



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

# MÉMOIRE DE MASTER

Domaine des Sciences de la Matière  
Filière de Chimie  
Spécialité Chimie Pharmaceutique

---

Présenté et soutenu par :

**AID AICHA**

**BEN SGHIER FATMA ZAHRA**

## Conception de nouvelles molécules à activité anti-oxydante par des études 3D QSAR

---

**Jury :**

Necira	Zelikha	M.C.B	Université Mohamed khider	Président
Lehraki	Faiza	M.A.A	Université Mohamed khider	Examineur
Khamouli	Saida	M.C.A	Université Mohamed khider	Encadreur

**Année universitaire : 2020-2021**

# *Remerciements*

*Tout d'abord, nous remercions ALLAH le Tout Puissant qui nous a accordé la volonté et le courage pour réaliser ce mémoire.*

*Nos plus sincères remerciements vont à **Khamouli Saida** notre encadreur de mémoire, pour sa disponibilité, ses corrections, ses conseils judicieux et pour nous avoir guidés tout au long de ce travail.*

*Nous adressons mes sincères remerciements à Monsieur **Necira Zelikha** Docteur à Université de Biskra d'avoir accepté de présider le jury de cette Mémoire.*

*Nos remerciements également à Monsieur **Lehraki Faiza** Docteur à Université de Biskra d'avoir accepté de juger mon travail.*

# *Dédicace*

*Je dédie ce modeste travail à:*

*A mes parents : **Mohamed Lamine, Souria**; aucun hommage ne pourrait être à la hauteur de l'amour dont ils ne cessent de me combler. Que dieu leur procure bonne santé et longue vie.*

*A celui que j'aime et qui m'a soutenue tout au long de ce mémoire:*

*Mon frères: **Tarek, kamal Eddine, Saif Eddine**.*

*A mon chère binôme: **Aid Aicha**.*

*Et mes beaux amis: **Loubna, Chafia, Amina, Ramla, Samar, Aroua**  
**rahimaha Allah**.*

*Et à tous ceux qui ont contribué de près ou de loin pour que ce mémoire soit possible, je vous dis merci.*

***Ben seghier fatma Zahra***

## *Dédicace*

*Je dédie ce modeste travail à:*

*A mes parent : **Aissa, Necira** rahimaha Allah; j'aimerais qu'elle soit avec moi mais je pense qu'elle est fière de moi. Mon père, que le Dieu procure bonne santé et longue vie.*

*A celui que j'aime et qui m'a soutenue tout au long de ce mémoire:*

*Mon frères: **Abd El Majid, Khaled, Sofiane, Nabil.***

*Mes sœurs: **Wafa, Samia.***

*A mon chère binôme: **Ben Seghier Fatma Zahra.***

*Et mes amis sont toujours à mes côtés **Nariman, Yasmin, Rabaa, Maria***

*Et mes amis proches **Souria, Chahinaz, Samah, Mariam.***

*A tous dans ma famille.*

*Et à tous ceux qui ont contribué de près ou de loin pour que ce mémoire soit possible, je vous dis merci.*

***Aid Aicha***

# *Sommaire*

---

## **Sommaire**

Liste des tableaux.....	i
Liste des figures.....	ii
Liste des Abréviations : .....	iii
Introduction générale .....	1
Référence	

### **Chapitre I : Les coumarines et l'activité antioxydant**

1. Historique : .....	4
2. Introduction :.....	4
2.1. Structure chimique et nomenclature :.....	5
2.2. Classification structurale :.....	6
2.2.2. Furanocoumarines : .....	7
2.2.3. Pyranocoumarines : .....	8
2.2.4. Dicoumarines (coumarines dimériques) :.....	9
2.2.5. Tricoumarine (coumarines trimériques) :.....	9
3. Propriétés physico-chimiques :.....	10
4. Activités biologiques et intérêts des coumarines :.....	10
5. Toxicité des coumarines : .....	11
6. Activité antioxydante :.....	11
7. Activité antioxydante des coumarines .....	12
Référence	

### **Chapitre II : Méthodologie de QSAR**

1. Histoire de la découverte des médicaments.....	19
---	----

## *Sommaire*

---

1.1. Choix d'une cible thérapeutique : .....	19
1.2. Identification des touches (hits): .....	20
1.3. Génération et optimisation de tête de série (leads) : .....	20
1.4 Tests précliniques et cliniques : .....	21
2. Relation s quantitatives structures activités (QSAR):.....	21
2.1. Historique:.....	21
2.2. Définition : .....	23
2.3. Principe de QSAR :.....	23
2.4. Outils et Méthodologie de QSAR : .....	24
2.4 .1. Paramètres biologiques :.....	24
2.4 .2. Les descripteurs moléculaires.....	24
2.5. QSAR 3D .....	26
2.5.1. Analyse comparative des champs moléculaires : CoMFA.....	26
2.5.2. Analyse comparative d'indices de similarité moléculaire : CoMSIA .....	27
2.6. Méthodes statistiques : .....	29
2.6.1. Régression linéaire simple (SLR):.....	29
2.6.2. Régression linéaire multiple (MLR):.....	29
2.6.3. L'analyse Partial Least Square (PLS): .....	30
2.6.4. Analyse par composantes principales (ACP) : .....	30
2.7. Coefficients et tests statistiques standards : .....	31
2.7.1. Coefficient de détermination ( $R^2$ ):.....	31
2.7.2. Coefficient de corrélation ( $R^2$ ): .....	32
2.7. 3. Limitation du coefficient de détermination $R^2$ : .....	32

## *Sommaire*

---

2.7.4. Test Fisher-Snedecor (F): .....	32
2.7.5. Coefficient de prédiction (Q <sup>2</sup> ): .....	33
2.8. Validation du modèle : .....	33
2.8.1. Validation interne : .....	34
2.8.2. Validation externe : .....	35
2.9. Application de QSAR : .....	36

Référence

### **Chapitre III : Etude 3D QSAR d'une série de coumarine**

1. Introduction: .....	41
2. Matériels et méthodes: .....	41
2.1. Base de données: .....	41
2.2. Minimisation et Alignement: .....	45
2.3. Analyse de CoMFA:.....	45
2.4. Construction et validation de modèle :.....	46
3. Résultats et discussion .....	46

Référence

Conclusion générale.....	52
--------------------------	----

Résumé

## *Liste des tableaux*

---

### Liste des tableaux

<b>Tableau I. 1 :</b> Coumarine simple et ses dérivés.....	7
<b>Tableau I. 2 :</b> représente des quelquesDicoumarines.....	9
<b>Tableau II. 1 :</b> Types de données biologiques utilisées dans l'analyse QSAR. ....	24
<b>Tableau III. 1 :</b> Structures chimique de molécules étudiées.....	42
<b>Tableau III. 2 :</b> Paramètres statistiques (PLS) de modèle CoMFA .....	47
<b>Tableau III. 3 :</b> Activités expérimentales et prédites de l'ensemble de données.....	47

## Liste des figures

---

### Liste des figures

<b>Figure I.1 :</b> Combinaison d'un benzène avec un pyrane en une coumarine.....	6
<b>Figure I.2:</b> Structure générale des coumarines.....	6
<b>Figure I.3:</b> Structures de quelques furanocoumarines [16].....	8
<b>Figure I.4 :</b> Structures de quelques Pyranocoumarine [17]. ....	8
<b>Figure I. 5 :</b> représente les structures des Dicoumarines.....	9
<b>Figure I. 6 :</b> Triumbéllatine.....	9
<b>Figure I.7 :</b> Quelques coumarines simples ou fusionnées, à activité antioxydante. ....	13
<b>Figure II.1 :</b> Etapes du processus de recherche et de développement des médicaments. ...	19
<b>Figure II.2 :</b> (a) Courbes des potentiels de Lennard-Jones (rouge) et Coulomb (bleu) utilisées dans les études CoMFA et définition d'une valeur limite supérieure. (b) La fonction gaussienne en forme de cloche des champs SEAL (orange) utilisée lors des études CoMSIA est une bonne approximation des potentiels de Lennard-Jones et Coulomb tout en présentant l'avantage d'être plus lisse. [41] Pour pallier ce problème, Klebe et ses collègues ont l'avantage d'être plus lissée.....	28
<b>Figure II.3 :</b> Aires de contour d'un modèle CoMSIA permettant de visualiser la contribution hydrophobique favorable (jaune) ou défavorables (gris), et la contribution des accepteurs de liaisons hydrogène favorable (violet) ou défavorable (vert) à l'activité. [44].....	29
<b>Figure III.1 :</b> Alignement des dérivés de coumarine.....	45
<b>Figure III. 2 :</b> Représentation graphique de corrélation entre les activités expérimentales et prédite des molécules étudiées.....	48
<b>Figure III. 3 :</b> Représentation graphique de corrélation entre les activités expérimentales et résiduelles des molécules étudiées.....	48
<b>Figure III.4 :</b> Carte de contour des champs stérique (a) et électrostatique (b) de COMFA.....	50
<b>Figure III.5 :</b> Diagramme de relation structure –activité des coumarines orientées vers l'activité antioxydant.....	50

## *Liste des abréviations*

---

### **Liste des abréviations :**

**ACP** : analyse en composantes principales.

**ADN** : Acide désoxyribonucléique.

**ARN** : Acide ribonucléique

**COMFA**: Comparative Molecular Field Analysis.

**COMSIA**: Analyse comparative indices de similarité moléculaire.

**ERO**: Espèces réactives de l'oxygène.

**LOO**: Leave One Out.

**MLR**: regression linéaire multiple.

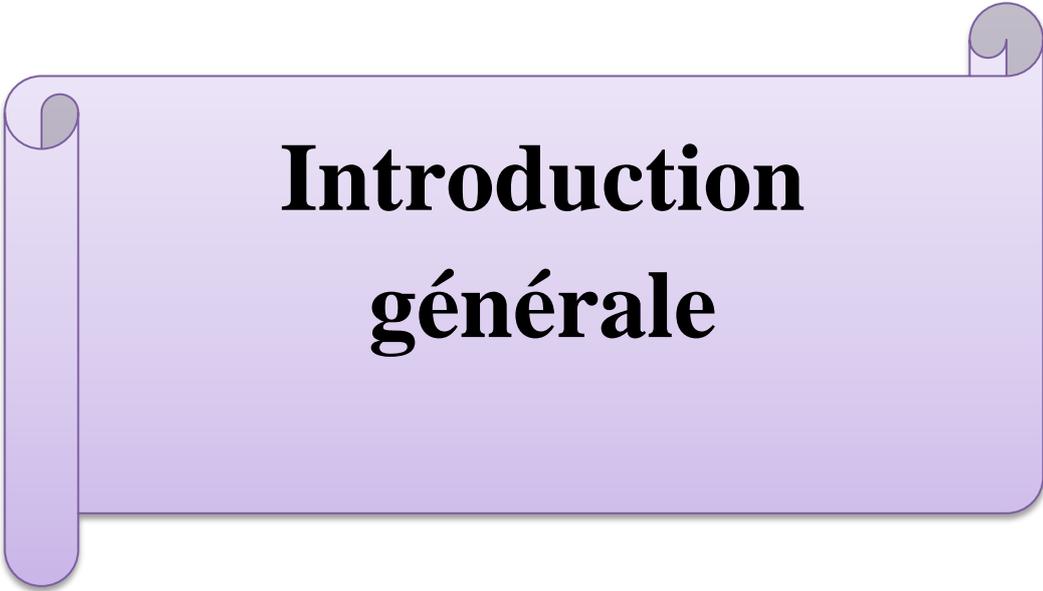
**PLS** : Partial Least Square.

**QSAR** : relation quantitative structure activité.

**SLR** : régression linéaire simple.

**UV/vis** : Ultraviolet-visible.

**VIH** : Virus de l'immunodéficience humaine.



**Introduction  
générale**

## *Introduction générale*

---

L'activité anti-oxydante demeure un axe très dynamique de la recherche scientifique multidisciplinaire. En 1956, Harman [1] a émis une hypothèse faisant le lien entre les radicaux libres générés dans l'organisme responsables de l'endommagement de nombreux composants cellulaires tels que les lipides, les protéines et les acides nucléiques ... causant des signes de vieillissement précoce. Depuis, le lien entre la propagation des radicaux libres communément connus sous le nom des Espèces Réactives Oxygénées (ERO) [2] et certaines pathologies plus moins graves ne cesse de se confirmer, dont les maladies cardiovasculaires [3], les maladies inflammatoires [4], et certains cancers [5], et d'autres.

Le corps humain se dispose de son propre dispositif de défense contre la génération et la propagation des ERO. Ce sont des systèmes enzymatiques endogènes comme le glutathion peroxydase, superoxyde dismutase, les catalases et autres [6]. Néanmoins, une génération excessive de ces espèces radicalaires prédatrices peut causer un déséquilibre entraînant un stress oxydatif. Pour la détoxification, le corps a besoin de certains composés connus par leur activité antioxydante apportée essentiellement par l'alimentation.

Une diversité structurelle a été notée dans cette famille de composés, qui ont attiré leur intérêt depuis longtemps en raison de leurs activités biologiques, notamment les activités anticoagulantes, antimicrobiennes, antivirales, anticancéreuses, activité d'inhibition d'enzymes, anti- inflammatoires, antioxydants, et sur le système nerveux central.

La modélisation QSAR (Quantitative Structure-Activity Relationship) constitue actuellement un outil fiable et performant pour combler les lacunes des études *in vivo* et *in vitro* et limiter le recours aux tests biologiques. C'est une méthode qui tente à identifier une relation quantitative entre l'activité mesurée expérimentalement et des descripteurs théoriques ou empiriques liés à la structure moléculaire.

L'activité anti-oxydante a été discutée dans plusieurs ouvrages [7, 8-11], et de nombreux auteurs ont essayé de construire des modèles QSAR afin de corréler l'activité anti-oxydante des dérivés de coumarine avec les propriétés moléculaires. Cependant, plusieurs travaux présentent certains inconvénients, tels que l'utilisation arbitraire de méthodes théoriques pour calculer les propriétés moléculaires.

## *Introduction générale*

---

Dans ce contexte, l'objectif de ce travail est d'élaborer des modèles QSAR robustes et fiables, en respectant toutes les étapes d'une étude QSAR pour expliquer et prédire l'activité anti-oxydante d'une série de coumarine constituée de 22 composés.

Notre travail est divisé en trois chapitres :

- Le premier chapitre présente un aperçu général sur les coumarines et l'activité antioxydant.
- Le deuxième chapitre est consacré à une étude bibliographique sur les approches de base, les méthodologies, le développement, les techniques de validation et les applications des méthodes QSAR.
- Le troisième chapitre a été consacré à la génération d'un modèle 3D QSAR par un ensemble des dérivés de coumarine.

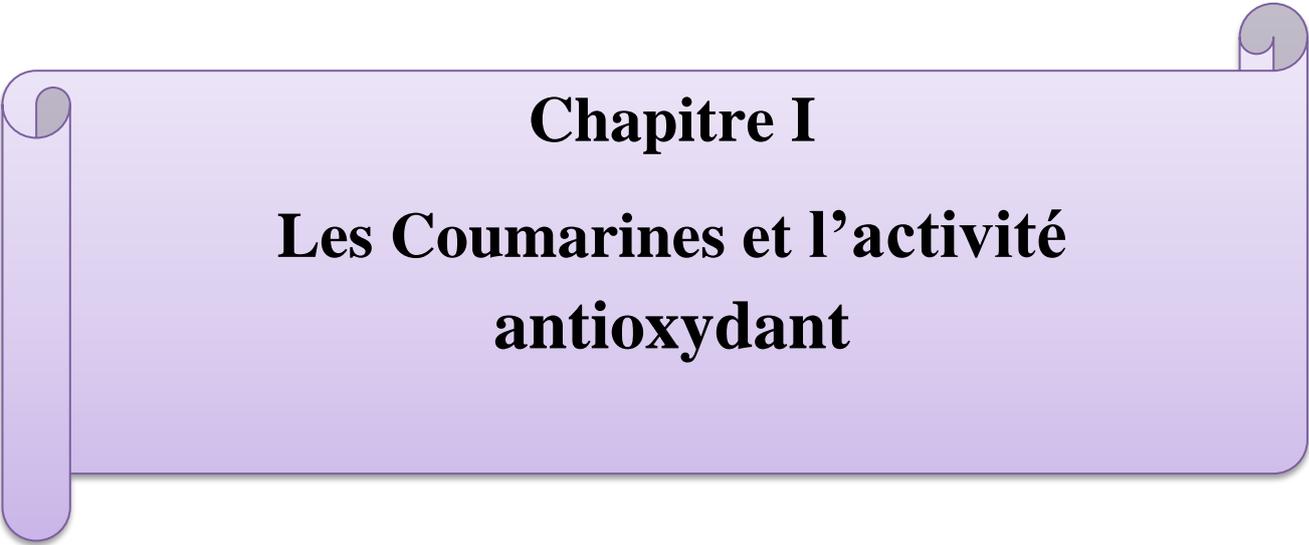
Enfin nous terminons ce travail par une conclusion générale récapitulative de notre travail.

