitée et autonome d'un clone cellulaire anormale. Cette prolifération cellulaire aboutit à la formation d'une

masse tumorale. Celle-ci détruit le tissu normale et envahit les tissus environnants et peu donner des métastases à distance [11]. En absence du traitement, le cancer finit par provoquer la mort du patient. La transformation d'une cellule normale en une cellule maligne résulte d'une accumulation d'altération au niveau des gènes de la cellule (mutation) ; ces altérations génétiques sont acquises progressivement au cours du temps [12].

### **I.1.6.1 Les propriétés de la cellule cancéreuses**

La cellule cancéreuse possède 5 propriétés capitales qui sont :

- Elle méprise tous les signaux provenant des autres cellules qui normalement lui interdisent de se diviser activement.
- Elle évite le suicide cellulaire, processus normal dans de nombreux tissus et que l'on appelle apoptose.
- Elle échappe au vieillissement normal d'une cellule.
- Elle est très instable sur le plan génétique, subissant de nombreux accidents chromosomiques.
- Elle peut s'échapper de son tissu d'origine pour survivre et proliférer dans des sites qui lui sont normalement étrangers, c'est le mécanisme de la métastase qui fait, par exemple, qu'un cancer du sein pourra envoyer une cellule dans le foie ou le poumon pour y créer une tumeur secondaire [13].

## **I.1.7 Classification des médicaments utilisés en chimiothérapie anticancéreuse**

Il existe cinq grandes classes de médicaments utilisés en chimiothérapie anticancéreuse. Il s'agit des agents alkylants, des agents intercalants, des agents antimétabolites, des inhibiteurs de topoisomérase et des poisons du fuseau mitotique [1].

Pour ce travail, on va expliquer brièvement les agents alkylants, les agents intercalants, Les agents antimétabolites et les inhibiteurs de topoisomérases.

### **I.1.7.1 Les agents alkylants**

Les agents alkylants sont des composés capables de fixer un groupe alkyl ( $R-CH_2$ ) sur des acides nucléiques (ADN) ou des protéines. Dans ce groupe on retrouve des médicaments capables de se fixer de façon irréversible à l'ADN, comme les sels de platine. Les alkylants forment des adduits sur les atomes d'azote et d'oxygène des bases d'ADN. Bi-fonctionnels, ils se fixent sur deux sites proches pour créer des ponts (intra ou inter-brins) dans des chaînes de l'ADN et créent des liaisons complexes, ce qui peut entraîner un arrêt du cycle cellulaire, une mort cellulaire par

apoptose et l'inhibition de la croissance tumorale. L'activité cytotoxique des agents alkylants, dits encore radiomimétiques, ne dépend pas du cycle cellulaire.

La position N-7 de la guanine est le site d'alkylation le plus courant, mais ce site varie selon le type de molécule utilisé. Les agents alkylants appartiennent à plusieurs familles chimiques comme les moutardes azotées (chlorambucil, busulfan, melphalan) et les dérivés du platine (cisplatine, carboplatine) [14]. Certains agents alkylants exercent un mécanisme d'action différent des agents alkylants « classiques » il s'agit par exemple du témozolomide qui induit une mort cellulaire du type autophagique (mort cellulaire programmée de type II) plutôt qu'apoptotique (mort cellulaire programmée de type I). Ceci explique en partie le bénéfice thérapeutique lié à cette molécule dans le cas du traitement des gliomes malins qui sont résistants au processus de mort pro apoptotique, et donc à la plupart des agents utilisés en chimiothérapie conventionnelle [15].

### **I.1.7.2 Les agents intercalants**

C'est une classe d'agents anticancéreux qui agit par insertion d'une molécule aromatique plane dans la double hélice de l'ADN. Il s'en suit un blocage de la réplication de l'ADN, l'arrêt du cycle cellulaire puis la mort de la cellule. On peut citer en exemple la doxorubicine. La plupart des agents intercalants sont toutefois des inhibiteurs de topoisomérases.

### **I.1.7.3 Les agents antimétabolites et les inhibiteurs de topoisomérases**

Les topoisomérases sont des enzymes qui modifient le degré d'enroulement de l'ADN en induisant les cassures transitoires et parfaitement régulées afin d'assurer la réplication de l'ADN. Les inhibiteurs de topoisomérases et les agents anti-métabolites agissent indirectement sur l'ADN en bloquant des enzymes impliquées dans la synthèse ou la réplication de l'ADN. Les inhibiteurs de ces enzymes empêchent la réparation des cassures, provoquant un arrêt du cycle cellulaire pouvant conduire à une mort cellulaire programmée de type I c'est-à-dire l'apoptose[1].

## I.2 Généralités sur les wogonines

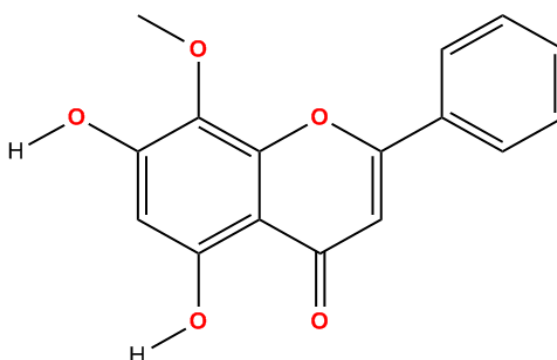
### I.2.1 Définition

Le wogonine est un composé végétal flavonoïde extrait de *Scutellaria baicalensis* (Huang-Qin ou scutellaire chinoise) (**figure I.4**), et a été étudié à fond par de nombreux chercheurs jusqu'à présent pour ses propriétés antivirales, anti-oxydantes, anti-cancéreuses et neuroprotectrices. Des nombreuses expériences menées in vitro et in vivo ont démontré les excellentes propriétés de la wogonine qui inhibitrices la tumeur [16].



**Figure I.4** : *Scutellaria baicalensis*.

### I.2.2 La structure

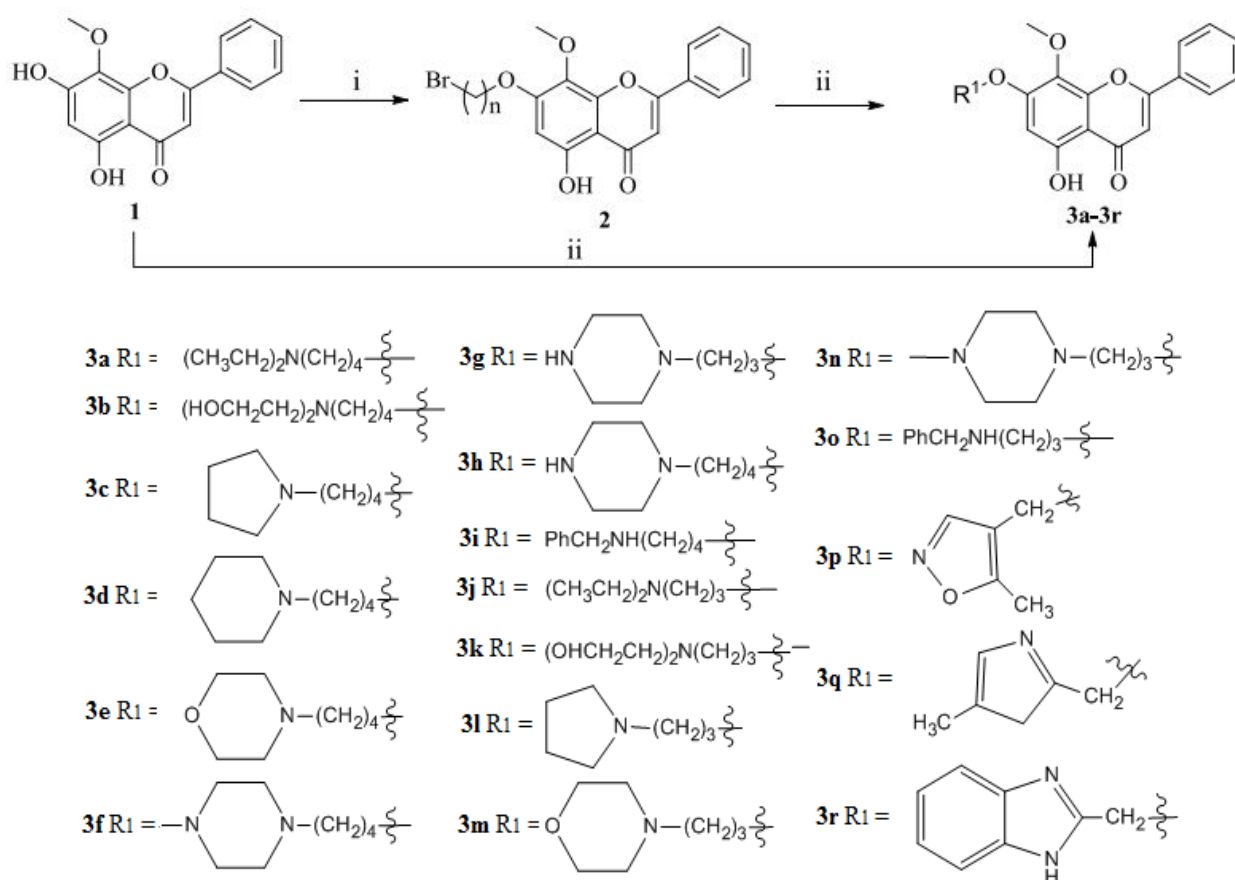


**Figure I.5** : Structure chimique de wogonine (5,7-dihydroxy-8-methoxyflavone).

### I.2.3 Synthèse des dérivés cytoxiques de wogonine

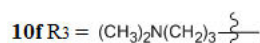
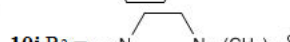
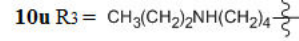
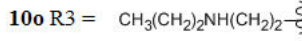
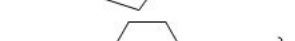
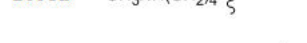
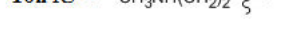
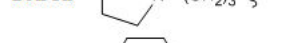
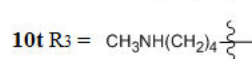
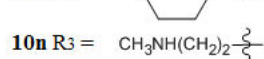
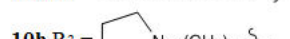
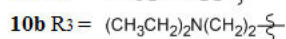
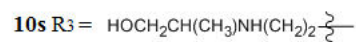
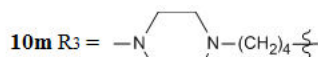
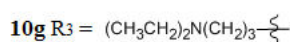
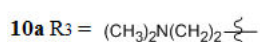
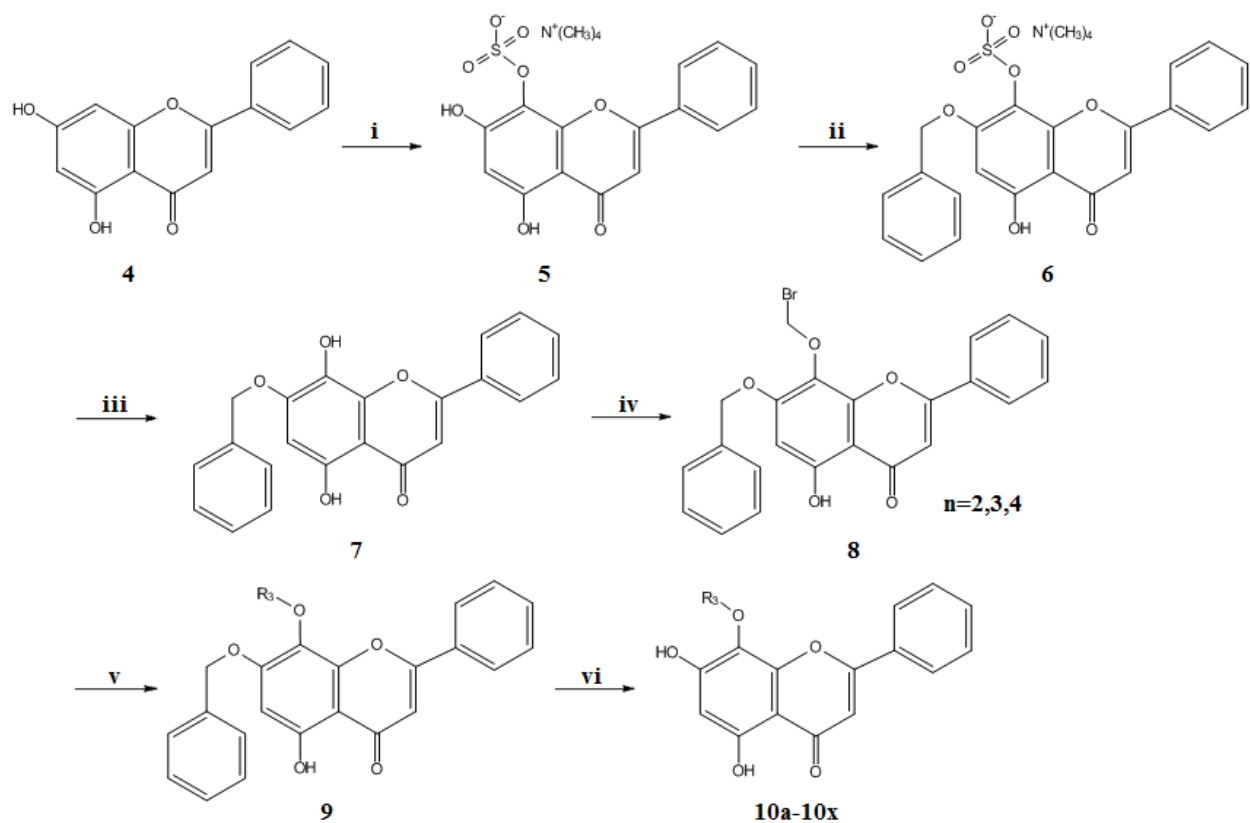
La synthèse des dérivés de wogonine a été décrite dans *le schéma I.1-I.3*. Tout d'abord, des dérivés substitués en 7-OH (**3a-3r**) ont été décrits dans *le schéma I.1*. A partir de wogonine (**1**), un lien à trois ou quatre carbones a été introduit dans 7-OH pour former l'intermédiaire **2** qui