## *Introduction générale*

---

### **Références:**

- [1] D. Harman, J. Gerontol, Aging, A theory based on free radical and radiationchemistry, 1956, 11, 298-300.
- [2] J. R. Roede, D.P. Jones, Environmental and Molecular Mutagenesis, 2010, 51, 380.
- [3] K. Sachidanandam, S. C. Fagan, A. Ergul, Cardiovascular Drug Reviews, 2005, 23, 115. [4] P. G. Winyard, D. R. Blake, C. H. Evans, Ed. Springer Basel AG, 2000. 11.
- [5] S. Kawanishi, Y. Hiraku, M. Murata, S. Oikawa, Am. cancer Soc., 2002, 32, 822.
- [6] F. ZERARGUI, Thèse-Doctorat, 2015, Université F. A. Sétif
- [7] V. Rastija, M. Medic, S. Eur, J. Med. Chem., 2009, 44, 400.
- [8] E. Sergediene, K. Jonsson, FEBS Lett, 1999, 462, 393.
- [9] D. Amic, D. Davidovic, D. Amic, N. Trinajstic, Croat. Chem. Acta., 2003, 76, 55.
- [10] D. Amic, D. Davidovic, V. Rastija, N. Trinajsti, C. Med. Chem, 2007, 14, 827.
- [11] M. Filipovic, Z. Markovic, C. R. Chimie, 2015, 18,492-498.



**Chapitre I**  
**Les Coumarines et l'activité  
antioxydant**

### Les coumarines

#### **1. Historique :**

En 1936, Von Werder a qualifié les coumarines comme d'agents thérapeutiques. En 1958, Bose a résumé la plupart des propriétés biologiques des coumarines naturelles. En 1964, Soine a publié une révision sur les effets biologiques et pharmacologiques des coumarines connus à ce jour, avec un accent particulier sur l'activité associée aux produits naturels. Plus récemment, en 1972, Scheel a publié une revue sur l'action biologique des coumarines. Depuis les années 1980, il y a eu un certain nombre d'excellentes revues sur les coumarines, principalement en ce qui concerne l'occurrence et les activités biologiques qui ont été publiées par Murray et al, et d'autres auteurs [1]. La présente étude vise à donner un aperçu de la recherche sur les coumarines simples et leurs analogues, et à décrire les progrès réalisés dans ce domaine au cours des dernières années, isolément. Ou en synthèse, et les propriétés biologiques les plus pertinentes associées à ce système.

#### **2. Introduction :**

L'étude des coumarines a commencé il y a plus de 200 ans [2]. Le nom de Cette famille chimique est dérivé de Coumarou na odorata Aube (*Dipteryx odorata*), dont on a isolé pour la première fois, le membre le plus simple de cette classe, la coumarine elle-même. La coumarine est un métabolite secondaire largement présent naturellement dans plusieurs familles des plantes et huiles essentielles, et qui a été utilisé comme parfum dans les produits alimentaires et cosmétiques. Le noyau coumarinique correspond à la benzo-pyrone (2H-1-benzopirane-2-one) dont la nomenclature systématique a été établie par IUPAC. [3]

Les Coumarines ont été grossièrement classées comme suit :

- a) Coumarines simples
- b) Furanocoumarines
- c) Pyranocoumarines
- d) Dicoumarines et Tricoumarines et bien que l'on puisse faire des distinctions selon leur structure chimique parmi les divers groupes de coumarines. [4]

## ***Chapitre I : Les coumarines et l'activité antioxydant***

---

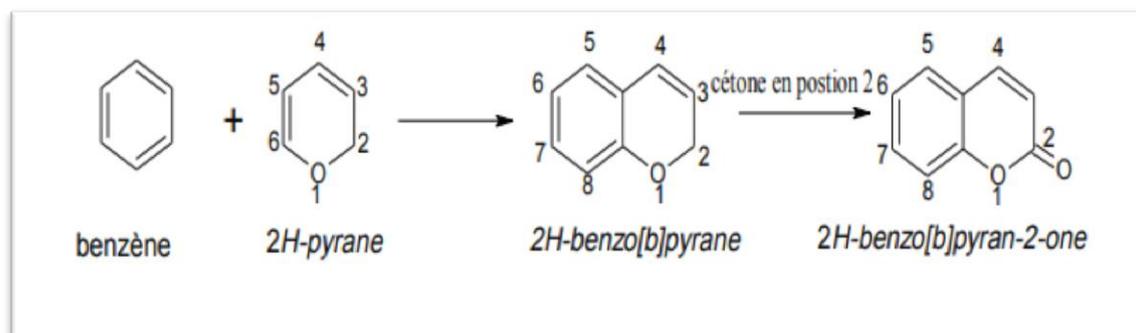
Au-delà des découvertes faites par l'isolement des coumarines à partir des centaines d'espèces de plantes et d'autres organismes, il existe des dérivés d'origine synthétique qui augmentent significativement le nombre de structures coumariniques connues jusqu'aujourd'hui. [5]

Les coumarines constituent une classe importante des produits naturels. La structure de la coumarine se trouve dans environ 150 espèces, appartenant à 30 familles de plants différents. Elle donne une odeur caractéristique semblable à celle du fion fraîchement fauché. [6,7] Les coumarines ont des effets différents sur le développement des plantes suivant leur concentration et aussi suivant l'espèce. Dans la cellule, les coumarines sont principalement présentés sous forme glycosylés. Cette glycosylation serait une forme de stockage permettant d'éviter les effets toxiques des coumarines. [8] Elles sont considérées comme des phytoalexines. Ces phytoalexines sont synthétisées au niveau des sites d'infection. Leur activité antimicrobienne est démontrée in vitro. [8] En général, les coumarines sont répandues dans tout le règne végétal. Les familles les plus riches en coumarines sont : Légumineuse, Rutacées, Apiacées, Astéracées et Thyméléacées. Elles se trouvent dans toutes les parties des plantes et également dans les fruits. [9]

### **2.1. Structure chimique et nomenclature :**

- **Structure chimique :**

La coumarine est un composé chimique hétérocyclique, produit par la combinaison d'un noyau benzène avec un pyrane, présentant une fonction cétone en position  $\alpha$  par rapport à l'oxygène.



**Figure I. 1 :** Combinaison d'un benzène avec un pyrane en une coumarine

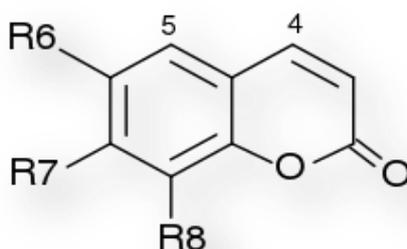
- **Formule brute :** C<sub>9</sub>H<sub>6</sub>O<sub>2</sub>
- **Noms chimiques:**
  - Nom IUPAC : 2H-chromen-2-one ou 2H-1-benzopyran-2-one
  - Autres: 1,2-benzopyrone ( $\alpha$ -benzopyrone) [10].

### 2.2. Classification structurelle :

La coumarine est le point de départ d'une famille des composés, qui se forment par une substitution sur un cycle aromatique. De ce fait et d'après la nature des substitutions on peut classer les coumarines en 5 catégories [11] :

#### 2.2.1. Coumarines simples :

Ces composés sont très répandus dans le règne végétal, plus de 700 structures sont déjà décrites [12]. La majorité d'entre eux, sont substituées en C7 par un hydroxyle: c'est la 7- hydroxycoumarine connue sous le nom d'Ombellifèrone qui constitue un précurseur des coumarines 6,7-di-et 6, 7,8-trihydroxylées. La structure générale des coumarines est reportée dans la figure ci-après.



**Figure I. 2 :** Structure générale des coumarines

## Chapitre I : Les coumarines et l'activité antioxydant

**Tableau I. 1 :** Coumarine simple et ses dérivés.

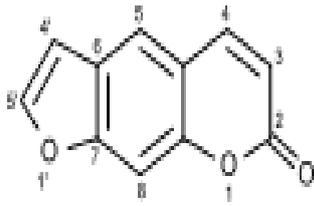
Composé	R6	R7	R8
Coumarine	H	H	H
Ombellifèrone	H	OH	H
Heriarine	H	OCH3	H
Esculétol	OH	OH	H
<u>Scopoléto</u>	OCH3	OH	H
Scopanone	OCH3	OCH3	OH
Fraxéto	OCH3	OH	OH

### 2.2.2. Furanocoumarines :

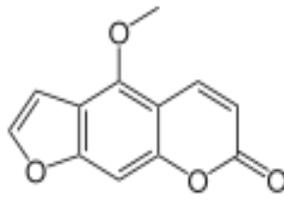
Les furocoumarines (appelées encore furanocoumarines) sont des molécules tricycliques, produites par la combinaison de deux hétérocycles (coumarine et furane). Ils constituent une classe abondante où chaque membre se distingue par la présence de divers groupements (hydroxy, alkoxy, géranyloxy...) sur les carbones 2, 5,6 et 8 [13].

Les Furanocoumarine constituent une famille de composés synthétisés par certaines espèces des végétaux supérieurs. Elles dérivent principalement de l'Ombelliféracée par condensation isopronoides en C5, et souvent liposolubles [14]. Le cycle furane peut être fusionné au cycle benzénique dans deux positions linéaires (dérivant de la molécule de psoralène), angulaire basées sur la structure de l'angélicine. [15]

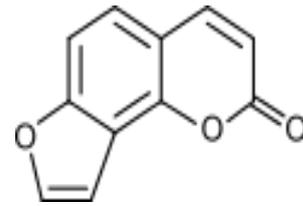
## Chapitre I : Les coumarines et l'activité antioxydant



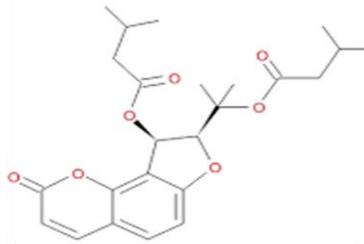
Psoralène



Bergaptène



Angélicine



Athamantine

Figure I. 3 : Structures de quelques furanocoumarines [16]

### 2.2.3. Pyranocoumarines :

Ils sont formés par la fusion d'un hétérocycle pyrane avec la coumarine :

- ❖ Série linéaire : xanthylétine.
- ❖ Série angulaire : séseline, visnadine. [17]

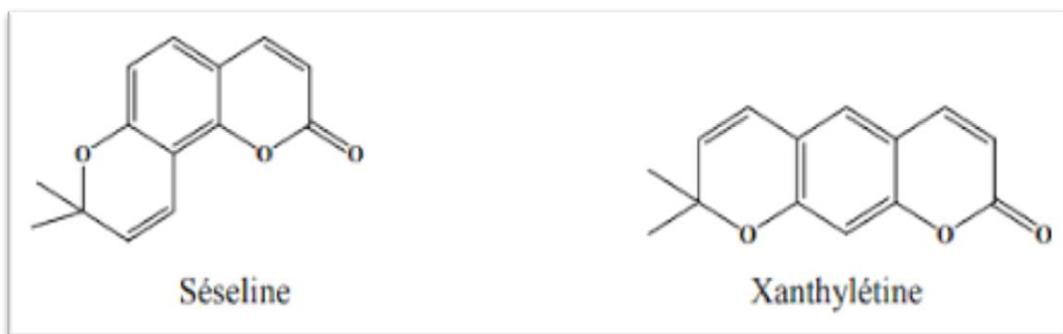


Figure I. 4 : Structures de quelques Pyranocoumarines [17].

## Chapitre I : Les coumarines et l'activité antioxydant

### 2.2.4. Dicoumarines (coumarines dimériques) :

Ce sont des composés formés par la liaison de deux unités coumariniques simples.

[18]

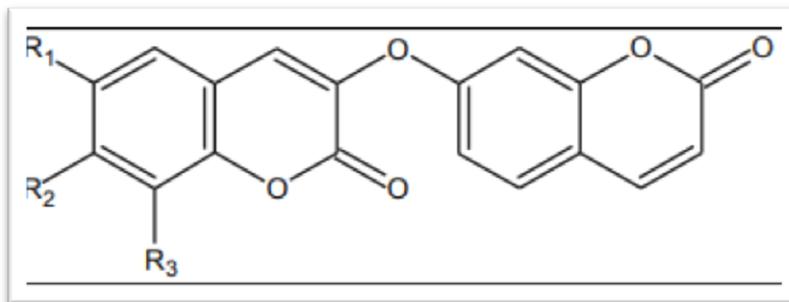


Figure I. 5 : représente les structures des Dicoumarines.

– Le tableau 2 représente des quelques Di coumarines :

Tableau I. 2 : représente des quelques Dicoumarines.

Nom	R1	R2	R3
Daphnorétine	OCH3	OH	H
Edgeworthine	OH	OH	H

### 2.2.5. Tricoumarine (coumarines trimériques) :

Ce sont des coumarines issues de l'union de trois entités coumariniques. [18]

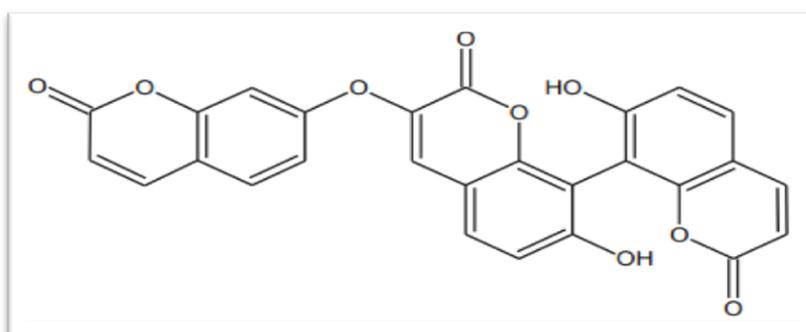


Figure I. 6 : Triumbéllatine.

### **3. Propriétés physico-chimiques :**

- Les coumarines sont des solides cristallisés blancs ou jaunâtres de saveur généralement amère ; certaines sont sublimables.
- Les coumarines hydroxylées possèdent une intense fluorescence bleue en lumière ultraviolette. Leur spectre U.V est également caractéristique et sert à leur identification.
- Les propriétés chimiques sont principalement dues à la fonction lactone insaturée, notamment l'ouverture de l'anneau lactonique en milieu alcalin.
- Avec les sels de plomb, on obtient des combinaisons insolubles [19].

### **4. Activités biologiques et intérêts des coumarines :**

Les coumarines manifestent diverses activités biologiques, qui varient selon la substitution sur le cycle benzopyrone [22-25] : anti-agrégant plaquettaire, anti-inflammatoire, anticoagulant, anti-tumorale, diurétiques, antimicrobien, antivirale et analgésique [26-32]. Elles sont indiquées dans le cas de lymphœdème du membre supérieure après traitement radiochirurgical du cancer du sein.

O'Neal et son équipe ont montré l'efficacité des coumarines pour bloquer le cancer induit chimiquement par les radiations ultraviolettes [33]. Ces molécules sont capables de prévenir la peroxydation des lipides membranaires et de capter les radicaux hydroxyles, superoxyde et peroxydes [34,35].

Les coumarines se révèlent être des composés thérapeutiques promoteurs dans l'amélioration du système immunitaire (action immunostimulante) : l'administration de la coumarine et de l'umbellifère par des malades atteints de cancers a raison de 100 mg par jour a provoqué une augmentation des lymphocytes T Helper dans la circulation sanguine [32,36].

Une activité antifongique significative a également été rapportée pour certaines coumarines. Dans leurs études *in vitro*, Degre et son équipe ont montré que les coumarines paralysent la croissance de *Saccharomyces servisiae*. Pour l'activité antibactérienne, les coumarines sont efficaces contre les bactéries à Gram positif [30, 37, 38].

Après stimulation par des rayonnements ultra-violet de grande longueur d'onde (365 nm), certaines furanocoumarines ont la propriété de stimuler fortement la synthèse de

## ***Chapitre I : Les coumarines et l'activité antioxydant***

---

mélanine par les mélanocytes cutanées, à travers une augmentation de leur activité mitotique. Cette propriété a été exploitée par les égyptiens pour le traitement de la dépigmentation [36]. Par ailleurs, des chercheurs espagnols ont évalué l'activité anti-VIH de onze composés dérivés de la 4-phénylcoumarine isolés de l'espèce *Marila pluricostata* [39].

Il convient d'ajouter que les furanocoumarines sont impliquées dans bien d'autres mécanismes biologiques que nous ne pouvons décrire ici, mais qui constituent des pistes de recherches pour le traitement de maladies comme la sclérose en plaque et la dépression [40].

### **5. Toxicité des coumarines :**

La coumarine n'est pas toxique en soi, elle peut être convertie par les champignons, en une toxine le dicoumarol qui est typiquement présent dans le foin moisi ; chez le bétail, le dicoumarol provoque des hémorragies fatales en inhibant la vitamine K qui est un facteur de coagulation du sang [41, 42, 30].