a ensuite réagi avec plusieurs amines structurellement diverses pour fournir des composés **3a-3o**. Les composés **3p-3r** ont été obtenus directement à partir de **1** avec un chlorohétérocycle correspondant par reflux dans l'acétonitrile. *Le schéma I.2* décrit la synthèse de dérivés substitués en 8-OH (**10a-10v**). Ces composés ont été synthétisés en six étapes à partir de la chryisine (**4**). L'hydroxyde de tétraméthylammonium ayant réagi avec **4** pour donner l'intermédiaire **5** avec un rendement de 30%. Ensuite, la benzylation de **5** a été effectuée par réaction avec du bromure de benzyle en présence de carbonate de potassium pour obtenir l'intermédiaire **6**. Le composé **7** a été obtenu par acidifié **6** en présence d'acide chlorhydrique. Suivi par introduit un lieu à deux ou trois ou quatre carbone à 8-OH à forme le composé **8**, qui réagit ensuite avec diverses amines pour donner des intermédiaires clés **9**. La débenzylation de **9a-9v** donne lieu aux composés cibles **10a-10x**. Les voies de synthèse pour les composés **19a-19e** étaient décrites dans *le schéma I.3*. La benzylation de **11** a été effectuée par réaction avec du bromure de benzyle dans  $K_2CO_3$  pour donner l'intermédiaire **12**. L'oxydation de **12** en utilisant 65% de  $HNO_3$  a fourni la quinone **13**, qui a été réduite par le thiosulfate de sodium en le phénol **14**. La méthylation de **14** avec diméthylsulfate a obtenu l'intermédiaire **15**, qui a ensuite débenzylé pour donner le composé **16**. Ensuite par une acylation de Friedel-Crafts avec des dérivés de chlorure de cinnamoyle en utilisant du  $BF_3 \cdot Et_2O$  comme catalyseur on obtient les intermédiaire **1a** et **17a-17e**. Par la suite, **1a** et **17a-17g** ont été cyclisés en flavones correspondantes (**1b** et **18a-18e**) médiées par l'iode à  $120^\circ C$ . Finalement, les composés cibles **1** et **19a-19e** ont été obtenus par déméthylation de l'intermédiaire **1b** et **18a-18e**. Tous les composés cibles ont été confirmés par les spectres RMN  $H^1$ , IR et HRMS (ESI), et la structure représentative **10o** a également été confirmée par spectre RMN  $C^{13}$  [17].

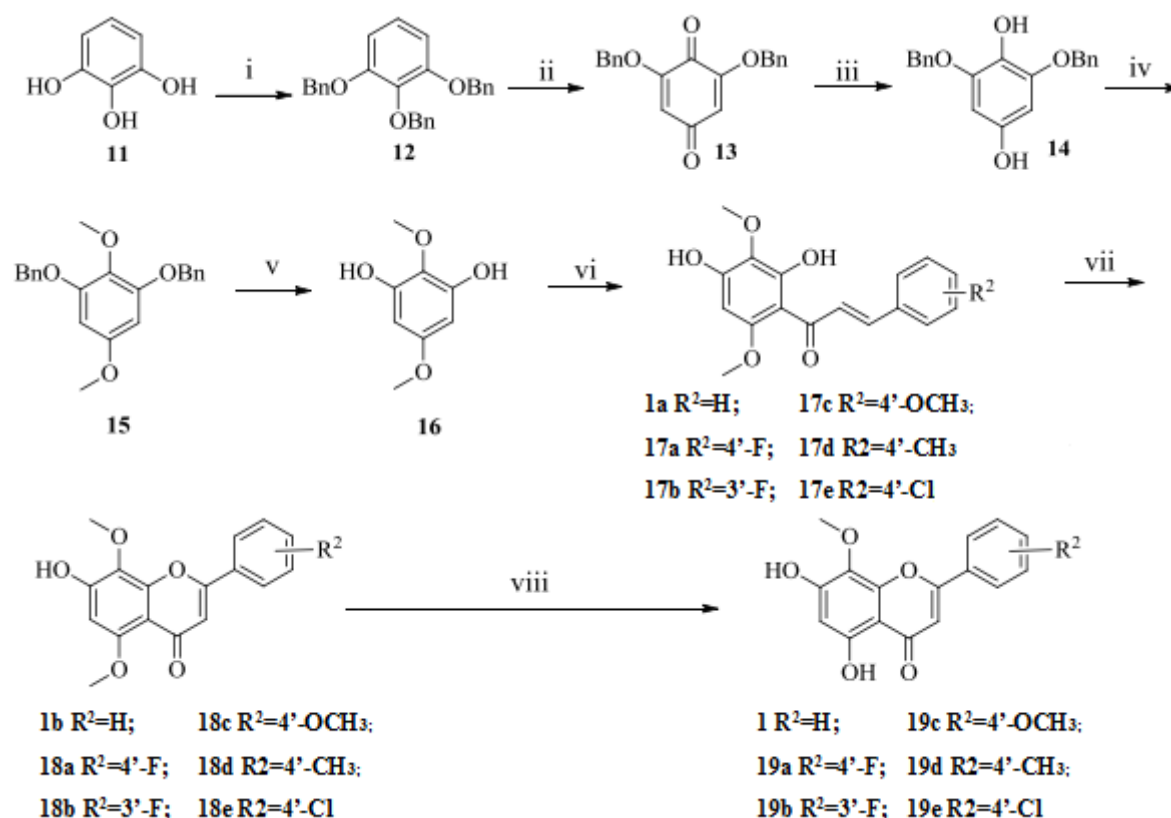


**Schéma I.1** : Synthèse des composés **3a-3r**. Réactifs et conditions : (i)  $\text{Br}(\text{CH}_2)_n\text{Br}$ ,  $\text{K}_2\text{CO}_3$ , acétone, reflux, 8 h, 70%; (ii) dérivés d'amino substitués,  $\text{CH}_3\text{CN}$ , reflux, 1-5 h, 30-60%.





**Schéma I.2** : Synthèse des composés **10a-10v**. Réactifs et conditions: (i) K<sub>2</sub>S<sub>2</sub>O<sub>8</sub>, (CH<sub>3</sub>)<sub>4</sub>NOH, H<sub>2</sub>O, r.t., 7 h, 30%; (ii) PhCH<sub>2</sub>Br, K<sub>2</sub>CO<sub>3</sub>, DMF, 50 ° C, 6 h; (iii) 6 M HCl, r.t., 8 h, 40%; (iv) Br(CH<sub>2</sub>)<sub>n</sub>Br, K<sub>2</sub>CO<sub>3</sub>, acétone, reflux, 8 h, 75%; (v) dérivés d'amino substitués, CH<sub>3</sub>CN, reflux, 1-5 h, 30-60%; (vi) H<sub>2</sub>, Pd / C, THF, r.t., 8 h, 80%.



**Schéma I.3** : Synthèse des composés **19a-19h**. Réactifs et conditions: (i) PhCH<sub>2</sub>Br, K<sub>2</sub>CO<sub>3</sub>, acétone, reflux, 48 h, 87%; (ii) AcOH, 65% de HNO<sub>3</sub>, 40 ° C, 4 heures, 45%; (iii) Na<sub>2</sub>S<sub>2</sub>O<sub>4</sub>, CH<sub>3</sub>COOC<sub>2</sub>H<sub>5</sub>, r.t., 1 h, 37%; (iv) (CH<sub>3</sub>)<sub>2</sub>SO<sub>4</sub>, NaOH, C<sub>2</sub>H<sub>5</sub>OH, r.t., 3 h, 85%; (v) H<sub>2</sub>, Pd /C, MeOH, r.t., 8h, 90%; (vi) Dérivés de chlorure de cinnamoyle, BF<sub>3</sub>-Et<sub>2</sub>O, CHCl<sub>3</sub>, reflux, 1,5-6 h, 35-80%; (vii) I<sub>2</sub>, DMSO, 120 ° C, 5-8 h, 30-40%; (viii) AlCl<sub>3</sub>, CH<sub>3</sub>CN, reflux, 8 h, 65-85%.

#### I.2.4 Mécanismes anticancéreux pour le wogonine « Perspectives moléculaires »

Les plus important cible dans le traitement du cancer est d'induire l'apoptose (la mort cellulaire programmé) et pendant le processus, noyau se condense et fragmente tout en corps apoptotiques sont formés sans dommages sur membrane plasmique. Ces cellules sont activement adressées à être défunt et sont naturellement éliminé par les phagocytes sans aucun signe de l'inflammation [18]. Un autre type de la mort cellulaire est une nécrose, qui se produit et affecte un grand domaine de cellules et est médiée par l'absence de l'approvisionnement énergétique ou par les dommages sur la membrane cellulaire.

Par conséquent, nécrose des résultats dans la libération de cytoplasmique contenu/ chimiotactique signaux en tissu et facilite de recruter des cellules inflammatoires. En raison de ces différences, l'apoptose est préféré et devient un des normes requises lorsque un nouvel

anticancéreux est développé. De nombreuses études ont été réalisées pour augmenter la cytotoxicité des médicaments sur les cellules cancéreuses au cours du traitement. Physiologiquement, la cytotoxicité peut conduire les cellules cancéreuses à passer par la nécrose, ce qui entraîne une détérioration de leurs membranes, conduisant à la lyse cellulaire ou à la suspension de la prolifération cellulaire ou à la promotion de l'apoptose. De nombreuses approches pour évaluer le niveau de cytotoxicité dans les cellules cancéreuses ont été étudiées, y compris des tests pour l'intégrité de la membrane cellulaire, le métabolisme cellulaire et la prolifération [19-21]. Pour le traitement du cancer, de nombreux agents chimiques induisent une cytotoxicité dans les cellules cancéreuses. Cependant, les principaux obstacles dans leur application clinique comprennent les effets secondaires et la résistance aux médicaments, conduisant à l'obsession chez les patients et à la rechute des tumeurs après traitement [16].

**Références bibliographiques**

- [1]. 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).
- [2]. A. Mihoubi, Effet des habitudes alimentaires sur les cancers du tube digestif au niveau de la wilaya de Batna Etude cas-témoins, Université colonel el hadj lakhdar-batna. (2009).
- [3]. F. Blasi, M. P. Stoppelli, Proteases and cancer invasion: from belief to certainty: AACR Meeting on Proteases and Protease Inhibitors in Cancer, Nyborg, Denmark, 14-18 June 1998. *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer* 1423 (1), 35-44, (1999).
- [4]. J. Condeelis, J. E. Segall, Intravital imaging of cell movement in tumours. *Nat. Rev. Cancer* 3: 921-930, (2003).
- [5]. J. Fidler Isaiah, L. Kripke Margaret. The challenge of targeting metastasis, *Cancer Metastasis Rev* 34:635–641, (2015)
- [6]. OMS, Le cancer dans le monde. IARC : Press Lyon, (2005).
- [7]. P. Mehlen, A. Puisieux, Metastasis: a question of life or death. *Nat Rev Cancer* 6(6), 449-458, (2006).
- [8]. Kh. Laadjal , Une contribution à l'étude épidémiologique du cancer du sein dans la région de Biskra, Université Mohamed Khider Biskra, (2013).
- [9]. Ag. Ludovic KAM, thèse de doctorat, tumeur abdominales de l'enfant : aspects épidémiologiques, diagnostiques, thérapeutiques et évolutifs au centre hospitalier national yalgado ouedraogo, Université de Ouagadougou BURKINA FASO, (2003).
- [10]. Kh. Medjdoub, M. Rahmani, A. Boudjarda, Curcuma : quelle stratégie adoptée en prévention du cancer?, Université des Frères Mentouri Constantine, (2016).
- [11]. S. Lechelache, Intérêts et limites de la classification moléculaire dans le cancer du sein. Master Bio signalisation cellulaire, moléculaire et immunologie. Université Saad Dahlab Blida, 53p, (2011).
- [12]. M. Abdelali, Génétique humaine, OPU, Algérie, pp. 157-62, 2006.
- [13]. P. Kamoun, Les maladies génétiques de l'hémophilie au cancer, Josette, Paris, pp. 109-28, 2006.
- [14]. E. Espinosa, P. Zamora, J. Feliu, M. Gonzalez-Baron, Classification of anticancer drugs, a new system based on therapeutic targets. *Cancer Treatment Reviews* 29(6), 515-523, (2003).

- [15]. F. Lefranc, N. Sadeghi, I. Camby, T. Metens, O. Dewitte, R. Kiss, Present and potential future issues in glioblastoma treatment. *Expert Review of Anticancer Therapy* 6(5), 719-732, (2006).
- [16]. Do Luong Huynh, et al, Anti-tumor activity of wogonin, an extract from *Scutellaria baicalensis*, through regulating different signaling pathways, *Chin J Nat Med*, 15(1): 15–40, (2017).
- [17]. J. Bian, T. Li, T. Weng, J. Wang, Y. Chen, Z. Li, Synthesis, evaluation and quantitative structure–activity relationship (QSAR) analysis of Wogonin derivatives as cytotoxic agents, *Bioorganic & Medicinal Chemistry Letters*, (2016).
- [18]. S. Elmore, Apoptosis: a review of programmed cell death [J]. *Toxicol Pathol*, 35(4): 495-516, (2007).
- [19]. T. Decker T, ML. Lohmann-Matthes, A quick and simple method for the quantitation of lactate dehydrogenase release in measurements of cellular cytotoxicity and tumor necrosis factor activity [J]. *J Immunol Methods*, 115: 61-69, (1988).
- [20]. AL. Niles, RA. Moravec, P. Eric-Hesselberth, et al, A homogeneous assay to measure live and dead cells in the same sample by detecting different protease markers [J]. *Anal Biochem*, 366(2): 197-206, (2007).
- [21]. TL. Riss, RA. Moravec, AL. Niles. Cytotoxicity testing: measuring viable cells, dead cells, and detecting mechanism of cell death [J]. *Methods Mol Biol*, 740: 103-114, (2011).