Chez l'être humain la consommation des espèces végétales qui renferment des furanocoumarines linéaires ou angulaires provoquent, s'il est accompagné d'une exposition à la lumière solaire, une dermatite phototoxique d'intensité variable allant du simple érythème jusqu'à l'apparition de bulles et de vésicules au niveau des zones exposées, quelques jours après le contact, les zones touchées présentent une hyperpigmentation [43]. Les furanocoumarines linéaires peuvent être le support de cycloaddition avec les bases pyrimidiques de l'ADN ou de l'ARN, la duplication des brins est alors bloquée ainsi que la traduction des ARN. Cette propriété est souvent invoquée pour expliquer les caractères mutagènes et carcinogènes de ces molécules. [43]

### **6. Activité antioxydante :**

Certaines recherches scientifiques récentes ont mis en évidence l'existence des facteurs communs responsables aussi bien du vieillissement que de maladies liées au stress oxydant comme le cancer, les maladies cardio-vasculaires, les maladies neurodégénératives (Parkinson et Alzheimer), et aussi certaines allergies et autres maladies chroniques. Ces diverses maladies auraient, entre autres, la même composante qui permet

## ***Chapitre I : Les coumarines et l'activité antioxydant***

---

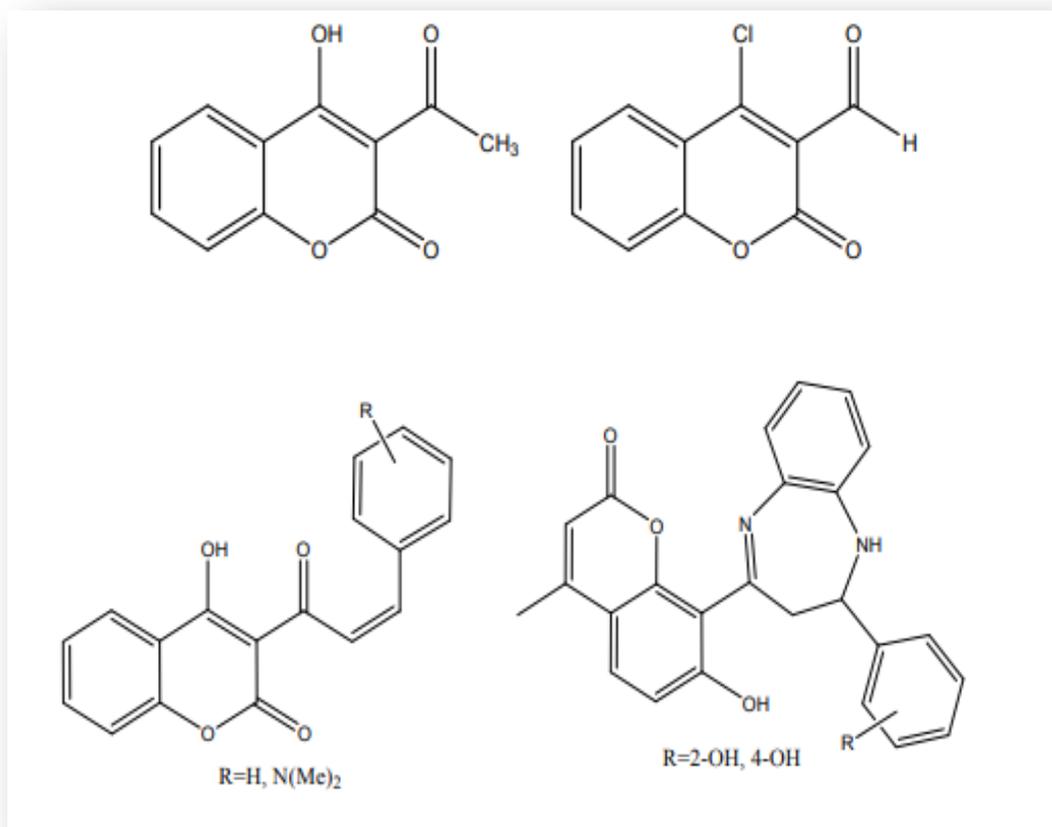
au bois de brûler, à l'huile de rancir, à l'aliment d'altérer ou au fer de rouiller. L'un des principaux acteurs de tout cela est l'oxydation de l'oxygène. [44]

Quand la cellule utilise de l'oxygène, il se passe, un grand nombre de réactions d'oxydation. Le résultat est la production d'énergie, mais aussi de différents sous-produits appelés espèces réactives de l'oxygène « ERO ». [45] L'organisme est équipé pour lutter contre ces ERO par un énorme système de défense constitué de système antioxydant. Cependant, ce système de défense est parfois dépassé, surtout quand les agressions sont multipliées sous l'effet des radicaux libres endogènes et exogènes.

### **7. Activité antioxydante des coumarines :**

Plusieurs coumarines ont montré des profils biochimiques bénéfiques par rapport aux processus physiopathologiques dépendants des espèces réactives de l'oxygène. Des analyses QSAR classiques et tridimensionnelles (3-D) de capteurs de radicaux, structurellement basés sur la coumarine, ont également été réalisées. [52] La prévention des dommages photodynamiques apportés par certaines hydroxycoumarines a été évaluée et comparée à celle de l'acide p-aminobenzoïque (PABA) comme écran solaire modèle. L'activité pourrait être liée à leur action antioxydante qui pourrait minimiser le photovieillissement cutané. [53]

Comme tout composé phénolique, les coumarines, de différents types, possèdent des propriétés très diverses. Ils sont capables de prévenir la peroxydation des lipides membranaires et de capter les radicaux hydroxyles, superoxydes et peroxyles [54]. Plusieurs coumarines simples, ou fusionnées à des chalcones ou à des benzodiazepines..., ont montré une activité antioxydante nettement supérieure ou similaire à celle de la vitamine E ou de la vitamine C. [55,56]



**Figure I. 7 :** Quelques coumarines simples ou fusionnées, à activité antioxydante.

### **Référence:**

- [1] Borges F, Roleira F, Milhazes N, Santana L, Uriarte E. Simple coumarins and analogues in medicinal chemistry: occurrence, synthesis and biological activity, *Current medicinal chemistry*, 12(8), 887-916, 2005.
- [2] Ojala T, *Biological screening of plant coumarins*. 2001.
- [3] Venugopala KN, Rashmi V, Odhav B. Review on natural coumarin lead compounds for their pharmacological activity, *BioMed research international*, 2013.
- [4] Wu L, Wang X, Xu W, Farzaneh F, Xu R, The structure and pharmacological functions of coumarins and their derivatives, *Current medicinal chemistry*, 16(32), 4236-60, 2009.
- [5] Detsi A, Kontogiorgis C, Hadjipavlou-Litina D. Coumarin derivatives: an updated patent review (2015-2016), *Expert opinion on therapeutic patents*, 27(11), 1201-26, 2017.
- [6] Booth, N.L., Dejan, N, Richard, B. et Stacie, E, *Clinical pharmacology and therapeutics*, 76, 511-516, 2004.
- [7] Deiana, M., Rosa, A., Casu, V., Cottiglia, F. et Bonsignore, L, *Journal-of-the American-oil-chemistry-Society*, 80,65-70, 2003.
- [8] Hofmann, L. *Thèse de doctorat: étude du métabolisme desphénylpropanoïdes*, Strasbourg, France, 2003
- [9] Guignard, J.L. *Biochimie végétale*, Masson, Paris, 1996.
- [10] John Roberts D, Marjorie Caserio C. *Chimie Organique Moderne*, Paris: Ediscience , 1968
- [11] Murray RD, *The naturally occurring coumarins*. *Fortschritte der Chemie organischer Naturstoffe/Progress in the Chemistry of Organic Natural Products*: Springer, p. 1-619, 2002.
- [12] Harbone J. B. *Chemicals from Plants Perspectives on Plant Secondary, Products*. London, Imperial, 1999

## *Chapitre I : Les coumarines et l'activité antioxydant*

---

- [13] Merzoug B. Contribution à l'étude phytochimique de deux plantes des familles des apiaceae: *Carum montanum* Coss. Constantine, faculté des Sciences exactes Département de Chimie, 2009
- [14].Neves M, Ferreira Neto P, Silva S, Araújo J, A new assay for detection of furocoumarins in *Brosimum gaudichaudii* Trec by melanin production in actinomycetes. *Revista Brasileira de Farmacognosia*, 12:53-4, 2002
- [15].Asif M. Pharmacological activities and phytochemistry of various plant containing coumarin derivatives. *Current Science*, 1(3):77-90, 2015.
- [16].Riveiro ME, De Kimpe N, Moglioni A, Vazquez R, Monczor F, Shayo C, et al. Coumarins: old compounds with novel promising therapeutic perspectives. *Current medicinal chemistry*, 17(13):1325-38, 2010.
- [17].Tosun A, Özkal N, Baba M, Okuyama T. Pyranocoumarins from *Seseligummiferum* subsp. *corymbosum* growing in Turkey. *Turkish Journal of Chemistry*, 29(3):327-34, 2005.
- [18] Dictionnaire des produits naturels, version 2003.
- [19] Bruneton J ,(pharmacognosie photochimie plantes médicinales), 2 éme edition, Paris (1999).
- [22] Barnard, D.L., Xu, Z.Q., Stowell, V.D., Yuan, H., Smee, D.F., Samy, R., Sidwell, R.W., Nielsen, M.K., Sun, L., Cao, H., Li, A., Quint, C., Deignan, J., Crabb, J et Flavin, M.T. Coumarins and pyranocoumarins, potential novel pharmacophores for inhibition of measles virus replication. *Antiviral Chemistry & Chemotherapy*, 13, 39–59 , (2002).
- [23] Kostova, I. Synthetic and Natural Coumarins as Cytotoxic Agents. *Curr.Med , Chem.-AntiCancer Agents*, 5, 29-46, (2005).
- [24] Sarker, S.D., Nahar, L., Rahman, M.N., Siakalima, Met Kumarasamy, Y. Bioactivity of umbelliprenin, the major component found in the seeds of *Angelica sylvestris* *Ars Pharm*, 46, 35- 41, (2005).
- [25] Kostova, I. Raleva, S. Genova, P et Argirova, R. Structure Activity Relationships of Synthetic Coumarins as HIV-1 Inhibitors, *Bioinorganic Chemistry and Applications*, 68274, 1-9, (2006).

## *Chapitre I : Les coumarines et l'activité antioxydant*

---

- [26] Ochocka, R.J. Rajzer, D. Kowalski, Lamparczyk H. Determination of coumarins from *Chrysanthemum segetum*, By capillary electrophoresis, *Journal of Chromatography A*, 709, 197- 202, (1995).
- [27] Taguchi G, Fujikawa S, Yazawa T, Kodaira R, Hayashida N, Shimosaka M, Okazaki M, Scopoletin uptake from culture medium and accumulation in the vacuoles after conversion to scopolin in 2,4-D-treated tobacco cells, *Plant Science*, 151, 153-161, (2000).
- [28] Ojala, T. Rames, S. Haansu P. Vuorela H. Hiltunen R, Haahtela K. Vuorela P. Antimicrobial activity of some coumarin containing hebal plants growing in Finland, *Journal of Ethnopharmacology*, 2000, 73, 299-305, (2000).
- [29] Chen C.N. Weng M.S. Wu C. Lin J. k, Comparison of radical scavenging activity, cytotoxic effects and apoptosis induction in human melanoma cells, *Food Chemistry*, 1(2), 175- 185, (2004).
- [30] Khan I. Kulkari M.V. Gopal M. Shahabuddin, Synthesis and biological evaluation of novel angularly fused polycyclic coumarins, *Bioorganic and Medicinal Chemistry Letters*, 15, 3584-3587, (2005).
- [31] Thati B. Noble, A. Rowan R. Creaven S.B. Walsh M. Egan d. Kavanagh K., Mechanism of action of coumarin and silver coumarin complexes against the pathogenic yeast *Candida albicans*, *Toxicology in vitro*, 21, 801-808, (2007).
- [32] Stefanova T. Nikolova, N. Michailova, A. Mitov, I. Iancovii. Zlabinger g.I. Neychev H. Enhanced resistance to *Salmonella enteric sero var typhimurium* infection in mice after coumarin treatment, *Microbes and infection*, 9, 7-14, 2007.
- [33] O'NEAL MA, GRIFFIN AC. *Cancer Res*, 17(9), 911-6, (1957).
- [34] Anderson C.M. Hallberg A. Hogberg T. Advances in the development of pharmaceutical antioxidant drug, *Food Chem*, 28, 65-180, (1996).
- [35] Hu S.G. Li L. He X.W. Solid-phase extraction of esculetin from the ash bark of chinese traditional medicine by using molecularly imprinted polymers, *Journal of Chromatography A*, 1062, 31-37, (2005).

## *Chapitre I : Les coumarines et l'activité antioxydant*

---

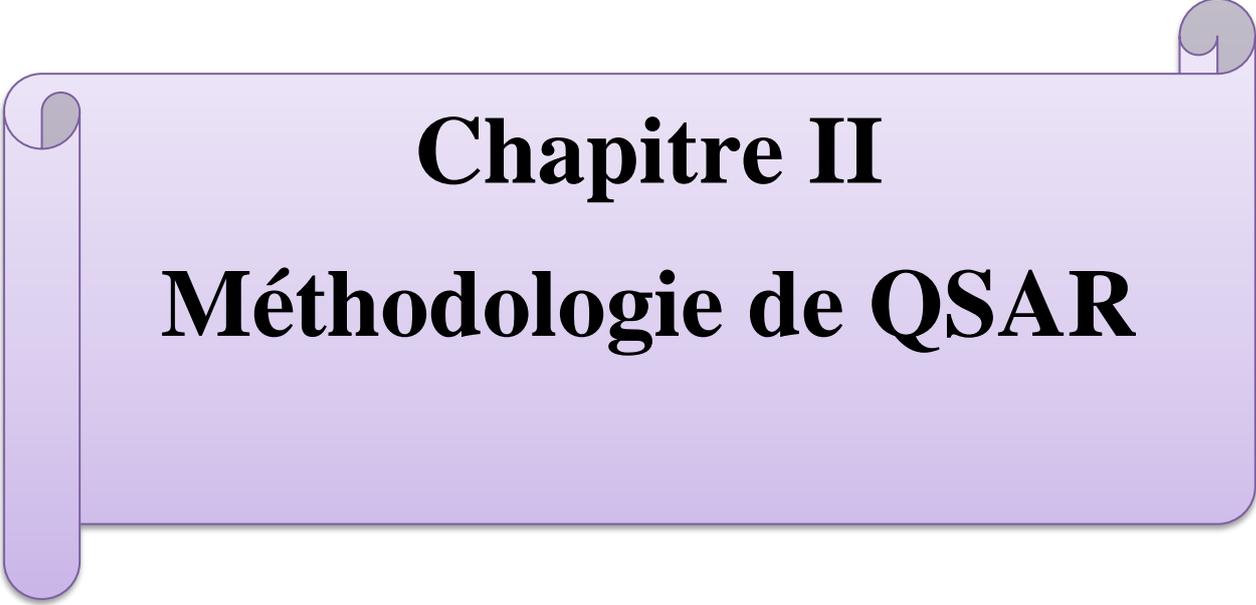
- [36] Bruneton J. Pharmacognosie, phytochimie des plantes médicinales. 2eme edition Tec et Doc (Ed), Paris, p 914, (1993).
- [37] Cottiglia F. Loy G. Garan, D, Floris C. Casu M. Pompei R. Bonsignore L. Antimicrobial evaluation of coumarins and flavonoids from the stems of *Daphne gnidium* L. *Phytomedecine*, 8(4), 302-305, (2001).
- [38] Laure F. Etude de la composition chimique et de la biodiversité du *Calophyllum urophyllum* de Polynésie française. Thèse de Doctorat, Nice, p 373 (2005).
- [39] Redoyal L.M. Beltram, M. Sauchó, R. Olmedo, D.A. *Bioorganic and medicinal chemistry Letters Fitoterapia*, 15, 4447-4450, (2005).
- [40] Veselovskaya, M.V. Garzed M.M. Vinogradova V.I. Khilia P.V. Modified coumarins furocoumarin derivatives of cytisin, *Chemistry of Natural Compound*, 42(3), 277-280, (2006).
- [41] Schorderet M, *Pharmacologie Des concepts fondamentaux aux applications thérapeutiques*, OPU, Alger, p918, (1992).
- [42] Repečak M. Imrich J. Fanekova M. Umbelliferone, a stress metabolite of *Chamomilla recutita* (L) Rauschert. *Plant Physiol*, 158, 1085, 1087, (2001).
- [43] Bruneton, J. *Pharmacognosie, Phytochimie et Plantes médicinales*, 3 ème Edition, Tec et Doc, Paris (1999).
- [44] Le Cren F, *Les antioxydants, la révolution du XXIe siècle*, 2ème édition, (2004).
- [45] Pokorny J, Yanishlieva N, Gordon M, *Antioxidants in food, Practical applications*, Woodhead Publishing Limited, (2001).
- [52] Detsi A, Kontogiorgis C, Hadjipavlou-Litina D. Coumarin derivatives an updated patent review (2015-2016), *Expert opinion on therapeutic patents*, 27(11), 1201-26, 2017.
- [53].Asif M. Pharmacological activities and phytochemistry of various plant containing coumarin derivatives. *Current Science*, 2015, 1(3), 77-90.
- [54] Madhavi D.L. et al, *Food antioxidants*, Ed: CRC PRESS, 361- 460, (1996).