# *Chapitre II*

---

*Aperçu théorique sur les  
méthodes de calcul  
utilisées (DFT, QSAR)*

## II.1 La théorie de la fonctionnelle de la densité (DFT)

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

### II.1.1 Aperçu historique

La théorie fonctionnelle de la densité a pour objet de décrire un système en considérant la densité  $p(r)$  comme variable de base. Ainsi le problème à  $n$  électrons est étudié dans l'espace de  $p(r)$  qui est de dimension 3 au lieu de l'espace de dimension  $3n$  de la fonction d'onde  $\psi$ . Les premiers à avoir exprimé l'énergie en fonction de la densité furent L. H. Thomas et E. Fermi en 1927. Dans leur modèle, les interactions électroniques sont traitées classiquement et l'énergie cinétique est calculée en supposant la densité électronique homogène. Ce modèle a été amélioré par P. A. Dirac en 1930 avec un terme d'échange.

Un peu plus tard, en 1951 J. C. Slater [1] proposa un modèle basé sur l'étude d'un gaz uniforme améliorée avec un potentiel local. Cette méthode, appelée Hartree-FockSlater ou  $X\alpha$ , fut essentiellement utilisée en physique du solide dans les années 70. Mais la DFT a véritablement commencé avec les théorèmes fondamentaux de Hohenberg et Kohnen 1964 [2] qui établissent une relation fonctionnelle entre l'énergie de l'état fondamental et sa densité. Alors que le premier réus applications de DFT pour la recherche sur la structure électronique moléculaire a commencé à apparaître dans les années 90 avec le développement des fonctionnels d'échange et de corrélation [1, 3,4], les plus précises et les plus rapides pour le calcul des propriétés électroniques de grands systèmes moléculaires ou a été introduite dans le code Gaussienne. Enfin nous signalons par le prix Nobel qui a été attribué à Kohn et à Pople [5,6] en 1998 dans le cadre de développement de cette méthode.

### II.1.2 Principes de la théorie DFT « théorèmes de Kohn Hohenberg »

Les deux théorèmes de Hohenberg et Kohn formulés en 1964 [2] ont permis de donner une cohérence aux modèles développés sur la base de la théorie proposée par Thomas et Fermi à la fin des années 30.

#### II.1.2.1 Première théorème (HK 1)

Le formalisme de la théorie de la fonctionnelle de la densité (DFT) est basé sur deux théorèmes de Hohenberg et Kohn [7]. Dans le premier théorème, les auteurs ont montré que l'énergie totale  $E$  d'un système à  $N$  électrons dans son état fondamental en présence d'un potentiel

externe  $V_{ext}(r)$  peut être déterminée uniquement par sa densité électronique  $p(r)$  et peut être écrite sous la forme :

$$E(p) = F(p) + \int p(r)V_{ext}(r)d^3r \quad (II-1)$$

Où :

$$F[(p)] = T [(p)] + E_{e-e}[(\rho)] \quad (II-2)$$

$F(\rho)$  Est une fonction universelle de la densité électronique  $\rho$ ,  $T$  est l'énergie cinétique et  $V_{e-e}$  l'énergie d'interaction électron-électron.

### II.1.2.2 Deuxième théorème (HK 2)

Dans le second théorème de Hohenberg et Kohn montrent que la vraie densité de l'état fondamental n'est que celle qui minimise l'énergie  $E(p)$  et toutes les autres propriétés sont également une fonctionnelle de cette densité.

$$E(p_0) = \min E(p)$$

$p_0$  Est la densité de l'état fondamental. Malheureusement, la fonctionnelle  $F(p)$  n'est pas connue et les équations correspondantes ne peuvent pas être résolues.

### II.1.3 Les équations de Kohn et Sham

En 1965 Kohn et Sham (KS) proposent une méthode pratique permettant d'utiliser la théorie de la fonctionnelle de la densité. Ces auteurs ont considéré l'équivalence entre un système d'électrons en interaction dans un potentiel extérieur  $V(r)$  et un système d'électrons sans interaction dans un potentiel effectif  $V_{eff}$ . Par conséquent, la fonctionnelle de l'énergie peut être exprimée par l'expression suivante :

$$E(p) = T_0 [p(\vec{r})] + E_H[p(r)] + E_{xc} [p(\vec{r})] + \int V_{ext}(r) p(\vec{r})d^3r \quad (II-3)$$

Où :  $T_0 [p(\vec{r})]$  est l'énergie cinétique du gaz électronique sans interaction,  $E_H[p(r)]$  désigne le terme de Hartree des électrons,  $E_{xc} [p(\vec{r})]$  est l'énergie d'échange-corrélation.

La différence entre l'énergie cinétique réelle et celle des électrons sans interaction ainsi que la différence entre l'énergie d'interaction réelle et celle de Hartree sont prises en compte dans l'énergie d'échange-corrélation  $E_{xc} [p(\vec{r})]$ .

Donc l'équation de Schrödinger à résoudre dans le cadre de l'approche de Kohn et Sham de façon cohérente est de la forme :

$$\left( -\frac{1}{2} \nabla^2 + V_{eff}(r) \right) \varphi_i(r) = \varepsilon_i \varphi_i(r) \quad (II-4)$$



Où le potentiel effectif est défini par :

$$V_{eff}(r) = V_H[\rho(r)] + V_{ext}[\rho(r)] + V_{xc}[\rho(r)] \quad (II-5)$$

Avec  $V_H[\mathbf{p}(r)] = \frac{1}{2} \int \frac{e^2}{4\pi\epsilon_0} \frac{\mathbf{p}(r')}{|r-r'|} d\mathbf{r}'$  est le potentiel de Hartree des électrons.

$V_{xc}[\mathbf{p}(r)] = \frac{\delta E_{xc}[\mathbf{p}(r)]}{\delta \mathbf{p}(r)}$  Est le potentiel d'échange et de corrélation.

Et  $\varphi_i(\mathbf{r})$  sont les états d'une seule particule et la densité électronique  $\mathbf{r}(\mathbf{r})$  est donnée par la somme des orbitales occupées comme suit :

$$p(r) = \sum_{occup} \varphi_i^*(r) \varphi_i(r) \quad (II-6)$$

La résolution des équations de Kohn et Sham permet de trouver la densité électronique exacte, à condition que la fonctionnelle d'échange et de corrélation  $E_{xc}[\overline{\mathbf{p}(r)}]$  soit déterminée exactement.

### II.1.4 Les différentes approximations

Comme c'est le problème de la fonction de corrélation et d'échange reste toujours délicat l'utilisation des approximations est toujours demandable et pour cela des nouvelles approximations ont été trouvées avec la DFT. D'après Ziegler les fonctionnels d'énergie d'échange-corrélation ont été classifiés en trois générations.

#### II.1.4.1 LDA ou approximation de la densité locale

L'approximation de la densité locale LDA (de l'anglais Local Density Approximation) proposée par Kohn et Sham, en 1965 [8] est basée sur le modèle du gaz uniforme d'électrons. L'énergie d'échange-corrélation est définie de la manière suivante :

$$E_{xc}^{LDA}[p] = \int_{xc}^g(p) dr \quad (II-7)$$

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

$$\epsilon_{xc} = \epsilon_x + \epsilon_c$$

On peut utiliser l'énergie d'échange proposée par Dirac [1] comme approximation de  $\epsilon_x$  :

$$K_D[p(r)]S_x = \int p(r)^{3/4} \quad (II-8)$$

La fonctionnelle de corrélation la plus utilisée a été développée par Vosko, Wilk et Nusair en 1980 [9]. Ces auteurs ont utilisé les résultats de calculs Monte Carlo effectués par Ceperley et Alder, pour ajuster une expression analytique de l'énergie de corrélation. Cette fonctionnelle est connue sous l'abréviation « VWN ».

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

électronique ainsi que son gradient. Alors que Le  $X\alpha$  une approximation, déjà citée, entre dans cette catégorie, depuis la corrélation est rapproché par un pourcentage donné de l'énergie d'échange[1].

#### II.1.4.2 GGA ou approximation des gradients généralisée

Cette deuxième génération de fonctionnelle introduit donc une information sur le gradient de la densité. La première amélioration que l'on puisse apporter à la méthode LDA consiste donc à exprimer la fonctionnelle d'énergie d'échange-corrélation en fonction de la densité électronique et de son gradient. Cette technique est appelée « approximation de l'expansion du gradient » (GEA).

Elle se révèle efficace pour les systèmes dont la densité électronique ne varie que lentement. Pour les systèmes chimiques, il s'avère qu'elle donne des résultats moins bons que LDA. La solution consiste alors à réécrire l'expression d'échange-corrélation sous une forme similaire à LDA :

$$E_{XC}^{LDA}[p] = \int f[p(r), \nabla p(r)] p(r) dr \quad (II-9)$$

De nombreuses fonctionnelles ont été développées depuis, tant pour l'échange que pour la corrélation. Parmi les plus connues et les plus utilisées on peut citer les fonctionnelles d'échange de Becke (B88) [10]. Pour la corrélation, on a les fonctionnelles de Perdew (P86) [11], de Lee, Yang et Parr (LYP) [12] et de Perdew et Wang (PW91) [13,1].

La troisième génération des fonctionnels d'énergie d'échange-corrélation est la génération des fonctionnels « au-delà du GGA ». Les plus importants est les fonctionnels hybrides.

Les fonctionnels hybrides sont des fonctionnels d'énergie qui contiennent les deux types : l'échange de DFT (c-à-d. une pièce de LDA et une pièce de GGA) et un type de Hartree-Fock d'échange calculé à partir des orbitales. L'utilisation de la partie d'échange HF associée aux fonctionnelles GGA fournit des résultats comparables à ceux de l'approximation des gradients généralisés. La première fonctionnelle de ce type a été proposée par Becke, et des gradients généralisés, et contient 50 % d'échange HF, la fonctionnelle de ce type actuellement la plus utilisée est celle connue sous le nom B3LYP [1, 14]. Pour nos calculs on a utilisé les deux fonctionnels GGA : BP86 (Becke (B88) Perdew (P86)) et la PW91 (Perdew Wang 91).

#### II.1.5 Succès et limites de la DFT

La DFT permet souvent d'obtenir, à plus faible coût, des résultats d'une précision proche de celle obtenue avec des calculs post-Hartree-Fock. De plus, elle peut être utilisée pour étudier des systèmes relativement gros, contenant jusqu'à plusieurs centaines d'électrons, que les calculs

post-Hartree-Fock ne peuvent pas traiter. Ceci explique que la DFT soit aujourd'hui très utilisée pour étudier les propriétés des systèmes moléculaires ou même biologiques [15].

Les nombreux travaux effectués ces dernières années montrent que les calculs DFT donnent de bons résultats sur des systèmes très divers (métalliques, ioniques, organométalliques, ...) pour de nombreuses propriétés (les structures moléculaires, les fréquences de vibration, les potentiels d'ionisation, ...), nous citons ici à titre d'exemple la revue de H. Chermette pour plus de détails[3].

Toutefois, ces méthodes souffrent encore de plusieurs défauts. Et par ailleurs, on ne comprend pas toujours les bons ou les mauvais résultats de la DFT sur certains systèmes, et il n'existe aucun véritable critère pour choisir une fonctionnelle plutôt qu'une autre. Il est en outre difficile de trouver des critères permettant d'améliorer un fonctionnelle donné, ce qui rend parfois l'utilisation de la DFT délicate. De plus, les états excités ne sont pas accessibles dans le formalisme développé. Cependant, des développements récents en utilisant un formalisme dépendant du temps pour une description des états excités ont été faits [16].

## II.2 Etude quantitative des relations structure-activité (QSAR)

### II.2.1 Introduction

Les premiers essais de modélisation des relations structure-activité ont commencé à la fin du 19<sup>ème</sup> siècle, lorsque *Crum-Brown* et *Frazer* ont postulé que l'activité biologique d'une molécule est une fonction de sa constitution chimique [17-19]. Mais ce n'est qu'au début des années 60 que les travaux de *Corwin Hanch* ont proposé un modèle mathématique reliant l'activité biologique à la structure chimique [20].

Aujourd'hui, l'utilisation de QSAR n'a cessé de progresser [21]. Elle est devenue indispensable en chimie pharmaceutique et pour la conception de médicaments, notamment dans le cas où la disponibilité des échantillons est limitée ou les mesures expérimentales sont dangereuses, longues et chères [22,23].

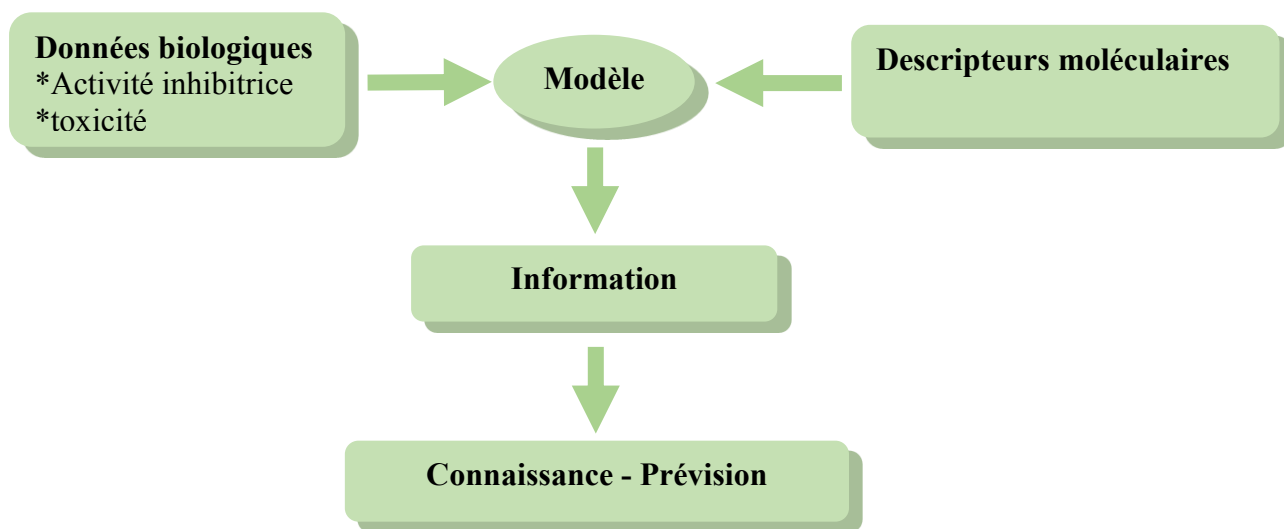
Sans l'utilisation de grands instruments analytiques, les résultats des études QSAR peuvent fournir des informations utiles pour obtenir une meilleure connaissance des structures moléculaires et probablement le mode d'action au niveau moléculaire. Ces informations peuvent être alors utilisées dans la prédiction des activités biologiques de nouveaux composés ainsi que dans la conception de nouvelles structures [24].

### II.2.2 Principe de QSAR

Le principe des méthodes QSAR est, comme leur nom l'indique, de mettre en place une relation mathématique reliant de manière quantitative la structure moléculaire, codée par des propriétés moléculaires appelées descripteurs, avec une activité en utilisant des méthodes d'analyse de données [25]. Par ces relations nous pouvons développer des modèles prédictifs de la forme générale suivante :

$$\text{Activité} = f(\text{descripteurs moléculaires})$$

L'objectif de ces méthodes est alors d'analyser les données structurales afin de déterminer les facteurs influençant l'activité mesurée [26].



**Figure 1 :** Présentation de la méthodologie de QSAR.

En pratique, le développement d'un modèle commence par la collecte de données expérimentales fiables et en nombre important, ensuite il est nécessaire de générer un nombre de descripteurs, caractérisant les structures moléculaires, utilisés dans le développement du modèle QSAR [27]. Une fois le modèle construit, il doit être évalué par des méthodes de validation (validations interne et externe) afin d'estimer sa robustesse et son pouvoir prédictif [28-30]. En fin, pour tout modèle, il est important de connaître quel type de molécules ce modèle est utilisable (connaître le domaine d'applicabilité) [31].

Il y a plusieurs raisons pratiques qui justifient l'utilisation des études QSAR [22].

- ✓ Pour prédire les propriétés de l'activité biologique et physico-chimiques par des moyens rationnels.
- ✓ Economie sur le coût de développement de produits (exemple de l'industrie pharmaceutique, des pesticides, des détergents,...etc.).
- ✓ Les prévisions pourraient réduire l'exigence de tests longs et coûteux pratiqués sur les animaux.

### II.2.3 Application de QSAR

Il existe un grand nombre d'applications des modèles QSPR et QSAR en milieu industriel, dans le domaine de la recherche universitaire, en économie, en prévision météorologique... etc [32]. Nous rapportons ci-dessous quelques applications possibles des modèles QSAR [33-35] :

- ✓ L'identification rationnelle de nouvelles pistes avec :
- ✓ l'optimisation de l'activité pharmacologique, biocide ou pesticide.

- ✓ la conception rationnelle de nombreux produits tels que des agents tensioactifs, des parfums, des colorants et des produits chimiques fins.
- ✓ l'identification des composés dangereux aux premiers stades de développement.
- ✓ la prédiction de la toxicité et les effets secondaires de nouveaux composés.
- ✓ la sélection des composés ayant des propriétés pharmacocinétiques optimales, que ce soit la stabilité ou la disponibilité dans les systèmes biologiques.

## II.2.4 Outils et Méthodologie de QSAR

### II.2.4.1 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.2 [36].

**Tableau II.1** : Types des données biologiques utilisées dans l'analyse QSAR.

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

### II.2.4.2 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 [37], permettent d'établir une relation entre une propriété et des descripteurs calculés à partir d'une conformation spatiale de la molécule définie.

#### II.2.4.2.1 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 [38]. Certains représentent la surface accessible au solvant (nommée « Connolly Accessible Surface »), le volume de solvant couvert par cette surface (« Connolly Solvent- Excluded Volume ») [39], le caractère hydrophile ou lipophile de la molécule généralement évalué à partir du coefficient de partage *octanol/eau* représenté par le *logP* [40].

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

#### II.2.4.2.2 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) [38].

#### II.2.4.2.3 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 [41].

### II.2.4.3 Modèles statistiques

L'objectif de l'analyse statistique est justement de « démêler » ces descripteurs et d'identifier ceux qui sont corrélés à la variable cible, qui produisent du signal, de ceux qui ne le sont pas, qui produisent du bruit. L'analyse statistique permet également d'identifier les descripteurs qui sont corrélés entre eux pour ne garder que les principaux et réduire ainsi la redondance d'informations [42].

Dans la terminologie des statistiques, la propriété étudiée est appelée variable cible ou variable dépendante ou encore variable Y. C'est la caractéristique que l'on cherche à expliquer et à optimiser.

Les descripteurs, en l'occurrence chimiques, sont appelés variables explicatives, variables indépendantes ou encore variables X.

L'analyse statistique détermine et quantifie les corrélations entre les descripteurs et la variable cible. Elle indique également la contribution relative de chaque descripteur dans l'explication globale de l'activité. Le modèle statistique est une équation donnant la valeur de la variable cible en fonction de la somme des valeurs pondérées des descripteurs. Les principaux outils statistiques pour obtenir un modèle sont :

- ✓ la régression linéaire multivariée (Multivariate Linear Regression - MLR) ;
- ✓ la régression en composante principale (Principale Component Regression - PCR) ;
- ✓ la régression des moindres carrés partiels (Partial Least Squares - PLS);
- ✓ les réseaux de neurones artificiels (Artificial Neural Network - ANN).

### II.2.5 Régression linéaire multiple

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 (II-10)

$$Y = Y_0 + a_1X_1 + a_2X_2 + \dots + a_nX_n \quad (\text{II-10})$$

On suppose que toutes les variables explicatives sont indépendantes les unes des autres [43]

#### II.2.5.1 Description de la méthode

La régression multilinéaire est la méthode la plus simple et la plus employée pour le développement de modèles prédictifs [44]. 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 :

$$Y = \beta_0 + \beta_1X_1 + \beta_2X_2 + \beta_3X_3 + \dots + \beta_pX_p + \varepsilon \quad (\text{II-11})$$



Où  $\beta_i$  sont les coefficients de la régression.

La détermination de l'équation (II-11) 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.

$$\begin{aligned}
 Y_1 &= \beta_0 + \beta_1 X_{1,1} + \beta_2 X_{2,1} + \dots + \beta_p X_{n,1} + \varepsilon_1 \\
 &\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\
 &\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\
 &\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\
 Y_2 &= \beta_0 + \beta_1 X_{1,2} + \beta_2 X_{2,2} + \dots + \beta_p X_{n,2} + \varepsilon_2 \\
 &\dots \\
 Y_n &= \beta_0 + \beta_1 X_{1,p} + \beta_2 X_{2,p} + \dots + \beta_p X_{n,p} + \varepsilon_n
 \end{aligned}$$

Où les résidus  $\varepsilon_i$  représentent l'erreur du modèle, constituée par l'incertitude sur la variable dépendante  $Y_i$  d'une part, sur les variables indépendantes  $X_i$  d'autre part, mais aussi par les informations contenues dans les variables indépendantes mais non exprimées via les variables dépendantes. Ce système d'équations peut être écrit sous la forme matricielle suivante :

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

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  $Y = X_b + \varepsilon$ .

La méthode consiste alors à choisir les coefficients du vecteur  $\beta$  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 supposées indépendants entre eux. Enfin, par nature, la dépendance de Y vis-à-vis des Xi est supposée linéaire. La valeur prédite de la variable dépendante est alors :

$$\hat{Y} = \beta_0 + \hat{\beta}_1 X_{1,i} + \dots + \beta_p X_{n,i} \quad (II-12)$$

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

$$\varepsilon = Y_i - \hat{Y}_i \quad (II-13)$$

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 [45].

### II.2.5.2 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.

$$R^2 = \frac{ESS}{TSS} = \frac{TSS - RSS}{TSS} = 1 - \frac{RSS}{TSS} \quad (II-14)$$

Où :

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  $0 < R < 1$ .

$$R = (ESS/TSS)^{0.5} = (1 - RSS/TSS)^{0.5} \quad (II-15)$$

#### c. Le Test de Fisher-Snedecor (F)

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.

$$F = [ESS/(K)]/[RSS/(n - K - 1)] \quad (II-16)$$

Où  $n$ , et  $(k-1)$  sont des degrés de liberté associés à ESS et RSS respectivement [43].

**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 :

$$S = \sqrt{\frac{RSS}{n-p-1}} \quad (II-17)$$

Dont  $p$  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 :

$$Q^2 = 1 - \frac{PRESS}{SSY} \quad (II-18)$$

**II.2.5.3 Validations des modèles**

Quel que soit le but ultime du modèle QSAR, il doit être obligatoirement validé avant d'être interprété ou utilisé à fins prédictives. Il existe différentes solutions pour s'assurer de validité d'un modèle. La validation croisée *LOO* (*leave-one-out*) est un processus qui permet de tester la précision prédictive d'un modèle, cette méthode basée sur le calcul de quelques paramètres statistiques tels que : la somme des carrés résiduels *RSS* (*PRESS*), Somme des carrés totaux *TSS*, qualité d'ajustement  $R_{adj}^2$ , coefficient de corrélation de validation croisée  $R_{CV}^2$ , la validation standard des erreurs de prédiction ( $S_{PRESS}$ ) et l'erreur de prédiction (*PE*). Ces paramètres statistiques sont calculés à partir des relations suivantes [45] :

$$PRESS = \sum(Y_{obs} - Y_{cal})^2 \quad (II-19)$$

$$TSS = \sum(Y_{obs} - \bar{Y})^2 \quad (II-20)$$

$$R_{adj}^2 = (1 - r^2) \left( \frac{n-1}{n-p-1} \right) \quad (II-21)$$

$$R_{CV}^2 = 1 - \frac{PRESS}{TSS} \quad (II-22)$$

$$S_{PRESS} = \sqrt{\frac{PRESS}{n}} \quad (II-23)$$

$$PE = 0.6745(1 - r^2)/\sqrt{n} \quad (II-24)$$

**II.2.6 Domaine d'applicabilité**

Le domaine d'applicabilité est la région de l'espace chimique définie par les molécules de l'ensemble d'apprentissage du modèle. Les modèles QSAR ne peuvent pas prédire des propriétés de manière fiable pour l'intégralité des composés chimiques existants. En effet, un modèle QSAR

n'est pas destiné à être employé en dehors de son domaine d'applicabilité, c'est-à-dire en dehors de l'espace chimique couvert par son ensemble d'apprentissage [46].

## II.3 L'activité biologique

### II.3.1 Découverte des médicaments

En général, une molécule est d'abord repérée pour son potentiel thérapeutique puis optimiser par modification structurale, le test retenu devra être simple rapide et fiable, de sorte que les tests devront être exécutés ; soit *in vitro* soit *in vivo*, en général les tests *in vitro* sont choisis de préférence aux tests *in vivo* [47].

### II.3.2 Les différentes cibles des médicaments

Le devenir des médicaments dans l'organisme à partir de la voie d'administration ciblé des récepteurs spécifiques tel que :

- ✓ Des récepteurs. Ex : les anti-sécrétoires gastriques sont des antagonistes des récepteurs histaminiques H<sub>2</sub>.
- ✓ Des enzymes. Ex : inhibition de l'enzyme de conversion de l'angiotensine par les IEC.
- ✓ Des protéines de transport : elles permettent le transport des ions et petites molécules à travers les membranes cellulaires (transport du glucose, des ions Na<sup>+</sup>...). Ex : inhibition de la Na<sup>+</sup>/K<sup>+</sup> ATPase par la digoxine, inhibition de la H<sup>+</sup>/K<sup>+</sup> ATPase (dite pompe à protons) par les inhibiteurs de la pompe à protons tels que l'oméprazole [48].

### II.3.3 Les tests biologiques

#### II.3.3.a Tests *in vivo*

L'exécution de tests *in vivo* implique l'induction d'un état pathologique chez l'animal en vue d'y créer des symptômes observables puis traités par des médicaments potentiels et on examine l'amélioration des systèmes, bref si le médicament semblé intéresse on vérifiait lequel des médicaments à l'essai guérissait le plus efficace.

Dans les tests *in vivo* on peut même remplacer des gènes de source par des gènes humains. Plusieurs problèmes sont toutefois associés aux tests *in vivo* comme par exemple la souffrance chez les animaux testés, de surcroît, diverses questions épineuses de pharmacocinétique émergée ainsi dans les résultats négatifs, comment pourrait-on savoir si cela est dû au fait que le médicament est incapable de se fixer sur sa cible ou s'il n'a tout simplement réussi à l'attendre.