## *Chapitre I : Les coumarines et l'activité antioxydant*

---

[55] Al-Ayed A. S., Synthesis of New Substituted Chromen [4, 3-c] pyrazol-4-ones and Their Antioxidant Activities Molecules, 16, 10292-10302, (2011).

[56] Patil R. B, Sawant S. D., Reddy K. V. Shirsat M. Synthesis, Docking Studies and Evaluation of Antioxidant Activity of Some Chromenone Derivatives, Research Journal of Pharmaceutical, Biological and Chemical Sciences, 6 (2), 381 -391, (2015).



**Chapitre II**  
**Méthodologie de QSAR**

### 1. Histoire de la découverte des médicaments :

Par le passé, un grand nombre de médicaments ont été découverts tout simplement grâce à l'identification de principes actifs extraits de substances naturelles historiquement utilisées dans la médecine non-conventionnelle, ou même par hasard. Par la suite, les avancées dans le domaine de la synthèse chimique et de la pharmacologie ont conduit à une démarche de recherche systématique permettant l'élaboration de nouveaux médicaments de plus grande efficacité [1]. Le processus de recherche et de développement de nouveaux médicaments est extrêmement long et coûteux. Les différentes étapes sont illustrées schématiquement dans la **Figure II. 1**. Au cours des phases successives, des milliers candidats de molécules doivent être triés et sélectionnés, afin d'obtenir un nombre très limité de. Pour chaque nouveau médicament, le temps passé depuis la première étape, au cours de laquelle une cible biologique pertinente dans un processus pathologique donné est identifiée, jusqu'à la mise sur le marché d'un composé tourne autour d'une moyenne d'environ 12 à 14 ans. Le coût global estimé est de 800 millions de dollars par composé mis sur le marché [1–3].



**Figure II. 1** : Etapes du processus de recherche et de développement des médicaments.

Les nouvelles méthodes permettant la découverte de nouveaux médicaments se doivent donc d'innover afin de mettre en évidence des molécules encore inconnues ayant un certain potentiel d'activité sur des cibles biologiques connues [2]. Les outils mis en place doivent être capables de guider les chimistes médicinaux dans le choix des molécules à cribler et à synthétiser [1,4].

#### 1.1. Choix d'une cible thérapeutique :

Le choix d'une cible thérapeutique revient à sélectionner, à l'aide de données issues de la littérature, une biomolécule (en général une protéine ou un complexe protéique) impliquée dans un certain processus pathologique. Certaines familles de protéines sont plus représentées dans le génome humain, ou plus souvent impliquées dans des processus

pathologiques, mais en contrepartie elles peuvent alors posséder un grand nombre de protéines homologues responsables d'effets indésirables. Certaines sont tout simplement plus aisées à cibler que d'autres, en particulier, il est plus facile pour de petites molécules de perturber la liaison d'un ligand endogène que d'avoir des interactions protéine. [5]. Ces paramètres régulent ce que l'on appelle la « druggabilité » de la cible. [6–8]

### 1.2. Identification des touches (hits):

La seconde étape consiste à identifier des touches (hits), c'est à dire des molécules ayant la capacité d'interagir avec la cible choisie et susceptibles de moduler ses effets sur le processus biologique en question. Ces molécules sont sélectionnées à partir de bibliothèques de ligands naturels, ou de chimiothèques, ciblées ou non, le plus souvent à l'aide des techniques de criblage à haut débit (HTS pour High Throughput Screening). Cependant, les tests expérimentaux, qui nécessitent le plus souvent une connaissance approfondie des systèmes étudiés, ne sont pas toujours aisés à mettre en oeuvre, surtout à grande échelle, en automatisant les procédures, sont très coûteux [9]. Ces difficultés ont un impact particulièrement important lorsque l'on se situe dans la phase d'identification des touches (hits) qui nécessite le criblage de milliers de molécules au minimum.

Les stratégies de criblage virtuel sont donc depuis quelques années employées en tant qu'alternative ou de façon complémentaire. Ces techniques sont en général assez faciles à mettre en place, pour un coût bien moindre que les criblages expérimentaux [10]. De plus, l'évolution technologique constante de ces dernières décennies a permis d'accélérer considérablement les temps de calculs nécessaires à la simulation de systèmes complexes ou de bases de données de plusieurs milliers de molécules. Le criblage virtuel est donc aujourd'hui employé dans de nombreux projets, afin de sélectionner, au sein de vastes bibliothèques de molécules, un nombre restreint de composés à cribler expérimentalement. [4].

### 1.3. Génération et optimisation de tête de série (leads) :

A ce niveau, un certain nombre de touches (*hits*) vont être transformés en leads, qui vont ensuite permettre de créer des séries de composés ayant une activité démontrée sur la cible choisie. Le choix porte sur des composés qui possèdent un motif structural de base sur lequel il est possible de greffer un certain nombre de substituant permettant de faire varier les propriétés des ligands formés. Cette étape est souvent considérée comme critique pour la sélection de composés de qualité [11]. Les séries de leads sont ensuite optimisées d'une part afin de présenter une activité vis-à-vis la cible la plus forte possible tout en

conservant une spécificité suffisante vis à vis d'autres protéines.[12], mais d'autre part afin de répondre à un certain nombre de critères régissant les propriétés ADMETox et « drug-like » de la molécule, qui déterminent la capacité de celle-ci à être un bon médicament [8,13,14]. Cette phase correspond à l'expression « drug design » (conception de médicaments) souvent utilisée à tort comme un équivalent de « drug discovery » (découverte de médicaments).

### 1.4 Tests précliniques et cliniques :

Enfin, les phases précliniques et cliniques, qui sont les plus longues et les plus coûteuses du processus de conception de médicament, ont pour but d'évaluer l'efficacité et la sûreté du composé directement à partir d'essais réalisés chez des patients. [3]. Au cours de ce processus, le rôle du chimiste médicinal consiste, en premier temps, à identifier les hits, puis les transformer en leads compounds ainsi que leur optimisation, afin d'obtenir des candidats médicaments susceptibles de devenir un jour un médicament commercialisé. [8].

## 2. Relation s quantitatives structures activités (QSAR):

### 2.1. Historique:

En 1868, Crum-Brown et Fraser. [15] une équation qui est considérée depuis comme la première formulation générale en QSAR. Leurs travaux sur différents alcaloïdes ont montrés que l'alkylation sur un atome d'azote basique, donnant un ammonium quaternaire, faisait varier l'effet biologique de façon significative. Ils en ont déduit que l'activité physiologique  $\Phi$  était fonction de la structure moléculaire chimique C.

$$\Phi = f(C) \quad (1)$$

Plus tard, Richet[16] a découvert que la toxicité de certains composés organiques était inversement proportionnelle à leur solubilité dans l'eau.

$$\Delta\Phi = f(\Delta C) \quad (2)$$

A la fin du 19ème siècle, Meyer. [17] et Overton.[18], ont indépendamment observé une relation linéaire entre la lipophile et l'activité biologique des narcotiques[19].

## Chapitre II : Méthodologie de QSAR

---

Fühner[20] a montré par la suite que l'activité, pour une série de narcotiques homologues, augmentait en fonction de la progression géométrique de la série de composés, ceci montrant l'importance de la contribution d'additivité de groupements fonctionnels pour l'activité biologique.

Les méthodes QSAR ont ensuite rapidement évolué dans les années 1950 avec Bruice, Kharasch et Winzler [21]. En 1962 Hansen [22] a donné une relation de type Hammett, montrant une corrélation entre la toxicité des acides benzoïques substitués et les constantes électroniques  $\sigma$  des substituants.

$$\text{Log } 1/C = 1.454 (\pm 0.42) \sigma + 1.787 (\pm 0.17) \quad (3)$$

L'année 1964 est considérée comme le début des méthodes QSAR modernes. Des travaux importants, de Hansch et al. Et Free *et al.* Apportent indépendamment deux nouvelles méthodes QSAR, qui seront appelées par la suite l'analyse de Hansch[23] . et l'analyse de Free Wilson [24]. La réelle percée de la méthode QSAR résultera de la combinaison des différents paramètres physicochimiques de manière linéaire comme par exemple cette équation :

$$\text{Log } 1/C = a \text{ Log } P + b \sigma + \dots + \text{constante} \quad (4)$$

Depuis 30 ans, on dénombre environ mille équations de Hansch et quelques 500 équations de Free Wilson. Ces méthodes ont connu un succès certain tout en étant limitées par le nombre de variables utilisées pour décrire les composés.

Par la suite, la méthode « partial least squares » (PLS).[25] a permis de traiter de nombreux descripteurs moléculaires. Les méthodes de régression classiques telle que la méthode de régression linéaire doivent tenir compte du fait que le nombre de variables doit être inférieur au nombre d'individus.

Maintenant, des méthodes 3D comme l'étude CoMFA (Comparative Molecular Field Analysis).[26] ou l'analyse des indices de similarité moléculaire avec le logiciel ASP (Automated Similarity Package). [27] permettent de traiter les relations structure-activité en trois dimensions ; c'est du 3D QSAR.

Les molécules biologiquement actives sur un récepteur donné se ressemblent, on parle même de mimétiques pour les dénommer. Il est donc fort utile de pouvoir comparer

de façon quantitative à quel point des molécules peuvent se ressembler ou différer, tant au point de vue stérique, que lipophile et électrostatique. En effet l'adéquation d'une molécule avec un site récepteur dépend du fait que géométriquement la molécule puisse s'insérer dans ce récepteur et qu'électrostatiquement elle vienne le compléter afin de s'y lier. Nous reviendrons de façon plus approfondie sur ces méthodes 3D QSAR par la suite dans ce chapitre.

### **2.2. Définition :**

Les méthodes QSAR sont basées sur l'hypothèse que l'activité ou la propriété d'un composé chimique est liée à sa structure, plus précisément cette approche affirme que l'activité et la structure d'un composé chimique sont liées d'un certain algorithme mathématique, cela est basé sur le postulat de base « les composés chimiques similaires ont des activités similaires ». De plus, lorsque les paramètres moléculaires sont exprimés par des chiffres, on peut proposer une relation mathématique, ou relation quantitative structure activité, entre les deux. Par définition, Une QSAR est un modèle mathématique qui associe un ou plusieurs paramètres quantitatifs dérivés de la structure chimique, à une mesure quantitative d'une activité. [28]

### **2.3. Principe de QSAR :**

Le principe d'une étude QSAR, consiste à trouver une relation mathématique reliant de manière quantitative une activité biologique, mesurée pour une série de composés similaires dans les mêmes conditions expérimentales, avec des descripteurs moléculaires à l'aide des méthodes statistiques. Ces études permettent d'analyser les données structurales afin de détecter les facteurs déterminants pour l'activité étudiée. Pour ce faire, différents types de méthodes statistiques peuvent être employées, les plus répondues sont : la méthode de moindres carrés partielles, la régression linéaire simple et multiple et les réseaux de neurones artificiels. [29]

La modélisation QSAR comporte généralement trois étapes [30] :

- Recueillir ou, si possible, concevoir un ensemble d'information sur la structure et les propriétés des produits chimiques.
- Choisir des descripteurs capables de relier correctement la structure chimique à l'activité biologique.
- Appliquer des méthodes statistiques qui corréleront les changements de structure avec les changements dans l'activité biologique.

### 2.4. Outils et Méthodologie de QSAR :

#### 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 I.1 [31].

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

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

#### 2.4 .2. Les descripteurs moléculaires:

Nous allons présenter les descripteurs moléculaires les plus courants, en

commençant par les descripteurs les plus simples, qui nécessitent peu de connaissances sur la structure moléculaire, mais véhiculent peu d'informations. Nous verrons ensuite comment les progrès de la modélisation moléculaire ont permis d'accéder à la structure 3D de la molécule, et de calculer des descripteurs à partir de cette structure.

**Les descripteurs 1D** sont accessibles à partir de la formule brute de la molécule (par exemple C<sub>6</sub>H<sub>6</sub>O pour le phénol), et décrivent des propriétés globales du composé. Il s'agit par exemple de sa composition, c'est-à-dire les atomes qui le constituent, ou de sa masse molaire. On peut remarquer que ces descripteurs ne permettent pas de distinguer les isomères de constitution.

**Les descripteurs 2D** sont calculés à partir de la formule développée de la molécule. Ils peuvent être de plusieurs types.

– Les **indices constitutionnels** caractérisent les différents composants de la molécule.

Il s'agit par exemple du nombre de liaisons simples ou multiples, du nombre de cycles...

– Les **indices topologiques** peuvent être obtenus à partir de la structure 2D de la molécule, et donnent des informations sur sa taille, sa forme globale et ses ramifications. Les plus fréquemment utilisés sont l'indice de Wiener [32], l'indice de Randić [33], l'indice de connectivité de valence de Kier-Hall [34] et l'indice de Balaban [35]. L'indice de Wiener permet de caractériser le volume moléculaire et la ramification d'une molécule : si l'on appelle distance topologique entre deux atomes le plus petit nombre de liaisons séparant ces deux atomes, l'indice de Wiener est égal à la somme de toutes les distances topologiques entre les différentes paires d'atomes de la molécule. L'indice de Randić est un des descripteurs les plus utilisés ; il peut être interprété comme une mesure de l'aire de la molécule accessible au solvant.

Ces descripteurs 2D reflètent bien les propriétés physiques dans la plupart des cas, mais sont insuffisants pour expliquer de façon satisfaisante certaines propriétés ou activités, telles que les activités biologiques. Des descripteurs, accessibles à partir de la structure 3D des molécules, ont pu être calculés grâce au développement des techniques instrumentales et de nouvelles méthodes théoriques.

Les **descripteurs 3D** d'une molécule sont évalués à partir des positions relatives de ses atomes dans l'espace, et décrivent des caractéristiques plus complexes; leurs calculs nécessitent donc de connaître, le plus souvent par modélisation moléculaire empirique ou *ab initio*, la géométrie 3D de la molécule. Ces descripteurs s'avèrent donc relativement coûteux en temps de calcul, mais apportent davantage d'informations, et sont nécessaires à la modélisation de propriétés ou d'activités qui dépendent de la structure 3D. [36]

### 2.5. QSAR 3D:

Les approches 3D-QSAR ont été développées pour corréler l'activité biologique d'une série de composés actifs de référence avec l'arrangement spatial de nombreuses propriétés de la molécule telles que les propriétés stériques, lipophiliques et électroniques. [37] L'analyse d'un modèle 3D-QSAR permet donc de fournir des indications pour l'optimisation par phamacomodulation et la conception de nouveaux composés avec des profils d'activité améliorés.

La première approche 3D-QSAR, proposée en 1979, décrivait des propriétés de champs moléculaires de composés, calculées sur une grille régulière [38] puis corrélées à leur activité biologique par analyse en composante principale (PCA). Cette méthode, plus tard dénommée DYLOMMS (Dynamic Lattice-Oriented Molecular Modeling System) ne prend en réalité son essor que grâce à l'application de la méthode des moindres carrés partiels (ou Partial Least Squares PLS) à la corrélation des propriétés à l'activité biologique.

Actuellement différentes méthodes 3D-QSAR sont utilisées parmi lesquelles CoMFA (Comparative Molecular Field Analysis), CoMSIA (Comparative Molecular Similarity Indices Analysis).

#### 2.5.1. Analyse comparative des champs moléculaires : CoMFA

L'analyse CoMFA (Comparative Molecular Field Analysis) est considérée comme une approche de corrélation structure-activité. Les caractéristiques de cette méthode sont :

- la représentation des molécules (ligands) par des champs stériques et électrostatiques, les champs ayant l'aspect de treillis tridimensionnel.
- les molécules sont alignées selon une méthode « field fit » et le RMS entre ces molécules sont minimisées.
- les données sont analysées par la méthode PLS.

- des représentations graphiques des parties communes des molécules en terme stérique et électrostatique sont données.