#### II.3.3.b Tests *in vitro*

Les tests *in vitro* permettant d'éviter de faire appel à des animaux vivants à leur place, on emploie des ensembles de cellules (tissus), des cellules isolées ou des enzymes particulières.

L'affinité des médicaments vis-à-vis de leur récepteur (c'est-à-dire la mesure de la force de leur ancrage) peut être établie à l'aide de radio-ligands. Dans ce cas, le tissu est mis en contact avec le médicament à tester.

En vue de donner ce dernier la possibilité de se lier au récepteur qu'il est censé cibler. Comme ça on peut avoir combien de récepteurs sont restés libres c'est-à-dire n'ont pas été occupés par le médicament testé [47].

### II.3.4 Etude de la liaison ligand-récepteur

La liaison du ligand au récepteur est une liaison spécifique qui déclenche un effet biologique ou au contraire bloque cet effet.

La liaison ligand-récepteur est une réaction réversible



Il existe deux types de méthodes d'étude de liaison ligand-récepteur [49] :

- ❖ A-méthode de saturation.
- ❖ B-méthode de déplacement : cette méthode permet de déterminer IC50.

**Références bibliographiques**

- [1]. C. Corminboeuf, F. Tran, J. Weber, *Journal of Molecular Structure: Theochem.* 762, (2006).
- [2]. P. Hohenberg, W. Kohn, *Phys. Rev.* 136, (1964).
- [3]. H. Chermette, *Coordination Chemistry Reviews* 699 :178-180, (1998).
- [4]. M. Roger, L. Belkhir, P. Thuéry, Thérèse Arliguie, M. Fourmigué, A. Boucekkine, M. Ephritikhine *Organometallics* 24 : 4941, (2005).
- [5]. W. Kohn, *Rev. Mod. Phys.* 71:1253, (1999).
- [6]. J. A. Pople, *Rev. Mod. Phys.* 71:1267, (1999).
- [7]. P. Hohenberg, W. Kohn: *Phys. Rev. B* 136 : 864, (1969).
- [8]. W. Kohn and L. J. Sham, *Phys. Rev.* 140, A1133, (1965).
- [9]. S. H. Vosko, L. Wilk and M. Nusair, *Can. J. Phys.* 58 : 1200, (1980).
- [10]. A. D. Becke, *Phys. Rev.* 38 : 3098, (1988).
- [11]. J. P. Perdew, *Phys. Rev. B*, 33 : 8822, (1986).
- [12]. C. Lee, W. Yang and R. G. Parr, *Phys. Rev. B*, 37 :785, (1988).
- [13]. J. P. Perdew, K. Burke, Y. Wang, *Phys. Rev. B*, 54 : 16533, (1996).
- [14]. A. D. Becke, *J. Chem. Phys.* 98 : 5648, (1993).
- [15]. F. Rabilloud, thèse de doctorat, Etude théorique et expérimentale des agrégats de bromure d'argent, Université de Paul SABATIER TOULOUSE III, (2000).
- [16]. M. van Faassen, K. Burke, *Chemical Physics Letters* 431, 410, (2006).
- [17]. A. Crum Brown, T. R. Fraser, *Trans. Roy. Soc.* 25, 151-203, (1868).
- [18]. Hansch, *Acc. Chem. Res.* 2, 232-239, (1969).
- [19]. Hansch, T. Fujita, *J. Am. Chem. Soc.* 86, 1616-1626, (1964).
- [20]. A. K. Debnath, *Mini Reviews in Medicinal Chemistry*. I, 187-195, (2001).
- [21]. M. Grover, B. Singh, M. Bakshi, S. Singh, *Pharm. Sci. Tech. Today*. 3, 50-57, (2000).
- [22]. T. I. Oprea, "Chemoinformatics in Drug Discovery" Ed. WILEY-VCH Verlag. Allemagne, 2005.
- [23]. E. A. Rekká ; P. N. Kourounakis "Chemistry and Molecular Aspects of Drug Design and Action" Ed. Taylor & Francis Group, LLC. Etats Unies, 2008.
- [24]. A. K. Ghose, V. N. Viswanadhan, "Combinatorial Library Design and Evaluation: Principles, Software Tools, and Applications in Drug Discovery" Ed. Marcel Dekker, Inc. Etats Unies, (2001).

- [25]. F. Li, H. Wu, L. Li, X. Li, J. Zhao, W. J. G. M. Peijnenburg, *Ecotoxicol. Environ. Saf.* 80, 273–279, (2012).
- [26]. A. Srinivas Reddy, S. Kumar, R. Garg, *J Mol Graph Model.* 28(8), 852–862, (2010).
- [27]. M. Zhao, Z. Li, L. Peng, Y. R. Tang, C. Wang, Z. Zhang, S. Peng, *Bioorg. Med. Chem.* 15, 2815–2826, (2007).
- [28]. A. A. Toropov, A. P. Toropova, E. Benfenati, G. Gini, D. Leszczynska, J. Leszczynski, G. De Nucci, *Biochem. Biophys. Res. Comm.* 432(2), 214–225, (2013)
- [29]. A. Golbraikh, A. Tropsha, *J. Mol. Graph. Model.* 20, 269-276, (2002).
- [30]. A. Tropsha, P. Gramatica, K.V. Gombar, *QSAR Comb. Sci.* 22, 69-77, (2003).
- [31]. P. Gramatica, E. Giani, E. Papa, *J. Mol. Graph. Mod.* 25, 755-766, (2007).
- [32]. T. Puzyn, J. Leszczynski, M. T. D. Cronin, « Recent advances in QSAR studies Methods and applications » Springer Dordrecht Heidelberg London New York, (2010).
- [33]. M. Tichý, *Science.Tot. Env.* 109, 407-410, 1991
- [34]. B. Peric, J. Sierra, E. Martí, R. Cruañas, M. A. Garau, *Ecotoxicol. Environ. Saf.* 115, 257-262, (2015).
- [35]. O. M. Bautista-Aguilera, G. Esteban, I. Bolea, K. Nikolic, D. Agbaba, I. Moraleda, I. Iriepa, A. Samadi, E. Soriano, M. Unzeta, J. Marco-Contelles, *Europ. J. Med. Chem.* 75, 82-95, (2014).
- [36]. 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).
- [37]. 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).
- [38]. R .Bosque, J.Sales, E.Bosch, M. Rosès, Garcia-Alvarez-Coque M.C., and TorresLapasio J.R. A QSPR study of the p solute polarity parameter to estimate retention in HPLC. *J. Chem. Inf. Comput. Sci.* 43 : 1240-1247, (2003).
- [39]. M.L. Connolly , Computation of molecular volume. *J. Am. Chem. Soc.* 107 : 118- 1124, (1985).
- [40]. 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).
- [41]. H.P. Schultz ,*Topological organic chemistry. 1. Graph theory and topological indices of alkanes.* *J. Chem. Inf. Comput. Sci.* 29 : 227-228, (1989).
- [42]. A. Fortuné, Thèse de doctorat, *Techniques de Modélisation Moléculaire Appliquées à l'Etude et à l'Optimisation de Molécules Immunogènes et de Modulateurs de la Chimiorésistance, Université Joseph Fourier – Grenoble I,* (2006).
- [43]. Z. Almi, Thèse de doctorat, *Etude qualitative et quantitative des relations structures-activités dans des hétérocycles à intérêt pharmaceutique, Université Mohamed Khider-Biskra,* (2015).
- [44]. I. Charif, *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,* Thèse de doctorat, Université de Tlemcen, (2012).
- [45]. N. Melkemi, Thèse de doctorat, *Etude des corrélations entre les paramètres structuraux et les propriétés bioactives de quelques composés organiques, Université Mohamed Khider-Biskra,* (2013).
- [46]. M. SAIHI Youcef, Thèse de doctorat, *Etude de la relation quantitative structure-activité inhibitrice des enzymes hydrolytiques : cas des alpha-glucosidases, Université Badji Mokhtar – Annaba,*( 2015).
- [47]. L. Patrick Graham, *Chimie pharmaceutique, De Doeck Supérieur,* 2002.
- [48]. C. Loichot, *Pharmacologie générale DCEM1, Mécanismes d'action des médicaments, Faculté de Médecine de Strasbourg,* (2006).
- [49]. Ariane-Valérie Faure, Martine Fontaine, Pascale Jolliet, *Pharmacologie soins infirmiers, Elsevier Masson,* 2007.

# *Chapitre III*

---

## *Résultats et discussion*

## **III.1 Etude quantitative des relations structure-activité (QSAR) des propriétés cytotoxiques d'une série de dérivés de wogonine**

### **III.1.1 Introduction**

La Relation Quantitative Structure et Activité (QSAR) est un outil 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, propriété physico-chimique, etc.), pour une série des molécules biologiquement actives à l'aide des méthodes d'analyse de données [1].

Dans ce travail, nous nous intéressons à l'étude des propriétés physico-chimiques et électroniques des dérivés de wogonine qui permettent de prévoir ces paramètres, qui influent sur l'activité biologique, afin de prédire l'activité biologique de nouvelles molécules. C'est pour cette raison, nous avons choisis une série de quarante-huit dérivés des différents descripteurs physico-chimiques et électroniques.

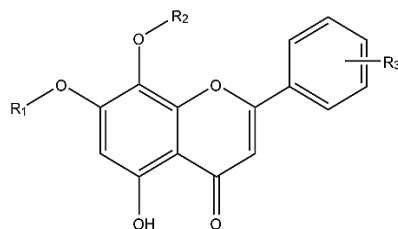
### **III.1.2 Méthodologie**

#### **III.1.2.1 Base de données**

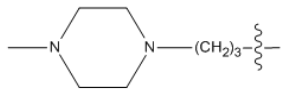

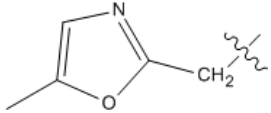
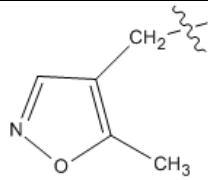
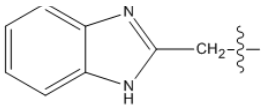


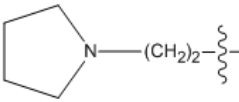
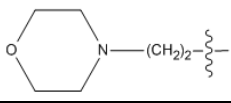
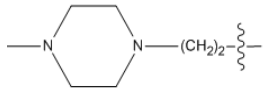


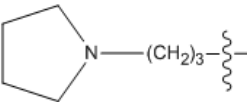
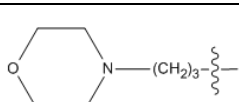
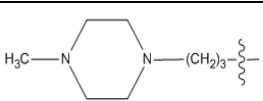

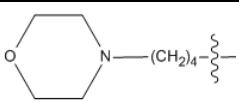
Une série de 48 dérivés de wogonine a été utilisée pour développer des modèles QSAR. Les structures et l'activité biologique expérimentale des dérivés de wogonine sont rapportées dans **le tableau III.1**. L'activité cytotoxique a été exprimée par  $pIC_{50}$ .

Avant de procéder l'étude QSAR, nous avons sélectionné l'ensemble de molécules (data set) en deux classes. La première classe comporte 35 molécules nommées « training set » utilisés pour développer des modèles QSAR. La seconde classe comporte 13 molécules nommées « test set » utilisés dans la validation des modèles QSAR obtenus des molécules de training set.

**Tableau III.1** : Structures chimique et activité biologique expérimentale des molécules utilisées dans l'étude QSAR.



N°	R <sub>1</sub>	R <sub>2</sub>	R <sub>3</sub>	pIC <sub>50</sub> (exp)
1 <sup>a,b</sup>	(CH <sub>3</sub> CH <sub>2</sub> ) <sub>2</sub> N(CH <sub>2</sub> ) <sub>4</sub> -	H	CH <sub>3</sub>	5.76
2	(HOCH <sub>2</sub> CH <sub>2</sub> ) <sub>2</sub> N(CH <sub>2</sub> ) <sub>4</sub> -	H	CH <sub>3</sub>	5.58
3	N-(CH <sub>2</sub> ) <sub>4</sub> -	H	CH <sub>3</sub>	5.44
4 <sup>a</sup>	N-(CH <sub>2</sub> ) <sub>4</sub> -	H	CH <sub>3</sub>	5.71
5	N-(CH <sub>2</sub> ) <sub>4</sub> -	H	CH <sub>3</sub>	5.48
6	N-(CH <sub>2</sub> ) <sub>4</sub> -	H	CH <sub>3</sub>	5.95
7	N-(CH <sub>2</sub> ) <sub>3</sub> -	H	CH <sub>3</sub>	5.04
8 <sup>a,b</sup>	N-(CH <sub>2</sub> ) <sub>4</sub> -	H	CH <sub>3</sub>	5.97
9	PhCH <sub>2</sub> NH(CH <sub>2</sub> ) <sub>4</sub> -	H	CH <sub>3</sub>	4.34
10	(CH <sub>3</sub> CH <sub>2</sub> ) <sub>2</sub> N(CH <sub>2</sub> ) <sub>3</sub> -	H	CH <sub>3</sub>	5.48
11	(HOCH <sub>2</sub> CH <sub>2</sub> ) <sub>2</sub> N(CH <sub>2</sub> ) <sub>3</sub> -	H	CH <sub>3</sub>	5.42
12	N-(CH <sub>2</sub> ) <sub>3</sub> -	H	CH <sub>3</sub>	4.81
13 <sup>a,b</sup>	N-(CH <sub>2</sub> ) <sub>3</sub> -	H	CH <sub>3</sub>	5.19