La méthode CoMFA génère une relation linéaire entre les différentes valeurs des champs électrostatiques et stériques environnant les molécules et leurs activités biologiques. Dans la méthodologie de la méthode CoMFA, l'alignement des molécules est très important et il influence beaucoup le résultat de l'analyse. L'analyse CoMFA est beaucoup moins indulgente que la méthode ASP vis-à-vis de l'alignement des molécules au départ.[39]

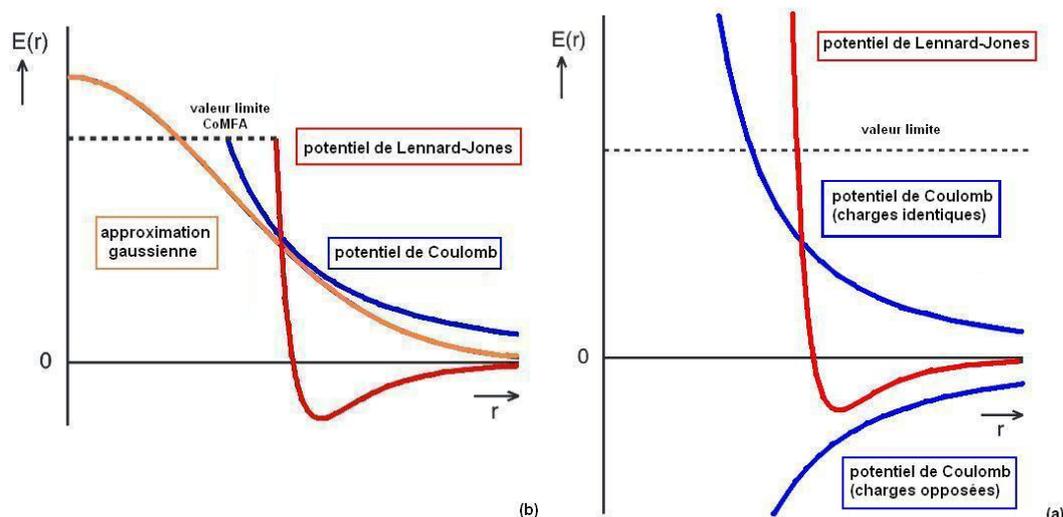
Lors de la comparaison des champs, une valeur de 1 est donnée pour des champs similaires, de -1 pour des champs complémentaires et de 0 pour des champs dissimilaires.

Un champ tenant compte des interactions hydrophobes[40] (hydrophobic interaction HINT) est aussi disponible pour l'analyse CoMFA.

### **2.5.2. Analyse comparative d'indices de similarité moléculaire :**

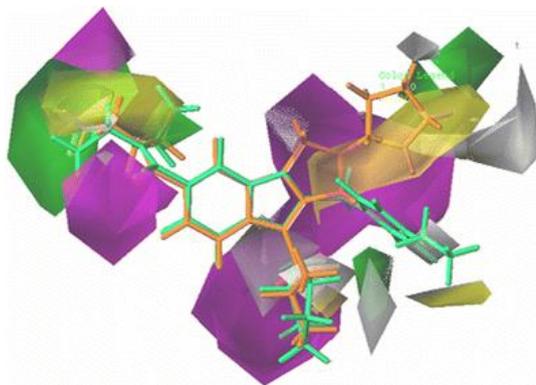
#### **CoMSIA**

Le problème majeur rencontré lors des études CoMFA réside dans la forme (pentes abruptes) des champs électrostatiques mais aussi et surtout stériques calculés par les potentiels de Coulomb et de Lennard-Jones (Figure 4). Ainsi, les variables pour le champ stérique peuvent parfois se réduire à des valeurs proches de 0 (lorsqu'il n'y a aucun atome autour) ou atteignant la valeur limite définie (atomes à l'intérieur de la molécule). Ceci conduit à des aires de contour fragmentées et difficilement interprétables.



**Figure II. 2 :** (a) Courbes des potentiels de Lennard-Jones (rouge) et Coulomb (bleu) utilisées dans les études CoMFA et définition d'une valeur limite supérieure. (b) La fonction gaussienne en forme de cloche des champs SEAL (orange) utilisée lors des études CoMSIA est une bonne approximation des potentiels de Lennard-Jones et Coulomb tout en présentant l'avantage d'être plus lissée. [41] Pour pallier ce problème, Klebe et ses collègues ont l'avantage d'être plus lissée.

Collègues [42] Proposent une méthode CoMFA basée sur des indices de similarité qu'ils nomment CoMSIA. Les indices de similarité stérique, électrostatique et hydrophobe sont obtenus à partir d'une forme modifiée de l'algorithme SEAL (Steric and Electrostatic Alignment) [43]. La fonction gaussienne utilisée pour décrire les molécules représente non seulement une bonne approximation des potentiels de Lennard-Jones et Coulomb (Figure 4) mais permet aussi d'obtenir des aires de contour facilitant l'interprétation des résultats [37]. Tout comme dans les études CoMFA, les molécules sont placées dans une grille cubique et des sondes sont situées sur chaque point de la grille. Les indices de similarité sont calculés à chaque point de la grille, portant principalement sur les propriétés stériques, électrostatiques, d'hydrophobie et de liaisons hydrogène. La corrélation de ces propriétés à l'activité biologique est ici encore représentée par des aires de contour (Figure II. 3).



**Figure II. 3 :** Aires de contour d'un modèle CoMSIA permettant de visualiser la contribution hydrophobique favorable (jaune) ou défavorables (gris), et la contribution des accepteurs de liaisons hydrogène favorable (violet) ou défavorable (vert) à l'activité. [44]

### 2.6. Méthodes statistiques :

Les techniques statistiques ou chimio-métriques constituent la base mathématique de la construction d'un modèle QSAR. La méthode la plus facile à interpréter est l'analyse de régression linéaire parmi diverses méthodes statistiques pour QSAR. Ces régressions représentent la corrélation directe des variables indépendantes (x) avec une variable dépendante (y). Ce modèle peut être considéré pour la prédiction de y à partir des données de variables x. Cela peut appartenir à un ensemble de systèmes qualitatifs ou quantitatifs. [45] :

#### 2.6.1. Régression linéaire simple (SLR):

que variable dépendante. Cette technique s'avère très prometteuse pour générer des relations de structure et d'activité en explorant certains des descripteurs les plus importants utilisés pour gouverner l'activité, alors que certaines des interactions de plusieurs descripteurs ont été négligées. La régression linéaire simple peut être exprimée par l'équation 1 :

$$y = a + bx. \quad (5)$$

Où y est la variable dépendante, x est la variable indépendante, a est la constante b, est le coefficient de régression.

#### 2.6.2. Régression linéaire multiple (MLR):

Cette méthode est l'extension de SLR à plus d'une dimension. Dans cette méthode, des calculs de régression multi variables standard sont effectués.

L'identification d'une propriété de médicament est effectuée sur tous les descripteurs sous enquête. La possibilité de corrélation est vérifiée par la valeur du coefficient de corrélation multiple ( $r$ ), valeur  $t$  à l'aide de la méthode « leave-one-out ». La corrélation est vérifiée par des valeurs  $r^2$  ou  $q^2$  qui sont généralement connues sous le nom de coefficient de corrélation croisé.

Cette méthode est également connue sous le nom de méthode de relation d'énergie libre linéaire (LFER). La relation est exprimée en équation linéaire unique à plusieurs termes comme suit [46] :

$$y = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + \dots + b_mx_m + e. \quad (6)$$

### 2.6.3. L'analyse Partial Least Square (PLS):

L'analyse Partial Least Squares (PLS) est dans son objectif identique à la régression linéaire mais toutes les variables sont prises en compte dans l'équation grâce à une procédure proche de l'analyse factorielle. Cette méthode d'analyse est aujourd'hui la plus utilisée, puisqu'elle permet de prendre en compte un grand nombre de descripteurs. Elle est particulièrement adaptée pour les descripteurs 3D issus d'analyses CoMFA ou d'analyses de similarité moléculaire. L'analyse PLS est une méthode qui utilise des «vecteurs» représentant l'ensemble des variables. Plusieurs critères permettent de valider un modèle. Le premier de ces critères est la «statistical significance». Le second critère est la valeur du PRESS qui doit être la plus faible possible pour le vecteur considéré.

$$\text{PRESS}(y) = \sum (Y_{\text{Théorique}} - Y_{\text{prédite}})^2 \quad (7)$$

Le calcul de ce critère est directement en relation avec la méthode qui détermine un autre critère qui est le «cross validation». Le dernier critère est le  $r^2$  qui donne l'aspect de la linéarité du modèle.

Nous avons utilisé cette méthode statistique dans toutes nos études pour corréler les valeurs biologiques avec les différents descripteurs. [47, 48,49].

### 2.6.4. Analyse par composantes principales (ACP) :

L'analyse en composantes principales (ACP), ou principal component

Analysis(PCA) en anglais, qui est sans doute la technique exploratoire la plus répandue pour décrire les données d'entrée, rechercher d'éventuels aberrations et à connaître les corrélations entre les variables d'entrée. Elle permet d'analyser et de visualiser un jeu de données concernant des individus décrits par plusieurs variables quantitatives. Elle est non supervisée, c'est-à-dire sans phase d'apprentissage [50-52].

C'est une méthode statistique qui permet d'explorer des données dites multivariées (données avec plusieurs variables). Chaque variable pourrait être considérée comme une dimension différente.

Si vous avez plus de 3 variables dans votre jeu de données, il pourrait être très difficile de visualiser les données dans un "hyper-espace" multidimensionnel. L'analyse en composantes principales est utilisée pour extraire et de visualiser les informations importantes contenues dans une table de données multivariées.

L'ACP Synthétise cette information en seulement quelques nouvelles variables appelées composantes principales. Ces nouvelles variables correspondent à une combinaison linéaire des variables originales. Le nombre de composantes principales est inférieur ou égal au nombre de variables d'origine. L'information contenue dans un jeu de données correspond à la variance ou l'inertie totale qu'il contient. L'objectif de l'ACP est d'identifier les directions (axes principaux ou composantes principales) le long desquelles la variation des données est maximale.

En d'autres termes, l'ACP réduit les dimensions de données multivariées à deux ou trois composantes principales, qui peuvent être visualisées graphiquement, en perdant le moins possible d'information. La compréhension des détails de l'ACP nécessite des connaissances de l'algèbre linéaire. [53]

### 2.7. Coefficients et tests statistiques standards :

#### 2.7.1. Coefficient de détermination ( $R^2$ ):

Le coefficient de détermination  $R^2$  est la mesure de degré de liaison entre  $Y_N$  et  $X_j$ .  $R^2$  déterminé par différentes relations suivantes :

$$R^2 = \frac{ESS}{TSS} = \frac{TSS-RSS}{TSS} = 1 - \frac{RSS}{TSS} \quad (8)$$

Un bon ajustement correspondra à un  $R^2$  proche de l'unité.

### 2.7.2. Coefficient de corrélation ( $R^2$ ):

C'est le coefficient de corrélation de Bravais Pearson entre  $Y$  et  $Y'$ , c'est à dire entre valeurs observées et prédites par le modèle de régression, il est noté  $R$ , sa valeur variée entre 0 et 1. [54]

### 2.7. 3. Limitation du coefficient de détermination $R^2$ :

$R^2$  qui varie entre 0 et 1, mesure la proportion de variation totale de  $Y$  autour de la moyenne expliquée par la régression. Plus la valeur de  $R^2$  sera proche de 1 (cas idéal) et plus les valeurs prédites et observées sont corrélées. Un  $R^2$  faible signifie que le modèle a un faible pouvoir explicatif et les descripteurs (certains d'eux) sont sans effet sur la réponse.

Le jugement sur la valeur de  $R^2$  est très subjectif. Bien que ce coefficient soit très facile à comprendre, il faut se garder d'y attaché trop d'importance car il est loin de fournir un critère suffisant pour juger de la qualité d'une régression. Il n'est pas recommandé d'utiliser  $R^2$  pour comparer des modèles avec un nombre différent de descripteurs, le coefficient  $R^2$  nous dira toujours de choisir le modèle avec le plus grand nombre de descripteurs car son  $R^2$  sera plus important (on projette sur un espace plus grand), même si les variables sont sans effets sur la réponse  $Y$ . [55]

### 2.7.4. Test Fisher-Snedecor (F):

Le test Fisher permet de justifier la liaison globale entre  $Y_n$  et  $X_n$ , une version dérivée de cet indicateur peut juger du degré de pertinence des variables du modèle. Il s'agit de vérifier pour chaque variable  $X_j$  si, lorsqu'on passe du modèle complet à  $p$  prédicateurs au modèle simplifié obtenu en  $t$  mesure en fait l'apport marginal de variable  $X_j$  à l'explication de  $Y_n$ .

On peut définir quelques paramètres utilisés dans la régression.

$$\text{Somme des carrés totaux : } TSS = \sum (Y_{\text{obs}} - \bar{Y})^2 \quad (9)$$

$$\text{Somme des carrés expliqués : } ESS = \sum (Y_{\text{cal}} - \bar{Y})^2 \quad (10)$$

$$\text{Somme des carrée résiduels : } RSS = \sum (Y_{\text{obs}} - Y_{\text{cal}})^2 \quad (11)$$

Ainsi,  $TSS = ESS + RSS$

F peut s'écrire comme suit :

$$F = \frac{ESS}{P} \frac{n-p-1}{RSS} \quad (12)$$

La forme de l'équation (12) représente le nombre de degrés de liberté associé

avec chaque paramètre. Le *ESS* associé avec  $p$  degrés de liberté et le *RSS* associé  $n-p-1$  degrés de liberté.

Le test de Fisher mesure le rapport entre la variance de la variable dépendante expliquée et non expliquée par le modèle de régression. En d'autres termes le test Fisher permet de tester l'hypothèse nulle selon laquelle chaque  $b$  est significativement différent de zéro, ce qui est signe d'une relation évidente entre la variable expliquée et les variables explicatives.

Intuitivement, nous rejeterons l'hypothèse nulle lorsque la somme des carrés expliquée par la régression est grande. En d'autres termes, la région critique de ce test est de la forme ( $F > \text{seuil}$ ). Si la quantité  $F$  observée dépasse le seuil, on rejette l'hypothèse  $H_0$  dans le cas contraire, on conserve  $H_0$ .

Pour éviter de raisonner sur  $F$ , le programme fournit la  $p$ -value associée au  $F$  observe. Le  $p$ -value est le niveau de significativité du test de Fisher-Snedecor, c'est-à-dire la probabilité de dépasser le  $F$  observe si l'hypothèse nulle est vraie. On compare la  $p$ -value au risque  $\alpha$  choisi (par exemple  $\alpha=0.05$ ).

### 2.7.5. 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 [54]:

$$Q^2 = 1 - \frac{PRESS}{SSY}. \quad (13)$$

### 2.8. Validation du modèle :

Après l'obtention de l'équation du modèle, outre la stabilité et la qualité de l'ajustement du modèle, il est également important d'estimer la puissance et la validité du modèle avant de l'utiliser pour prédire l'activité biologique. La validité consiste à établir la fiabilité et la signification de la méthode pour un usage particulier. Par conséquent, la validation d'un modèle QSAR doit être effectuée [56].

Il existe trois types de validation : validation interne / validation externe / Coefficients et tests statistiques standards.

### 2.8.1. Validation interne :

La méthode de la validation croisée, une méthode de validation interne, est la plus utilisée dans la plupart du temps dans QSAR. Elle est réalisée par différentes procédures : Leave- One-Out et Leave-n-Out où « n » représente le nombre de composés éliminés.

#### 2.8.1. a. La procédure Leave-n-Out:

Correspond à un découpage en plusieurs parties de l'ensemble de données. A tour de rôle, une partie de l'ensemble de données est attribuée pour un ensemble de test interne. Les autres constituent l'ensemble d'entraînement. Les molécules de chaque groupe éliminé sont prédites par le modèle formé. Cette procédure est répétée p fois pour prédire les propriétés de toutes les molécules, p est le nombre des groupes de molécules éliminés.

#### 2.8.1. b. La procédure Leave-One-Out:

Retire successivement une molécule de l'ensemble d'apprentissage (TSET). Un modèle QSAR est construit alors sur un ensemble de m-1 de composés et la molécule retirée est prédite par le modèle formé, « m » représente le nombre de molécules utilisées dans la construction du modèle (TSET). Cette procédure est répétée m fois afin de prédire les propriétés de toutes les molécules [57].

La performance des modèles de régression est estimée avec les paramètres statistiques de la validation croisée  $S_{cv-loo}$  (l'écart type de la validation croisée), et  $Q^2_{cv-loo}$  (coefficient de détermination de validation, il représente la capacité de la prévision du modèle QSAR).

L'écart type de la validation croisée SCV-LOO:

$$S_{CV-LOO} = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - y_{icv})^2}$$

$Q^2_{cv-loo}$  est calculé à partir de l'équation suivante :

$$Q_{cv-loo}^2 = \frac{\sum_{i=1}^N (y_i - \bar{y})^2 - \sum_{i=1}^N (y_i - y_{icv})^2}{\sum_{i=1}^N (y_i - \bar{y})^2}$$

Avec :

$y_{icv}$  est la valeur prévue de la propriété  $y$  par la méthode de la validation croisée.  $\bar{y}$  est la moyenne arithmétique des  $y$ .

$\bar{y}$  : La valeur moyenne de l'activité.

$y_i$  : Les valeurs observées de l'activité.

$Q_{cv-loo}^2$  est employé en tant qu'outil de diagnostic pour évaluer le pouvoir prédictif ou la qualité de la prévision d'un modèle.

### 2.8.2. Validation externe :

Cette méthode consiste à prédire la propriété/activité d'une série de molécules appelée généralement série de test qui ne sont pas dans la série de développement du modèle, cette validation est caractérisée par les paramètres  $R^2$  (test)  $R^2_{cv}$  (test).