14		H	CH <sub>3</sub>	4.50
15	PhCH <sub>2</sub> NH(CH <sub>2</sub> ) <sub>3</sub> - 	H	CH <sub>3</sub>	5.09
16 <sup>a</sup>		H	CH <sub>3</sub>	4.69
17		H	CH <sub>3</sub>	4.66
18		H	CH <sub>3</sub>	4.53
19	H	H	(CH <sub>3</sub> ) <sub>2</sub> N(CH <sub>2</sub> ) <sub>2</sub> - 	4.40
20	H	H	(CH <sub>3</sub> CH <sub>2</sub> ) <sub>2</sub> N(CH <sub>2</sub> ) <sub>2</sub> - 	4.17
21 <sup>b</sup>	H	H		4.16
22	H	H		4.40
23 <sup>a,b</sup>	H	H		4.25
24	H	H	(CH <sub>3</sub> ) <sub>2</sub> N(CH <sub>2</sub> ) <sub>3</sub> - 	4.38
25	H	H	(CH <sub>3</sub> CH <sub>2</sub> ) <sub>2</sub> N(CH <sub>2</sub> ) <sub>3</sub> - 	4.17
26	H	H		4.23
27	H	H		4.68
28 <sup>a,b</sup>	H	H	H <sub>3</sub> C-N 	4.37
29 <sup>a,b</sup>	H	H	(CH <sub>3</sub> ) <sub>2</sub> N(CH <sub>2</sub> ) <sub>4</sub> - 	4.45
30	H	H		4.82

31	H	H		4.82
32 <sup>a,b</sup>	H	H	$\text{CH}_3\text{NH}(\text{CH}_2)_2$ -	4.19
33 <sup>a,b</sup>	H	H	$\text{CH}_3(\text{CH}_2)_2\text{NH}(\text{CH}_2)_2$ -	4.01
34	H	H	$(\text{CH}_3)_2\text{CHNH}(\text{CH}_2)_2$ -	4.85
35 <sup>a,b</sup>	H	H	$\text{CH}_3(\text{CH}_2)_3\text{NH}(\text{CH}_2)_2$ -	4.24
36	H	H	$\text{HO}(\text{CH}_2)_2\text{NH}(\text{CH}_2)_2$ -	4.24
37	H	H	$\text{HOCH}_2\text{CH}(\text{CH}_3)\text{NH}(\text{CH}_2)_2$ -	4.36
38 <sup>b</sup>	H	H	$\text{CH}_3\text{NH}(\text{CH}_2)_4$ -	4.01
39	H	H	$\text{CH}_3(\text{CH}_2)_2\text{NH}(\text{CH}_2)_4$ -	5.00
40	H	H	$(\text{CH}_3)_2\text{CHNH}(\text{CH}_2)_4$ -	4.52
41 <sup>a,b</sup>	H	H	$\text{CH}_3(\text{CH}_2)_3\text{NH}(\text{CH}_2)_4$ -	4.41
42	H	H		4.96
43	H	4'-F	$\text{CH}_3$	4.36
44 <sup>a,b</sup>	H	3'-F	$\text{CH}_3$	4.23
45	H	4'-OCH <sub>3</sub>	$\text{CH}_3$	4.37
46	H	4'-CH <sub>3</sub>	$\text{CH}_3$	4.60
47	H	4'-Cl	$\text{CH}_3$	4.47
48	H	H	$\text{CH}_3$	4.72

a : molécule test set dans la phase gazeuse

b : molécule test set dans la phase aqueuse

### III.1.2.2 Sélection des descripteurs et méthodes de calcul

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 [2].

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

Les molécules de la série de wogonines ont été pré-optimisées au moyen de la mécanique moléculaire, champ de force (MM+) en utilisant le logiciel HyperChem 8.07 [3]. Les structures minimisées obtenus précédemment ont été optimisées en utilisant la méthode DFT avec une fonctionnelle hybride B3LYP [4] et la base 6-31G (d, p) de logiciel Gaussien 09[5].

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 (qO1, qO2, qO3, qO4, qO5) (tableau III.3 et III.4).

Le module "propriétés QSAR" de l'HyperChem 8,07 [3] a été utilisé pour calculer les paramètres suivants : le coefficient de partage octanol/eau (logP), le volume moléculaire (V), la surface moléculaire(SAG) et en utilise le molinspiration [6] pour calculé la liaison hydrogène donneurs (HBD) et la liaison hydrogène accepteurs (HBA) (tableau III.2).

**Tableau III.2 :** Valeurs des descripteurs physicochimique utilisés dans l'étude QSAR.

N° de composé	V (Å <sup>3</sup> )	SAG (Å <sup>2</sup> )	Log P	HBD	HBA
1	1205.92	707.22	-0.86	1	6
2	1253.64	739.75	-2.43	3	8
3	1157.83	671.23	-1.22	1	6
4	1196.65	688.82	-0.82	1	6
5	1178.93	676.74	-1.88	1	7
6	1228.99	702.71	-1.74	1	7
7	1135.62	656.83	-2.55	2	7
8	1182.36	684.07	-2.10	2	7
9	1180.83	648.41	-0.89	2	6
10	1153.60	680.44	-1.31	1	6

<b>11</b>	1201.13	711.37	-2.88	3	8
<b>12</b>	1108.92	639.48	-1.67	1	6
<b>13</b>	1125.38	651.05	-2.34	1	7
<b>14</b>	1187.91	676.85	-2.19	1	7
<b>15</b>	1119.51	613.66	-1.34	2	6
<b>16</b>	997.52	569.34	-2.88	1	7
<b>17</b>	977.73	557.24	-2.59	1	7
<b>18</b>	1076.65	612.91	-2.82	2	7
<b>19</b>	921.62	536.96	-2.08	1	6
<b>20</b>	1010.15	573.55	-1.39	1	6
<b>21</b>	983.04	567.02	-1.75	1	6
<b>22</b>	994.66	572.26	-2.42	1	7
<b>23</b>	1051.07	600.47	-2.27	1	7
<b>24</b>	980.58	570.38	-2.02	1	6
<b>25</b>	1079.65	620.64	-1.34	1	6
<b>26</b>	1037.69	602.18	-1.70	1	6
<b>27</b>	1060.43	610.13	-2.37	1	7
<b>28</b>	1111.09	635.51	-2.22	1	7
<b>29</b>	1031.60	594.85	-1.57	1	6
<b>30</b>	1076.04	591.67	-1.92	1	7
<b>31</b>	1133.37	625.62	-1.77	1	7
<b>32</b>	885.05	524.79	-2.44	2	6
<b>33</b>	996.11	584.76	-1.63	2	6
<b>34</b>	983.39	575.12	-1.68	2	6
<b>35</b>	1049.90	615.69	-1.23	2	6
<b>36</b>	963.79	568.35	-2.88	3	7
<b>37</b>	1005.19	589.16	-2.47	3	7
<b>38</b>	993.19	583.81	-1.93	2	6
<b>39</b>	1103.24	648.61	-1.12	2	6
<b>40</b>	1094.26	637.49	-1.18	2	6
<b>41</b>	1158.23	678.27	-0.73	2	6
<b>42</b>	1196.07	683.75	-0.35	2	6



<b>43</b>	755.15	457.64	-2.66	2	5
<b>44</b>	755.24	458.44	-2.66	2	5
<b>45</b>	824.63	494.19	-3.05	2	6
<b>46</b>	799.88	477.14	-1.90	2	5
<b>47</b>	790.77	476.51	-2.28	2	5
<b>48</b>	739.33	445.54	-2.06	2	5

**Tableau III.3 :** Valeurs des descripteurs quantique utilisés dans l'étude QSAR dans la phase gazeuse.

N° de composé	MD (Debyes)	qO1	qO2	qO3	qO4	qO5
<b>1</b>	4.8562	-0.201	-0.521	-0.491	-0.408	-0.286
<b>2</b>	6.4436	-0.105	-0.518	-0.500	-0.391	-0.361
<b>3</b>	4.7490	-0.118	-0.519	-0.504	-0.333	-0.334
<b>4</b>	4.7317	-0.135	-0.521	-0.502	-0.321	-0.335
<b>5</b>	4.4074	-0.132	-0.519	-0.505	-0.317	-0.334
<b>6</b>	5.2720	-0.128	-0.518	-0.501	-0.329	-0.347
<b>7</b>	4.6718	-0.177	-0.521	-0.505	-0.382	-0.336
<b>8</b>	5.3751	-0.140	-0.520	-0.504	-0.318	-0.334
<b>9</b>	6.0632	-0.193	-0.528	-0.490	-0.395	-0.280
<b>10</b>	5.5115	-0.108	-0.516	-0.503	-0.384	-0.338
<b>11</b>	3.5248	-0.126	-0.523	-0.506	-0.365	-0.343
<b>12</b>	5.0799	-0.179	-0.526	-0.501	-0.385	-0.348
<b>13</b>	3.5309	-0.175	-0.519	-0.505	-0.379	-0.336
<b>14</b>	4.4510	-0.177	-0.527	-0.497	-0.365	-0.341
<b>15</b>	6.2260	-0.149	-0.520	-0.509	-0.394	-0.333
<b>16</b>	3.4648	-0.208	-0.521	-0.494	-0.284	-0.255
<b>17</b>	4.6613	-0.211	-0.522	-0.503	-0.376	-0.277
<b>18</b>	5.0478	-0.198	-0.512	-0.486	-0.354	-0.286
<b>19</b>	4.2737	-0.196	-0.582	-0.550	-0.450	-0.382
<b>20</b>	4.2292	-0.172	-0.582	-0.558	-0.467	-0.367

<b>21</b>	4.2685	-0.133	-0.580	-0.551	-0.468	-0.334
<b>22</b>	4.3299	-0.174	-0.575	-0.550	-0.500	-0.303
<b>23</b>	3.7864	-0.140	-0.580	-0.548	-0.455	-0.330
<b>24</b>	4.0865	-0.170	-0.588	-0.553	-0.463	-0.353
<b>25</b>	4.1479	-0.180	-0.585	-0.556	-0.469	-0.345
<b>26</b>	4.0350	-0.194	-0.581	-0.545	-0.526	-0.311
<b>27</b>	2.6338	-0.188	-0.580	-0.545	-0.497	-0.313
<b>28</b>	3.9456	-0.195	-0.584	-0.551	-0.514	-0.319
<b>29</b>	4.1569	-0.203	-0.584	-0.553	-0.459	-0.330
<b>30</b>	2.6214	-0.179	-0.579	-0.548	-0.462	-0.316
<b>31</b>	3.8799	-0.145	-0.580	-0.552	-0.482	-0.291
<b>32</b>	4.1172	-0.140	-0.578	-0.548	-0.456	-0.403
<b>33</b>	4.2188	-0.150	-0.581	-0.549	-0.462	-0.410
<b>34</b>	4.3623	-0.177	-0.588	-0.558	-0.457	-0.392
<b>35</b>	4.2501	-0.148	-0.597	-0.544	-0.458	-0.416
<b>36</b>	4.0423	-0.167	-0.583	-0.551	-0.465	-0.402
<b>37</b>	3.6288	-0.188	-0.583	-0.549	-0.476	0.379
<b>38</b>	4.0972	-0.190	-0.586	-0.556	-0.467	-0.351
<b>39</b>	4.1696	-0.201	-0.582	-0.554	-0.465	-0.339
<b>40</b>	4.1001	-0.206	-0.579	-0.556	-0.469	-0.338
<b>41</b>	4.3600	-0.206	-0.585	-0.557	-0.467	-0.345
<b>42</b>	3.3985	-0.180	-0.581	-0.545	-0.503	-0.357
<b>43</b>	5.4105	-0.127	-0.574	-0.547	-0.516	-0.365
<b>44</b>	5.4308	-0.122	-0.578	-0.558	-0.515	-0.357
<b>45</b>	8.2277	-0.134	-0.582	-0.560	-0.517	-0.370
<b>46</b>	7.1951	-0.138	-0.579	-0.556	-0.521	-0.354
<b>47</b>	4.8691	-0.125	-0.572	-0.548	-0.514	-0.359
<b>48</b>	4.5865	-0.146	-0.582	-0.550	-0.533	-0.266

**Tableau III.4 :** Valeurs des descripteurs quantique utilisés dans l'étude QSAR dans la phase aqueuse.

N° de composé	MD (Debyes)	qO1	qO2	qO3	qO4	qO5
1	6.7663	-0.128	-0.604	-0.549	-0.425	-0.373
2	8.7491	-0.142	-0.605	-0.554	-0.410	-0.370
3	6.0276	-0.186	-0.601	-0.552	-0.355	-0.323
4	6.3029	-0.208	-0.603	-0.543	-0.322	-0.316
5	6.2282	-0.204	-0.601	-0.547	-0.336	-0.316
6	6.4830	-0.189	-0.601	-0.548	-0.330	-0.313
7	6.7620	-0.175	-0.606	-0.549	-0.376	-0.375
8	6.5612	-0.201	-0.603	-0.548	-0.311	-0.315
9	7.4839	-0.151	-0.607	-0.555	-0.412	-0.341
10	7.8510	-0.135	-0.606	-0.552	-0.399	-0.355
11	5.2929	-0.124	-0.602	-0.552	-0.384	-0.361
12	6.7944	-0.177	-0.608	-0.535	-0.365	-0.369
13	5.1057	-0.171	-0.605	-0.545	-0.375	-0.372
14	5.8337	-0.165	-0.606	-0.535	-0.353	-0.383
15	7.3314	-0.222	-0.608	-0.545	-0.472	-0.196
16	4.3349	-0.198	-0.601	-0.544	-0.280	-0.320
17	7.7946	-0.182	-0.604	-0.552	-0.310	-0.380
18	2.3373	-0.166	-0.605	-0.530	-0.308	-0.393
19	5.5573	-0.141	-0.631	-0.585	-0.507	-0.428
20	5.7952	-0.173	-0.630	-0.590	-0.505	-0.383
21	5.9049	-0.135	-0.628	-0.583	-0.518	-0.368
22	4.7547	-0.132	-0.627	-0.580	-0.518	-0.371
23	5.1919	-0.130	-0.628	-0.588	-0.508	-0.377
24	5.3924	-0.237	-0.646	-0.597	-0.487	-0.406
25	5.5213	-0.240	-0.638	-0.594	-0.481	-0.403
26	5.3814	-0.192	-0.630	-0.580	-0.540	-0.369
27	3.8358	-0.181	-0.629	-0.587	-0.512	-0.380

28	5.3381	-0.200	-0.635	-0.583	-0.552	-0.355
29	4.7999	-0.151	-0.629	-0.588	-0.541	-0.407
30	4.3162	-0.163	-0.628	-0.584	-0.500	-0.314
31	5.3629	-0.177	-0.633	-0.587	-0.515	-0.328
32	5.8568	-0.148	-0.634	-0.590	-0.478	-0.443
33	5.8934	-0.166	-0.633	-0.586	-0.479	-0.443
34	5.9193	-0.161	-0.635	-0.588	-0.490	-0.442
35	5.9547	-0.165	-0.632	-0.587	-0.479	-0.447
36	5.6585	-0.161	-0.637	-0.589	-0.485	-0.447
37	4.8505	-0.158	-0.634	-0.588	-0.506	-0.431
38	5.8261	-0.183	-0.635	-0.592	-0.495	-0.390
39	5.8558	-0.191	-0.634	-0.590	-0.498	-0.382
40	5.8195	-0.190	-0.634	-0.591	-0.495	-0.373
41	5.8804	-0.173	-0.631	-0.586	-0.496	-0.389
42	4.5881	-0.257	-0.633	-0.577	-0.533	-0.405
43	7.5010	-0.119	-0.627	-0.585	-0.540	-0.386
44	7.4060	-0.114	-0.629	-0.594	-0.540	-0.384
45	11.0632	-0.131	-0.636	-0.596	-0.543	-0.388
46	9.6424	-0.137	-0.633	-0.593	-0.539	-0.379
47	6.7161	-0.120	-0.626	-0.587	-0.541	-0.378
48	6.3959	-0.143	-0.635	-0.588	-0.567	-0.303

### III.1.3 Développement des modèles 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 [7].

Les corrélations entre l'activité biologique et les descripteurs utilisées sont exprimées par les relations mathématiques suivantes :

**1- Dans la phase gazeuse**

$$pIC_{50} = 8.538 - 0.005 V + 0.007 SAG + 0.501 \log P + 0.437 HBA + 4.692 qO1 + 22.273 qO2 - 16.301 qO3 \quad (1)$$

$$N = 35 \quad R = 0.863 \quad R^2 = 0.744 \quad F = 11.204 \quad S = 0.272$$

**2- Dans la phase aqueuse**

$$pIC_{50} = 5.907 - 0.012 V + 0.021 SAG + 0.374 \log P + 0.342 HBA + 0.036 DM + 1.864 qO4 + 5.729 qO5 \quad (2)$$

$$N = 35 \quad R = 0.877 \quad R^2 = 0.769 \quad F = 12.805 \quad S = 0.259$$

$B_0$  (8.538 et 5.907) est l'ordonnée à l'origine ou de la valeur de base de  $pIC_{50}$ .

$R$  : coefficient de corrélation.

$R^2$  : coefficient de détermination.

$S$  : erreur standard de l'estimation.

$F$  : statistique Fischer.

$N$  : indique le nombre de composés.

- Les valeurs de la fraction de la variance peuvent varier entre **0** et **1**. Le modèle **QSAR** doit considérer un  $R^2 > 0.6$  pour qu'il soit valide [8].
- Dans notre modèles QSAR obtenus  $R^2 = 0,744$  et  $R^2 = 0.769$  dans la phase gazeuse et dans phase aqueuse respectivement.
- La valeur  $F$  est jugée la signification statistique au niveau de 95%, pour toutes les valeurs de  $F$  calculées sont supérieures par rapport aux valeurs lus dans la table du Fischer.
- Dans les modèles (1) et (2), le coefficient négatif de  $V$  montre que toute augmentation dans la valeur de ce paramètre entraine une diminution de l'activité biologique. Donc les molécules volumineuses possèdent une faible activité cytotoxique.
- Dans l'équation (1), les coefficients positives de  $SAG$ ,  $\log P$ ,  $HBA$ ,  $qO1$  et  $qO2$  montrent que toute augmentation dans les valeurs de ces paramètres entraine une augmentation de l'activité cytotoxique. En effet, toute augmentation de paramètre  $SAG$  conduit à l'augmentation de surface du contact des molécules avec le récepteur biologique.
- La lipophilie est un paramètre important dans le mécanisme d'action des molécules médicamenteuses. Elle permet aux molécules de pénétrer facilement des membranes lipidiques pour arriver au site d'action.

- Les paramètres HBA, qO1 et qO2 dans le modèle 1 et qO4 et qO5 dans le modèle 2 sont responsables des interactions électrostatiques des dérivés de wogonine avec les résidus d'acides aminés dans les sites actifs de récepteur biologique.
- L'apparition de moment dipolaire dans le modèle 2, montre l'importance de ce paramètre dans la phase aqueuse. En effet, l'eau est un solvant physiologique protique polaire, joue un rôle important dans polarisation et la stabilité des dérivés de wogonine ce qui permet de renforcer les interactions électrostatiques avec les récepteurs biologiques.

Les matrices de corrélation entre les descripteurs obtenus par l'analyse MLR et l'activité biologique sont reportée dans les tableaux III.5 et III.6.

**Tableau III.5** : matrice de corrélation du 1<sup>er</sup> modèle

	pIC50	V	SAG	Log P	HBA	qO1	qO2	qO3
pIC50	1							
V	0.561	1						
SAG	0.601	0.984	1					
Log P	0.060	0.357	0.316	1				
HBA	0.424	0.679	0.677	-0.355	1			
qO1	0.498	-0.022	0.042	-0.121	-0.043	1		
qO2	0.659	0.567	0.565	-0.137	0.459	0.261	1	
qO3	0.570	0.604	0.595	-0.088	0.468	0.167	0.967	1

**Tableau III.6** : matrice de corrélation du 2<sup>ème</sup> modèle

	pIC50	V	SAG	Log P	HBA	DM	qO4	qO5
pIC50	1							
V	0.582	1						
SAG	0.625	0.984	1					
Log P	0.161	0.396	0.365	1				
HBA	0.356	0.630	0.623	-0.383	1			
DM	0.132	-0.241	-0.213	-0.040	-0.278	1		
qO4	0.523	0.572	0.549	-0.112	0.548	-0.118	1	
qO5	0.500	0.231	0.167	0.132	0.029	0.148	0.277	1

Les tableaux III.5 et III.6 montrent que les descripteurs obtenus sont corrélés entre eux et avec l'activité biologique. Le tableau III.5 montre que le descripteur qO2 est le plus important dans la corrélation entre les descripteurs sélectionnés et l'activité cytotoxique des dérivés de wogonine dans la phase gazeuse.

Le **tableau III.6** montre que le descripteur SAG est le paramètre le plus important dans la corrélation entre les descripteurs sélectionnés et l'activité cytotoxique des wogonines dans la phase aqueuse.

### III.1.4 Validation des modèles QSAR

#### III.1.4.1 Validation interne

Pour tester la validité de pouvoir prédictif des modèles QSAR sélectionnés (dans la phase gazeuse et dans la phase aqueuse), la technique de validation croisée *LOO* (*leave-oneout*) a été utilisée. Les modèles développés ont été validés par le calcul des paramètres statistiques suivants : La somme des carrés résiduels (PRESS), la somme des carrés totaux (TSS), les erreurs de prédiction (SPRESS), coefficient de corrélation de validation croisée ( $R^2_{cv}$ ), qualité d'ajustement ( $R^2_{adj}$ ) et l'erreur de prédiction (PE) (**Tableau III.7**).

**Tableau III.7** : Paramètres de validation croisée

Modèle	PRESS	SSY	PRESS/SSY	SPRESS	$R^2_{cv}$	$R^2_{adj}$	6PE
1	1.993	7.782	0.256	0.239	0.744	0.678	0.175
2	1.812	7.827	0.232	0.228	0.769	0.708	0.158

- La somme résiduelle prédictive des carrés (PRESS, Predictive Residual Sum of Squares) est une mesure des capacités prédictives d'un modèle. Sa valeur est inférieure à SSY (somme des carrés des distances) sur ce modèle prédit mieux que le hasard et peut être considérée comme statistiquement significative. La petite valeur de PRESS signifie la meilleure prévisibilité du modèle [9]. En outre, pour un modèle raisonnable de QSAR, le rapport PRESS/SSY doit être inférieur à 0,4 [8].
- Le tableau III.6 représente que ce rapport est 0.256 pour le modèle dans la phase gazeuse et 0.232 pour le modèle dans la phase aqueuse.
- Les valeurs élevées du  $R^2_{cv}$  et  $R^2_{adj}$  sont des critères essentiels pour une meilleure qualification des modèles QSAR [9].
- Dans notre étude on constate que  $R^2_{cv} = 0.744$  et  $R^2_{adj} = 0.678$  (dans la phase gazeuse), et  $R^2_{cv} = 0.769$  et  $R^2_{adj} = 0.708$  (dans la phase aqueuse).
- Pour les deux modèles la condition  $r > 6PE$  est satisfaite, ce qui indique que les deux modèles 1 et 2 possèdent un bon pouvoir prédictif [10].

**III.1.4.2 Validation externe**

Afin de tester de manière fiable le pouvoir prédictif du modèle QSAR, il est nécessaire d'utiliser la validation externe. Une fois l'ensemble de validation internes mis en place, il suffit alors d'appliquer le modèle QSAR aux molécules qui le composent et de déterminer la corrélation existant entre les activités calculées et celles expérimentales. Plus cette corrélation est importante, plus le modèle est capable de prédire les activités pour des molécules hors l'ensemble d'apprentissage [11,12].

Une valeur du test  $q^2$  supérieure à 0,6 peut être considérée comme un indicateur de bonne prévisibilité externe.

$$q^2 = 1 - \frac{\sum(Y_{\text{pred}} - Y_{\text{exp}})^2}{\sum(Y_{\text{exp}} - Y_{\text{mean}})^2}$$

Avec :

–  $Y_{\text{pres}}$  et  $Y_{\text{exp}}$  sont des activités prédite et expérimentale des molécules de test set.

–  $Y_{\text{mean}}$  est la moyenne de l'activité expérimentale des molécules de training set.

Après le calcul la valeur de  $q^2$  dans la phase gazeuse est égal 0.625 et dans la phase aqueuse est égal 0.683. Les résultats obtenus de validation interne et externe confirme la fiabilité des modèles prédictifs obtenus.

**III.1.5 Prédiction de l'activité**

Les activités cytotoxiques expérimentales, prédites et résiduelles de wogonine sont représentés dans le **tableau III.8**.

Le résidu est dû d'une part à l'incertitude de la mesure expérimentale et d'autre part à l'imperfection du modèle, donc si l'activité prédite est plus proche de l'activité expérimentale, le modèle peut être applicable [13].