Récemment plusieurs études [58, 59] ont montré l'insuffisance des paramètres  $R^2$ ,  $R^2_{cv}$  pour vérifier le pouvoir prédictif des modèles QSAR. Par conséquent, d'autres paramètres doivent être vérifiés pour cet objectif. Ces paramètres sont connus sous le nom « critères de validation externe » ou souvent appelés « critères de Tropsha » (Tropsha criteria). [58]

Critères de validation Externe (série de test):

- $R^2 > 0.6$  (critère 1)
- $R^2_{cv} > 0.6$  (critère 2)
- $\frac{R^2 - R^2_o}{R^2} < 0.1$  et  $0.85 \leq K \leq 1.15$  (critère 3)
- $\frac{R^2 - R'^2_o}{R^2} < 0.1$  et  $0.85 \leq K' \leq 1.15$  (critère 4)
- $|R^2 - R^2_o| \leq 0.3$  (critère 5)

Avec

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

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

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

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

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

### **2.9. Application de QSAR :**

La capacité de prédire une activité biologique est précieuse dans un grand nombre d'industries.

Alors que certaines SAR semblent être un peu plus que des études universitaires, il existe un grand nombre d'applications de ces modèles au sein de l'industrie, du milieu universitaire et des organismes gouvernementaux (de réglementation).

Un petit nombre d'utilisations potentielles sont énumérées ci-dessous [60] :

- L'identification rationnelle des nouvelles pistes avec une activité pharmacologique, biocide ou pesticide.
- L'optimisation de l'activité pharmacologique, biocide ou pesticide.
- La conception rationnelle de nombreux autres produits tels que les agents tensio-actifs, les parfums, les colorants et les produits chimiques fins.
- L'identification des composés dangereux aux premiers stades du développement du produit ou le criblage des inventaires des composés existants.
- La conception de la toxicité et des effets secondaires dans les nouveaux composés.
- La prédiction de la toxicité pour l'homme par une exposition délibérée, occasionnelle et professionnelle.
- La prédiction de la toxicité pour les espèces environnementales.
- La sélection de composés ayant des propriétés pharmacocinétiques optimales, que ce soit la stabilité ou la disponibilité dans les systèmes biologiques.
- La prédiction d'une variété de propriétés physico-chimiques des molécules (qu'il s'agisse de produits pharmaceutiques, de pesticides, de produits personnels, de produits de chimie Fine, etc.).
- La prédiction du devenir des molécules libérées dans l'environnement.
- La rationalisation et la prédiction des effets combinés des molécules, que ce soit dans des mélanges ou des formulations.

### Références:

- [1] J.G. Lombardino, J.A. III Lowe, THE ROLE OF THE MEDICINAL CHEMIST IN DRUG DISCOVERY — THEN AND NOW, *Nat Rev Drug Discov*, (2004), 3, 853– 862.
- [2] J.A. Dimasi, R.W. Hansen, H.G. Grabowski, The price of innovation: new estimates of drug development costs, *J Heal. Econ*, (2003), 22, 151–185.
- [3] R.E. Hubbard, 3D structure and the drug-discovery process, *Mol. BioSyst.* (2005), 1, 391–406.
- [4] W.P. Walters, M.T. Stahl, M.A. Murcko, Virtual screening – an overview, *Drug Discov. Today*, (1998), 3, 160–178.
- [5] K.H. Bleicher, H. Bohm, K. Muller, A.I. Alanine, HIT AND LEAD GENERATION: BEYOND HIGH-THROUGHPUT SCREENING, (2003), 2, 369–378.
- [6] A.C. Cheng, R.G. Coleman, K.T. Smyth, Q. Cao, P. Soulard, D.R. Caffrey, A.C. Salzberg, E.S. Huang, Structure-based maximal affinity model predicts small-molecule druggability, *Computational-Biology*, (2007), 25, 71–75.
- [7] J. Owens, Determining druggability, *Nat. Rev. Drug Discov*, (2007), 6, 187–187.
- [8] G. Vistoli, A. Pedretti, B. Testa, Assessing drug-likeness – what are we missing, *Drug Discov. Today*, (2008), 13, 285–294.
- [9] T. Tuccinardi, Docking-based virtual screening: recent developments. *Comb Chem High Throughput Screen* , (2009), 12, 303–14.
- [10] N. Moitessier, P. Englebienne, D. Lee, J. Lawandi, C.R. Corbeil, Towards the development of universal , fast and highly accurate docking / scoring methods : a long way to go, *Br. J. Pharmacol*, (2008), 153, 7–26.
- [11] T.I. Oprea, H. Matter, Integrating virtual screening in lead discovery, *Curr. Opin. Chem. Biol.* (2004), 8, 349–358.
- [12] M.J. Keiser, V. Setola, J.J. Irwin, C. Laggner, A.I. Abbas, S.J. Hufeisen, N.H. Jensen, M.B. Kujjer, R.C. Matos, T.B. Tran, R. Whaley, R. a Glennon, J. Hert, K.L.H. Thomas, D.D. Edwards, B.K. Shoichet, B.L. Roth, Predicting new molecular targets for known drugs, *Nature*, (2009), 462, 175–181.
- [13] C.A. Lipinski, F. Lombardo, B.W. Dominy, P.J. Feeney, Experimental and computational approaches to estimate solubility and permeability in drug discovery and developmental settings, *Adv. Drug Deliv. Rev*, (1997), 23, 3–25.
- [14] P.D. Leeson, B. Springthorpe, The influence of drug-like concepts on decision-making in medicinal chemistry, *Nat. Rev. Drug Discov*, (2007), 6, 881–890.

## *Chapitre II : Méthodologie de QSAR*

---

- [15] Crum Brown, A.; Fraser, T.R. *Trans. Roy. Soc. Edinburgh* 1868-69, 151-203, 693-739.
- [16] Richet, M.C. *Compt. Rend. Soc. Biol. (Paris)* 1893, 45, 775-776.
- [17] Meyer, H. *Arch. Exp. Path. Pharm.* 1899, 42, 109-118.
- [18] Overton, E. *Studien über die Narkose, zugleich ein Beitrag zur allgemeinen Pharmakologie*, Fischer, G., Jena, 1901.
- [19] a) Lipnick, R. L. *Trends Pharmacol. Sci.* 1986, 7,161-164.  
b) Lipnick, R. L. *Trends Pharmacol. Sci.* 1989, 10, 265-269.
- [20] Fühner, H.; Neubauer, E. *Arch. Exp. Path. Pharm.* 1907, 56, 333-345.
- [21] Bruice, T.C.; Kharasch, N.; Winzler, R.J. *Arch. Biochem. Biophys*, 1956, 62, 305-317.
- [22] Hansen, O.R. *Acta Chem. Scand.* 1962, 16, 1593-1600.
- [23] Hansch, C.; Fujita, T. *J. Am. Chem. Soc.* 1964, 86, 1616-1626.
- [24] Free Jr, S.M.; Wilson, J.W. *J. Med. Chem.* 1964, 7, 395-399.
- [25] Wold, S.; Ruhe, A.; Wold, H.; Dunn III, W.J. *SIAM J. Sci. Stat. Comput.* 1984, 5, 735-743.
- [26] Cramer III, R. D. Patterson, D.E. Bunce, J.D. *J. Am. Chem. Soc.* 1988, 110, 5959-5967.
- [24] ASP V3.0 Oxford Molecular Ltd, The magdalen Centre, Oxford Science Park, Sandford on Thames, Oxford OX4 4GA, United Kingdom.
- [28] C. Hansch and E.J. Lien, —Structure-activity relationships in antifungal agents. A survey, *Journal of Medicinal Chemistry*, 14(8), 1971, 653–670
- [29] KEBAIRI Maroua, OMRI Maroua. *Modélisation par les réseaux de neurones artificiels : application QSAR*. Université Mohamed Khider de Biskra.
- [30] Benazzouz Hicham, Khebiza Ayoub. *Relation Structure Activité : Etude Qualitative et Quantitative et Développement de Recherche sur les Coumarines*. Université Abou Bekr Belkaïd Tlemcen.
- [31] C.D. Selassie, *History of quantitative Structure-Activity relationships*", in Book "History of quantitative Structure-Activity relationships", D. J. Abraham, Editor., John Wiley & Sons, Inc., California, 2003.
- [32] Wiener, H. *Structural determination of paraffin boiling points*. *Journal of Chemical Information and Computer Sciences*, 1947, 69, p. 17-20.

## *Chapitre II : Méthodologie de QSAR*

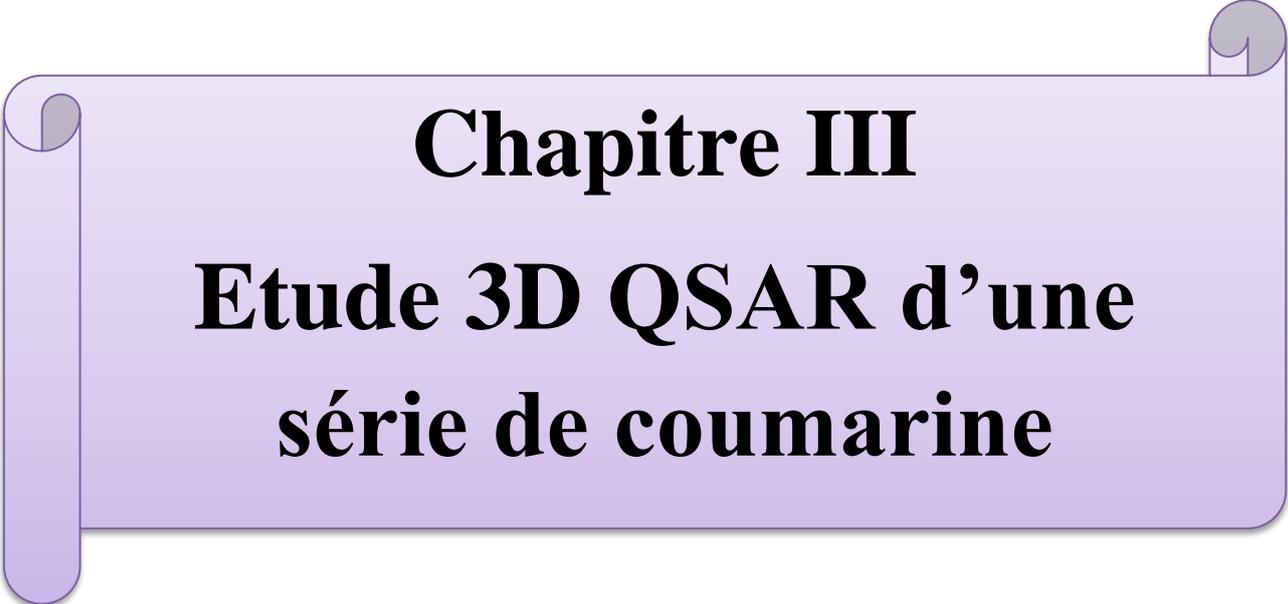
---

- [33] Randić, M On characterization of molecular branching. *Journal of the American Chemical Society*, 1975, 97, p. 6609-6614.
- [34] Kier, L.B, et Hall, L.H. *Molecular connectivity in chemistry and drug research*, New-York, Academic Press, 1976.
- [35] Balaban, A.T. Highly discriminating distance-based topological index, *Chemical Physics Letters*, 1982, 89, p. 399-404.
- [36] Heritage, T.W., et al. EVA: A novel theoretical descriptor for QSAR studies. *Perspectives in Drug Discovery and Design*, 1998, 9-11 (0), p. 381-398.
- [37] Sippl, W. 3D-QSAR - Applications, recent advances, and limitations, in *Recent Advances in QSAR Studies Methods and Applications*, T.L. Puzyn, K.; Cronin, M.T.D Editor. 2010. p. 414.
- [38] Cramer, R.D.; Milne, M., The lattice model: A general paradigm for shape-related structure/activity correlation, in *American Chemical Society Meeting, Computer Chemistry Section* 1979.
- [39] Richards, W. G. The dominant Role of Shape Similarity and Dissimilarity in QSAR, *10th European Symposium on SAR: QSAR and Molecular Modelling*. Barcelone, 4-9 Septembre, 1994.
- [40] Kellogg, G.E. Semus, S.F.; Abraham, D.J. *J. Comput Aided Mol.Design* 1991, 5, 545-552.
- [41] Bohm, H.J.; Klebe, G.; Kubinyi, H. *Wirkstoffdesign*. Spektrum-Akademischer Vlg ed, 1996.
- [42] Klebe, G. Abraham, U. Mietzner, T. Molecular similarity indices in a comparative analysis (CoMSIA) of drug molecules to correlate and predict their biological activity, *J Med Chem*, 1994. 37(24): p. 4130-46.
- [43] Kearsley, S.K. Smith, G.M. An alternative method for the alignment of molecular structures: Maximizing electrostatic and steric overlap, *Tetrahedron Computer Methodology*, 1990. 3(6): p. 615-633.
- [44] Cichero, E. Cesarini, S. Mosti, L. et al. CoMFA and CoMSIA analyses on 1, 2, 3, 4-tetrahydropyrrolo[3,4-b]indole and benzimidazole derivatives as selective CB2 receptor agonists. *J Mol Model*, 2010, 16(9): p. 1481-98.
- [45] Puzyn T, Leszczynski J, Cronin MT. *Recent advances in QSAR studies: methods and Applications* Springer Science & Business Media, 2010.

## Chapitre II : Méthodologie de QSAR

---

- [46] Geladi P, Kowalski BR. Partial least-squares regression: a tutorial. *Analytica chimica acta* 1986, 185, 1-17.
- [47] Stahle, L.; Wold, S. *Progress in Medicinal Chemistry*, ed. Ellis, G.P.; West, G.B., Elsevier, 1988.
- [48] Kubinyi, H. *3D QSAR in Drug Design. Theory, Methods, and Applications*, ESCOM Science Publishers, Leiden, 1993.
- [49]. Dunn III, W.J. Wold, S. Edlund, U.; Hellberg, S. Gasteiger, J. J. *Quant. Struct.-Act, Relat.* 1984, 3, 131-137.
- [50] Abdi H. & Williams L.J. *Principal component analysis*. Wiley interdisciplinary reviews: computational statistics, 2(4), (2010), 433-459.
- [51] Husson, F. Lê, S. & Pagès, J. *Exploratory multivariate analysis by example using R*. CRC press, (2017).
- [52] Jolliffe, I.T. 2002. *Principal Component Analysis*. 2nd ed, New York: Springer-Verlag. <https://goo.gl/SB86SR>.
- [53] TOUHAMI, M. *Modélisation de l'activité biologique de composés hétérocycles*. Le diplôme de Doctorat LMD en « CHIMIE COMPUTATIONNELLE ». Université de TAHAR MOULAY – SAIDA-. Algérie. (2019).
- [54] MEHELLOU Mohammed Nadjib. *Etude des relations structure/activité quantitatives (QSAR/2D) d'une série de dérivés de Triazolothiadiazoles*, mémoire de master. Université ECHAHIDE HAMMA LAKHDER D'EL-OUED
- [55] BELLIFA, K. *Etude des relations quantitatives structure-toxicité des composés chimiques à l'aide des descripteurs moléculaires « Modélisation QSAR »*, Thèse de Doctorat, Université ABOU BEKR BELKAÏD DE TLEMCEN, Algérie. (2015).
- [56] Shao, J. *Linear model selection by cross-validation*, *Journal of the American statistical Association*, 1993, 88(422), 486-94.
- [57] P. Bultinck; H. D. Winter, W. Langenaeker; J. P. Tollenaere: "Computational Medicinal Chemistry for Drug Discovery" Ed, Marcel Dekker, Etats Unies, 2004.
- [58] A. Golbraikh, A. Tropsha, *Beware of  $q_2$ !* *J. Mol. Graph, Model*, 2002, 20, 269–276.
- [59] T. M. Martin, P. Harten, D. M. Young, E.N. Muratov, A. Golbraikh, H. Zhu and A. Tropsha, *Does Rational Selection of Training and Test Sets Improve the Outcome of QSAR Modeling?* *J, Chem. Inf. Model*, 2012, 52, 2570–2578.
- [60] Puzyn T, Leszczynski J, Cronin MT, *Recent advances in QSAR studies methods and applications* Springer Science & Business Media; 201.



**Chapitre III**  
**Etude 3D QSAR d'une**  
**série de coumarine**

### **1. Introduction:**

Le terme 3D-QSAR décrit le développement d'un modèle de régression mathématique entre les caractéristiques structurales 3D des molécules (la variable indépendante notée X) et leurs valeurs d'activité biologique correspondantes (variable dépendante notée Y), par Analyse partielle des moindres carrés (PLS). Le modèle 3D-QSAR généré pourrait être utilisé pour extraire les caractéristiques structurelles vitales des molécules nécessaires pour améliorer leur activité biologique et pour la prédiction de l'activité des molécules nouvellement conçues. [1–3]

Les outils tridimensionnels (3D), comme les analyses de similarité moléculaire ou les études CoMFA (Comparative Molecular Field Analysis) sont aujourd'hui des méthodes très employées pour les études QSAR. Elles permettent de comparer les molécules à travers leur aspect tridimensionnel. Les comparaisons se font sur des critères stériques, électrostatiques, phénomènes responsables des interactions dans les récepteurs. [4,5]

Notre objectif dans cette étude est d'établir un modèle 3D QSAR quantitatif des relations structure-activité antioxydant par une série de coumarine.

### **2. Matériels et méthodes:**

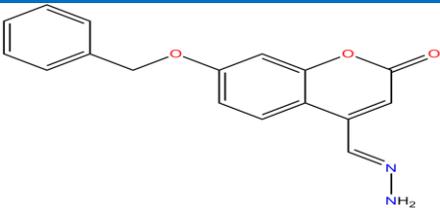
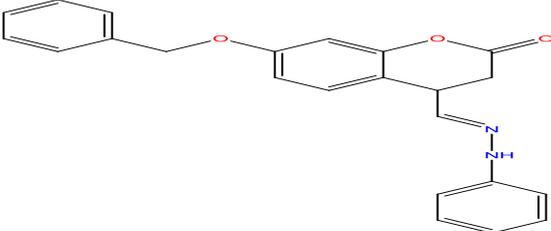
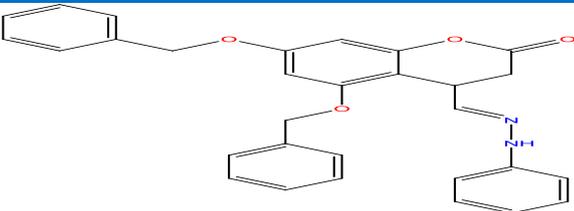
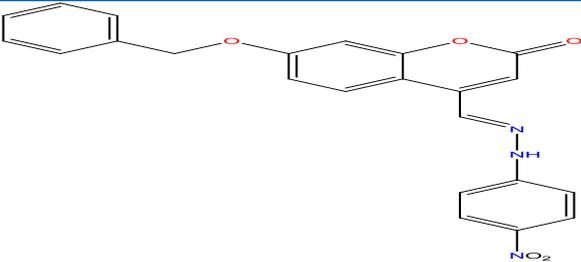
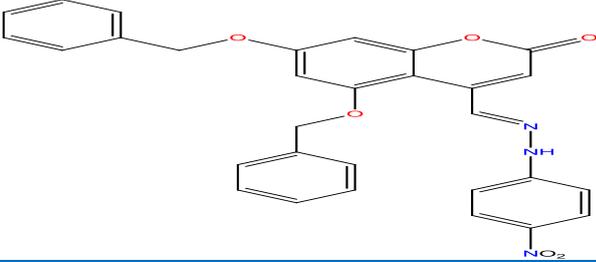
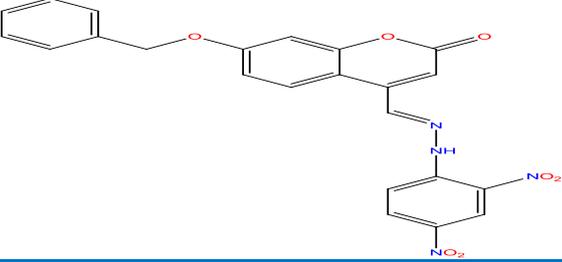
#### **2.1. Base de données:**

Dans la présente étude, un ensemble de données de 22 composés des dérivés de coumarine a été choisi à partir de la littérature avec leurs activités inhibitrices in vitro [6].

L'activité biologique IC<sub>50</sub> ont été converties en pI<sub>50</sub> ( $pI_{50} = -\text{Log}(IC_{50})$ ). L'ensemble de données a été utilisé pour générer le modèle 3D-QSAR.

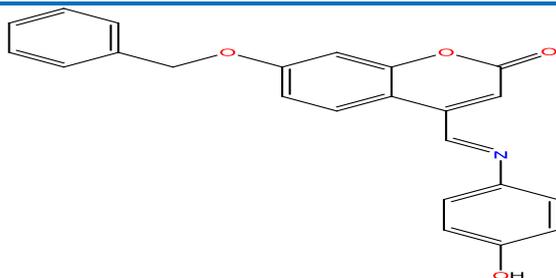
## Chapitre III : Etude 3D QSAR d'une série de coumarine

Tableau III. 1 : Structures chimique de molécules étudiées.

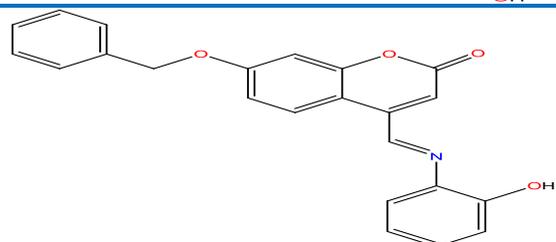
No.	Structures
1	
2	
3	
4	
5	
6	

### Chapitre III : Etude 3D QSAR d'une série de coumarine

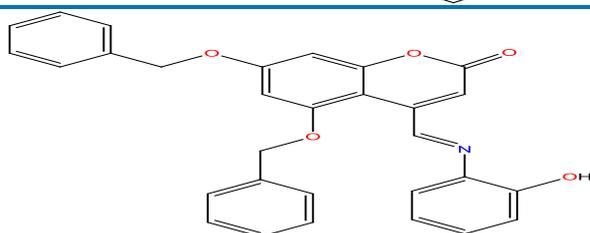
7



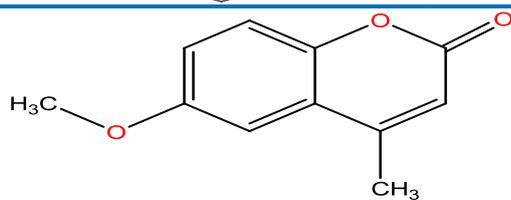
8



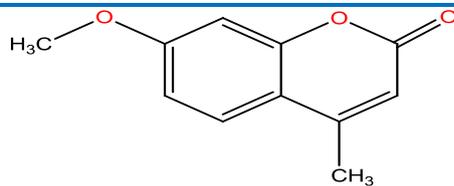
9



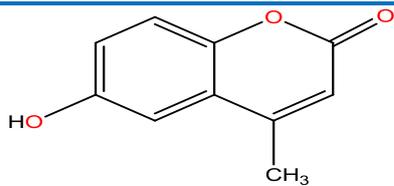
10



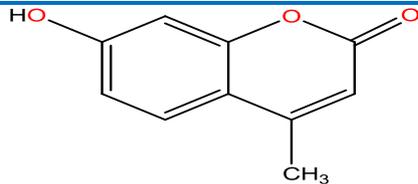
11



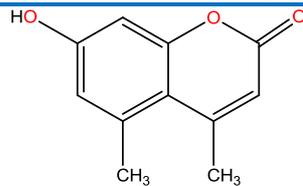
12



13

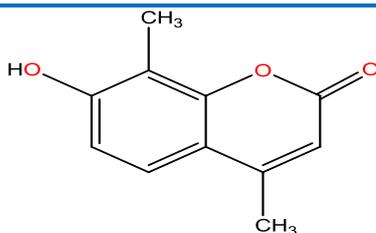


14

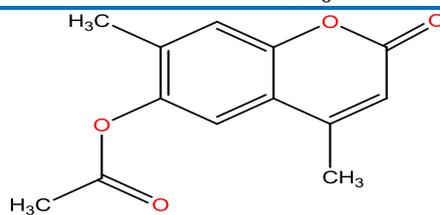


### Chapitre III : Etude 3D QSAR d'une série de coumarine

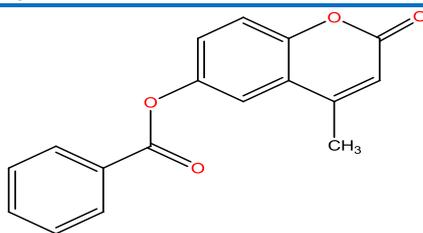
15



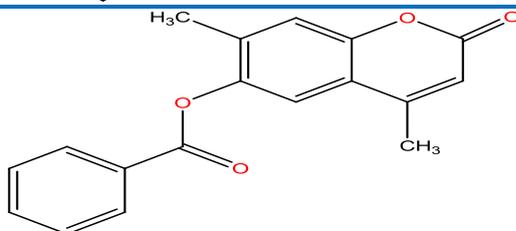
16



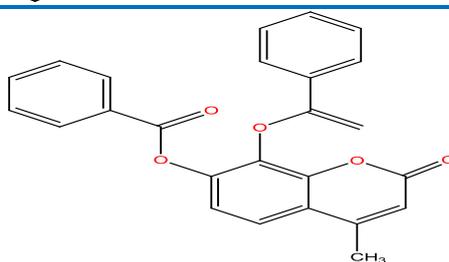
17



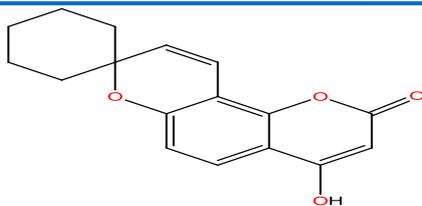
18



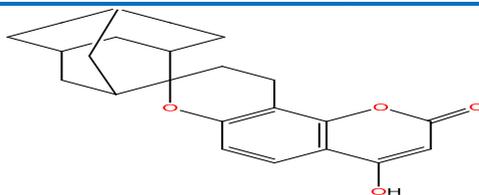
19



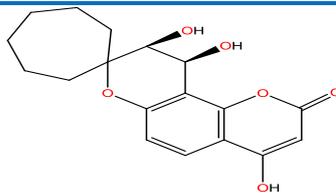
20



21



22



### 2.2. Minimisation et Alignement:

Toutes les structures moléculaires (Tableau 1) ont été construites dans le logiciel SYBYL X-2.0. L'optimisation de la géométrie des molécules a été effectuée à l'aide du champ Tripos avec Powell algorithme de gradient (0.01 kcal / mol) et les charges atomiques partielles ont été calculées en utilisant la méthode Gasteiger Huckel.

La technique d'alignement moléculaire est considérée comme étape les plus importantes pour le développement de modèle CoMFA. Le composé 6 a été sélectionné pour visualiser les cartes de contour en raison de son activité la plus élevée. Les structures superposées de l'ensemble de données alignés sont illustrés dans la Figure III.1.

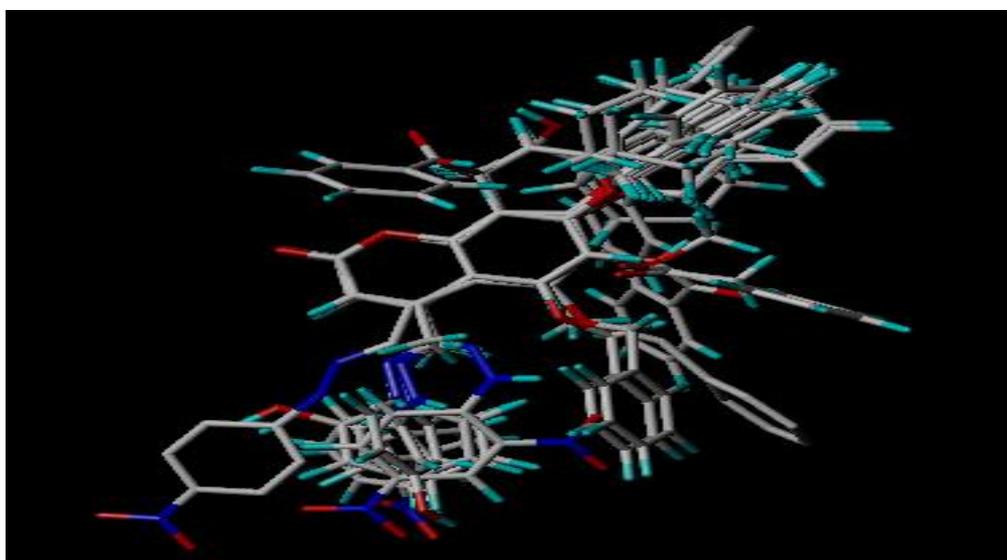


Figure III. 1 : Alignement des dérivés de coumarine

### 2.3. Analyse de CoMFA:

Les molécules alignées ont été placées dans une boîte 3D avec un espacement de grille de 2Å. Ensuite, les champs stérique et électrostatiques ont été calculés à chaque point de la grille avec le champ de force Tripos en utilisant une sonde à atome de carbone avec  $sp^3$  hybridation ( $C_{sp^3}$ ) avec un rayon de Van Der Waal 1.52 et la charge nette 1 a été utilisée comme sonde.

### **2.4. Construction et validation de modèle :**

Le modèle 3D-QSAR a été réalisé par une analyse partielle des moindres carrés (PLS) des descripteurs de type CoMFA. Les descripteurs utilisés sont les champs d'interaction stérique et électrostatique.

L'analyse de modèle 3D-QSAR dépend de paramètres statistiques, comme le coefficient de corrélation à validation croisée ( $Q^2$ ), le coefficient de corrélation non croisée ( $R^2$ ), l'ONC était le nombre de composants, l'erreur standard d'estimation (SEE) et le test de Fisher (F), sont généralement utilisées pour l'évaluation interne d'un modèle. Le modèle favorable devrait avoir des valeurs élevées de  $Q^2$ ,  $R^2$  et F et une faible valeur de SEE.

### **3. Résultats et discussion :**

L'analyse (PLS) est utilisée pour modéliser la relation structure-activité. Cette technique mathématique minimise la différence entre les valeurs réelles et prévues.

La valeur de coefficient de détermination ( $R^2 = 0.85 > 0.6$ ), nous permet d'indiquer fermement la corrélation entre les différents descripteurs utilisés et l'activité biologique dans (PIC50), la valeur de test de Fisher  $F(53.787)$  est élevée et l'erreur standard d'estimation est faible  $SEE = 0.477$ , le nombre optimal de composants principaux ONC (2). Les contributions des champs stérique et électrostatique au modèle CoMFA étaient respectivement de 89.4% et 10.6%.

Pour la validation de modèle on utilise la méthode de validation croisée, est une méthode d'estimation de fiabilité d'un modèle fondé sur une technique d'échantillonnage. Le leave-one-out (LOO) a été utilisé pour cette proposition dans laquelle un composé est retiré du jeu de données et reconstruit le modèle, le coefficient  $Q^2 = 0.718 > 0.5$  cela indique le meilleur modèle PLS prédictif.

### Chapitre III : Etude 3D QSAR d'une série de coumarine

Tableau III. 2 : Paramètres statistiques (PLS) de modèle CoMFA

Statistiques	CoMFA
Q <sup>2</sup>	0.718
R <sup>2</sup>	0.850
SEE	0.477
F	53.787
ONC	2
Répartition des champs (%)	
Champ stérique	89.4%
Champ électrostatique	10.6%.

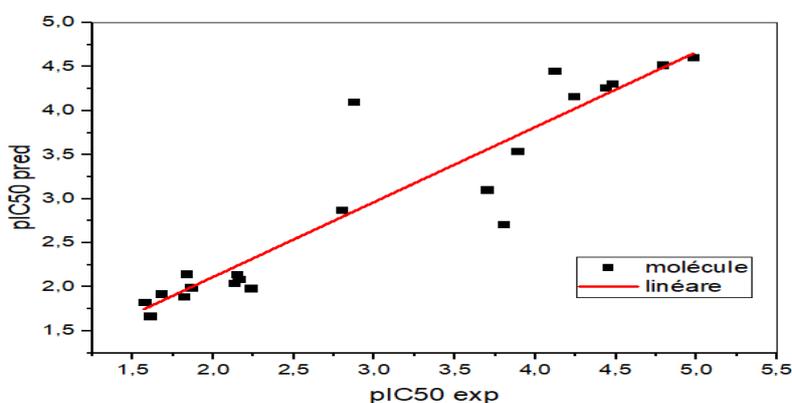
Tableau III. 3 : Activités expérimentales et prédites de l'ensemble de données.