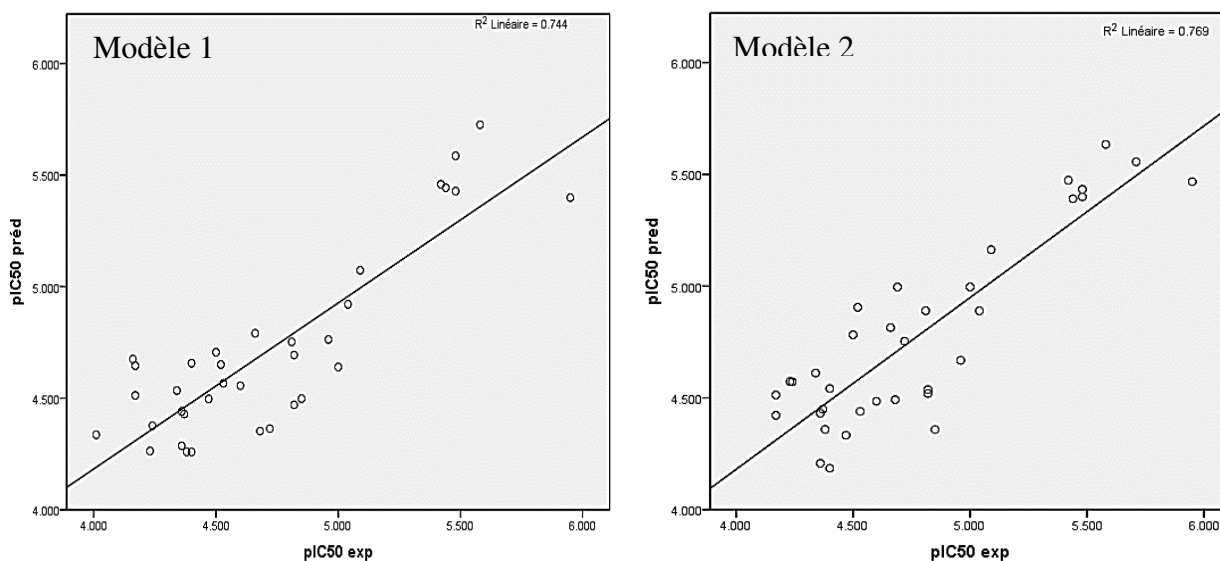


**Tableau III.8 :** Activité cytotoxique expérimentale, prédite et résiduelle des dérivés de wogonine.

<i>Dans la phase gazeuse</i>				<i>Dans la phase aqueuse</i>			
<b>Composé</b>	<b>pIC<sub>50</sub> expérimentale</b>	<b>pIC<sub>50</sub> prédite</b>	<b>pIC<sub>50</sub> résiduelle</b>	<b>Composé</b>	<b>pIC<sub>50</sub> expérimentale</b>	<b>pIC<sub>50</sub> prédite</b>	<b>pIC<sub>50</sub> résiduelle</b>
<b>2</b>	5.580	5.726	-0.146	<b>2</b>	5.580	5.634	-0.054
<b>3</b>	5.440	5.443	-0.003	<b>3</b>	5.440	5.391	0.049
<b>5</b>	5.480	5.428	0.052	<b>4</b>	5.710	5.556	0.154
<b>6</b>	5.950	5.399	0.551	<b>5</b>	5.480	5.432	0.048
<b>7</b>	5.040	4.921	0.119	<b>6</b>	5.950	5.467	0.483
<b>9</b>	4.340	4.535	-0.195	<b>7</b>	5.040	4.890	0.150
<b>10</b>	5.480	5.586	-0.106	<b>9</b>	4.340	4.611	-0.271
<b>11</b>	5.420	5.458	-0.038	<b>10</b>	5.480	5.400	0.080
<b>12</b>	4.810	4.752	0.058	<b>11</b>	5.420	5.474	-0.054
<b>14</b>	4.500	4.706	-0.206	<b>12</b>	4.810	4.890	-0.080
<b>15</b>	5.090	5.074	0.016	<b>14</b>	4.500	4.782	-0.282
<b>17</b>	4.660	4.791	-0.131	<b>15</b>	5.090	5.163	-0.073
<b>18</b>	4.530	4.567	-0.037	<b>16</b>	4.690	4.996	-0.306
<b>19</b>	4.400	4.259	0.141	<b>17</b>	4.660	4.815	-0.155
<b>20</b>	4.170	4.646	-0.476	<b>18</b>	4.530	4.440	0.090
<b>21</b>	4.160	4.676	-0.516	<b>19</b>	4.400	4.185	0.215
<b>22</b>	4.400	4.657	-0.257	<b>20</b>	4.170	4.422	-0.252
<b>24</b>	4.380	4.259	0.121	<b>22</b>	4.400	4.542	-0.142
<b>25</b>	4.170	4.512	-0.342	<b>24</b>	4.380	4.358	0.022
<b>26</b>	4.230	4.263	-0.033	<b>25</b>	4.170	4.513	-0.343
<b>27</b>	4.680	4.352	0.328	<b>26</b>	4.230	4.573	-0.343
<b>30</b>	4.820	4.470	0.350	<b>27</b>	4.680	4.492	0.188

<b>31</b>	4.820	4.693	0.127	<b>30</b>	4.820	4.537	0.283
<b>34</b>	4.850	4.499	0.351	<b>31</b>	4.820	4.520	0.300
<b>35</b>	4.240	4.377	-0.137	<b>34</b>	4.850	4.358	0.492
<b>37</b>	4.360	4.441	-0.081	<b>35</b>	4.240	4.572	-0.332
<b>38</b>	4.010	4.336	-0.326	<b>37</b>	4.360	4.431	-0.071
<b>39</b>	5.000	4.640	0.360	<b>39</b>	5.000	4.996	0.004
<b>40</b>	4.520	4.651	-0.131	<b>40</b>	4.520	4.906	-0.386
<b>42</b>	4.960	4.763	0.197	<b>42</b>	4.960	4.668	0.292
<b>43</b>	4.360	4.286	0.074	<b>43</b>	4.360	4.207	0.153
<b>45</b>	4.370	4.429	-0.059	<b>45</b>	4.370	4.449	-0.079
<b>46</b>	4.600	4.556	0.044	<b>46</b>	4.600	4.485	0.115
<b>47</b>	4.470	4.497	-0.027	<b>47</b>	4.470	4.333	0.137
<b>48</b>	4.720	4.363	0.357	<b>48</b>	4.720	4.753	-0.033

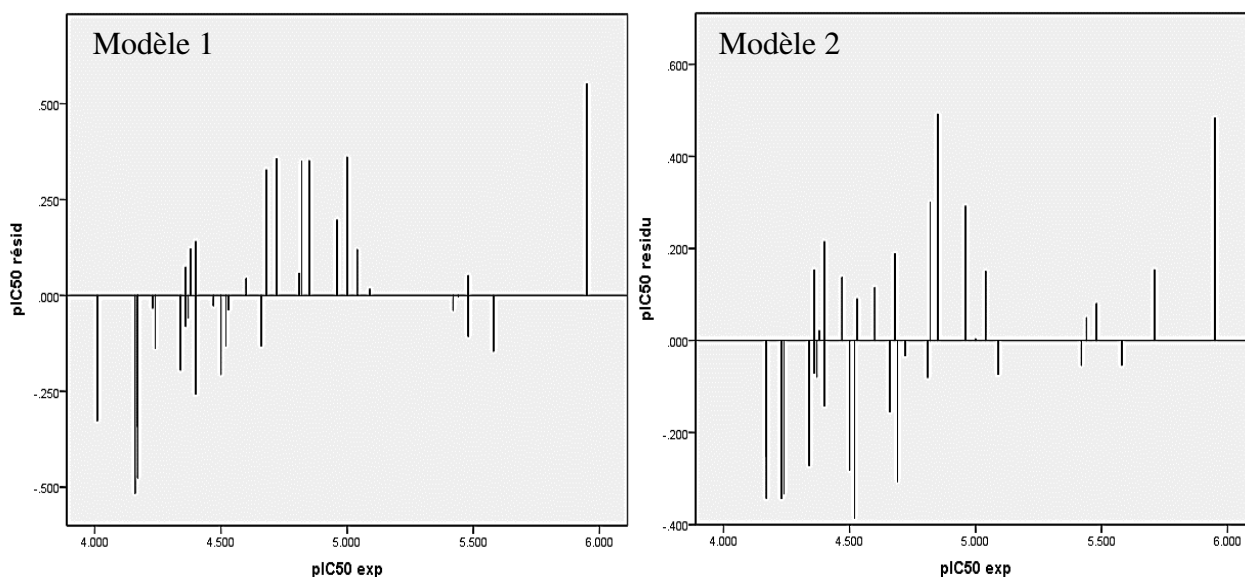
La **figure III.1** présente la corrélation entre les valeurs prédites et expérimentales pour l'activité cytotoxique des dérivés de wogonine. Du point de vue des performances statistiques, ces modèles présentent des coefficients de corrélation  $R^2$  significatifs de valeur. Les valeurs de  $R^2$  est égale à 0,744 et 0,769, pour les modèles **1** et **2** respectivement. Ils indiquent que les modèles peuvent être appliqués avec succès pour prédire l'activité biologique.



**Figure III.1 :** Corrélation entre l'activité biologique expérimentale et prédite pour les modèles (1) et (2).

Pour rechercher la présence d'une erreur systématique dans l'élaboration des modèles QSAR, les valeurs résiduelles de l'activité biologique ont été tracées en fonction des valeurs expérimentales comme montre la **Figure III.2**.

La répartition des résidus dans les deux côtés de l'axe des abscisses indique l'absence de l'erreur systématique [14]. Il indique que ces deux modèles peuvent être appliqués avec succès pour prédire l'activité cytotoxique pour la série de molécules utilisée dans le développement des modèles QSAR.



**Figure III.2 :** Corrélation entre les valeurs résiduelles et l'expérimentales des modèles (1) et (2).

**Références bibliographiques**

- [1]. G. Fayet, Développement de modèles QSAR pour la prédiction des propriétés d'explosibilité des composés nitroaromatiques, Thèse de doctorat de l'université pierre et marie curie, Paris VI, 2010.
- [2]. H. Gonzalez-Diaz, E. Olazabal, L. Santana, E. Uriarte, Y. Gonzalez-Diaz, N. Castanedo, *Bioorg. Med. Chem.*, 15, 962 (2007).
- [3]. HyperChem (Molecular Modeling System) Hypercube, Inc., 1115 NW, 4th Street, Gainesville, FL 32601, USA, (2008).
- [4]. F. Jensen; J. Wiley, *Introduction to Computational Chemistry*, Chichester, Angleterre, (1999).
- [5]. M.J. Frisch; G.W. Trucks; H.B. Schlegel; G.E. Scuseria; M.A. Robb; J.R. Cheeseman; G. Scalmani; V. Barone; B. Mennucci; G.A. Petersson; H. Nakatsuji; M. Caricato; X. Li; H.P. Hratchian; A.F. Izmaylov; J. Bloino; G. Zheng; J.L. Sonnenberg; M. Hada; M. Ehara; K. Toyota; R. Fukuda; J. Hasegawa; M. Ishida; T. Nakajima; Y. Honda; O. Kitao; H. Nakai; T. Vreven; J.A. Montgomery; J.J.E. Peralta; F. Ogliaro; M. Bearpark; J.J. Heyd; E. Brothers; K.N. Kudin; V.N. Staroverov; R. Kobayashi; J. Normand; K. Raghavachari; A. Rendell; J.C. Burant; S.S. Iyengar; J. Tomasi; M. Cossi; N. Rega; J.M. Millam; M. Klene; J.E. Knox; J.B. Cross; V. Bakken; C. Adamo; J. Jaramillo; R. Gomperts; R.E. Stratmann; O. Yazyev; A.J. Austin; R. Cammi; C. Pomelli; J.W. Ochterski; R.L. Martin; K. Morokuma; V.G. Zakrzewski; G.A. Voth; P. Salvador; J.J. Dannenberg; S. Dapprich; A.D. Daniels; Ö. Farkas; J.B. Foresman; J.V. Ortiz; J. Cioslowski. D. J. Fox, Gaussian, Inc., Wallingford CT, (2009).
- [6]. [www.molinspiration.com/cgi-bin/properties](http://www.molinspiration.com/cgi-bin/properties).
- [7]. H.A. Carlson. *Curr. Op. Chem. Biol. S*, 447-452, 2002.
- [8]. K. Ajeet, Bijander, *Int. J. Pharm. & Pharm. Sci.*, 4, 445, (2012).
- [9]. S.O. Podunavac-Kuzmanović, D.D. Cvetković, D.J. Barna, *Int. J. Mol. Sci.*, 10, 1670, (2009).
- [10]. A.K. Srivastava, N. Shukla, *J. Saudi Chem. Soc.*, 17, 321 (2013).
- [11]. LM. Shi, H. Fang, W. Tong, *J. Chem. Inf. Comput. Sci.* 41, 186–195, (2001).
- [12]. N. Nicola Chirico, P. Gramatica, *J. Chem. Inf. Model.* 51(9), 2320–2335, (2011).

- [13]. M.Clark, and R. D. Cramer, III, The probability of chance correlation using partial Least squares (PLS). *Quantitative Structure-Activity Relationships*, 12 (2): pp. 45,137, (1993).
- [14]. M. Jalali-Heravi and A. Kyani, *J. Chem. Inf. Comput. Sci.* 44, 1328, (2004).

*Conclusion*

*générale*

### Conclusion générale

Dans ce travail nous avons réalisé une étude quantitative de la relation structure-activité (QSAR) sur une série des dérivés de wogonine. Cette étude effectuée dans la phase gazeuse et dans la phase aqueuse. Plusieurs descripteurs ont été utilisés dans le développement des modèles QSAR à savoir : logP, SAG, V, HBA, DM et les charges atomiques qO1, qO2, qO3, qO4, qO5.

Deux modèles QSAR ont été établis en utilisant la méthode de régression multilinéaire (MLR). Les modèles obtenus dans les deux phases peuvent être utilisés pour prédire l'activité cytotoxique de wogonine.

Le pouvoir prédictif des modèles obtenus a été confirmé par la méthode de validation croisée LOO. La validité du modèle a été établie par la détermination des paramètres statistiques appropriés. Les résultats obtenus de validation interne et externe confirme la fiabilité des modèles prédictifs obtenus.

Les valeurs prédites sont proche que les valeurs expérimentales pour les deux activités biologiques.

Une forte corrélation a été observée entre les valeurs expérimentales et prédites de l'activité cytotoxiques, ce qui indique la validité et la qualité des modèles QSAR obtenus.

L'effet de solvant est confirmé par la comparaison entre les modèles 1 et 2 dans la phase gazeuse et la phase aqueuse respectivement.

# Annexe

## Table du F (de Fischer)

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	8,58
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	2,18
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	2,08
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	1,68
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 :**

**U**ne étude QSAR a été effectuée sur quarante-huit dérivés de wogonine dans la phase gazeuse et dans la phase aqueuse en utilisant les descripteurs suivants : SAG, V, log P, HBD, HBA, DM et les charges atomiques. La régression linéaire multiple (MLR) a été utilisée pour quantifier les relations entre les descripteurs moléculaires et l'activité cytotoxique des dérivés de wogonine. Le pouvoir prédictif des modèles obtenus a été confirmé par la méthode de validation croisée LOO. Une forte corrélation a été observée entre les valeurs expérimentales et prédites de l'activité biologique, ce qui indique la validité et la qualité des modèles QSAR obtenus.

**Mots Clés :** Wogonine, Activité cytotoxique, QSAR, MLR, LOO.

### **Abstract :**

**A** QSAR study was performed on forty-eight wogonine derivatives in the gas phase and the aqueous phase using the following descriptors: SAG, V, log P, HBD, HBA, DM and atomic charges. Multiple linear regression (MLR) was used to quantify the relationships between molecular descriptors and the cytotoxic activity of wogonin derivatives. The predictive power of the obtained models was confirmed by the LOO cross validation method. A strong correlation was observed between the experimental and predicted values of biological activity, indicating the validity and quality of the QSAR models obtained.

**Keywords :** Wogonin, Cytotoxic activity, QSAR, MLR, LOO.