No. Structures	PIC50 exp	PIC50 pré	Résiduel
1	2.980	3.532	-0.552
2	4.239	4.163	0.076
3	4.117	4.453	-0.336
4	2.871	4.103	-1.232
5	2.794	2.875	-0.081
6	4.978	4.607	0.371
7	4.436	4.261	0.175
8	4.478	4.307	0.171
9	4.790	4.522	0.268
10	1.86	1.995	-0.135
11	1.833	2.149	-0.316
12	2.232	1.985	0.247
13	2.128	2.045	0.083
14	2.144	2.144	0.000
15	2.156	2.088	0.068
16	1.815	1.892	-0.077
17	1.604	1.670	-0.066
18	1.572	1.825	-0.253
19	1.672	1.921	-0.249
20	3.697	3.103	0.594
21	3.886	3.541	0.345
22	3.802	2.711	1.091

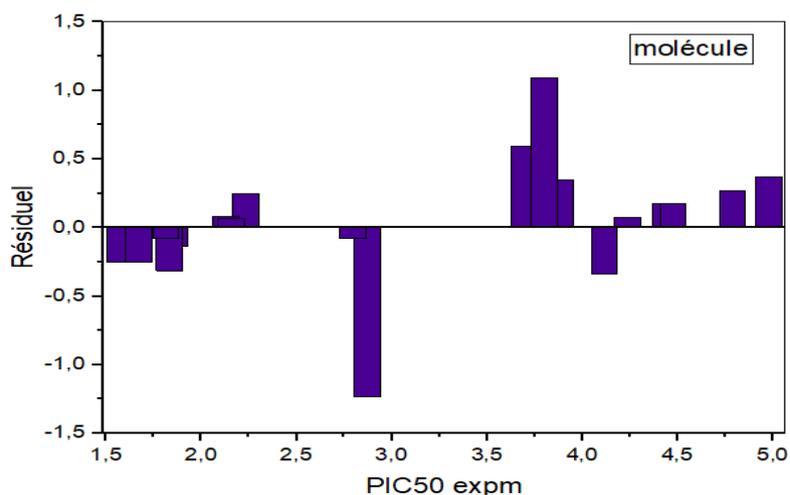
### Chapitre III : Etude 3D QSAR d'une série de coumarine

La Figure III. 2 ci-dessous montre la courbe de corrélation obtenue par PLS entre les valeurs prédites par rapport aux les valeurs expérimentales de l'activité inhibitrice PIC50 des coumarines. La courbe obtenue pour le modèle présente un coefficient de corrélation  $R^2$  (0.85) de valeur significative. Il indique que ce modèle, peut être appliqué avec succès pour prédire l'activité antioxydant.

Les valeurs résiduelles de l'activité biologique (PIC50) ont été tracées en fonction des valeurs expérimentales (Figure III. 3). La répartition des résidus dans les deux côtés de l'axe des abscisses indique l'absence de L'erreur systémique [7]. Il indique que ce modèle peut être appliqué avec succès pour prédire l'activité biologique antioxydant pour des dérivés de coumarine utilisée dans le développement du modèle QSAR.



**Figure III. 2 :** Représentation graphique de corrélation entre les activités expérimentales et prédite des molécules étudiées.



**Figure III. 3 :** Représentation graphique de corrélation entre les activités expérimentales et résiduelles des molécules étudiées.

## *Chapitre III : Etude 3D QSAR d'une série de coumarine*

---

### • Contours CoMFA :

Ce champ est le plus riche en information. C'est celui qui contribue le plus à expliquer l'activité avec un poids relatif de 89,4%.

Afin de mieux comprendre les relations structure-activité de notre série de composés nous allons analyser les graphes de contour SD\*Coefficient. Pour le champ stérique, nous allons visualiser le niveau 20 % qui correspond aux régions défavorables et le niveau 80 % qui correspond aux régions favorables. Pour les champs stérique et électrostatique, nous utilisons les niveaux 20 % et 80 %. Ces niveaux sont représentés avec différents codes de couleurs. . Les cartes des contours stériques et électrostatiques de CoMFA dans la Figures III. 4

#### **a. Le champ stérique:**

Dans le champ stérique, les régions en vert sont favorables à l'encombrement et doivent être occupées alors que les régions jaune sont défavorables à l'encombrement, le composé 9 qui porte le plus long groupement ( $pIC_{50}= 4.790$ ) et le composé 10 qui porte courte groupement ( $pIC_{50}=1.86$ ).

#### **b. Le champ électrostatique:**

Dans le champ électrostatique, les régions en bleu sont favorables au groupement électropositif et les régions en rouge sont favorables au groupement électronégatif.

La région électronégative au niveau des substituants indique que ces substituants renforcent l'activité s'ils sont riches en électrons. Le groupement phényle et les atomes électronégatifs comme l'oxygène, nitrogène sont bénéfiques dans cette région, par exemple l'activité de composé 6 ( $pIC_{50}=4.978$ ) est élevée par rapport au composé 3( $pIC_{50} = 4.117$ ).

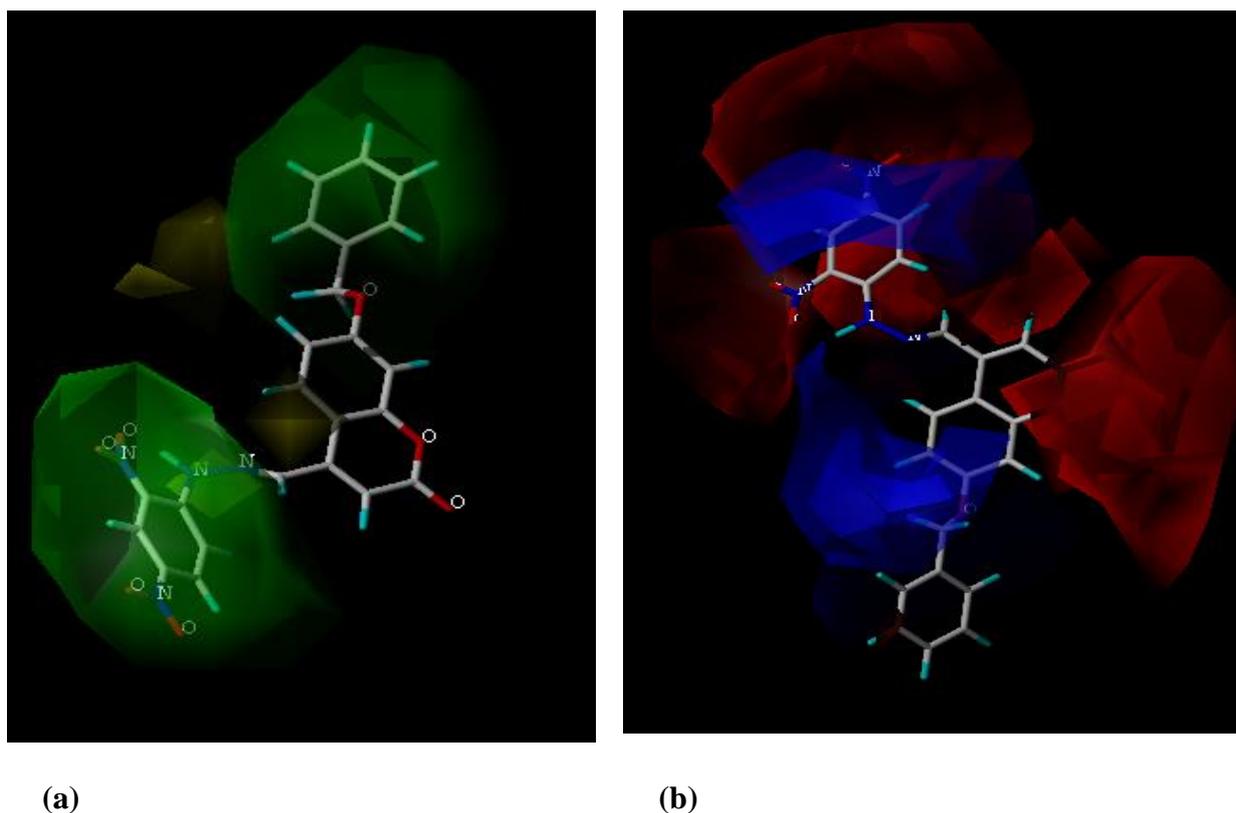


Figure III. 4: Carte de contour des champs stérique (a) et électrostatique (b) de COMFA.

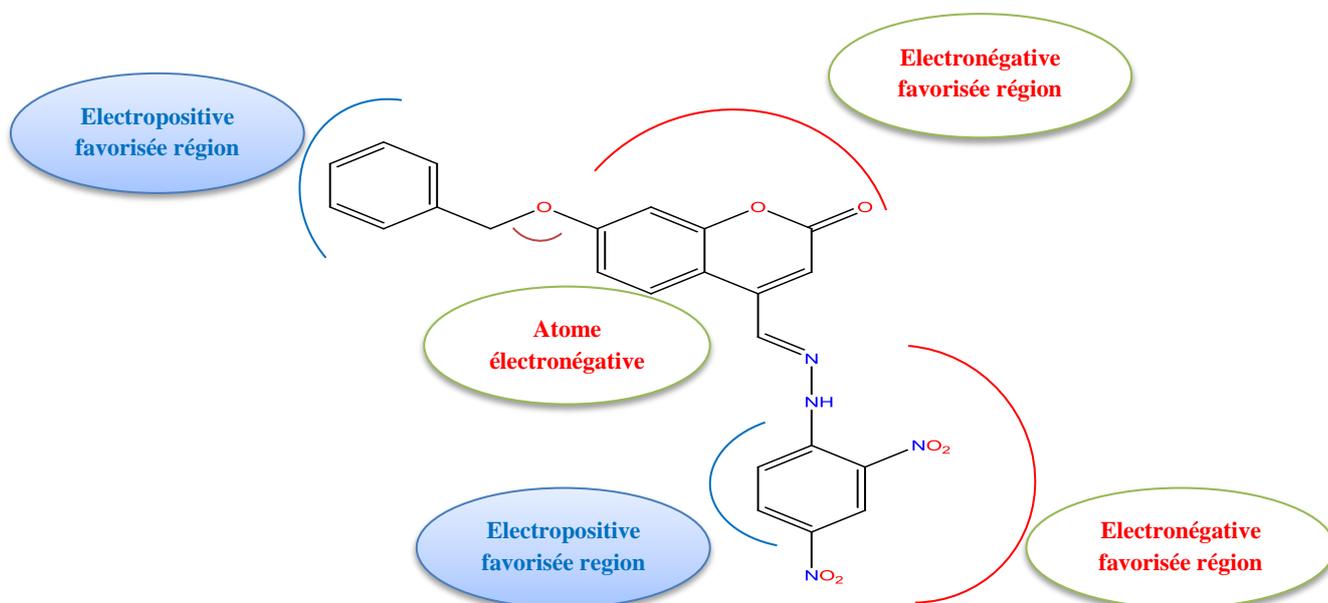
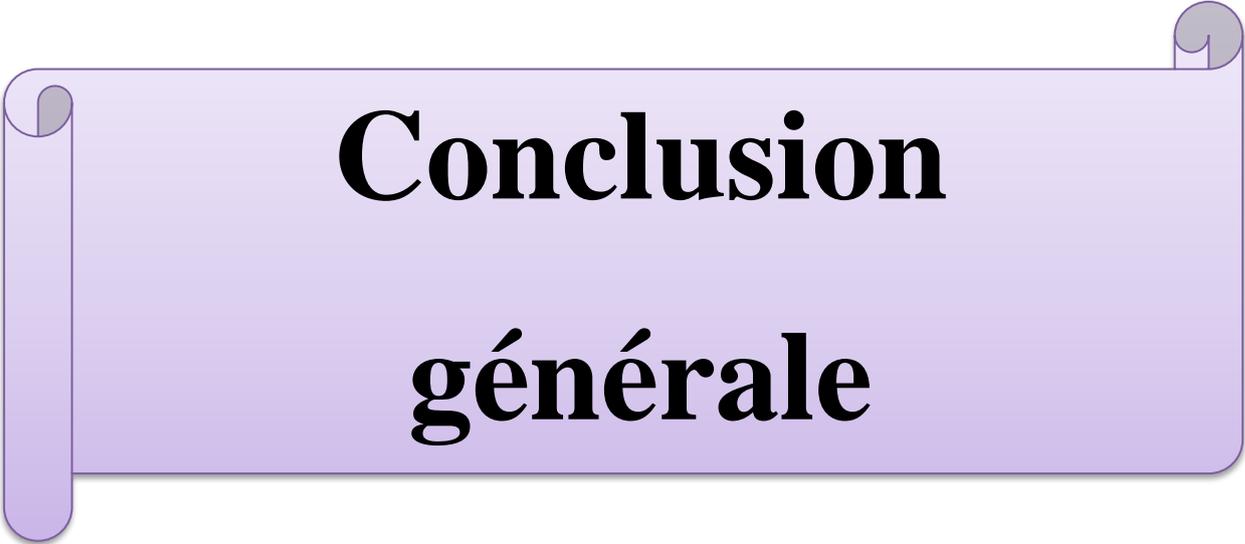


Figure III.5 : Diagramme de relation structure –activité des coumarines orientées vers l'activité antioxydant.

#### **Références:**

- [1] PHASE, version 3.4, Schrödinger, LLC, New York, NY.
- [2] S.L. Dixon, A.M. Smondyrev, E.H. Knoll, S.N. Rao, D.E. Shaw, R.A. Friesner, PHASE: a new engine for pharmacophore perception, 3D QSAR model development, And 3D database screening 1, Methodology and preliminary results, *J. Comput Aided Mol. Des.*, 20 (2006) 647–671.
- [3] O. Silakari, S. Chand, M.S. Bahia, Structural Basis of Amino Pyrimidine Derivatives For Inhibitory Activity of PKC- $\theta$ : 3D- QSAR and Molecular Docking Studies, *Mol. Inform.* 31 (2012) 659–668.
- [4] Simon, Z.; Badileuscu, I.; Racovitan, T. *J. Theor. Biol.* **197**, 66, 485.
- [5] Simon, Z.; Dragomir, N.; Plauchithiu, M.G.; Holban, S.; Glatt, H.; Kerek, F. *Eur. J. Med. Chem.* **1980**, 15, 521.
- [6] I. Mitra, A. Saha, and K. Roy," Predictive Modeling of Antioxidant Coumarin Derivatives Using Multiple Approaches: Descriptor-Based QSAR, 3D-Pharmacophore Mapping and HQSAR".81(2012)64-68
- [7] M. Jalali-Heravi and A. Kyani, *J. Chem. Inf. Comput. Sci.*, 44, 1328 (2004).



**Conclusion**  
**générale**

## *Conclusion générale*

---

Notre travail de mémoire a été consacré à la modélisation de la relation quantitative structure activité antioxydant d'une série constituée de 22 composés de coumarine, pour construire de modèle de QSAR fiable, stable, capable de prédire efficacement cette activité.

QSAR se trouve à l'intersection de la chimie, les statistiques et la biologie, pour cette raison les deux premiers chapitres de cette thèse ont été consacrés à la présentation de L'activité anti-oxydante, la méthodologie QSAR/QSPR ainsi tous les outils nécessaires pour la mise en place d'un modèle QSAR.

Une étude de la relation quantitative structure-activité (3D QSAR) est appliquée à un ensemble de 22 des dérivés de coumarine, afin de prédire l'activité biologique exprimée des composés à tester et de trouver une corrélation entre les différents paramètres moléculaires de ces composés et son activité biologique, en utilisant la méthode de moindre carrée PLS dans le développement du modèle QSAR. Les valeurs des paramètres statistiques du modèle :  $R^2= 0,850$  et  $Q^2= 0,718$  montrent que ce dernier est statistiquement significatif possédant un bon pouvoir prédictif et une bonne stabilité. La validation de modèle a été utilisée pour déterminer la qualité statistique et le pouvoir prédictif du QSAR de modèle PLS.

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 des activités biologiques, ce qui indique la validité et la qualité de modèle QSAR obtenus.

En fin de compte, nous pouvons dire que nous avons atteint l'objectif que nous mentionné à travers les résultats obtenus utilisant des techniques informatiques et méthodes statistiques lors du développement des modèles QSAR.

## Résumé

### Résumé:

Dans ce travail, nous avons analysé l'activité de 22 composés de coumarine en utilisant 3D-QSAR pour calculer des descripteurs de type COMFA. Des modèles QSAR ont été développés à l'aide de technique d'analyse de moindres carrés (PLS). afin d'étudier la relation entre la structure moléculaire de coumarine et leur activité biologique, la prédiction des modèles obtenus a été confirmée par validation interne. Il a été observé qu'il existe une forte corrélation entre les valeurs expérimentales et valeurs prédites des activités biologiques, ce qui indique la validité et la qualité des modèles QSAR obtenus.

Mots clés : 3D QSAR, Coumarine, antioxydant.

### Abstract:

In this work, we analyzed the activity of 22 compounds using 3D-QSAR to calculate COMFA type descriptors. QSAR models were Partial Least Square (PLS) technique, in order to studying the relationship between the molecular structure of Coumarin and its biological activities, the prediction of the obtained models was confirmed by internal validation. It was observed that there is a strong correlation between the experimental values and predicts values of biological activities, and this indicates the validity and quality of the obtained QSAR models.

Keywords: 3D QSAR, Coumarin, antioxidant.

### المخلص:

في هذا العمل, قمنا بتحليل نشاط 22 مركب من الكومارين باستخدام 3D-QSAR لحساب واصفات من نوع COMFA تم تطوير نماذج QSAR باستخدام تقنية تحليل المربعات الصغرى الجزئية (PLS), من اجل دراسة العلاقة القائمة بين البنية الجزيئية للكومارين وانشطتها البيولوجية تم التنبؤ بالنماذج التي تم الحصول عليها من خلال التحقق الداخلي. و لوحظ وجود ارتباط قوي بين القيم التجريبية و القيم المتوقعة للأنشطة البيولوجية, وهذا يدل على صحة و نوعية نماذج QSAR التي تم الحصول عليها.

الكلمات المفتاحية: 3D QSAR, كومارين, مضاد اكسدة.